

Assessment of coal deposit using multivariate statistical analysis techniques

Suheyla Yerel Kandemir

To cite this article: Suheyla Yerel Kandemir (2016) Assessment of coal deposit using multivariate statistical analysis techniques, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 38:7, 1002-1006, DOI: [10.1080/15567036.2010.540635](https://doi.org/10.1080/15567036.2010.540635)

To link to this article: <https://doi.org/10.1080/15567036.2010.540635>



Published online: 18 Apr 2016.



Submit your article to this journal [↗](#)



Article views: 163



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)



Assessment of coal deposit using multivariate statistical analysis techniques

Suheyla Yerel Kandemir

Department of Civil Engineering, Engineering Faculty, Bilecik Seyh Edebali University, Bilecik, Turkey

ABSTRACT

Energy is one of the indispensable factors regarding the economic development of Turkey. Examination of calorific values in coal deposits has importance in the production of energy, and thus planning of the coal deposit. In this article, multivariate statistical analysis techniques, including cluster analysis and discriminant analysis, were applied to calorific values obtained from boreholes. Cluster analysis grouped borehole locations into two clusters based on the similarity of calorific values. Afterwards, discriminant functions were supported to cluster analysis and developed a linear discriminant function. Based on the locations of the boreholes, it was concluded that calorific values are a highly central part of the coal deposit. Thus, this article shows the usefulness of multivariate statistical analysis techniques for understanding spatial variation of coal deposits and effective deposit management.

KEYWORDS

Calorific value; cluster analysis; coal; discriminant analysis; multivariate statistical analysis techniques; Turkey

1. Introduction

Energy is one of the indispensable factors regarding the assurance of social prosperity and economic development of a country. The developing countries struggle to obtain energy sources reliable in the long term to complete their economic development, and the developed countries struggle to get them to keep their present prosperity levels. For many years now, coal continues its indispensable position among the other energy sources because of its production in more than 50 countries, the small effect it has from the price fluctuation, and its usage lifespan of more than 200 years. The countries that are aware of coal's benefits have investments in mining fields in the countries rich in coal reserves, and they continue investing considerable amounts of money (Yilmaz and Uslu, 2006; Yilmaz, 2009).

Turkey is situated at the meeting point of three continents (Asian, Europe, and Africa) and stands as a bridge between Asia and Europe. The country is located in southeastern Europe and southwestern Asia. Its size is 779,452 square kilometers. Turkey's population is 65 million. Further, 56% of the inhabitants of the country live in cities. The average annual income per capital is US \$2,970 (Demirbas, 2001; Balat, 2006).

As Turkey's economy has expanded in recent years, the consumption of primary energy has dramatically increased. Although Turkey's primary energy resources seem limited, indigenous energy production meets nearly 48% of the total primary energy demand (TPC, 2002; BCC, 2003). Turkey's main energy resource is coal, which has been produced for years domestically, and its share of the country's total energy consumption is about 24%. It is used mainly for power generation, cement production, and steel manufacturing (Akdeniz et al., 2002; Canyurt and Ozturk, 2008).

In this study, multivariate statistical analysis techniques were used to interpret the boreholes data collected from a coal deposit. The objectives of this article were:

- to determine similarity of the borehole locations using multivariate statistical analysis techniques;
- to investigate the spatial variations of calorific values of borehole locations; and
- to present the usefulness of the multivariate statistical analysis techniques to planning the coal deposits.

2. Multivariate statistical analysis techniques

2.1. Cluster analysis

Cluster analysis is an unsupervised pattern recognition method that divides a large group of cases into smaller groups of clusters of relatively similar cases, which are dissimilar to other groups (McKenna, 2003; Lattin et al., 2003; Yerel et al., 2010). These analyses are divided into several groups, such as hierarchical, non-hierarchical, fuzzy, medoid, regression clustering, etc. (Praus and Praks, 2010). Hierarchical clustering is the most common approach in which clusters are formed sequentially (Johnson and Wichern, 2002; Alvin, 2002; Alkarkhi et al., 2009). Hierarchical clustering analysis was carried out by Ward's method, an extremely powerful grouping mechanism. In this article, hierarchical clustering analysis was substantiated using Ward's method.

The similarities are quantified through square Euclidean distance measurements. This distance, among the pairs of clusters, is calculated using Eq. (1) (Everitt, 1993):

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

where d_{ij}^2 denotes the square Euclidean distance, z_{ik} and z_{jk} are values of variable k for object i and j , respectively, and m is the number of variables. The square Euclidean distances and the Ward's method are used to construct a dendrogram. Dendrograms in cluster analysis provide a useful graphical tool determining the number of clusters, which describe the underlying process that leads to spatial variation (Boyacioglu and Boyacioglu, 2007).

2.2. Discriminant analysis

Discriminant analysis involves the derivation of linear combinations of the two or more independent variables that will discriminate between the previously established groups. It operates on raw data and the technique constructs a discriminant function for each group. A simple linear discriminant function transforms on the original set of measurements on a sample into a single discriminant score (Ratha and Venkataraman, 1997; Singh et al., 2004; Sanchez Lopez et al., 2004). Furthermore, discriminant analysis helps in grouping samples sharing common properties (Shrestha and Kazama, 2007).

In summary, discriminant analysis is based on the linear model of the familiar matrix notation form given by Eq. (2):

$$D = \lambda^\lambda, \quad (2)$$

which may be expanded and is followed by Eq. (3):

$$D_t = \lambda_{t_0} + \lambda_{t_1}x_1 + \lambda_{t_2}x_2 + \dots + \lambda_{t_n}x_n, \quad (3)$$

where D_t is the predicted discriminant score for group t ; t is the number of groups differentiated by the t discriminant functions; x is the measured values of the p independent variables used to predict group membership (Smith, 2008; Boyacioglu and Boyacioglu, 2010).

3. Result and discussions

Seyitomer basin is located within the borders of the Kutahya province in western Turkey (Gulec et al., 2001). The basin is rich in coal deposits. In this article, coal deposits are investigated using the multivariate statistical analysis techniques, such as cluster analysis and discriminant analysis. In assessing coal deposits, calorific values have a significance parameter. In this study, calorific values were obtained from boreholes. Afterwards, the borehole locations were grouped using Ward's method, which is one of the cluster analysis techniques. This technique accepts the calorific values in borehole locations that are in the same row as a cluster and evaluates the similarity of a cluster with the other clusters.

The dendrogram resulting from Ward's method of measured data set is presented in Figure 1. The figure shows that the boreholes were formed into two clusters. Cluster 1 consists of 19 borehole locations. These borehole locations are located at the center part of the coal deposit. Cluster 2 is formed by 20 borehole locations.

Discriminant analysis was used to find linear discriminant functions that best separate the two known groups. Thus, discriminant analysis has provided information over cluster analysis in evaluation of the spatial differences between borehole locations. Hence, discriminant analysis shows that there are differences between two groups (Cluster 1 and Cluster 2), which are expressed in terms of eight coal samples. The linear discriminant functions of groups are given by Eqs. (4) and (5):

$$D_1 = -8.55 + 0.08(S1) + 0.15(S2) + 0.01(S3) + 0.27(S4) - 0.01(S5) + 0.08(S6) - 0.51(S7) + 0.12(S8), \quad (4)$$

$$D_2 = -7.83 + 0.02(S1) + 0.05(S2) + 0.21(S3) + 0.32(S4) - 0.45(S5) + 0.01(S6) - 0.12(S7) + 0.02(S8). \quad (5)$$

According to the equations, D_1 is presented as a linear discriminant function of cluster 1, D_2 is shown as a linear discriminant function of cluster 2, and $S1, S2, \dots, S8$ is presented as calorific values. Each of the samples obtained from the borehole were transformed to discriminant scores using these two linear discriminant functions. Predicted maximum values of linear discriminant functions were presented for that original group of borehole. Then, the mean of calorific values were used to compare the aspects of the variation in coal samples collected for 39 borehole locations. The mean of calorific values were shown in Figure 2. Based on the locations of the boreholes, it was concluded that calorific values are a highly central part of the deposit. Thus, the mean of calorific values supported the multivariate analysis techniques.

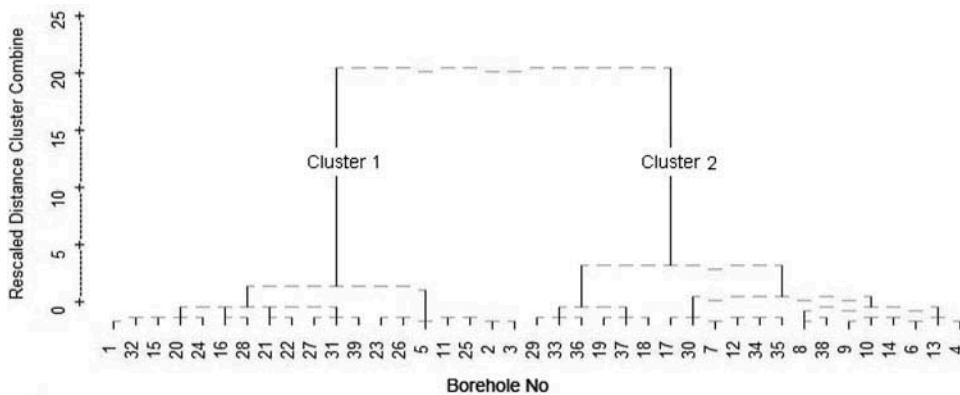


Figure 1. Dendrogram of Ward's method.

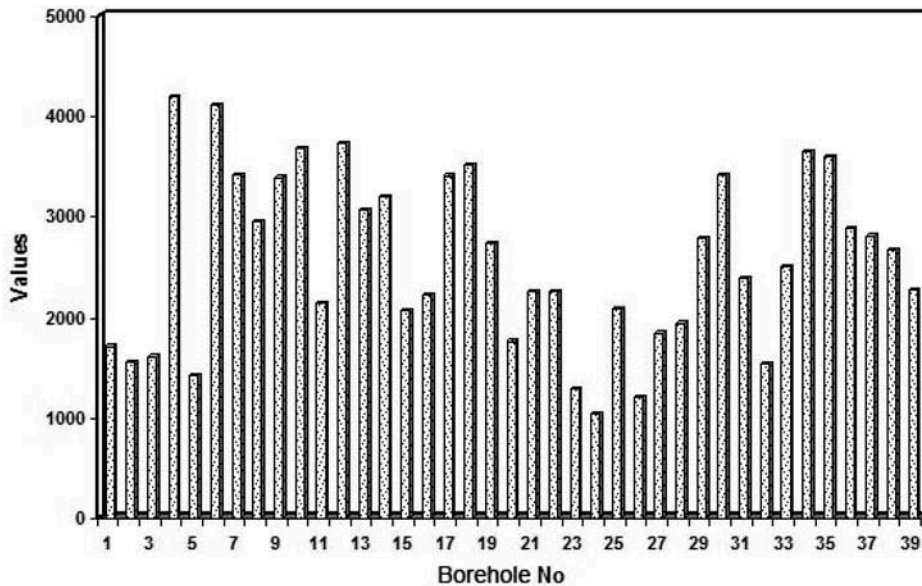


Figure 2. Borehole no. vs. mean of calorific values.

4. Conclusions

Multivariate statistical analysis techniques can successfully be used to acquire the relevant information from dataset. In this study, multivariate statistical analysis techniques, such as cluster analysis and discriminant analysis, were used to investigate coal deposits. Cluster analysis grouped the boreholes obtained from a coal deposit. These boreholes were separated into two different clusters with similar calorific values. Discriminant analysis was applied to develop the linear discriminant functions that best separate the two known clusters. In addition, mean of the calorific values were assessed according to borehole locations and it was determined that calorific values are highly central part of the coal deposit. Thus, this study illustrates the usefulness of multivariate statistical analysis techniques for the interpretation of coal deposits and understanding of variations in calorific values for effective deposit management.

References

- Akdeniz, F., Caglar, A., and Gullu, D. 2002. Recent energy investigation on fossil and alternative nonfossil resources in Turkey. *Energy Convers. Manage.* 43:575–589.
- Alkarkhi, A. F. M., Ahmad, A., and Easa, A. M. 2009. Assessment of surface water quality selected estuaries of Malaysia: Multivariate statistical techniques. *Environmentalist* 29:255–262.
- Alvin, C. R. 2002. *Methods of Multivariate Analysis*. New York: Wiley Inc.
- Balat, M. 2006. Turkey's coal potential and future appearances. *Energy Sources Part B* 1:137–146.
- Boyacioglu, H., and Boyacioglu, H. 2007. Surface water quality assessment by environmetric methods. *Environ. Monit. Assess.* 131:371–376.
- Boyacioglu, H., and Boyacioglu, H. 2010. Detection of seasonal variations in surface water quality using discriminant analysis. *Environ. Monit. Assess.* 162:15–20.
- Brinkster Communication Corporations (BCC). 2003. *Summary report of Turkish energy sector*. Available at: www.brinkster.com.
- Canyurt, O. E., and Ozturk, H. K. 2008. Application of genetic algorithm (GA) technique on demand estimation of fossil fuels in Turkey. *Energy Policy* 36:2562–2569.
- Demirbas, A. 2001. Energy balance, energy sources, energy policy, future developments, and energy investments in Turkey. *Energy Convers. Manage.* 42:1239–1258.
- Everitt, B. S. 1993. *Cluster Analysis*. New York: John Wiley & Sons Inc.

- Gulec, N., Gunal, B., and Erler, A. 2001. Assessment of soil and water contamination around an ash-disposal site: A case study from the Seyitomer coal-fired power plant in western Turkey. *Environ. Geol.* 40:331–343.
- Johnson, R. A., and Wichern, D. W. 2002. *Applied Multivariate Statistical Analysis*. London: Pearson Education International.
- Lattin, J., Carroll, D., and Green, P. 2003. *Analyzing Multivariate Data*. New York: Duxbury.
- McKenna, J. E. 2003. An enhanced cluster analysis program with bootstrap significance testing for ecological community analysis. *Environ. Model. Software* 18:205–220.
- Praus, P., and Praks, P. 2010. Hierarchical clustering of RGB surface water images based on MIA-LSI approach. *Water SA* 36:143–150.
- Ratha, D. S., and Venkataraman, G. 2007. Application of statistical methods to study seasonal variation in the mine contaminations in soil and groundwater of Goa, Indis. *Environ. Geol.* 29:253–262.
- Sanchez Lopez, F. J., Gil Garcia, M. D., Martiez Vidal, J. L., Aguilera, P. A., and Garrido Frenich, A. 2004. Assessment of metal contamination in Donana National park (Spain) using crayfish (*Procambarus clarkii*). *Environ. Monit. Assess.* 93:17–29.
- 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. Modell. Software* 22:464–475.
- Singh, K. P., Malik, A., Mohan, D., and Sinha, S. 2004. Multivariate statistical techniques for the evolution of spatial and temporal variations in water quality of Gomti River (India): A case study. *Water Res.* 38:3980–3992.
- Smith, S. 2008. *Discriminant Analysis Tutorial*. Available at: <http://marketing.byu.edu/htmlpages/tutorials/discriminant.html>.
- Turkish Petroleum Corporation (TPC). 2002. *Energy in Turkey*. Ankara, Turkey: Ajans-Turk Basin and Basim Co.
- Yerel, S., Ankara, H., and Ozbay, N. 2010. Determination of water quality parameters of Yesilirmak river, Turkey. *Asian J. Chem.* 22:3959–3965.
- Yilmaz, A. O. 2009. Present coal potential of Turkey and coal usage in electricity generation. *Energy Sources Part B* 4:135–144.
- Yilmaz, A. O., and Uslu, T. 2006. The role of coal in energy production—Consumption and sustainable development of Turkey. *Energy Policy* 35:1117–1128.