

RESEARCH

Open Access



# Assessment of soil quality for agricultural purposes around the Barapukuria coal mining industrial area, Bangladesh: insights from chemical and multivariate statistical analysis

Md. Mustafizur Rahman, M. Farhad Howladar\* and Mohammed Omar Faruque

## Abstract

**Background:** Barapukuria coal mine which has been operating from 2002 is situated in such a location that is dependent on agriculture. Pollution from Coal mining poses a huge risk to the ecosystem and surrounding soil. So, there could be some impact of the Barapukuria coal mine on the surrounding soil. The main aim of this study is to identify the soil quality of the study area and also see how much the soil is deviated from its standard reference value.

**Results:** From the result it has been identified that pH of the soil is relatively low near the coal mine due to the acid mine drainage. Moreover, Total Nitrogen of the soil samples is also below than the standard value. On the other hand, Iron (Fe), Potassium (K), and Organic matter (OM) and Copper (Cu) are relatively higher than the Standard Reference value for agriculture. So, it is clear that Iron pyrite ( $\text{FeS}_2$ ) and chalcopyrite ( $\text{CuFeS}_2$ ) which are released during mining operations is the pivotal reason for the degradation of the soil of the surrounding area. Various statistical analytical tools have been used in this study for analyzing the data. Correlation matrix and factor analysis show that degradation of soil is encountered because of anthropogenic causes i.e. because of the coal mine. Besides, cluster analysis entails that which soil samples are deviated from their standard condition and also classifies the soil samples according to their homogeneity. Moreover, one way ANOVA identifies the soil quality controlling factors and the spatial variability of soil samples.

**Conclusion:** From the agricultural point of view, this study reveals that there are significant factors involved in the degradation of the soil and this degradation is occurring rapidly. The nutrients in the soil for plant growth such as Nitrogen (N), Potassium (K), Iron (Fe), Copper (Cu) and Organic matter (OM) are deviated from the standard reference value due to mining activities which ultimately affect the annual paddy production of the study area.

**Keywords:** Soil quality, Agriculture, Mining operations, Acid mine drainage, Statistical analysis

## Introduction

Agriculture is the driving force to boost up the economy of a developing country like Bangladesh and soil quality is the most important factor in this regard. Industrialization has an adverse effect on agriculture throughout the whole

world. Mining brought new potential hazards and risks to the environment (Lazareva and Pichler 2007; Othman and Al-Masri 2007; Li et al. 2014). In addition, mining waste land is an inevitable byproduct which caused a great mass of soils being spoiled away (Liu et al. 2003; Li et al. 2014). On the contrary, coal is the most abundant fossil fuel on the earth (Rashid et al. 2014) that comprises about 75% of the total fuel resources (Elliott 1981; Rashid et al. 2014). So, coal mine is necessary to meet the energy

\*Correspondence: farhadpme@gmail.com  
Department of Petroleum & Mining Engineering, Shahjalal University of Science and Technology, Sylhet 3114, Bangladesh

demand of a country and coal mines are being excavated frequently in the world. But coal mining has severe environmental, ecological, and human-health consequences. If not done properly, coal mining has potential to damage landscape, soils, surface water, groundwater, air during all phases of exploration (Martha 2001).

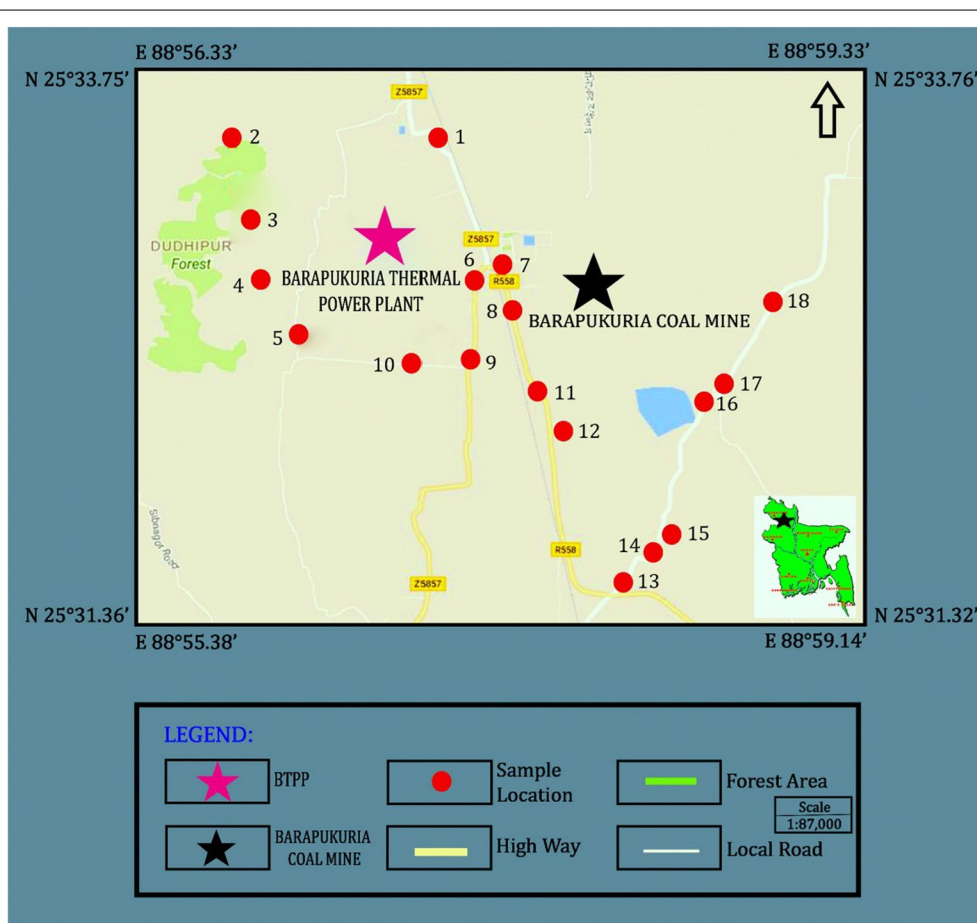
Barapukuria coal mine is the only mine of Bangladesh which has been running since 2002 by using underground multi slice long wall mining method. This coal mine has been extracted 9,240,718.745 MT coal by the year 2016–2017 (Petrobangla 2017). In this mining operation underground sump water is released to the surface without proper treatment. Moreover, there is no proper infrastructure for coal storage so coal dust can get easily mix with the surrounding soil and pollute the soil which ultimately effect the vast paddy land around the Barapukuria coal mine area. However, the impact of the coal mine on the soil quality and paddy production have not been considered previously. Therefore, the main objective of this study is to determine the quality of surface soil samples and find out how much soil is degraded with

respect to the agricultural point of view. This study also entails about the chemical reaction process which are involved in the soil degradation.

#### Location and geology of the study area

Barapukuria coal mine is situated in North-Eastern part of Bangladesh. It is 299.4 km away from the capital of the country, Dhaka. Figure 1 is showing the index map of the collected sample of the study area. The Barapukuria coal deposit has a proved area of 5.25 sq. km. (approx.), with a total reserve of 390 million ton. Long wall multi slices coal mining method has been using in this mine (Imam 2013). Recently, the newly adopted technique Long wall Top Coal Caving (LTCC) is being used in the mine. Figure 2 is showing the coal stock pile and drainage exit location of Barapukuria coal mine. These pictures are clearly indicating about the lackings of proper infrastructure and improper disposal of mine water.

Based on geological framework, Bangladesh is divided into two main divisions namely (Imam 2013): (1) The Precambrian Indian platform and (2) The basin or



**Fig. 1** Index map showing the collected samples location



geosyncline. Aligned from north-east to south-west, a narrow zone between the above a mentioned division is called “the Hinge zone”. Further the Precambrian Indian platform is subdivided into: (1) Rangpur Saddle, and (2) Bogra shelf; whereas the basin or geosyncline is into: (1) Bengal for deep, and (2) Folded Belt (Imam 2013). The Barapukuria coal basin is located in the tectonic unit Rangpur saddle of the stable platform in the northwest Bangladesh. The coal bearing sedimentary rocks of the basin lie unconformably on the Precambrian crystalline basement. The basin is elongated in north south direction and has a length of about 4.5 km and a width of about 1.5 km. It is a half graben type asymmetric basin bounded in the east by a major fault known as Eastern Boundary fault. The surface geology over the entire study area comprises the Tertiary Dupi Tila formation, which unconformably overlays the Gondwana (Permian) coal-bearing sediments. These are folded into an asymmetric syncline or basin, whose axis strikes approximately north-south. A major fault forms the eastern limit of the deposit, beyond which, Archaean basement rocks (Pre-Cambrian) are

present immediately below the Dupi Tila. Several lesser faults were identified within the coal basin by geophysical seismic survey. The Gondwana sequence comprises predominantly sandstones, with subordinate siltstones and mudstones, which contain up to six coal seams in the centre of the basin. The lowest of these, Seam VI is the principal target seam of the Barapukuria coal mine and this seam has an average thickness of 36 m (Imam 2013).

## Methods

### Soil sampling

Eighteen soil samples were randomly collected from the Barapukuria Coal mine area and all soil samples were surface soil samples. Random sampling involves the arbitrary collection of samples within a defined area. Random sampling was done in order to eliminate the tendency of biasing the samples. Samples were collected during summer in the year of 2015. Most of the soil samples were collected from the agricultural lands near the coal mine and some samples were collected little away from the mine area in order to find the natural condition of the



surface soil of the study area. United States Department of Agriculture (USDA) standard of soil sampling for nutrient management was maintained during sampling (USDA 2007).

### Chemical analysis

Each soil sample was reduced to 200 g by quartering. The samples were dried naturally without sun light and then sieved through a 2-mm sieve. Each sample was separated for the measurement of 11 chemical properties, including pH, Organic Matter (OM), Potassium (K), Sulfur (S), Nitrogen (N), Phosphorus (P), Iron (Fe), Arsenic (As), Copper (Cu), Nickel (Ni) and Total alkalinity (as  $\text{CaCO}_3$ ). The pH value of soil sample was determined by using digital pH meter. Organic matter (OM) was determined following potassium chromate method for determination of soil organic matter [NY/T 1121.6-2006 (2006)]. Total Nitrogen was analyzed using Semi-micro Kjeldahl method (Fu et al. 1999; Li et al. 2014). Potassium (K) was determined by Potassium Strip Test, Sulfur (S) was determined by using Ion Chromatography, Iron (Fe), Copper (Cu) and Nickel (Ni) were determined by Spectrophotometer. Arsenic (As) was determined by Hach EZ Arsenic test kit and total alkalinity (as  $\text{CaCO}_3$ ) was determined by Titration Method.

### Correlation matrix and principal component (PCA)/factor analysis

Because of the simplicity of the method, PCA is most popular method all over the World. PCA are widely applied to analyze interrelationship among different sets of groundwater hydro- chemical and soil sample data to extract the most significant factors and to reduce the data with minimum loss of information (Mustapha and Aris 2012; Schaefer and Einax 2010; Howladar and Rahman 2016).

In this research, PCA is applied to soil sample data which are collected from the Barapukuria Coal mine area to determine the principal factors corresponding to the different sources of variation in the data.

In PCA, each manifest variable is a linear function of principal components, with no separate representation of unique variance

$$Y = Z_c K'_c$$

where,  $Z_c$  is a  $N \times p$  matrix of standardized component scores and  $K_c$  is a  $p \times p$  matrix of component loading. PCA is a linear orthogonal transformation of the variables in  $Y$  to principal components in such a way that the first component has the largest possible variance; the second component has the largest possible variance of the remaining data, etc., with the total  $p$  components explaining 100% of the variance.

When  $S$  is a  $p \times p$  sample correlation matrix,  $S$  can be decomposed as follows:

$$S = WMW'$$

with  $W$  is a  $p \times p$  matrix of eigenvectors and  $M$  is a  $p \times p$  diagonal matrix of eigenvalues of  $S$ . The component scores  $Z_c$  are:

$$Z_c = YW$$

The  $p \times p$  matrix of component loadings (i.e., the correlation coefficients between  $Z_c$  and  $Y$ ) is calculated as follows (De Winter and Dodou 2016; Howladar and Rahman 2016):

$$K_c = W.$$

### Cluster

Cluster analysis is a group of multivariate technique whose primary aim is to assemble objects based on the characteristics they possesses (Shrestha and Kazama 2007; Oketola et al. 2013; Howladar et al. 2017). Hierarchical clustering joins the most similar observations. The levels of similarity at which observations are merged are used to construct a dendrogram (Oketola et al. 2013; Howladar et al. 2017). Cluster analysis is the unsupervised classification of patterns (observations, data items, or feature vectors) into groups (clusters). The Euclidean distance ( $E^2$ ) is the geometric between two objects and can be calculated by the given formula:

$$E^2 = (Z_{OM1} - Z_{OM2})^2 + (Z_{K1} - Z_{K2})^2 + (Z_{pH1} - Z_{pH2})^2 + (Z_{AS1} - Z_{AS2})^2 + \dots$$

(Howladar et al. 2017)

### ANOVA

“One way between subject effect analysis of variance” (one-way ANOVA) compares the variance (variability in scores) between the different groups with the variability within each of the group (Rogerson 2010, Howladar and Rahman 2016). The analysis of variance (ANOVA), also known as the F-test, is a method to determine the variation of the means of a group of data or variables to evaluate statistical significance. This method, when comparing two means, is similar to the  $t$  test for independent samples. The single factor ANOVA test assumes a null hypothesis,  $K_0$ , which states there is no difference between the groups within the population, as shown in Equation.

$$K_0: a_1 = a_2 = \dots \dots a_q = 0$$

(modified after Howladar and Rahman 2016).

If the analysis is found to be statistically significant, then the null hypothesis is rejected for the alternative hypothesis. The alternative hypothesis states that the means of the groups in the population are different. For this research, a P-value of  $\leq 0.05$  was used to determine statistical significance. In this research One way ANOVA was performed by using IBM SPSS Statistics 22 software for 18 soil samples.

## Results and discussion

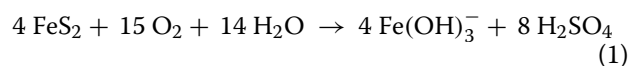
### Chemical analysis results in the context of soil quality

The result of analyzed variables e.g. pH, Total Nitrogen, Organic Matter (OM), Potassium (K), Phosphorus (P), Sulfur (S), Iron (Fe), Arsenic (As), Copper (Cu), Nickel (Ni), and Total alkalinity is summarized in the Table 1. The table entails about the maximum, minimum, average, and standard deviation of each chemical parameter. From the table it is seen that Sulfur (S), Phosphorus (P), and Total alkalinity have relatively high standard deviation. High standard deviation indicates that the data points are spread out over a wider range of values. So the soil samples have a lot of variability regarding Sulfur (S), Phosphorus (P), and Total alkalinity. On the other side, standard deviation of Arsenic and Total Nitrogen are very low that means most of the soil samples contain relatively similar amount of Arsenic (As) and Total Nitrogen.

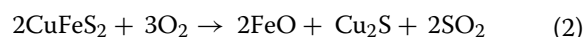
The experimental value of different chemical parameters has been shown in Table 1. Moreover, the standard reference value for each parameter has been presented also in the same table based on Regional Soil Resource Development Institute of Bangladesh (RSRDIB 2015). The table clarify that the mean value of pH, Total N are lower than the Standard reference value. On the other hand, Organic matter (OM), Potassium (K), Iron (Fe) and Copper (Cu) are higher than the Standard reference value

(Fig. 3). Higher value of these chemical parameters is not favorable for the production of paddy.

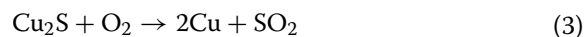
Only Sulfur (S) and Phosphorus (P) have mean value within the optimum limit for soil. This is a clear indication that soil quality of the study area is degraded to a large extent from their standard reference value. Higher value of Iron (Fe) indicates that that Iron pyrite ( $\text{FeS}_2$ ) and chalcocopyrite ( $\text{CuFeS}_2$ ) which are released during coal mining operations are liable for high Iron (Fe) and Copper (Cu) in the soil samples. Again the high Fe quantity of soil make lower the pH value of the soil and that's why pH value of the soil samples in surrounding area of mine found low. The chemical reactions involved for lowering the pH and Copper (Cu), and Iron (Fe) generation are shown below:



Iron pyrites + Oxygen + water = Iron(iii) hydroxide ion + Sulfuric acid



Cu formation:



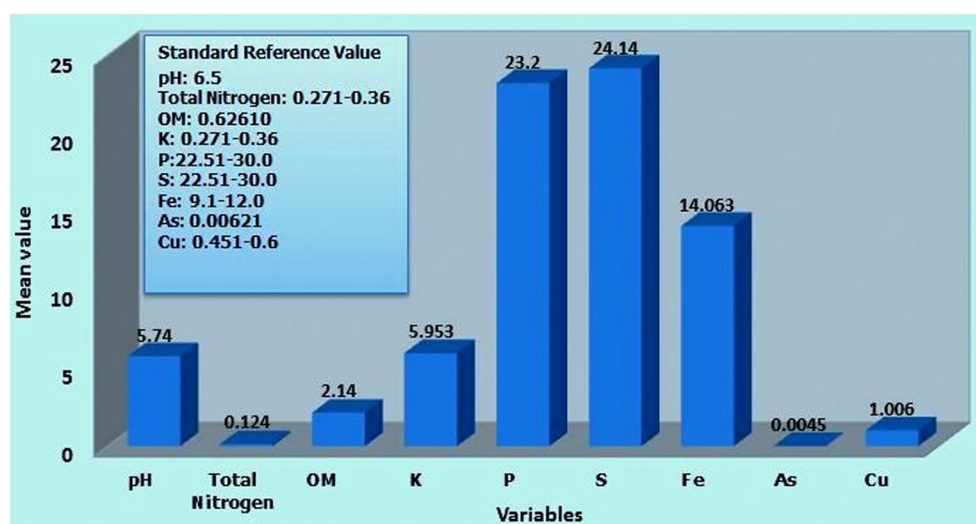
Sulfuric Acid formation from  $\text{SO}_2$ :



Equation (1) indicates the formation of  $\text{H}_2\text{SO}_4$  from Iron pyrites. During mining operation of coal iron pyrites is released and mixed with very much available Oxygen and water and finally form  $\text{H}_2\text{SO}_4$ . Equation (2) indicates why Cu and Fe are much higher in the study area. Chalcocopyrite ( $\text{CuFeS}_2$ ) is also released during coal extraction

**Table 1** The value of different chemical parameters and their max, min, mean, and standard deviation as well as standard reference value for respective parameter

Variables	Max	Min	Mean	Standard deviation	Standard reference value
pH	6.83	4.97	5.74	0.53	6.50
'Total Nitrogen' %	0.21	0.06	0.12	0.04	0.27–0.36
'Organic Matter' %	3.71	1.08	2.14	0.63	0.63
K, ppm	12.81	1.72	5.95	2.45	0.27–0.36
P, ppm	65.29	2.84	23.20	14.55	22.51–30.00
S, ppm	127.75	2.26	24.14	23.48	22.51–30.00
Fe, ppm	24.37	1.27	14.06	7.72	9.10–12.00
As, ppm	0.03	0.00	0.01	0.01	0.01
Cu, ppm	2.14	0.25	1.01	0.44	0.45–0.60
Ni, ppm	1.28	0.10	0.49	0.24	–
Total alkalinity (as $\text{CaCO}_3$ )	150	10	59.67	33.78	–



**Fig. 3** The mean value of the chemical parameters with their Standard reference value

and firstly it reacts with oxygen and forms Copper (I) sulphide, Ferrous oxide and Sulfur di-oxide. Copper (I) sulphide further react with oxygen and remain as Cu metal in the soil. Again,  $\text{SO}_2$  reacts with oxygen and produce  $\text{SO}_3$  and finally  $\text{SO}_3$  reacts with water and produce  $\text{H}_2\text{SO}_4$ .

Coal is stock piled in open location which is very close to the agricultural land. Coal of Barapukuria mine (seam VI) has total sulfur percentage 0.52–0.61% averaging 0.57% (Imam 2013). So, surrounding soil of coal mine is also affected by the S through the precipitation of coal dust in the agricultural land.

#### Application of multivariate statistics to elucidate the spatial differences of soil quality of the study area

##### Correlation matrix and factor analysis

Correlation matrix is used for determining the dependence between multiple variables at the same time. Classification of principal components is thus “strong”, “moderate” and “weak”, corresponding to absolute loading values of  $> 0.75$ ,  $0.75 - 0.50$  and  $0.50 - 0.30$ , respectively (Liu et al. 2003). However, positive correlation between two variables indicate that one variable will increase or decrease with respect to other changes and negative correlation between two variables indicate one variable will increase when other will decrease and vice versa.

Table 2 is showing the correlation matrix coefficient of all soil samples. High positive correlation values have been found for Total Nitrogen: OM, and pH: Alkalinity. Again, high negative correlation values have been found for pH: OM, and OM: Alkalinity. An eigenvalue gives a measure of the significance of the factor and factors with

the highest eigenvalues are the most significant (Howladar et al. 2017). Eigenvalues of 1.0 or greater are considered significant (Shrestha and Kazama 2007; Howladar et al. 2017). From Table 3 it is obvious that Factor 1 has eigenvalue of 2.8, Factor 2 has eigenvalue of 2.5 and Factor 3 has eigenvalue of 1.35. Here from the Table 3 it has been found that three factors describes 61% that means they describes most of the variability so these factors are good and so these factors can be used as the replacement of other factors. Table 4 is showing three factors while loading is obtained with unrotated factor matrix. Factor 1 has high positive and negative values but Factor 2 has positive values for most of the variables and Factor 3 has negative values for most of the variables.

From the Fig. 4 of factor 1 vs. factor 2 it is clear that Total Nitrogen: OM and pH: Alkalinity is contaminated from the same source. Again, factor 2 has dominance over factor 1 which means contamination occurs for anthropogenic causes.

First factor loading explained 25.4% of the total variance. The factor shows strong positive loading for Total N and OM (Table 4). The results of first factor loading provide evidence of natural factor impact in the study area (Belkhiri and Narany 2015). Second factor loading explained 22.9% of the total variance. The factor has positive loadings for pH, K, P, S, Fe, and Total alkalinity. It is clear that the results of Factor analysis satisfy the experimental value of chemical analysis of the soil samples as well as gives the clear indication about the degraded parameters. The results of factor loading show clear evidence of anthropogenic influence in the study area that has been supported by similar scientific results done by Belkhiri and Narany (2015). However, factor 1 has also

**Table 2 Correlation matrix coefficient of 18 soil samples**

Variables	pH	Total nitrogen	OM	K	P	S	Fe	As	Cu	Ni	Total alkalinity as CaCO <sub>3</sub>
pH	1.0										
Total Nitrogen	-0.313	1.0									
OM	-0.496	0.631	1.0								
K	0.001	0.441	0.42	1.0							
P	0.21	0.124	0.351	0.436	1.0						
S	0.136	0.278	0.037	0.213	0.431	1.0					
Fe	0.018	0.047	-0.086	-0.034	0.002	-0.016	1.0				
As	-0.178	-0.052	-0.116	-0.342	-0.438	0.124	-0.236	1.0			
Cu	-0.1	-0.215	0.043	-0.108	-0.192	-0.363	-0.182	0.142	1.0		
Ni	0.146	0.126	-0.04	0.449	0.018	0.112	0.147	-0.09	0.120	1.0	
Total Alkalinity(as CaCO <sub>3</sub> )	0.969	-0.347	-0.552	-0.02	0.214	0.13	-0.059	-0.202	-0.081	0.192	1.0

*Italic values indicate the significant correlation between the various chemical parameters*

**Table 3 Total Variance Explained**

Factor	Eigenvalues	Total variance (%)	Cumulative Eigenvalues	Cumulative variance (%)
1	2.8	25.4	2.8	25.4
2	2.58	22.9	5.38	48.3
3	1.4	12.3	6.7	60.6

**Table 4 Factor loading obtained with unrotated factor Matrix**

Variables	Factor 1	Factor 2	Factor 3
pH	− 0.734044	<b>0.575550</b>	− 0.003379
Total Nitrogen	<b>0.758232</b>	0.222330	0.107436
OM	<b>0.871169</b>	<b>0.729553</b>	− 0.084570
K	0.497128	<b>0.620121</b>	− 0.322719
P	0.229370	0.062053	0.109550
S	0.152452	− 0.394920	0.1397270
Fe	0.005454	<b>0.611775</b>	− 0.121557
As	− 0.068721	− 0.546154	0.390191
Cu	− 0.084452	<b>0.514229</b>	− 0.620729
Ni	0.026929	0.389459	− 0.539698
Total Alkalinity(as CaCO <sub>3</sub> )	− 0.764570	<b>0.570653</b>	− 0.027939

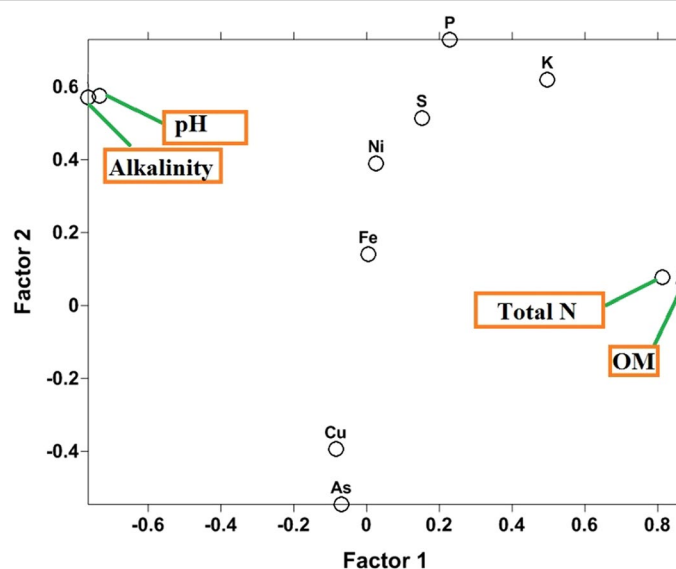
Bolditalic indicates strong positive loadings and Italic shows strong negative loadings

little dominance that means some natural cause is also encountered in the study area. So, major reason of soil degradation of the study area is because of anthropogenic causes i.e. the coal mining activities.

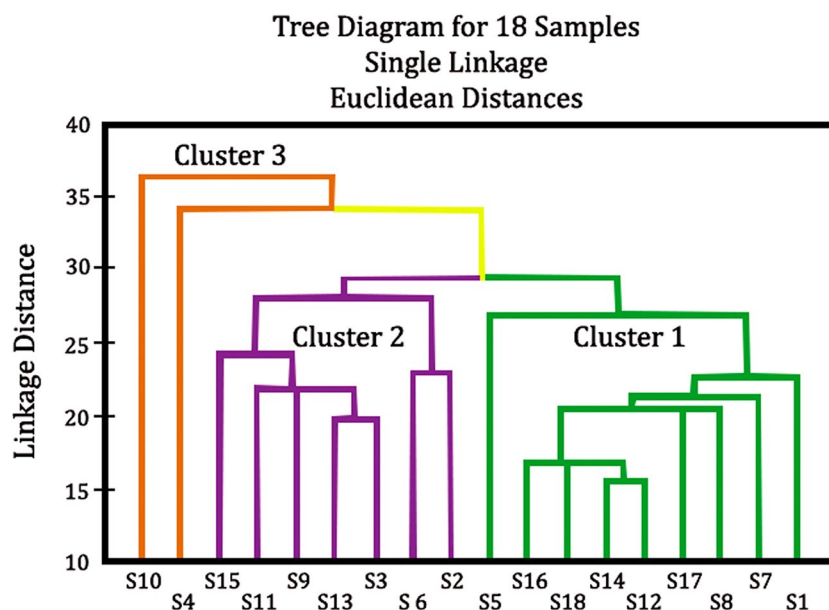
### Cluster

Cluster analysis is performed by using Statistica software (Version 08). The dendrogram of the 18 soil samples is shown in Fig. 5. There are three cluster groups in the dendrogram in which cluster 1 group possess 50% of total samples, cluster 2 has 38.9% samples, and cluster 3 has 11.1% soil samples. Table 5 shows the cluster group, sample numbers with respect to cluster group and percentage of sample belongs to each cluster group. In order to find out the similarity of grouped soil samples, their grouped min, max, and mean value of chemical parameters are shown collectively in Table 6. Cluster 1 group has high Fe, and OM value and low pH value than the Standard Reference value of RSRDIB. Again, cluster 2 has high K value than the Standard Reference value. In addition, all three cluster groups have low value of Total Nitrogen than the standard reference value. Moreover, Copper (Cu) has a higher value in all three cluster groups than the standard reference value.

Figure 6 is showing the grouped soil samples location with respect to the Barapukuria coal mine area. Cluster 1 group is relatively degraded as Fe, K and OM is very high, and pH is low. High Fe value is the main reason of reducing pH of the study area which is described previously by using the chemical reactions. From Fig. 6 it is clear that cluster 1 samples are near to the coal mine. So, there are direct impacts of the coal mine on the soil samples of the surrounding area of the mine and with the increasing distance from the coal mine, Fe and OM becomes low and pH becomes high that is the clear indication of anthropogenic effect. Moreover, Total Nitrogen in the study area for samples of all

**Fig. 4** The bi-plot analysis for factor 1 & factor 2





**Fig. 5** The dendrogram of the cluster analysis

**Table 5** Cluster analysis of the soil samples

Cluster	Sample number	Sample percentage
Cluster 1	1,5,7,8,12,14,16,17,18	50
Cluster 2	2,3,6,9,11,13,15	38.9
Cluster 3	4,10	11.1

cluster groups that indicate the as usual condition of this parameter in this study area. However, Cu is high in the sample of all cluster groups than that of standard

reference value. But, Cluster 1 has more high value of Cu than other Clusters, this higher value of Cu in Cluster 1 prove the release of chalcopyrite ( $\text{CuFeS}_2$ ) because of the coal mining.

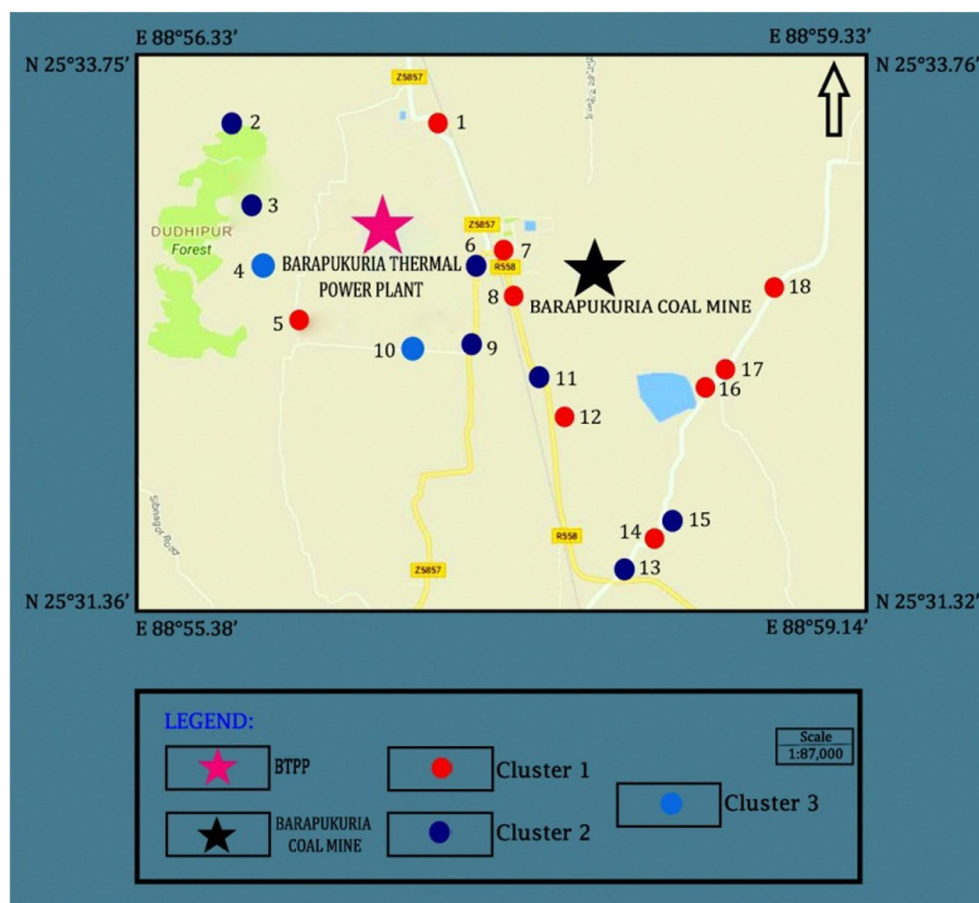
#### One way ANOVA

The main aim of determining one way ANOVA is to compare the differences between the different sampling points for each parameter. In One way ANOVA, one independent variable is selected and with respect to this independent variable the Sum of Squares (SS), Mean

**Table 6** Chemical parameters of different clustered groups with their min, max, and mean value

Clusters	pH	Total nitrogen	OM	K	P	S	Fe	As	Cu	Ni	Total alkalinity
Cluster 1											
Min	<b>4.97</b>	<b>0.06</b>	2.36	0.27	22.8	3.79	1.27	0	1.54	0.1	10
Max	<b>6.65</b>	<b>0.21</b>	2.63	0.48	44.6	44.52	24.37	0.025	2.138	1.28	100
Mean	<b>5.67</b>	<b>0.13</b>	2.495	0.35	29.2	20.01	15.13	0.004	1.854	0.49	54.2
Cluster 2											
Min	6.13	<b>0.13</b>	0.58	4.6	15.8	14.26	1.891	0.01	0.837	0.41	85
Max	6.77	<b>0.15</b>	0.7	9.6	38.4	25.39	2.92	0.01	1.25	0.42	120
Mean	6.45	<b>0.14</b>	0.58	7.1	29.1	24.83	2.41	0.01	1.04	0.415	102.5
Cluster 3											
Min	5.82	<b>0.07</b>	0.48	0.28	21.3	2.262	7.83	0	0.454	0.54	65
Max	6.83	<b>0.1</b>	0.68	0.31	25.3	4.32	10.72	0	1.231	0.57	150
Mean	6.33	<b>0.09</b>	0.58	0.295	23.3	3.29	9.05	0	0.84	0.56	107.5

Italic marks in the table indicate higher value of chemical parameters than the Standard reference value of RSRDIB and bolditalic marks indicate lower value of chemical parameters than the Standard reference value of RSRDIB



**Fig. 6** The status of the quality of soil around the mining area

Square (MS), Degrees of Freedom (DF), F-ratio, and p-level are determined (Howladar and Rahman 2016).

The one-way ANOVA test revealed that there are not significant mean differences between Total N, K, P, S, As and Ni in study area at  $P < 0.05$  (Table 7). The mean differences between cluster one, two and three show significantly differences for some soil quality parameters including pH, OM, Fe, Cu, and Alkalinity (Table 7). This shows that these parameters have high variation in terms of their spatial distribution in the study area that satisfies similar scientific merit done by Belkhiri and Narany, 2015. These parameters are found to have higher factor loading and are also grouped in the second factor (Table 4) of the rotated components and contributed about 23% of the total variance of soil quality. It clearly shows that the degradation of the soil is due to the anthropogenic effects based on the result of factor analysis and ANOVA.

#### Soil quality with respect to agricultural purposes

Pyrite ( $\text{FeS}_2$ ) is the main mineral found in coal. When pyrite gets in contact with oxygenated waters,

sulphates, iron and acidity are released to the environment. The estimation of soil quality around the mining areas and impacts of natural or anthropogenic activities on the soil chemistry is a comparatively complicated work because of the interruption of physical properties of soil horizons and very high heterogeneity of contaminant concentrations in mine soils (Hudson et al. 1997). The regulatory standards in different countries, including Bangladesh, vary for particular class of soils depending on the type of area management and anthropogenic influence such as residential soils, agricultural soils, technogenic soils, etc. (Galuszka et al. 2015). Barapukuria coal mine is surrounded with paddy fields and other agricultural land and the upland products are produced in study area. A details description about soil quality with respect to the optimum RSRDIB standard for upland crops is shown in Table 1. In addition, the degraded condition of each parameter in every sample has been clearly presented by Table 8 and Fig. 7. Table 8 shows that the percentages of highly polluted, slightly polluted, and

**Table 7 The ANOVA test showing the mean differences between three clusters**

Parameters	SS	D.F	MS	F-ratio	p-level
pH					
Between groups	3.380	2	1.690	16.342	0.000
Within groups	1.551	15	0.103		
Total	4.932	17			
Total N					
Between groups	0.005	2	0.002	2.711	0.099
Within groups	0.013	15	0.001		
Total	0.018	17			
OM					
Between groups	0.890	2	0.445	1.488	0.0001
Within groups	4.483	15	0.299		
Total	5.373	17			
K					
Between groups	5.897	2	2.948	0.332	0.722
Within groups	133.023	15	8.868		
Total	138.920	17			
P					
Between groups	415.918	2	207.959	2.046	0.164
Within groups	1524.407	15	101.627		
Total	1940.325	17			
S					
Between groups	1318.413	2	659.207	6.291	0.060
Within groups	1571.684	15	104.779		
Total	2890.097	17			
Fe					
Between groups	186.264	2	93.132	1.348	0.0013
Within groups	1036.359	15	69.091		
Total	1222.623	17			
As					
Between groups	0.000	2	0.000	0.042	0.959
Within groups	0.001	15	0.000		
Total	0.001	17			
Cu					
Between groups	0.784	2	0.392	3.771	0.047
Within groups	1.559	15	0.104		
Total	2.343	17			
Ni					
Between groups	0.020	2	0.010	0.408	0.672
Within groups	0.373	15	0.025		
Total	0.393	17			
Alkalinity					
Between groups	14,749.865	2	7374.933	20.852	0.000
Within groups	5305.079	15	353.672		
Total	20,054.944	17			

unpolluted soil samples and Fig. 7 showing the same result graphically. The criteria of ranking the pollution level are taken from the standard of RSRDIB. Here,

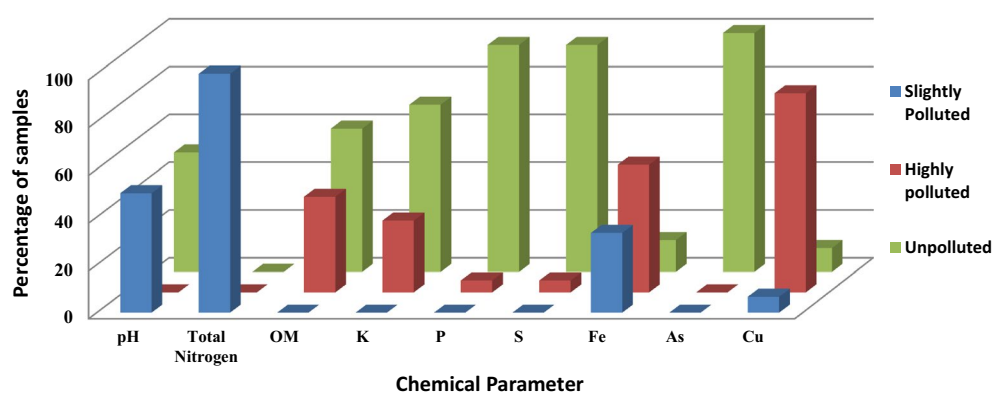
**Table 8 The percentage of polluted and unpolluted soil samples**

Chemical parameters	Percentage of slightly polluted samples (low)	Percentage of highly polluted samples (high)	Unpolluted sample (%)
pH	50.00	0	50.00
Total Nitrogen	100	0	0
OM	0	40	60
K	0	30	70
P	0	5	95
S	0	5	95
Fe	33.33	53.34	13.33
As	0	0	100
Cu	6.67	83.33	10

Criteria of ranking pollution level are taken from the standard reference value of RSRDIB

parameters which maintain the standard reference value of RSRDIB are ranked unpolluted, which are highly deviated from the standard reference values are ranked highly polluted and which parameters are not deviated much from the standard reference values are ranked slightly polluted.

It is clear that, almost 50% soil samples have low pH value than the standard reference value. In fact, the pH is a good measure of acidity and alkalinity of soil water suspension and it provides a good identification of soil chemical nature and the desired pH for good vegetation ranges from 5.5 to 6.8 (Sharma 2008). In addition, the noticeable amount of degradation has been seen for Fe, which might be the reason for low pH value. Moreover, Arsenic (As), S, and P are not found much degraded. Only a few samples reflect the slight contamination of these parameters. The main degraded parameters of the soil samples are Total Nitrogen, Iron (Fe), Copper (Cu), Organic matter (OM), and Potassium (K). For soil quality Total Nitrogen is one of the key as paddy growth largely depends on the availability of Total Nitrogen. In particular, Total Nitrogen is vital for chlorophyll, which allows plants to carry out photosynthesis (Feeco 2017). In addition, Potassium (K) one of the quality elements that contributes to the growth and development of the paddy. It has a large contribution to the characteristics of the plants such as size, shape; color (Feeco 2017). So, deviation of K from the standard value can decrease the production of paddy. Moreover, high Fe, Cu, and OM value can also decrease the paddy production. So, Agriculture of the study area is facing problem now but this situation may become more worsen in near future.



**Fig. 7** The percentage of polluted and unpolluted soil samples

## Conclusion

Barapukuria coal mine is an asset for Bangladesh. It has been playing a vital role for fulfilling the increasing demand of energy in Bangladesh since 2002. But the lack of poor infrastructure for coal storage and inadequate treatment of underground mine water is polluting the surrounding soil to a large extent. From the analysis of soil samples it has been identified that several important chemical parameters that is necessary for agriculture are deviated from the Standard Reference value. It is also clear that soil samples are degraded because of the poor infrastructure of coal stock pile and improper disposal of mine water. Iron pyrite ( $\text{FeS}_2$ ) and chalcopyrite ( $\text{CuFeS}_2$ ) released during mining activities are mixing with the soil and consequently degrading the soil quality. Principal Component Analysis (PCA) also suggests that the soil samples are degraded mostly because of the anthropogenic causes i.e. because of the coal mine. Besides, from the cluster analysis it has been found that soil samples which are located near to the coal mine have been mostly degraded. Moreover, one way ANOVA also confirms about the spatial variability which due to the three different cluster groups and the location of the samples. With the agricultural point of view a few chemical parameters of the soil is deviated from the standard reference value and are not suitable for the agriculture. Without proper initiatives, degradation may lead to a severe damage to the soil quality and it can be brought disaster on the agriculture of the study which ultimately affects the vast paddy land around the Barapukuria coal mine area.

## Authors' contributions

MFH contributed to conception, design, and acquisition of data, data analysis and interpretation of data. MMR participated to acquisition of data, data analysis and interpretation of data as well as involved in drafting the primary manuscript. MOF made a contribution to drafting, checking graphs, and revising the manuscript critically to fulfill the reviewer comments as well as maintaining the scientific merits with syntactic correctness for final submission of the manuscript. All authors read and approved the final manuscript.

## Acknowledgements

The authors would like to thank to the Ministry of Education, Bangladesh for the partial financial support to the successful completion of the research work, otherwise it was beyond our reach.

## Competing interests

The authors declare that they have no competing interests.

## Ethics approval and consent to participate

Not applicable.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 8 July 2017 Accepted: 12 November 2017

Published online: 02 December 2017

## References

- Belkhir L, Narany TS (2015) Using multivariate statistical analysis, geostatistical techniques and structural equation modeling to identify spatial variability of groundwater quality. *Water Resour Manag* 29:2073–2089. <https://doi.org/10.1007/s11269-015-0929-7>
- De Winter JCF, Dodou D (2016) Common factor analysis versus principal component analysis: a comparison of loadings by means of simulations. *Commun Stat Simul Comput* 45:299–321. <https://doi.org/10.1080/03610918.2013.862274>
- Elliott MA (1981) *Chemistry of coal utilization: second supplementary volume*. Wiley, New York, p 6001785
- Feeco (2017) NPK Fertilizer: What Is It And How Does It Work. <http://feeco.com/npk-fertilizer-what-is-it-and-how-does-it-work/>. Accessed 21 Sept 2017
- Fu B, Ma K, Zhou H, Chen L (1999) The effect of land use structure on the distribution of soil nutrients in the hilly area of the Loess Plateau China. *Chin Sci Bull* 44(8):732–736
- Gałuszka A, Migaszwski ZM, Gowska SD, Michalik A, Czernikiewicz AD (2015) Geochemical background of potentially toxic trace elements in soils of the historic copper mining area: a case study from Miedzianka Mt, Holy Cross Mountains, south-central Poland. *Environ Earth Sci* 74:4589–4605. <https://doi.org/10.1007/s12665-015-4395-6>
- Howladar MF, Rahman MM (2016) Characterization of underground tunnel water hydrochemical system and uses through multivariate statistical methods: a case study from Maddhapara Granite Mine, Dinajpur,



- Bangladesh. *Environ Earth Sci* 75:1501. <https://doi.org/10.1007/s12665-016-6309-7>
- Howladar MF, Numanbakth MAA, Faruque MO (2017) An application of Water Quality Index (WQI) and multivariate statistics to evaluate the water quality around Maddhapara Granite Mining Industrial Area, Dinajpur, Bangladesh. *Environ Syst Res* 6:13. <https://doi.org/10.1186/s40068-017-0090-9>
- Hudson TL, Borden JC, Russ M, Bergstrom PD (1997) Controls on As, Pb, and Mn distribution in community soils of an historic mining district, southwestern Colorado. *Environ Geol* 33:25–42
- Imam B (2013) Energy resources of Bangladesh, 2nd edn. pp. 19, 20, 215, 216, 226, 227, 231. ISO 10390
- Lazareva O, Pichler T (2007) Naturally occurring arsenic in the Miocene Hawthorn Group, southwestern Florida: potential implication for phosphate mining. *Appl Geochem* 22(5):953–973
- Li GX, Zhang J, Shao JP, Zhou B, Bi B, Xie KM, Fang XJ, Wang YZ (2014) Chemical properties of soil layers of restoration sites in phosphate mining area, China. *Environ Earth Sci* 73:2027–2030
- Liu WX, Li XD, Shen ZG, Wang DC, Wai OWH, Li SY (2003) Multivariate statistical study of heavy metal enrichment in sediments of the Pearl River Estuary. *Environ Pollut* 121(3):377–388. [https://doi.org/10.1016/S0269-7491\(02\)00234-8](https://doi.org/10.1016/S0269-7491(02)00234-8)
- Martha K (2001) Cradleto Grave: the Environmental Impacts from Coal. Clean Air Task Force, Boston
- Mustapha A, Aris AZ (2012) Multivariate statistical analysis and environmental modeling of heavy metals pollution by industries. *Pol J Environ Stud* 21(5):1359–1367
- NY/T 1121.6-2006 (2006) Soil testing part 6: method for determination of soil organic matter. <https://www.chinesestandard.net/List-PDF/NY.aspx?PageNumber=11>. Accessed 25 March 2017
- Oketola AA, Adekolurejo SM, Osibanjo O (2013) Water quality assessment of river Ogun using multivariate statistical techniques. *J Environ Prot* 4:466–479
- Othman I, Al-Masri MS (2007) Impact of phosphate industry on the environment: a case study. *Appl Radiat Isot* 65(1):131–141
- Petrobangla (2017) Production Statistics. [http://www.bcmcl.org.bd/index.php?page=production\\_statistics](http://www.bcmcl.org.bd/index.php?page=production_statistics), Accessed 21 Sept 2017
- Rashid HO, Hossain MS, Urbi Z, Islam MS (2014) Environmental impact of coal mining: a case study on the Barapukuria coal mining industry, Dinajpur, Bangladesh. *MiddleEast J Sci Res* 21(1):268–274
- Rogerson PA (2010) A statistical method for the detection of geographic clustering. *Geogr Anal* 33(3):216–237
- RSRDIB (2015) Regional Soil Resource Development Institute Bangladesh
- Schaefer K, Einax J (2010) Analytical and chemometric characterization of the Cruces River in South Chile. *Environ Sci Pollut Res* 17(1):115–123
- Sharma HP (2008) Study of the impact of municipal solid waste dumping on soil quality in Guwahati city. *Poll Res* 27(2):327–330
- Shrestha S, Kazama F (2007) Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin. *Jpn Environ Model Softw* 22(4):464–475
- USDA (2007) Sampling soils for nutrient management manure as a resource series. Department of Agriculture, Washington

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)