Multivariate Statistical Analysis of Coal in Dahor Formation, Borneo Island, Indonesia: A Comparative Study Utilizing Principal Component Analysis (PCA)

Ratih Hardini Kusuma Putri *

Facult of Mineral and Marine Technology, Institut Teknologi Adhi Tama Surabaya, Indonesia
Natural Resources and Environmental Studies, National Dong Hwa University, Hualien, Taiwan
*e-mail: rhkputri@itats.ac.id

| Article info	Abstract
	This study focuses on the multivariate statistical analysis of coal samples from the Dahor Formation in Borneo Island, Indonesia, utilizing Principal Component Analysis (PCA). The objective is to discern the differences in coal samples based on their geological history and chemical properties. The Indonesian coal industry heavily relies on domestic consumption, with plans to increase coal usage for domestic purposes. Coal classification systems play a vital role in evaluating coal quality and determining its economic value. The geological background of the study area in Borneo Island is discussed, emphasizing relevant formations and sedimentation processes. The application of PCA in multivariate statistical analysis is explained, along with the generation of a correlation matrix to explore relationships between variables. The results reveal the presence of four distinct quadrants in the processed coal data, indicating the influence of volatile matter, total moisture, ash content, and fixed carbon. Further analysis demonstrates relatively insignificant differences in the proximate analysis parameters, except for ash content, which indicates the presence of impurity minerals. The concentration of volatile matter, ash, and moisture impacts the fixed carbon content, while volatile matter influences the combustion process. These findings provide valuable insights into coal quality and utilization. The utilization of PCA and multivariate analysis enhances understanding of coal characteristics and facilitates decision-making in coal-related industries. The study concludes with a call for further investigations and analyses to enhance our understanding of coal deposits, improve resource estimation, and develop more sustainable and efficient coal-based processes. Continued research in this field will contribute to advancements in coal science and the development of strategies for responsible coal utilization.

1. Introduction

Coal, being a vital energy source for centuries, has garnered extensive knowledge regarding its use. Its utility as a heat source and the diverse array of by-products derived from coal have been well understood. However, the comprehensive examination of coal's mineral composition, except for sulfur and iron, has only recently received attention. The continued and increasingly large-scale use of coal in the United States and numerous other industrialized and developing nations has led to heightened concerns about known hazards and speculation regarding potential additional hazards to environmental quality and human health. Consequently, there is still much to be discovered about both the harmful and beneficial attributes of coal, with the aim of minimizing harm to humans and the environment, while maximizing its utility for the betterment of society.

Presently, the Indonesian coal industry heavily relies on domestic consumption, with nearly 70% of the country's coal production being utilized by the state-owned Electricity Company as fuel for electricity generation. Approximately 10% of the coal is used in cement manufacturing, while the remainder is employed as industrial or process fuels in metallurgy. In line with national energy policies, the Indonesian government has announced plans to increase the usage of coal for domestic purposes while reducing coal exports. By 2025, Indonesian coal is projected to constitute approximately 33% of the country's total energy [1].
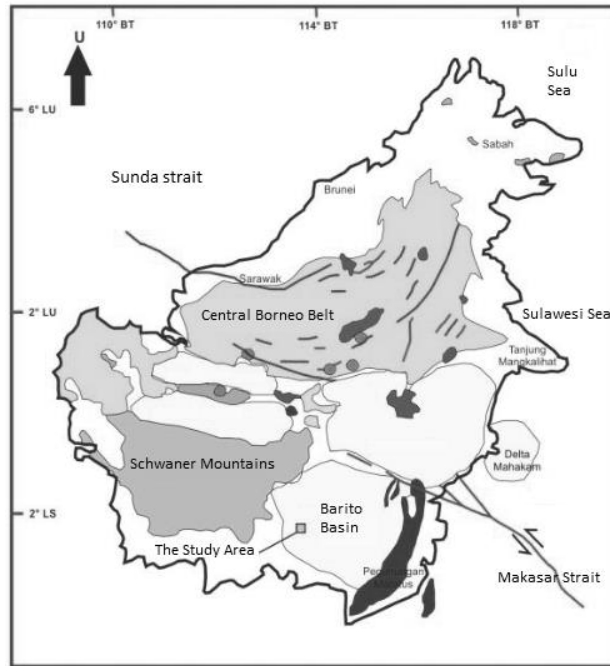
Figure 1. Location of the investigation area The Barito Basin, comprising the Warukin Formation and Dahor Formation, is the geological product of interest in this study. While both formations are coal-bearing, this report specifically focuses on the Dahor Formation. Previous investigations conducted by [2] encompassed petrography and geochemical analyses of the Dahor Formation.

The primary purpose of coal classification systems is to differentiate coal based on its physical characteristics, enabling the evaluation of its quality and economic value for various utilization purposes. Coal classification also provides information regarding specific coal properties that can serve as cutoff values for estimating coal resources and reserves. These properties include ash yield, calorific value, and total sulfur content. Coal can be classified based on its scientific properties, such as elemental composition, as well as commercial properties that determine its market value for specific coal uses, such as burning or carbonization. These commercial properties encompass coking or caking properties, calorific value, durability, grind ability, water content, and others.
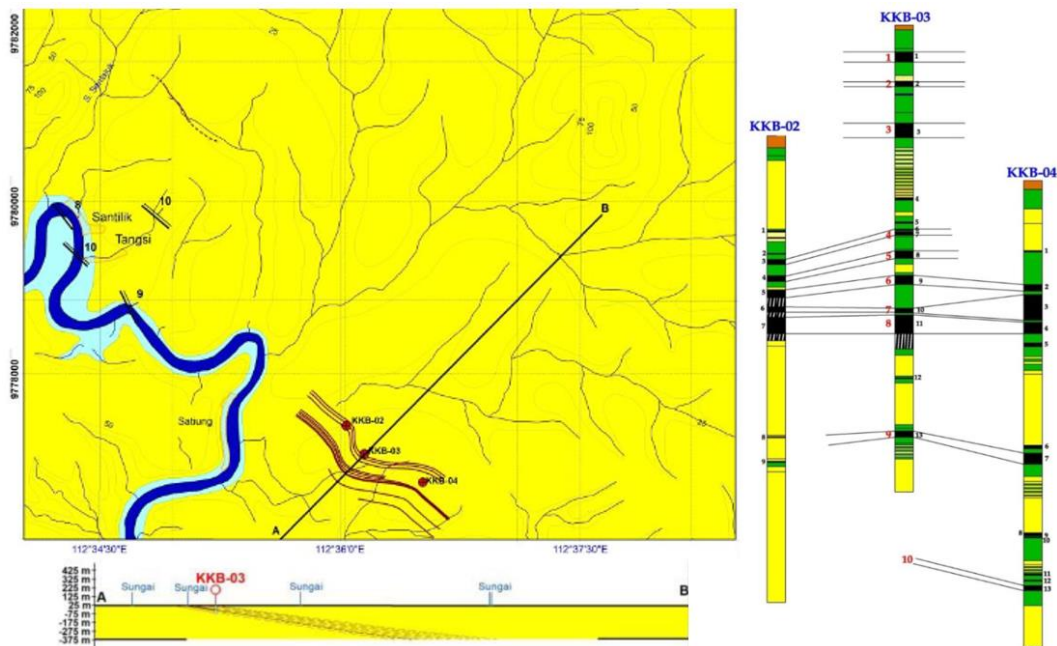
Figure 2. Geological map that showing the alternating sandstone-claystone units (top left), geological section illustrating the coal seam continuity (bottom left), and borehole correlation (right) [2].
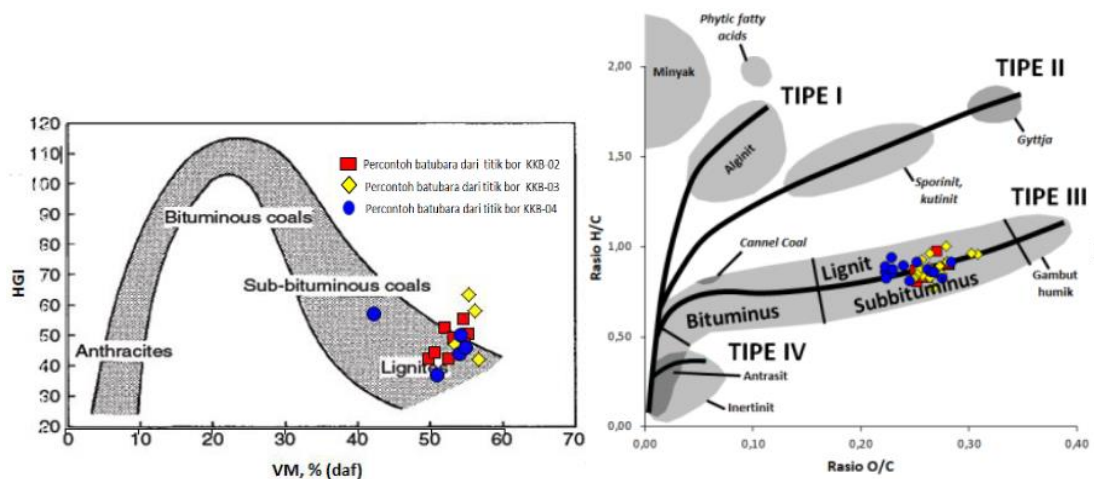
Figure 3. Summary of coal geochemical analysis based on the hardness grain index vs volatile matter (left) and the ratio of H/C vs O/C (right) [2].

Proximate analysis is a straightforward method employed to assess coal properties. In the literature, the terms "ash" and "inorganic mineral substances" are often used interchangeably. Ash refers to the residue left behind after coal combustion, and its composition varies based on the number and chemical characteristics of the mineral substances present in the coal. These mineral substances undergo thermal transformations during heating [3]. Hence, the analysis and testing of coal samples are necessary to gain insights into the quality of coal in the study area.

In this study, multivariate statistical analysis, specifically the Principal Component Analysis (PCA) method, is employed for the statistical analysis of coal samples. Multivariate statistical analysis is well-suited for datasets with multiple correlated variables. The PCA method allows for the identification of the key variables contributing to sample quality. Parameters such as total moisture, volatile matter, fixed carbon, and ash content are examined. The application of PCA in statistical analysis has been demonstrated in various fields, including hydrogeology [4].

Given the aforementioned background, the objective of this study is to discern the differences in coal samples using multivariate statistical analysis. Each variable contributes to the overall quality of the coal samples, and their relative influences facilitate the classification and interpretation of samples based on their geological history and chemical properties. The outcomes of this study will contribute to a better understanding of coal characteristics, aiding efforts to reduce its harmful effects on human health and the environment while maximizing its beneficial applications.

2. Geological background

The study area is situated in Borneo, also known as Kalimantan, which is part of the micro-Sunda continental plate. In the past, it was a component of the Eurasian plate [5]. The basement rock of Kalimantan consists of Mesozoic granite. During the Cenozoic era, the island underwent subduction processes, resulting in the formation of Cenozoic magmatic products [6]. Since the Paleocene period, weathering processes have occurred on the island, leading to the generation of various sedimentary products. The abundant sedimentation has caused the slopes to flatten. From the Eocene epoch onwards, tidal environments have developed in certain parts of the island. Particularly from the Miocene period, the island experienced a regression process that continued until the Pleistocene [7].

The datasets used in this study were sourced from the publication by the Ministry of Energy and Mineral Resources, Indonesia, focusing on the petrography and geochemistry of coal samples from the Barito Basin on Borneo Island [2] (refer to Figure 2). Based on the geochemical analysis conducted by Wibisono et al. (2019) [2], it was determined that the majority of the samples belonged to the Lignite-Sub-bituminous class in terms of maturity (see Figure 3). The coal outcrops in the Barito Basin are associated with the Dahor Formation, which was deposited during the middle Miocene to early Pleistocene period [8]. This formation, along with the Warukin Formation, is known for its abundant

coal deposits, which were formed during a regression sedimentation phase of the Barito Basin [9]. The regression sedimentation phase provided an ideal environment for peat formation and subsequent coalification.

3. Methodology and the datasets

To determine differences in coal samples, various parameters such as total moisture, volatile matter, fixed carbon, and ash content were used as data. Proximate analysis was employed as a method to characterize coal and understand its characteristics. It is well-known that Kalimantan Island, as a prominent coal producer in Indonesia, exhibits distinct coal characteristics across different regions. By comprehending these coal characteristics, qualitative and statistical differences can be observed through visual calculations using statistical analysis, specifically the PCA method.

Coal originates from the formation of peat in mires that provide favorable conditions for peat preservation. These waterlogged environments, often filled with standing or slowly moving water, promote prolific plant growth. When plants fall into the mire, they become submerged, slowing down or preventing rapid decomposition. This results in the accumulation of plant matter and the formation of peat. The organic matter composing peat can either accumulate in place or, less commonly, be transported by flowing water and accumulate nearby. Most coal beds are formed from plant material that accumulates in situ. The process of coal formation involves two primary processes: peatification and coalification (refer to Figure 3).

Coal rank plays a significant role in determining coal quality. Rank refers to the stages in the gradual and natural process of "coalification," during which buried plant matter undergoes transformations, becoming denser, drier, richer in carbon, and harder. The major coal ranks, ranked from lowest to highest, include lignite (also known as "brown coal" in some regions), sub-bituminous coal, bituminous coal, and anthracite. An example of coal ranking can be found in Table 1, as provided by ASTM [10], [11]

Coal geochemistry analysis involves two main types: ultimate analysis and proximate analysis. The ultimate analysis aims to determine the chemical elements present in coal, including carbon, hydrogen, oxygen, nitrogen, and sulphur. On the other hand, proximate analysis is conducted to determine the coal rank based on parameters such as total moisture, fixed carbon, volatile matter, and ash content. According to the Encyclopedia of Coal (2009) [12], the proximate analysis consists of the following components; Proximate analysis of coal involves the determination of several parameters. The first parameter, total moisture, is divided into three types. Free moisture refers to water that sticks to the coal's surface, filling cracks and capillary holes. Its quantity can be determined by comparing the weight of the coal before and after drying it at room temperature or under sunlight at 30°-40°C for 1-2 hours until the weight stabilizes. Inherent moisture, on the other hand, represents water trapped within the internal pores of the coal during its formation. Lower-ranked coals generally have higher inherent moisture content. The total moisture is the combined weight of free moisture and inherent moisture. Another parameter, volatile matter, refers to the portion of coal that vaporizes when heated, including gases like CH_4 (methane) and others. Ash content is the inorganic residue that remains after the complete combustion of coal. Lastly, fixed carbon is calculated by subtracting the total percent of moisture, ash content, and volatile matter from 100%. These parameters are essential in characterizing the composition and properties of coal.

Table 1. Classification of Coal Rank [10], [11]].

Coal Rank	Moisture Content (%)	Volatile Matter (%)	Fixed Carbon (%)	Ash Content (%)
Lignite	25 - 40	35 - 45	20 - 35	5 - 25
Subbituminous coal	15 - 30	25 - 35	35 - 45	5 - 20
Bituminous coal	5 - 15	20 - 30	45 - 60	5 - 15
Anthracite	2 - 8	5 - 15	80 - 95	2 - 6
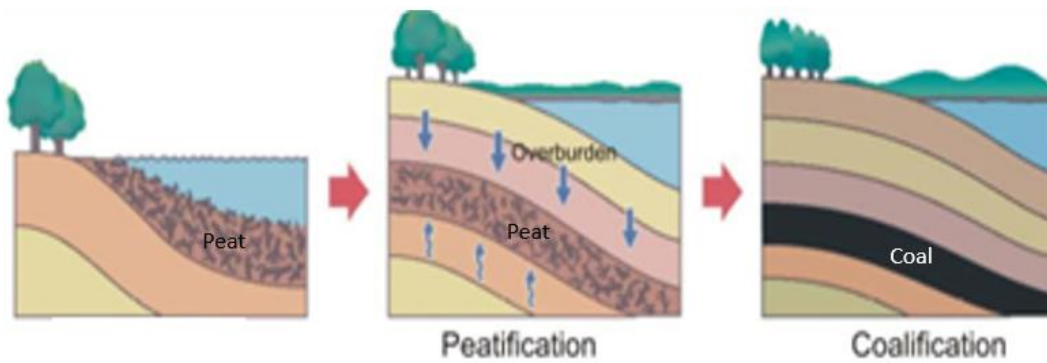
Figure 4. Early stages of coal formation.

Proximate analysis is a crucial method used to evaluate the characteristics of coal. It involves the determination of several key parameters. Total moisture, one of the parameters, includes free moisture and inherent moisture. Free moisture refers to water that adheres to the coal's surface, filling cracks and capillary holes. Its quantity can be measured by weighing the coal before and after drying it at room temperature or under controlled conditions. Inherent moisture, on the other hand, represents water trapped within the internal pores of the coal during its formation. Another parameter, volatile matter, refers to the fraction of coal that evaporates when heated. It includes gases such as methane and other volatile components. Ash content is the inorganic residue that remains after the complete combustion of coal, representing minerals and other non-combustible substances. Lastly, fixed carbon is calculated by subtracting the percentages of moisture, volatile matter, and ash from 100%. The proximate analysis provides valuable insights into the composition and characteristics of coal, assisting in its classification and utilization in various industries.

The formula used for calculating proximate analysis values, according to ASTM (1999), is as follows:
%TM = (initial weight + sample) - final weight x 100% / Sample weight
%VM = {(initial weight + sample) - final weight x 100%} - %TM / Sample weight
%Ash = (final weight) (g) x 100% / (Sample weight) (g)
%FC = 100% - (%TM - %ash - %VM)

Abbreviations:
FC = Fixed Carbon % (adb)
VM = Volatile Matter % (adb)
TM = Moisture % (adb)
A = Ash % (adb)

The results of the proximate analysis were obtained from 42 core drill samples taken from 10 coal layers within the study area, as shown in Table 3. The analysis, based on a dry air basis (adb), reveals volatile matter content (%VM) ranging from 29.19% to 47.72%, moisture (%M) ranging from 7.37% to 13.74%, fixed carbon (%FC) ranging from 24.85% to 50.39%, and ash (%Ash) ranging from 3.41% to 38.6%. Additionally, calorific value (CV), HGI (Hardgrove Grindability Index), and density were determined through combustion and physical properties analysis. The results, based on adb, indicate a range of calorific values (3,539 cal/g to 5,803 cal/g), density (1.4% to 1.75%), and HGI (38 to 64).

Multivariate analysis [13], specifically Principal Component Analysis (PCA), is a powerful statistical method used to analyze complex data sets with multiple variables. In the context of coal analysis, multivariate statistical analysis helps uncover relationships and patterns among parameters such as total moisture, volatile matter, fixed carbon, and ash content. PCA, as a specific technique within the multivariate analysis, reduces the dimensionality of the data by transforming it into a new set of uncorrelated variables called principal components. These components capture the most significant sources of variation in the data [14]. By applying PCA to coal analysis, it becomes possible to identify the key factors influencing coal quality and to visualize the data in a simplified manner. This

methodology enables a deeper understanding of the interrelationships between different coal properties, aiding in quality evaluation, resource estimation, and decision-making related to coal utilization.

Table 2. Results of proximate, physical, and combustion properties analysis [2].

| Sample	seam	depth	thick	FM	TM	M	VM	FC	Ash	TS	HGI	SG	CV cal/g
S02-1	-	20.3	0.3	43.8	49.7	10.6	38.1	36.3	15.1	0.4	-	1.6	4.8
S02-2	-	25.2	0.2	40.6	46.3	9.6	34.9	32.4	23.3	0.4	53.0	1.6	4.2
S02-3	4	26.5	1.1	41.5	47.1	9.8	36.4	32.3	21.8	0.3	50.0	1.6	4.4
S02-4	5	30.1	1.2	46.6	52.6	11.2	42.0	42.5	4.3	0.3	43.0	1.6	5.6
S02-5	6	33.0	1.7	45.4	51.3	10.8	38.9	38.1	12.1	0.6	45.0	1.5	5.0
S02-6	7	36.6	1.2	40.4	45.1	8.0	34.2	28.7	29.1	2.5	56.0	1.6	4.1
S02-7	8	38.9	1.1	44.3	49.7	9.8	41.3	33.5	15.4	0.7	51.0	1.7	4.9
S02-8	8	40.0	1.0	45.5	51.6	11.2	43.1	39.2	6.5	0.3	43.0	1.5	5.4
S02-9	8	41.0	1.4	45.2	51.0	10.6	44.3	36.8	8.4	0.3	56.0	1.5	5.4
S03-1	1	5.7	1.3	47.0	52.0	9.4	43.6	37.8	9.2	0.3	51.0	1.5	5.4
S03-2	1	7.0	0.9	42.3	47.2	8.5	35.5	28.6	27.5	0.3	43.0	1.5	4.0
S03-3	2	12.1	1.1	45.0	50.4	9.8	43.7	34.4	12.2	0.2	48.0	1.7	5.2
S03-4	3	21.0	1.0	46.0	51.6	10.3	46.6	35.8	7.4	0.2	-	1.5	5.5
S03-5	3	22.0	1.0	47.1	53.0	11.2	47.7	37.7	3.4	0.2	-	1.5	5.6
S03-6	3	23.0	1.0	47.3	53.2	11.1	46.0	39.4	3.4	0.2	59.0	1.5	5.7
S03-7	-	37.0	0.5	45.2	50.8	10.1	40.0	30.8	19.1	0.3	-	1.4	4.6
S03-8	4	44.2	0.8	38.6	44.5	9.2	34.5	30.2	26.1	0.3	-	1.6	4.1
S03-9	5	48.3	0.9	50.3	55.8	11.1	41.1	40.3	7.6	1.4	-	1.7	5.3
S03-10	5	49.2	0.8	44.3	50.3	10.9	39.1	39.9	10.2	0.6	43.0	1.5	5.1
S03-11	6	53.6	1.0	45.3	51.5	11.2	41.5	41.3	6.0	0.4	48.0	1.6	5.4
S03-12	6	54.6	1.0	41.0	46.7	9.6	35.8	32.4	22.2	0.7	-	1.5	4.4
S03-13	7	60.6	1.1	37.5	42.7	8.3	33.8	27.4	30.4	2.6	-	1.6	4.0
S03-14	8	62.1	0.9	41.1	46.4	8.9	35.2	31.6	24.3	2.9	-	1.7	4.4
S03-15	8	63.0	1.0	36.9	41.5	7.4	29.2	24.9	38.6	0.9	-	1.8	3.5
S03-16	8	64.0	1.0	44.6	50.7	11.1	40.1	40.6	8.3	0.6	-	1.5	5.4
S03-17	8	65.0	1.1	43.7	49.5	10.4	36.0	35.6	18.1	0.4	-	1.5	4.7
S03-18	9	87.0	1.0	46.9	52.8	11.1	41.4	41.9	5.6	0.6	-	1.4	5.6
S03-19	9	88.0	0.3	46.2	53.6	13.7	38.4	43.3	4.6	0.2	-	1.4	5.3
S04-1	-	15.0	0.3	50.0	55.5	11.0	40.2	41.2	7.6	0.2	45.0	1.4	5.4
S04-2	6	22.3	1.4	42.0	47.4	9.2	39.9	34.2	16.7	0.3	58.0	1.5	4.8
S04-3	7	24.4	1.1	44.0	49.3	9.4	36.5	50.4	3.7	0.1	-	1.4	5.8
S04-4	7	25.4	1.1	45.0	50.4	9.8	45.4	40.8	4.1	0.1	-	1.4	5.7
S04-5	7	26.4	1.1	44.1	49.5	9.7	45.4	40.8	4.2	0.1	-	1.4	5.7
S04-6	7	27.4	1.1	45.7	51.2	10.1	45.5	40.2	4.2	0.1	51.0	1.5	5.7
S04-7	7	28.4	1.1	44.7	49.7	9.2	40.8	34.8	15.3	0.2	-	1.4	4.9
S04-8	8	30.3	1.1	45.3	50.8	10.0	44.5	39.4	6.1	0.2	-	1.4	5.6
S04-9	8	31.3	1.5	42.5	47.8	9.2	38.7	32.7	19.4	0.1	-	1.5	4.6
S04-10	-	34.8	0.8	42.9	48.6	10.0	42.3	35.6	12.1	0.2	-	1.5	5.2
S04-11	-	56.7	0.8	44.4	51.2	12.3	38.0	39.7	10.1	0.2	-	1.5	5.0
S04-12	9	58.4	1.2	44.7	50.9	11.3	42.9	41.6	4.2	0.2	38.0	1.5	5.7
S04-13	9	59.6	1.2	43.4	49.5	10.7	41.3	36.2	11.8	0.2	-	1.5	5.2
S04-14	10	86.8	1.1	39.6	45.2	9.3	37.7	31.3	21.8	0.3	47.0	1.5	4.6
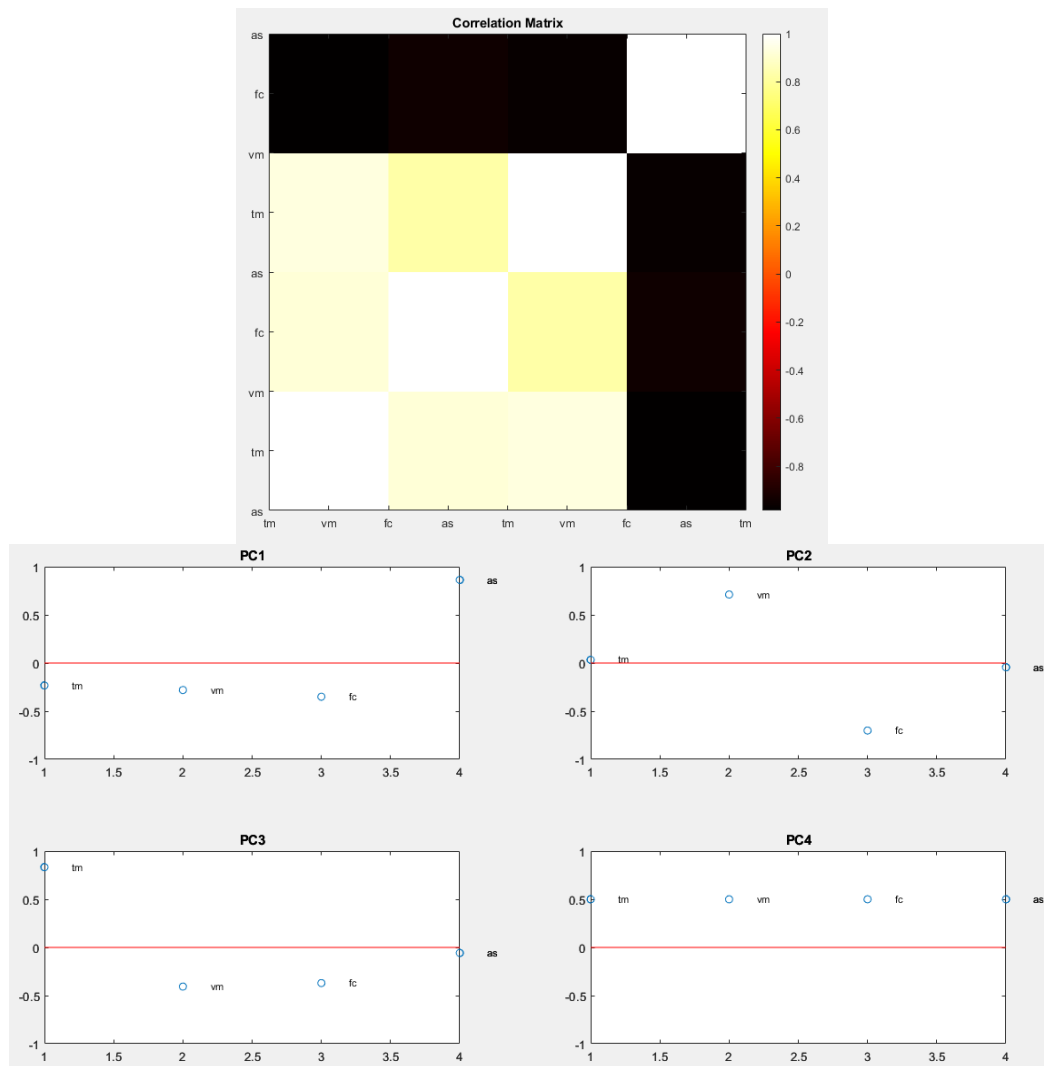
Figure 5. The Correlation Matrix (top) and Principal Component Analysis (bottom).

4. Results and discussion

The multivariate analysis encompasses statistical methods that analyze multiple measurements (variables) for each object simultaneously [15]. It can be categorized into independent variables (variable interdependence) and dependent variables (variable dependency). Principal Component Analysis (PCA) is employed to reduce the dimensions of a data set with numerous dependent variables of the data. It accomplishes this by transforming the data into a smaller set while preserving uncorrelated data that contribute significantly to the variation of the original data [14].

By applying PCA to the coal analysis in this study, assumptions are made to simplify and provide clear results. These assumptions are based on linearity, which simplifies the problem in two ways: by limiting the potential set of bases and assuming the data set is continuous [14]. The primary goal of PCA is to transform the linear matrix P, derived from the m x n X matrix data set, into new data represented on a Y basis. This transformation involves finding the coefficients of the principal components that capture the most significant variation in the X data set [14].

During the analysis, a correlation matrix is generated to illustrate the correlation coefficients between the variables. This matrix serves multiple purposes, including summarizing the data, providing input for more advanced analyses, and serving as a diagnostic tool [13]. It allows researchers to gain insights into the relationships between different variables and identify potential patterns or associations.
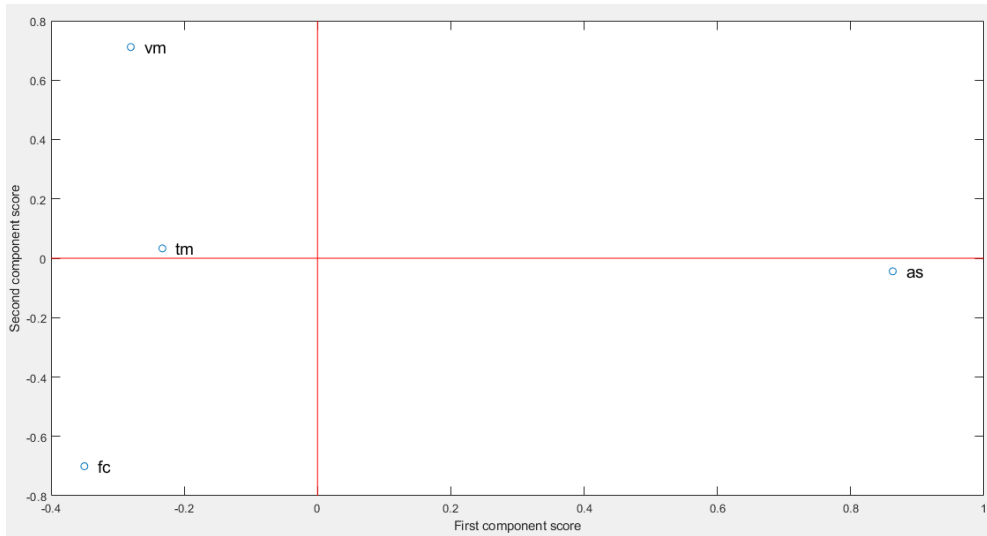
Figure 6. First component score vs. second component score.

The results obtained from PCA analysis reveal the presence of four distinct quadrants in the processed coal data. Quadrant I indicated that coal within this quadrant is primarily influenced by Volatile Matter and Total Moisture. Quadrants II and III show that coal in these quadrants is primarily affected by Ash content. On the other hand, Quadrant IV represents coal samples that are predominantly impacted by Fixed Carbon [4]. These findings provide valuable insights into the interrelationships between the different parameters and help in understanding the factors that contribute to variations in coal quality.

In the proximate analysis of the coal samples, which involved examining parameters such as total moisture, volatile matter, and fixed carbon, the observed differences were relatively insignificant, except for the ash content [16]. The variation in ash content serves as an indicator of impurity minerals that accumulate within the coal seam, thus exerting an influence on the overall quality of the coal. Furthermore, the value of fixed carbon is influenced by the levels of volatile matter, ash content, and moisture. Higher concentrations of volatile matter, ash, and moisture correspond to lower fixed carbon content. Volatile matter, consisting of flammable gases like carbon monoxide and methane, plays a pivotal role in the combustion process. Coal with a higher volatile matter content tends to burn more rapidly, resulting in a lower coal rank [16-17].
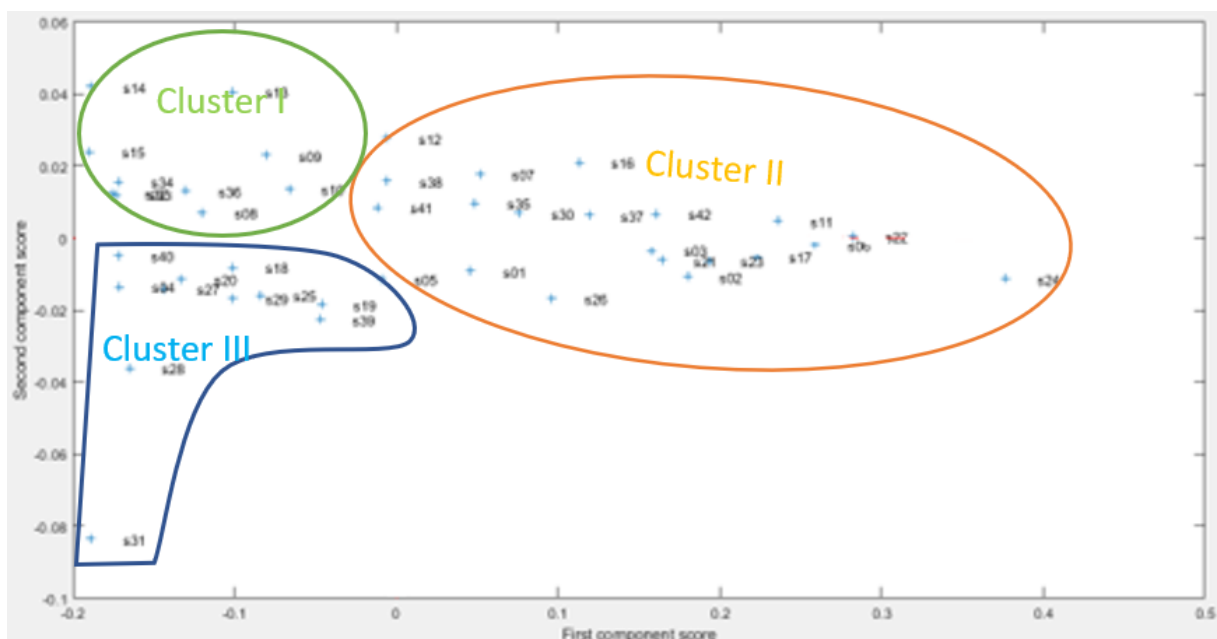
Figure 7. The clustering results.

These findings offer valuable insights into the interrelationships between different parameters in coal analysis and their implications for coal quality and utilization. By employing multivariate analysis, specifically Principal Component Analysis (PCA), researchers are able to gain a deeper comprehension of the intricate dynamics among various coal properties. This empowers them to make well-informed decisions regarding coal quality evaluation, resource estimation, and the formulation of effective strategies for coal utilization [18-22]. The application of PCA and other multivariate techniques provides a comprehensive understanding of coal characteristics and aids in the development of optimal approaches for harnessing the potential of coal resources in a sustainable and efficient manner. Continued research and analysis in this field will contribute to advancements in coal science and the improvement of coal-related industries.

5. Conclusion

The utilization of Principal Component Analysis (PCA) on the coal dataset has yielded intriguing results, unveiling the existence of three distinctive clusters. Cluster 1 predominantly encompasses data characterized by elevated values of Volatile Matter (VM) and Total Moisture (TM). Cluster 2, on the other hand, is predominantly composed of data exhibiting high Ash content. Lastly, Cluster 3 showcases data points exhibiting notable values of Fixed Carbon (FC) [4].

The analysis further indicated that the majority of data points were concentrated within Clusters 1 and 2, signifying the prevalence of high VM, TM, and Ash values. This combination of parameters contributes to a diminished Fixed Carbon content and consequently leads to lower calorific values. Consequently, the coal ranking within the study area is relatively low, specifically classified as lignite to sub-bituminous.

These findings underscore the significance of incorporating VM, TM, Ash, and FC parameters in coal analysis due to their substantial impact on coal quality and calorific value. Understanding the distinct characteristics and ranking of coal holds paramount importance for facilitating effective utilization and informed decision-making processes within the coal industry.

Further investigations and comprehensive analyses can be conducted to explore additional factors that may influence coal quality and ranking, thereby enhancing our understanding of the coal deposits present in the study area. Such insights have the potential to contribute significantly to improved resource estimation, the development of enhanced utilization strategies, and the establishment of more efficient and sustainable coal-based processes [23], [24], [25]. By expanding our knowledge in these areas, we can pave the way for advancements in coal science and technology, promoting more responsible and optimized utilization of this vital energy resource.

References:
[1]   S. Mujiyanto and G. Tiess, "Secure energy supply in 2025: Indonesia's need for an energy policy strategy," Energy Policy, vol. 61, pp. 31-41, 2013.
[2]   S. A. Wibisono, E. P. Dwitama, and I. O. Prahesti, "Petrografi dan geokimia batubara di daerah pahirangan dan sekitarnya, kabupaten kotawaringin timur, provinsi kalimantan tengah," Buletin sumber daya geologi, vol. 14, no. 1, pp. 65-78, 2019.
[3]   R. Kurniawan and S. Managi, "Coal consumption, urbanization, and trade openness linkage in Indonesia," Energy Policy, vol. 121, pp. 576-583, 2018.
[4]   M. F. Septiana and A. Darul, "Perbandingan analisis statistik multivariat PCA pada batubara di Kab. Tabalong, Kalimantan Selantan."

[5] P. Bird, "An updated digital model of plate boundaries," Geochemistry, Geophysics, Geosystems, vol. 4, no. 3, 2003.

[6] R. W. Van Bemmelen, "The geology of Indonesia, vol. 1A," Government Printing Office, The Hague, 1949.

[7] H. Darman, "An outline of the geology of Indonesia," Lereng Nusantara, 2000.

[8] M. I. Kusuma and A. N. Nafi, "Prospek Hidrokarbon Formasi Warukin Di Cekungan Barito Kalimantan," Kumpulan Makalah Pertemuan Ilmiah Tahunan XIV IAGI, Jakarta, pp. 105-124, 1986.

[9] M. S. Siregar and R. Sunaryo, "Depositional environment and hydrocarbon prospects, Tanjung Formation, Barito Basin, Kalimantan," 1980.

[10] ASTM Committee E-20 on Temperature Measurement, "Manual on the use of Thermocouples in Temperature Measurement (Vol. 470)," ASTM International, 1981.

[11] B. J. Wood and K. M. Sancier, "The mechanism of the catalytic gasification of coal char: a critical review," Catalysis Reviews Science and Engineering, vol. 26, no. 2, pp. 233-279, 1984.

[12] P. Kearey (Ed.), "The Encyclopedia of the solid earth sciences," John Wiley & Sons, 2009.

[13] I. Olkin and A. R. Sampson, "Multivariate analysis: overview."

[14] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences, vol. 374, no. 2065, 2016.

[15] W. R. Dillon and M. Goldstein, "Multivariate analysis: Methods and applications," New York, NY: Wiley, 1984.

[16] E. Malaidji, A. Anshariah, and A. A. B. A. A. Budiman, "Analisis Proksimat, Sulfur, Dan Nilai Kalor Dalam Penentuan Kualitas Batubara Di Desa Pattappa Kecamatan Pujananting Kabupaten Barru Provinsi Sulawesi Selatan," Jurnal Geomine, vol. 6, no. 3, pp. 131-137, 2018.

[17] F. R. Widiatmoko, "Pendekatan analisa geokimia dengan multivariate analysis untuk mengetahui tipe mata air panas: studi kasus lapangan panas bumi mapos, nusa tenggara timur," Jurnal IPTEK, vol. 23, no. 2, pp. 71-78, 2019.

[18] M. Tripathi and S. K. Singal, "Use of principal component analysis for parameter selection for development of a novel water quality index: a case study of river Ganga India," Ecological Indicators, vol. 96, pp. 430-436, 2019.

[19] T. J. Algeo and J. Liu, "A re-assessment of elemental proxies for paleoredox analysis," Chemical Geology, vol. 540, p. 119549, 2020.

[20] K. J. Bergen, P. A. Johnson, M. V. de Hoop, and G. C. Beroza, "Machine learning for data-driven discovery in solid Earth geoscience," Science, vol. 363, no. 6433, p. eaau0323, 2019.

[21] C. Lei et al., "A comparison of random forest and support vector machine approaches to predict coal spontaneous combustion in gob," Fuel, vol. 239, pp. 297-311, 2019.

[22] L. Sun et al., "Levels, sources, and spatial distribution of heavy metals in soils from a typical coal industrial city of Tangshan, China," Catena, vol. 175, pp. 101-109, 2019.

[23] Z. Ni et al., "Research on the co-pyrolysis of coal slime and cellulose based on TG-FTIR-MS, artificial neural network, and principal component analysis," Fuel, vol. 320, p. 123960, 2022.

[24] H. Bi et al., "Influence of lignin on coal gangue pyrolysis and gas emission based on multi-lump parallel reaction model and principal component analysis," Science of the Total Environment, vol. 820, p. 153083, 2022.

[25] N. K. Mohalik et al., "TGA/DSC study to characterise and classify coal seams conforming to susceptibility towards spontaneous combustion," International Journal of Mining Science and Technology, vol. 32, no. 1, pp. 75-88, 2022.