Application of multivariate analysis to investigate the trace element contamination in top soil of coal mining district in Jorong, South Kalimantan, Indonesia

Arie Pujiwati1,2, K Nakamura1, N Watanabe1 and T Komai1
1 Graduate School of Environmental Studies, Tohoku University, 6-6-20 Aramaki Aoba, Aoba-ward, Sendai, Miyagi 980-8579, Japan
2 General Secretariat of National Energy Council, Ministry of Energy and Mineral Resources, Republic of Indonesia, Jl. Jend. Gatot Subroto Kav. 49 Jakarta Selatan 12950, Indonesia
E-mail: pujiwati.arie.r6@dc.tohoku.ac.jp

Abstract. Multivariate analysis is applied to investigate geochemistry of several trace elements in top soils and their relation with the contamination source as the influence of coal mines in Jorong, South Kalimantan. Total concentration of Cd, V, Co, Ni, Cr, Zn, As, Pb, Sb, Cu and Ba was determined in 20 soil samples by the bulk analysis. Pearson correlation is applied to specify the linear correlation among the elements. Principal Component Analysis (PCA) and Cluster Analysis (CA) were applied to observe the classification of trace elements and contamination sources. The results suggest that contamination loading is contributed by Cr, Cu, Ni, Zn, As, and Pb. The elemental loading mostly affects the non-coal mining area, for instances the area near settlement and agricultural land use. Moreover, the contamination source is classified into the areas that are influenced by the coal mining activity, the agricultural types, and the river mixing zone. Multivariate analysis could elucidate the elemental loading and the contamination sources of trace elements in the vicinity of coal mine area.

1. Introduction
A huge amount of coal deposits are infilled in Kalimantan Islands as the consequence of its position in the Cenozoic sedimentary coal-bearing basins of Southeast Asia, in particular the Neogene Southern Sundaland [1]. Due to particular tectonic setting and coal development processes, the area contains not only extensive and thick coal deposits but also low sulphur content coal [2–4]. The rife coal deposits and the coal cheap price have been driving the extensive coal use as the Indonesian primary energy resources [5,6].

Despite the considerable use of coal as an important energy resource, coal development faces abundant environmental issues including metal contamination [7]. In particular to metal contamination into terrestrial environment in Kalimantan Island, several studies have observed the metal contamination from mine drainage of coal mining area in Asam-asam Basin, South Kalimantan. pH and metal content of particular pit lakes were found exceeding the coal mining effluent standard. Furthermore, some remediation and mitigation strategies have been implemented to overcome the metal contamination issues in the surrounding environment [8–10]. However, the loading and association among trace elements have not been investigated yet. Moreover, since the soils are influenced by their rich metal deposits and utilized as other land functions, it is also important to investigate the original contamination source.
This study investigates the geochemistry of selected trace elements (As, Ba, Cd, Co, Cu, Cr, Ni, Pb, Sb, Zn, V) in top soils and their contributory association with the contamination source as the influence of coal mines in Jorong District, South Kalimantan, Indonesia. Furthermore, this study will determine (1) the loading of selected trace elements and (2) the contamination source types. To achieve the aims, this study applies multivariate analysis (PCA and CA).

2. Material and Methods

2.1. Study area

The present study is carried out in soils in the vicinity of coal mines in Jorong District, Tanah Laut Regency, South Kalimantan Province, Indonesia. The mining area is associated with Asam-asam River which accepts the mine drainage from the mining [10]. The river flows into the Java Sea in the southern part of the province. The mines are located around the hilly surface of the river’s middle section. Numerous land uses lie surrounding the mining, for instances plantation, industrial forest, farming, settlement, and mangrove [11].

The northern part of the study area is constituted by Warukin formation, developed during the Miocene period. The coal was deposited in a paralic depositional environment, intercepted with the quartz sandstone and sandy claystone. Dahor formation is to the south of Warukin formation, consisting of lignite, kaoline, and limonite minerals. Lignite dominates the particular coal rank, having low Sulphur and ash content. The alluvium soils dominate the coastal area in the southern part. The Kalimantan coal is observed as having low trace elements, compared to the worldwide coal range and average [3,12–15].

2.2. Methods

2.2.1. Sample collection. The total number of top soil samples (0-20 cm depth) were twenty (20) around the Asam-asam River Basin, as described in figure 1. The soil samples were stored in polyethylene bags for immediate transport and storage. In the laboratory, the soil samples were oven-dried at 40 °C, pulverized, sieved through a 2-mm sieve, and then stored in sealed polyethylene bags until analyses.

2.2.2. Geochemical analysis. Bulk analysis was performed to determine the total concentration. The representative samples were ground for 5 min in a planetary ball mill at 450 rpm and then pressed into 32-mm internal diameter pellets using a hydraulic press (20 tonnes pressure). The selected trace elements were analyzed by the Energy Dispersive X-Ray Fluorescence (EDXRF) spectrometry (PAnalytical Epsilon 5) [16]. pH and EC were measured based on [17] method.

2.2.3. Geostatistical analysis. The statistical methods were applied in terms of correlation and classification among the selected trace elements. Basic statistical analyses were applied to determine minimum, median, maximum, 1st quartile, 3rd quartile, average, standard deviation, coefficient of variation (CV), and skewness. CV is classified as weak variability if CV < 10%, and strong variability if CV >100%. CV ranged within 10-100% is classified as moderate variability [18].

Numerous statistical analyses are applied to elucidate the correlation and classification of the studied elements. Pearson correlation (r < 0.05) is applied to investigate the inter-elemental correlation for all selected elements derived from the linear relationship of two variables [19].

Moreover, multivariate analysis is applied to determine the contamination source and the elemental classification. All data are normalized by means of z-scores to equate the variables. Two methods of multivariate analysis are applied in this study. The first method is Principal Component Analysis (PCA) which has been widely applied to describe the relationship among variables by reducing the data dimension and drawing up the incorporated variables. Furthermore, the data will be transfigured into a new data set delineating the factor loading (eigenvector) and score (eigenvalues) [17,20]. The second method is Cluster Analysis (CA), which develops a partitioning of multivariate data into significant subgroups or clusters. The well-developed clusters elucidate as much as possible similarities within each cluster, and as huge as possible differences with other clusters [21,22]. This study applies the PCA Ward’s method and the hierarchical clustering with the Euclidean’s distance.
3. Results and discussion

3.1. Geochemical analysis

Table 1 shows the geochemical results. The mean values of As, Cu, Cr, Ni, Pb, and V were higher than those in the worldwide upper continental crust by [23]. All elements have greater mean values instead of their median values, corresponding with their positive skewness. In particular, the mean values of Cr and Ni are much higher than those found in other studies [24–27]. Cr and Ni also perform larger variation than other elements due to their higher CV values. The maximum values of Cr and Ni were found in the estuary (A19) and near the farming area (A16), respectively. Furthermore, the maximum value of Ba, Sb, and V were observed near the area closed to the coal mines. In addition, the maximum value of As was observed near the plantation area (A1) which is located on the upstream side of coal mines.

3.2. Correlation analysis

Pearson correlation is performed with the confidence level 95%, as described in table 2. Among the total concentration results, several elements have strong correlation with others. Strong positive correlation occurs among Cd, Co, Cu, and Ni. Meanwhile, Sb has insignificant correlation with others.
Table 1. Geochemical analysis results.

|         | Min  | Q1   | Median | Q3   | Max  | Mean | SD   | CV   | skewness |
|---------|------|------|--------|------|------|------|------|------|----------|
| pH      | 2.00 | 3.28 | 3.85   | 4.53 | 7.20 | 4.06 | 1.28 | 31.58 | 0.82     |
| EC (µS/cm) | 0.27 | 1.68 | 54.50  | 76.25 | 156.00 | 52.67 | 46.50 | 88.27 | 0.50     |
| Bulk analysis (mg/kg) | | | | | | | | | |
| As      | 7.00 | 12.16 | 16.25 | 20.09 | 27.76 | 16.38 | 5.57 | 34.00 | 0.30     |
| Ba      | 35.14 | 81.76 | 121.29 | 180.13 | 269.79 | 135.22 | 71.03 | 52.53 | 0.43     |
| Cd      | N/D  | N/D  | N/D   | 0.07 | 0.33 | 0.00 | 0.12 | N/D  | 1.12     |
| Co      | 2.51 | 12.90 | 21.04 | 32.44 | 59.12 | 24.30 | 15.48 | 63.71 | 0.80     |
| Cu      | 9.40 | 17.27 | 32.45 | 44.97 | 131.49 | 40.35 | 33.50 | 83.01 | 1.99     |
| Cr      | 87.56 | 126.32 | 288.03 | 943.95 | 5649.81 | 748.42 | 1211.20 | 161.83 | 3.54     |
| Ni      | 14.58 | 20.56 | 35.25 | 137.47 | 533.03 | 121.18 | 163.21 | 134.68 | 1.87     |
| Pb      | 8.41 | 13.03 | 16.55 | 19.59 | 35.29 | 17.28 | 6.78 | 39.25 | 1.18     |
| Sb      | 0.36 | 0.64 | 0.73 | 1.03 | 1.50 | 0.80 | 0.29 | 36.36 | 0.49     |
| V       | 40.20 | 99.75 | 118.22 | 167.35 | 209.07 | 124.40 | 45.02 | 36.19 | 0.14     |
| Zn      | 21.31 | 40.85 | 49.77 | 71.33 | 186.77 | 62.60 | 37.52 | 59.94 | 2.01     |

Table 2. Pearson correlation (r < 0.05) of elemental total concentration.

|     | Cd | V  | Co | Ni | Cu | Zn | As | Pb | Cr | Sb | Ba |
|-----|----|----|----|----|----|----|----|----|----|----|----|
| Cd  | 1.00 | | | | | | | | | |
| V   | 0.56 | 1.00 | | | | | | | | |
| Co  | 0.45 | 0.56 | 1.00 | | | | | | | |
| Ni  | 0.91 | 0.66 | 0.49 | 1.00 | | | | | | |
| Cu  | 0.88 | 0.65 | 0.44 | 0.92 | 1.00 | | | | | |
| Zn  | 0.82 | 0.62 | 0.42 | 0.86 | 0.82 | 1.00 | | | | |
| As  | -0.31 | -0.07 | 0.28 | -0.26 | -0.32 | -0.06 | 1.00 | | | |
| Pb  | 0.39 | 0.28 | 0.36 | 0.39 | 0.49 | 0.67 | 0.16 | 1.00 | | |
| Cr  | 0.03 | 0.22 | 0.05 | 0.13 | -0.02 | 0.15 | 0.05 | -0.31 | 1.00 | |
| Sb  | -0.47 | -0.66 | -0.59 | -0.51 | -0.48 | -0.46 | -0.08 | -0.22 | -0.33 | 1.00 |
| Ba  | 0.72 | 0.67 | 0.43 | 0.75 | 0.71 | 0.77 | -0.17 | 0.60 | -0.18 | -0.47 | 1.00 |

3.3. Principal component analysis

Figure 2 shows PCA results of total concentration. Three significant components of Total concentration account for 80.3% of total variance. PC1 accounts for 53.1% of total variance. It is characterized by the positive loadings of most of the elements. Cd, Ni, Cu, Zn, and Ba have higher positive loadings, showing the similar patterns among all sampling points. On the other hand, only Sb and As show negative loadings. Those results correspond to the strong linear correlation of Cd, Ni, Cu, and Zn, and strong inverse correlation of Sb with most elements. Furthermore, areas near the settlements (A15) and farming (A16) show very high PC1 score. They correspond with the enrichment of several elements in the points, for instances Cu (130-132 mg/kg), Cr (657-1049 mg/kg), Ba (250-254 mg/kg), Pb (18-36 mg/kg), and Zn (124-187 mg/kg).

PC2 accounts for 14.2% of total variance. Only Cr, As, Co, and V perform positive eigenvectors. It reflects a higher positive loading of Cr and As. On the other hand, Sb shows very low negative loading which corresponds with the strong inverse correlation between Sb and either Cr or As. Moreover, PC2
score describes very high loading elements in the estuary (A19), which corresponds with the maximum value of Cr (5649.8 mg/kg) and low value of Sb (0.6 mg/kg).

PC3 accounts for 12.9% of total variance. It describes high positive eigenvectors of As and Pb. Cr shows high negative eigenvector. PC3 score shows high loading in the plantation area (A1). The maximum values of As (27.76 mg/kg) is found in this area. On the other hand, Cr show low value (105.9 mg/kg) below its Q1 (126.3 mg/kg). It could be concluded that Cr, Cu, Ni, Zn, As, and Pb contribute to the high elemental loadings of the soil total concentration. In addition, the high elemental loadings are found in the area near the settlement (A15), the farming (A16), the upstream plantation (A1), and the estuary (A19).

PCA could elaborate the elemental loading and inter elemental classification of the selected elements. In addition, it is also able to illustrate the contamination sources that have particular elemental loadings. It conforms to the PCA application in previous studies [28,29].

3.4. Cluster analysis
Hierarchical cluster analysis results are shown in the figure 3 by dendrogram visualization. CA of total concentration remains four significant groups; group 1: Pb-Cr, group 2: Cd-Co-Sb, group 3: Ni-As-Ba, and group 4: V-Cu-Zn. The sources of Pb and Cr are immensely different with the sources of other elements.

CA is also performed to classify the sampling points, as described in the fig 7. CA obtains four significant groups. The first group (A9, A16, A15, A8) is classified as the river flow-mixing area, which may increase the elemental concentration [30]. The upstream part of the rivers are mixed in these points. The second group (A19, A17, A13) and the third group (A18, A3, A1) have similarities so that they are classified as the plantation affected area. Soils in the third group are mostly representing the natural vegetation types rather than those in the second group. The last group is strongly represented the coal mining affected area.
Cluster analysis is not intended to elucidate the elemental loadings. Nevertheless, the association of the elements and the contamination source types could be described well by the application of cluster analysis.

4. Conclusions
It is concluded that multivariate statistical methods can be used to identify the association of the trace elements and the contamination source types in soils of the study area. PCA shows that Cr, Cu, Ni, Zn, As, and Pb contribute to the elemental loading in the top soils. Moreover, the origin of the contamination sources could clearly be classified by the cluster analysis as three sources, the river flow-mixing area, the plantation, and the coal mining affected area. Therefore, the coal mining activity in Jorong District is not solely affecting the trace element contamination in soil of the study area.

Acknowledgments
Any opinions, findings, conclusions, or recommendations revealed in this study are those of the author(s) and do not necessarily represent the view or policy of the Ministry of Energy and Mineral Resources, Republic of Indonesia. We acknowledge comments from the reviewer and the editor of GCGE2017.

References
[1] Hall R 2002 Cenozoic geological and plate tectonic evolution of SE Asia and the SW Pacific: Computer-based reconstructions, model and animations J. Asian Earth Sci. 20 353–431
[2] Esterle J S and Ferm J C 1994 Spatial Variability in Modern Tropical Peat Deposits from Sarawak, Malaysia and Sumatra, Indonesia - Analogs for Coal Int. J. Coal Geol. 261–41 ST–Spatial Variability in Modern Tropical
[3] Belkin H E, Tewalt S J, Hower J C, Stucker J D and O’Keefe J M K 2009 Geochemistry and petrology of selected coal samples from Sumatra, Kalimantan, Sulawesi, and Papua, Indonesia Int. J. Coal Geol. 77 260–8
[4] Friederich M C, Moore T A and Flores R M 2016 A regional review and new insights into SE Asian Cenozoic coal-bearing sediments: Why does Indonesia have such extensive coal deposits? Int. J. Coal Geol. 166 2–35
[5] Geological Resource Center 2016 Update of Energy Resources Data and Balance, Executive Summary Minist. Energy Miner. Resour. Repub. Indon.
[6] Dutu R 2016 Challenges and policies in Indonesia’s energy sector Energy Policy 98 513–9
[7] Finkelman R B 1999 Trace elements in coal: environmental and health significance. Biol. Trace Elem. Res. 67 197–204
[8] Gautama R S and Hartaji S 2004 Improving the accuracy of geochemical rock modelling for acid rock drainage prevention in coal mine Mine Water Environ. 23 100–4
[9] Gautama R S, Novianti Y S and Supringgo E 2014 Review on In-pit Treatment of Acidic Pit Lake in Jorong Coal Mine, South Kalimantan, Indonesia An Interdisciplinary Response to Mine Water Challenges - Sai, Sun & Wang (eds) pp 645–9
[10] Edraki M, Baumgartl T, Sayoga G R, Julu K G, Ali M, Lana S and Sujatmiko 2015 Mitigating Acid and Metalliferous Drainage in the Asam-Asam Basin, South Kalimantan, Indonesia
[11] MEF R of I 2016 South Kalimantan land use Minist. For. Environ.
[12] Kusnama 2008 Warukin formation of coal in Sampit and its vicinity, Central Kalimantan Indones. J. Geosci. 3 11–22
[13] Milligan E N, Friederich M C and Lim M S W 1996 Coal Exploration and Development in Southeastern Kalimantan, Indonesia Fifth Circum-Pacific Energy Miner. Ressources Conf. 2009
[14] Widodo S, Oschmann W, Bechtel A, Sachsenhofer R F, Anggayana K and Puettmann W 2010 Distribution of sulfur and pyrite in coal seams from Kutai Basin (East Kalimantan, Indonesia): Implications for paleoenvironmental conditions Int. J. Coal Geol. 81 151–62
[15] Friederich B M C, Langford R P and Moore T A 1999 The geological setting of Indonesian coal deposits AusIMM Proc. 23–9
[16] Matsunami H, Matsuda K, Yamasaki S ichi, Kimura K, Ogawa Y, Miura Y, Yamaji I and Tsuchiya N 2010 Rapid simultaneous multi-element determination of soils and environmental samples with polarizing energy dispersive X-ray fluorescence (EDXRF) spectrometry using pressed powder pellets Soil Sci. Plant Nutr. 56 530–40
[17] Nakamura K, Kuwatani T, Kawabe Y and Komai T 2016 Extraction of heavy metals characteristics of the 2011 Tohoku tsunami deposits using multiple classification analysis Chemosphere 144 1241–8
[18] Jin Z, Li Z, Li Q, Hu Q, Yang R, Tang H, Li M, Huang B, Zhang J and Li G 2015 Canonical correspondence analysis of soil heavy metal pollution, microflora and enzyme activities in the Pb–Zn mine tailing dam collapse area of Sidi village, SW China Environ. Earth Sci. 73 267–74
[19] Yildirim G and Tokalioglu S 2016 Heavy metal speciation in various grain sizes of industrially contaminated street dust using multivariate statistical analysis Ecotoxicol. Environ. Saf. 124 369–76
[20] Li X and Feng L 2012 Multivariate and geostatistical analyses of metals in urban soil of Weinan industrial areas, Northwest of China Atmos. Environ. 47 58–65
[21] Scott A A J and Knott M 1974 A Cluster Analysis Method for Grouping Means in the Analysis of Variance Biometrics 30 507–12
[22] Tempf M, Filzmoser P and Reimann C 2008 Cluster analysis applied to regional geochemical data: problems and possibilities Appl. Geochemistry 23 2198–213
[23] Wedepohl K H 1995 The composition of the continental crust Geochim. Cosmochim. Acta 59 1217–32
[24] Halim M A, Majumder R K and Zaman M N 2015 Paddy soil heavy metal contamination and uptake in rice plants from the adjacent area of Barapukuria coal mine, northwest Bangladesh Arab. J. Geosci. 8 3391–401
[25] Candelas C, Melo R, Avila P F, Da Silva E F and Salgueiro A R 2014 Heavy metal pollution in mine-soil-plant system in S. Francisco de Assis - Panasqueira mine (Portugal) Appl. Geochemistry 44 12–26
[26] Masto R E, Sheik S, Nehru G, Selvi V A, George J and Ram L C 2015 Assessment of environmental soil quality around Sonepur Bazari mine of Raniganj coalfield, India Solid Earth 6 811–21
[27] Keshav Krishna A and Rama Mohan K 2016 Distribution, correlation, ecological and health risk assessment of heavy metal contamination in surface soils around an industrial area, Hyderabad, India Environ. Earth Sci. 75 1–17

[28] Loska K and Wiechuła D 2003 Application of principal component analysis for the estimation of source of heavy metal contamination in surface sediments from the Rybnik Reservoir Chemosphere 51 723–33

[29] Horák J and Hejcman M 2016 800 years of mining and smelting in Kutná Hora region (the Czech Republic) — spatial and multivariate meta-analysis of contamination studies J. Soils Sediments 1584–98

[30] Miller J R 1997 The role of fluvial geomorphic processes in the dispersal of heavy metals from mine sites J. Geochemical Explor. 58 101–18