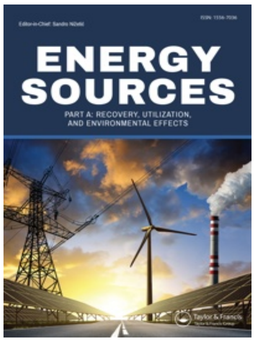

Energy Sources, Part A: Recovery, Utilization, and Environmental Effects

ISSN: 1556-7036 (Print) 1556-7230 (Online) Journal homepage: www.tandfonline.com/journals/ueso20

The Investigation of Coal Washing Plant Equipment by Using Multivariate Statistical Analysis

S. Yerel

Energy Sources, Part A, 32:1829–1836, 2010

ISSN: 1556-7036 print/1556-7230 online
DOI: 10.1080/15567030902883008

Taylor & Francis
Taylor & Francis Group

The Investigation of Coal Washing Plant Equipment by Using Multivariate Statistical Analysis

S. YEREL[1]

[1]Vocational School of Bozuyuk, Bilecik University, Bozuyuk/Bilecik, Turkey

Abstract *Coal washing plant equipment, regardless of its type, is frequently broken down in accordance with use and age. Maintenance and repair durations for broken equipment account for a large ratio among the total working period. In this study, the repair data for coal sieves used in a fine coal washing plant has been analyzed using multivariate statistical techniques, including cluster analysis, principal component analysis, and factor analysis. Through the results from the multivariate statistical analyses, it is determined that there are two groups of coal sieves according to their repair durations. These analyses are believed to assist the management of coal washing plants and determine priorities to improve plant issues.*

Keywords cluster analysis, coal sieves, factor analysis, multivariate statistical analyses, principal component analysis, repair durations

1. Introduction

A mine production system consists of many subsystems. The optimization of each subsystem in relation to one another is imperative to make the system profitable and viable for operation (Blischke and Murthy, 2003). The effectiveness of the coal washing plant equipment is mainly influenced by the availability, reliability, and maintainability of the system, and its capability to perform as expected. To maintain the designed reliability, availability, and maintainability characteristics and to achieve the expected performance, an effective maintenance program is a must. Effective maintenance is characterized by high equipment performance.

Maintenance is widely considered to be an effective strategy for reducing the number of system failures, thus lowering the overall maintenance durations (Okogbaa and Peng, 1996). Scheduled maintenance activities include equipment checks and partial or complete overhauls at specified periods. Therefore, a focus on maintainability is critical for the improvement of equipment performance (Barabady and Kumar, 2008).

Some studies conducted in the mining industry to improve the maintainability of different mining equipment by using statistical analysis are as follows: (1) Kumar et al. (1989) made a comprehensive estimate of the operational reliability of load haul dump machines, located items or assemblies that needed an improved design to enhance the reliability, and to decide the duration of the optimal preventive maintenance. (2) Ankara (1997) carried out a study in which variations of the Truck-Shovel System have been used to conduct availability analysis. (3) Hosseini (1999) discussed the equipment reliability issue and addressed the need for the analysis of various performance measures through

Address correspondence to Dr. Suheyla Yerel, Vocational School of Bozuyuk, Bilecik University, Bozuyuk/Bilecik, Turkey. E-mail: syerel@anadolu.edu.tr

"what-if" examination of various business conditions and operation scenarios. This was achieved by developing comprehensive reliability, availability, and maintenance optimization models, which were integrated with the operational characteristics of operating systems. (4) Samanta et al. (2004) presented the reliability, availability, and maintainability of a load haul dumper machine using repair data utilizing Markov modeling. (5) Yerel et al. (2007) determined a preventive maintenance policy for a mineral processing plant by using a mean quality control chart and the Kolmogorov-Smirnov statistical test. Their work indicated that a four-month preventive maintenance policy could be better than preventive maintenance performed once a year.

Cluster analysis (CA), principal component analysis (PCA), and factor analysis (FA) are multivariate statistical techniques used to depict the associations between two or more categorical variables (Reghunath et al., 2002; Simeonov et al., 2004; Kazi et al., 2009). The application of different multivariate statistical techniques help the interpretation of complex data matrices to better understand the systems studied, allows the identification of possible factors that influence the systems, and offers a valuable tool for the reliable management of coal washing plants as well as providing rapid solutions to problems.

This article is structured as follows: Section 2 introduces a basic concept and a methodology for multivariate statistical analysis. Section 3 presents a case study describing the multivariate statistical techniques used of coal sieves. Section 4, finally, concludes the manuscript.

2. Multivariate Statistical Techniques

2.1. Cluster Analysis

CA is an exploratory data analysis tool for solving classification problems. Its objective is to sort cases into groups or clusters, so that the degree of association is strong between members of the same cluster and weak between members of different clusters. Each cluster thus describes, in terms of the data collected, the class to which its members belong; and this description may be abstracted through use from the particular to the general class type (Einax et al., 1998; Kowalkowski et al., 2006).

In the case of CA, the similarities-dissimilarities are quantified through Euclidean distance measurements, the distance between two objects, i and j, is given as

$$d_{ij}^2 = \sum_{k=1}^{m} (z_{ik} - z_{jk})^2, \tag{1}$$

where d_{ij}^2 donates the Euclidean distance, z_{ik} and z_{jk} are the values of variable k for object i and j, respectively, and m is the number of variables (Sharma, 1996). Euclidean distance and the Ward method are used to obtain dendrograms.

2.2. Principal Component Analysis and Factor Analysis

PCA is designed to transform the original variables into new, uncorrelated variables, called the principal components, which are linear combinations of the original variables. A principal component provides information on the most meaningful parameters, which describes a whole data set, affording data reduction with a minimum loss of the original information (Helena et al., 1999; Shrestha and Kazama, 2007). The principal component can be expressed as

$$z_{ij} = a_{i1}x_{1j} + a_{i2}x_{2j} + a_{i3}x_{13} + \ldots + a_{im}x_{mj}, \tag{2}$$

where z is the component score, a is the component loading, x is the measured value of the variable, i is the component number, j is the sample number, and m is the total number of variables.

FA follows the principal component analysis. The main purpose of factor analysis is to reduce the contribution of less significant variables and to simplify even more of the data structure coming from principal component analysis. This purpose can be achieved by rotating the axis defined by principal component analysis according to well-established rules, and constructing new variables, also called varifactors. As a result, a small number of factors will usually account for approximately the same amount of information as do the much larger set of original observations (Shrestha and Kazama, 2007). The FA can be expressed as

$$z_{ji} = a_{f1}f_{1i} + a_{f2}f_{2i} + a_{f3}f_{3i} + \ldots + a_{fm}f_{mi} + e_{fi}, \tag{3}$$

where z is the measured variable, a is the factor loading, f is the factor score, e is the residual term accounting for errors or other sources of variation, i is the sample number, and m is the total number of factors.

3. Case Study

3.1. Dataset

In the fine coal washing plant, one cyclone pump engine, one cyclone pump, three cyclones, three coal sieves fed by the overflow of the cyclone (1, 2, and 3) and two coal sieves fed by the underflow of the cyclone (4 and 5) are utilized (Ankara et al., 2007). Daily production is largely affected by unexpected coal sieves breakdown and operation may be interrupted. In this study, multivariate statistical techniques, including CA, PCA, and FA, were performed to repair durations belonging to coal sieves. Multivariate statistical calculations were performed using the “SPSS for Windows” and “Minitab statistical software.” Coal sieves belonging to a heavy dense cyclone circuit in a fine coal washing plant are seen in Figure 1.

3.2. Results of CA

CA organizes sampling entities into discrete groups, such that within group similarity is maximized and among group similarity is minimized according to some objective criteria (McGarial et al., 2000). In this study, coal sieves belonging to a heavy dense cyclone circuit in a fine coal washing plant were examined by the use of CA and dendrogram generated. This method accepts the repair durations for each coal sieve that are in the same row as a cluster and evaluates the similarity of a cluster with the other clusters. The number of clusters for coal sieves was selected as 5 and the number of observations for these clusters was determined as 23 when using the Ward method. Then, for each cluster pair, distances belonging to coal sieves were calculated and distance matrices were formed. Cluster pairs that have the highest similarity were chosen and were joined

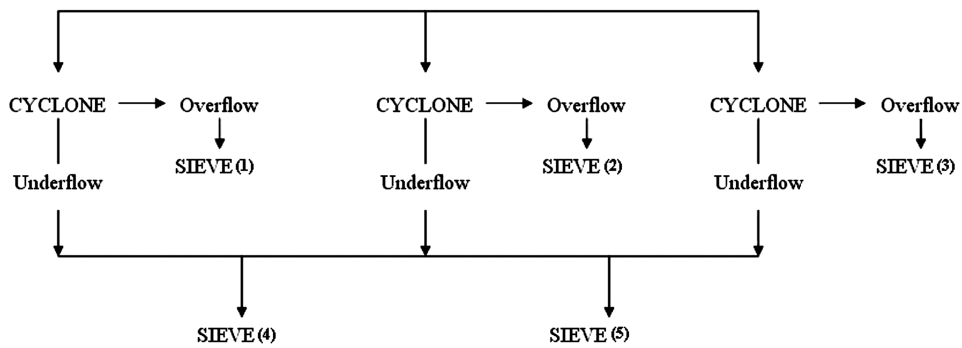

Figure 1. Coal sieves.

to form a new cluster. The steps were repeated five times for coal sieves and all clusters were joined in one cluster (Table 1).

A dendrogram was constructed using the clusters joined versus the distance values. The dendrogram shows that all the coal sieves may be generally grouped into two main clusters (Figure 2). Based on the results of CA, the following are concluded:

- Cluster I: This cluster is formed by coal sieves 1, 2, 3, and the cluster fed by the overflow of the cyclone circuit in the fine coal washing plant. Coal sieves 1, 2, and 3 show relatively similar performance and repair characteristics in comparison with the other cluster.
- Cluster II: This cluster consists of coal sieves 4, 5, and the cluster fed by the underflow of the cyclone circuit in the fine coal washing plant. These sieves (4 and 5) show relatively similar performance and repair characteristics.

3.3. *Confirmation of CA Results by PCA and FA*

PCA and FA were applied to the dataset to confirm the results of CA. The PCA was carried out by diagonalization of the correlation matrix, so the problem of different numeric ranges of the original variables was avoided since all variables were scaled to variance unit and contributed equally (Chen et al., 2007). A screen plot given in Figure 3 shows the sorted eigenvalues from large to small as a function of the PCA. There were several criteria to identify the number of principal components to be retained in order to understand the underlying data structure. Eigenvalues greater than 1 were taken as

Table 1
Distance level, clusters joined, and new cluster for coal sieves

Step	Number of clusters	Distance level	Clusters joined	New cluster	Number of new cluster
1	4	16,200.00	1–2	1	2
2	3	23,033.33	1–3	1	3
3	2	43,100.00	4–5	4	2
4	1	1.2223E+06	1–4	1	5

Figure 2. Dendrogram of the CA according to Ward.

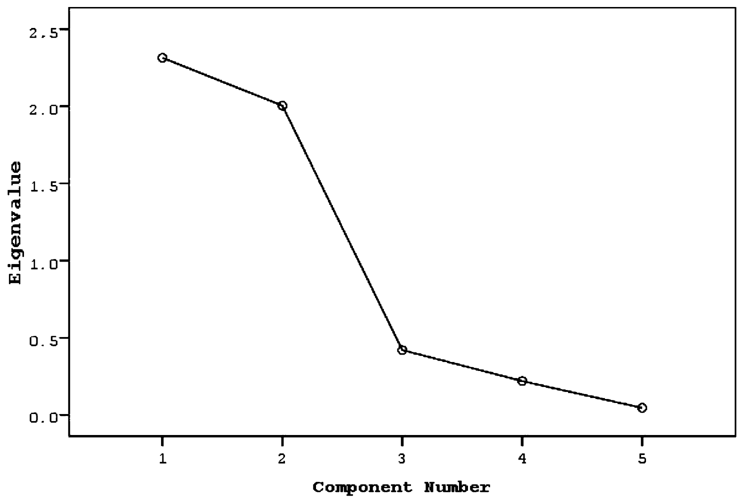

Figure 3. Screen plot of the eigenvalues.

criterion for extraction of the principal components required to explain the sources of variance in the data (Iscen et al., 2008). The principal component weights for the two components from the PCA of the dataset are given in Table 2. PCA weights revealed that PCA1 was formed with coal sieves 1, 2, and 3, the sieves fed by the overflow of the cyclone circuit in the fine coal washing plant. The second component (PCA2) was

Table 2
Principal component analysis (PCA) weights

Coal sieves no.	PCA 1	PCA 2
1	−0.583	
2	−0.525	
3	−0.579	
4		−0.672
5		−0.678

Table 3
Extracted values of various FA parameters for coal sieves

	Total variance explained before rotation			Total variance explained after rotation		
	Extraction sums of squared loadings			Extraction sums of squared loadings		
	Initial eigenvalues			Eigenvalues		
Component	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	2.31	46.28	46.28	2.30	45.94	45.94
2	2.00	40.08	86.36	2.02	40.42	86.36

Table 4
Factor loading matrix

Coal sieves no.	Factor 1	Factor 2
1	0.926	
2	0.851	
3	0.845	
4		0.984
5		0.986

associated with coal sieves 4 and 5, the sieves fed by the underflow of the cyclone circuit in the fine coal washing plant.

FA is applied to repair duration's datasets belonging to the coal sieves. The correlation matrix of variables was generated and factors were extracted by the centroid method, rotated by varimax. The results of FA including eigenvalues, variances, and the factor loading matrix are presented in Tables 3 and 4. This analysis led to the explanation of 86.36% of the variance in the data.

The factor loading for the two components from the FA of the dataset are given in Table 3. The first factor (Factor 1) explains 45.94% of the total variance and is related to the coal sieves 1, 2, and 3. While the coal sieves fed by the overflow of the cyclone circuit in the fine coal washing plant has strong positive loadings with this factor, Factor 1 shows the relation between the coal sieves 1, 2, and 3 and all the repair durations. Factor 2 accounts for 40.42% of the total variance and has strong positive loadings for coal sieves (4 and 5), which are fed by the underflow of the cyclone circuit in the fine coal washing plant. This factor represents the relation between the sieves (4 and 5) and all the repair durations. Thus, PCA and FA findings confirmed the results of CA.

4. Conclusions

Multivariate statistical methods can successfully be used to derive information from the variables about the possible influence of repair durations for coal sieves and also to determine groupings in the dataset. In this study, multivariate statistical techniques including cluster analysis (CA), principal component analysis (PCA), and factor analysis (FA) provided a presentation of the association between coal sieves and repair durations by giving a comprehensive view of the datasets for effective interpretation.

CA grouped five coal sieves into two clusters of similar repair durations. PCA and FA were used to explain the correlations between repair durations in terms of the underlying factors that are not directly observable. The results of PCA and FA findings confirmed the results of CA. Examining Cluster I and Cluster II for the dendrogram, coal sieves 1, 2, and 3, and coal sieves 4 and 5, respectively, were shown to have similar performances.

This study shows that multivariate statistical methods provide useful information for the coal washing plant managers in helping them plan their maintenance activities. These methods are believed to assist decision makers assessing coal washing plants maintenance activities in order to improve the efficiency of maintenance facilities.

References

Ankara, H. 1997. *Availability analysis for truck-shovel system in surface mining*. Ph.D. Thesis, Middle East Technical University, Turkey.

Ankara, H., Uygun, M., and Yerel, S. 2007. Comparison of screens in coal washing plant using the Kolmogorov-Smirnov test: Application of Omerler washing plant. *Proceedings of Maden Makinalari Sempozyuma*, Kutahya, Turkey, May 10–12, pp. 305–308.

Barabady, J., and Kumar, U. 2008. Reliabiliy analysis of mining equipment: A case study of a crushing plant at Jajrm Bauxide Mine in Iran. *Reliability Eng. & Sys. Safety* 93:647–653.

Blischke, W. R., and Murthy, D. N. 2003. *Case Studies in Reliability and Maintenance*. Hoboken, NJ: Wiley.

Chen, K., Jiao, J. J., Huang, J., and Huang, R. 2007. Multivariate statistical evaluation of trace elements in groundwater in a coastal area in Shenzhen, China. *Environ. Pollution* 147:771–780.

Einax, J. W., Truckenbrodt, D., and Kampe, O. 1998. River pollution data interpreted by means of chemometric methods. *Microchem. J.* 58:315–324.

Helena, B. A., Vega, M., Barrado, E., Pardo, R., and Farnandez, L. 1999. A casu of hydrochemical characterization of an alluvial aquifer influenced by human activities. *Water Air Soil Pollut.* 112:365–387.

Hosseini, M. 1999. Reliability revolution is a new shift in paradigm occurring. *Canadian Mining J.* 120:33–35.

Iscen, C. F., Emiroglu, O., Ilhan, S., Arslan, N., Yilmaz, V., and Ahiska, S. 2008. Application of multivariate statistical techniques in the assessment of surface water quality in Uluabat Lake, Turkey. *Environ. Monit. Assess*. 144:269–276.

Kazi, T. G., Arain, M. B., Jamali, M. K., Jalbani, N., Afridi, H. I., Sarfraz, R. A., Baig, J. A., and Shah, A. Q. 2009. Assessment of water quality of polluted lake using multivariate statistical techniques: A case study. *Ecotoxicol. & Environ. Safety* 72:301–309.

Kowalkowski, T., Zbytniewski, R., Szpejna, J., and Buszewski, B. 2006. Application of chemometrics in river water classification. *Water Res.* 40:744–752.

Kumar, U., Klefsjö, B., and Granholm, S. 1989. Reliability investigations for a fleet of load haul dump machines in a Swedish mine. *Reliability Eng. & Sys. Safety* 26:341–361.

McGarial, K., Cushman, S., and Stafford, S. 2000. *Multivariate Statistics for Wildlife and Ecology Research*. New York: Springer.

Okogbaa, O. G., and Peng, X. 1996. Methodology for preventive maintenance analysis under transient response. *Proceedings of the Annual Reliability and Maintainability Symposium*, Las Vegas, NV, January 22–25, pp. 335–340.

Reghunath, R., Murthy, T. R. S., and Raghavan, B. R. 2002. The utility of multivariate statistical techniques in hydrochemical studies: An example from Karnataka, India. *Water Res*. 36:2437–2442.

Samanta, B., Sarkar, B., and Mukherjee, S. K. 2004. Reliability modeling and performance analyses of an LHD system in mining. *J. South African Inst. Mining & Metallurgy* 104:1–8.

Sharma, S. 1996. *Applied Multivariate Techniques*. New York: John Wiley & Sons Inc.

Shrestha, S., and Kazama, F. 2007. Assessment of surface water quality using multivariate statistical techniques. A case study of the Fuji river basin, Japan. *Environ. Model. & Software* 22:464–475.

Simeonov, V., Simeono, P., and Tsitourido, R. 2004. Chemometric quality assessment of surface waters: Two case studies. *Chem. Eng. Ecol*. 11:449–469.

Yerel, S., Ankara, H., Konuk, A., and Ozdag, H. 2007. Preventive maintenance policy supported by a quality control chart and Kolmogorov Smirnov tests: Emet Colemanite Mineral Processing Plant, Turkey. *Minerals & Metallurgical Proc*. 24:152–156.