Menemui Matematik (Discovering Mathematics)
Vol. 34, No. 1: 49 – 56 (2012)

A Correlation Analysis of the Influence of Geothermal Environment to the Quality of Porcelain Insulator
(Analisis Korelasi terhadap Pengaruh Persekitaran Geothermal kepada Kualiti Isolator Tanah Liat)

Wan Nur Syahidah Wan Yusoff[1] & Maman Abdurachman Djauhari[2]

[1]Department of Sciences, Faculty of Industrial Sciences & Technology, UMP, 26300, Gambang, Pahang.
[2]Department of Mathematics, Faculty of Science, UTM, 81310, Skudai, Johor.

[1]syazul88@yahoo.com , [2]maman@utm.my

ABSTRACT

It is assumed that polluted environment reduces the quality of porcelain insulator in terms of leakage current. In this paper, we conduct an analysis to understand the effect of geothermal environment to the correlation structure among the variables representing the quality of porcelain insulator. Our approach is based on minimum spanning tree developed for this analysis in econophysics. An example will be presented and discussed.

Keywords: Correlation Structure, Degree Centrality, Dot Plot Matrix, Jennrich's Test, Minimum Spanning Tree

ABSTRAK

Diandaikan bahawa persekitaran yg tercemar akan mengurangkan kualiti isolator tanah liat dari segi kebocoran arus. Dalam makalah ini, kami mengendalikan satu analisis untuk memahami kesan persekitaran geothermal kepada struktur korelasi di antara pembolehubah yang mewakili kualiti isolator tanah liat. Pendekatan yang kami gunakan adalah berdasarkan minimum spanning tree yang dikembangkan untuk analisis ini dalam econophysics. Satu contoh akan dibentangkan dan dibincangkan.

Katakunci: Degree Centrality, Dot Plot Matrik, Minimum Spanning Tree, Struktur Korelasi, Ujian Jennrich

INTRODUCTION

Geothermal area as an electric power production has gained popularity in recent years, for example, it can be seen in Cid-Fernandez and Araujo (2007), Sogut *et al.* (2010) and Georgsson and Karlsdottir (2010). The transmission lines are used to transmit electric energy from geothermal power plant to a switchyard of distribution. Their performance depends on the insulation system and we assume that the performance of the insulator is influenced by the environmental pollution. In this regards, according to Forrest (1941), an important indication of the performance of an insulator is given by the leakage current. See also, for example, Reddy and Nagabhushana (2003) and Dixit and Gopal (2007) and some references therein.

Geothermal energy does have some environmental impact, most of which are associated with the exploitation of high temperature geothermal systems and toxicity of the waste fluid. Hunt (2000) mentioned that the gas CO_2 which is contained in the fluid can increase the temperature. Bw'Obuya (2002) also mentioned that oil, grease and diesel which are used extensively in the drilling at geothermal area will pose serious environmental problems when they leak. In this paper we study the influence of geothermal environment to the quality of porcelain insulator by analyzing the correlation structure among variables that determine the quality. According to Rivera (2007), temperature and humidity are among the parameters that can seriously affect the lifetime and reliability of that insulator. It is very sensitive to the environment where it operates.

In this paper, fifteen variables which determine the quality of insulator are studied and their correlation structures of clean insulator and polluted one are analyzed by testing the equality of the corresponding correlation matrices. There are many different methods available in the literature to test that equality. See, for example, Jennrich's test (Jennrich (1970), Tang (1998) and Schindler (2009)) and Box'M test (Tang (1998)). Here, we use the most commonly used method, namely, Jennrich's test (Jennrich (1970)). In case the two correlation matrices are not equal, we use minimum spanning tree (MST) to study how they differ to each other.

The rest of the paper is organized as follows. In the next section, we present the result of Jennrich's test and then, in Section 3, we use MST to analyze which variables that are strongly influenced by geothermal environment. This paper will be closed with the conclusion in the fourth section and recommendation in the last section.

TESTING THE EQUALITY OF CORRELATION MATRICES

Waluyo *et al.* (2010) has reported their experimental correlation matrix of clean porcelain insulator in Table 1 and of polluted one in Table 2. The tables are obtained based on $n_1 = 82$ and $n_2 = 39$ observations, respectively.

Table 1: Correlation matrix of clean porcelain insulator

Variables	Vmax	T	H	P	Imax	Pha	Pf	H1	H3	H5	H7	H9	H11	H13	THD
Vmax	1	-0.06	-0.07	0.12	0.73	0.22	-0.21	0.76	0.59	0.93	0.88	0.39	0.92	0.88	0.44
T	-0.06	1	-0.37	0.08	-0.29	0.24	-0.25	-0.28	-0.33	-0.20	-0.22	-0.37	-0.08	-0.09	0.31
H	-0.07	-0.37	1	0.10	0.56	-0.88	0.87	0.54	0.59	0.24	0.24	0.54	0.00	0.19	-0.83
P	0.12	0.08	0.10	1	0.14	-0.06	0.05	0.15	0.08	0.13	0.16	0.00	0.10	0.15	-0.03
Imax	0.73	-0.29	0.56	0.14	1	-0.47	0.47	1.00	0.94	0.92	0.91	0.80	0.74	0.81	-0.26
Pha	0.22	0.24	-0.88	-0.06	-0.47	1	-1.00	-0.43	-0.55	-0.11	-0.12	-0.60	0.13	-0.05	0.88
Pf	-0.21	-0.25	0.87	0.05	0.47	-1.00	1	0.43	0.56	0.11	0.12	0.61	-0.13	0.05	-0.89
H1	0.76	-0.28	0.54	0.15	1.00	-0.43	0.43	1	0.92	0.94	0.92	0.78	0.76	0.83	-0.22
H3	0.59	-0.33	0.59	0.08	0.94	-0.55	0.56	0.92	1	0.82	0.81	0.93	0.53	0.62	-0.38
H5	0.93	-0.20	0.24	0.13	0.92	-0.11	0.11	0.94	0.82	1	0.98	0.65	0.89	0.89	0.12
H7	0.88	-0.22	0.24	0.16	0.91	-0.12	0.12	0.92	0.81	0.98	1	0.67	0.86	0.86	0.08
H9	0.39	-0.37	0.54	0.00	0.80	-0.60	0.61	0.78	0.93	0.65	0.67	1	0.36	0.43	-0.49
H11	0.92	-0.08	0.00	0.10	0.74	0.13	-0.13	0.76	0.53	0.89	0.86	0.36	1	0.94	0.38
H13	0.88	-0.09	0.19	0.15	0.81	-0.05	0.05	0.83	0.62	0.89	0.86	0.43	0.94	1	0.23
THD	0.44	0.31	-0.83	-0.03	-0.26	0.88	-0.89	-0.22	-0.38	0.12	0.08	-0.49	0.38	0.23	1

Source: (Waluyo *et al.*, 2010)

Table 2: Correlation matrix of polluted porcelain insulator

Variables	Vmax	T	H	P	Imax	Pha	Pf	H1	H3	H5	H7	H9	H11	H13	THD
Vmax	1	0.22	-0.41	-0.26	0.23	0.27	-0.34	0.22	0.20	0.28	0.31	0.23	0.24	0.18	0.13
T	0.22	1	-0.85	-0.08	0.59	-0.24	0.25	0.56	0.43	0.79	0.86	0.77	0.85	0.45	-0.28
H	-0.41	-0.85	1	0.09	-0.33	-0.13	0.12	-0.31	-0.13	-0.58	-0.70	-0.69	-0.65	-0.33	-0.03
P	-0.26	-0.08	0.09	1	-0.09	0.01	-0.01	-0.09	0.00	-0.12	-0.11	-0.01	-0.08	-0.30	-0.03
Imax	0.23	0.59	-0.33	-0.09	1	-0.76	0.73	1.00	0.87	0.94	0.86	0.65	0.88	0.32	-0.78
Pha	0.27	-0.24	-0.13	0.01	-0.76	1	-1.00	-0.78	-0.67	-0.57	-0.43	-0.20	-0.51	-0.11	0.97
Pf	-0.34	0.25	0.12	-0.01	0.73	-1.00	1	0.74	0.62	0.55	0.41	0.19	0.50	0.12	-0.96
H1	0.22	0.56	-0.31	-0.09	1.00	-0.78	0.74	1	0.87	0.93	0.84	0.63	0.87	0.30	-0.80
H3	0.20	0.43	-0.13	0.00	0.87	-0.67	0.62	0.87	1	0.78	0.68	0.43	0.76	0.20	-0.63
H5	0.28	0.79	-0.58	-0.12	0.94	-0.57	0.55	0.93	0.78	1	0.98	0.82	0.97	0.46	-0.58
H7	0.31	0.86	-0.70	-0.11	0.86	-0.43	0.41	0.84	0.68	0.98	1	0.88	0.97	0.49	-0.44
H9	0.23	0.77	-0.69	-0.01	0.65	-0.20	0.19	0.63	0.43	0.82	0.88	1	0.82	0.37	-0.21
H11	0.24	0.85	-0.65	-0.08	0.88	-0.51	0.50	0.87	0.76	0.97	0.97	0.82	1	0.36	-0.52
H13	0.18	0.45	-0.33	-0.30	0.32	-0.11	0.12	0.30	0.20	0.46	0.49	0.37	0.36	1	-0.07
THD	0.13	-0.28	-0.03	-0.03	-0.78	0.97	-0.96	-0.80	-0.63	-0.58	-0.44	-0.21	-0.52	-0.07	1

Source: (Waluyo *et al.*, 2010)

In those tables, the variables that determine the quality of insulator consist of applied voltage amplitude (*Vmax*), temperature (*T*), humidity (*H*), pressure (*P*), leakage current amplitude (*Imax*), phase angle (*Pha*), power factor (*Pf*), seven types of relative humidity *H1*, *H3*, *H5*, *H7*, *H9*, *H11* and *H13*, and total harmonic distortion (*THD*). There are some correlations that are equal to 1 and 0. Since in those tables the significant number is until two digits behind decimal point, they must be read as close to 1 and 0, respectively.

Let P_1 and P_2 are the correlation matrices of clean and polluted insulators, respectively. To test the hypothesis $H_0 : P_1 = P_2$ versus $P_1 \neq P_2$, we use Jennrich's statistic (Jennrich (1970)),

$$J = \frac{1}{2} tr(Z^2) - dg'(Z) S^{-1} dg(Z) \qquad (1)$$

where

i. $Z = \bar{R}^{-1}(R_1 - R_2)\sqrt{\frac{n_1 n_2}{n_1 + n_2}}$,

ii. $\bar{R} = (\bar{r}_{ij}) = \frac{n_1 R_1 + n_2 R_2}{n_1 + n_2}$ is the pooled correlation matrix,

iii. $S = (s_{ij}) = (\delta_{ij} + \bar{r}_{ij}\bar{r}^{ij})$.

In (1), R_1 and R_1 represent sample correlation matrix of clean insulator and polluted insulator, respectively; δ_{ij} is the Kronecker delta, i.e., δ_{ij}=1 for $i = j$, otherwise $\delta_{ij} = 0$ and $\bar{r}^{ij}$ are the elements of $\bar{R}^{-1}$, the inverse of $\bar{R}$ Jennrich (1970) shows that the statistical test (1) is asymptotically χ^2 distributed with degree of freedom $k = p(p-1)/2$ where p is dimension of the correlation matrix. Therefore, $H_0 : P_1 = P_2$ is rejected at level of significance α if J exceeds, $\chi^2_{\alpha,k}$, the $(1-\alpha)$quantile of χ^2 distribution.

Based on the above tables, we obtain $J = 3737.175$ and $\chi^2_{\alpha,k} = 146.07$ for $\alpha = 0.05$. Evidently, we reject the null hypothesis which means that geothermal environment influence significantly the quality of porcelain insulator. Since J is too far greater than the critical point $\chi^2_{\alpha,k}$, we conclude that the influence is very strong in reducing the performance of porcelain insulator. In the next section, we analyze how those correlation matrices differ to each other and which variables that are strongly influenced.

INTERPRETATION AND DISCUSSION

Since the null hypothesis is rejected, by using MST (Mantegna (1999)), we analyze those two sample correlation matrices to explain why the two population correlation matrices are different. This analysis, in general, is started by transforming correlation matrix into distance matrix (Mantegna (1999)). Based on distance matrix, we contruct a MST, as suggested Kruskal Jr (1956), by using Kruskal's algorithm provided in Matlab. From MST, we construct the adjacent matrix to obtain the dot plot matrix and the network topology of all variables. To visualize that network, we use Pajek software (Ohta (2006) and Li and Ma (2008)). The interpretation of that network will be delivered by using the degree centrality measure (Xu *et al*. (2009) and Abbasi and Altmann (2011)).

Dot Plot Matrix

From Tables 1 and 2, by using Matlab version 7.8.0.347 (R2009a) we obtain dot plot matrix in Fig. 1(a) for clean insulator and Fig. 1(b) for polluted insulator. In Fig. 1, blank and dot cells represent the entries 0 and 1 of the adjacent matrix, respectively. In Fig. 1 we see three levels of correlations: low (black), high (green) and very high (red). However our concern is on the high and very high correlation. From Fig. 1, we learn that (i) *H5&Vmax*, *H11&Vmax*, *H5&H1*, *H13&H11*, *H9&H3*, *H5&Imax*, *H7&T*, *H1&Pf*, *H11&H5* and *H9&H7* are the correlations that occur only in one of the clean insulator or polluted one and (ii) correlation between *Pf* and *H* shifted from high correlation to very low correlation. From that figure, we also learn that the variables in clean insulator are more dispersed than the variables in polluted insulator. This justifies our conclusion that geothermal environment strongly influences the performance of porcelain insulator. This conclusion is general and will be clarify in the next sub section.

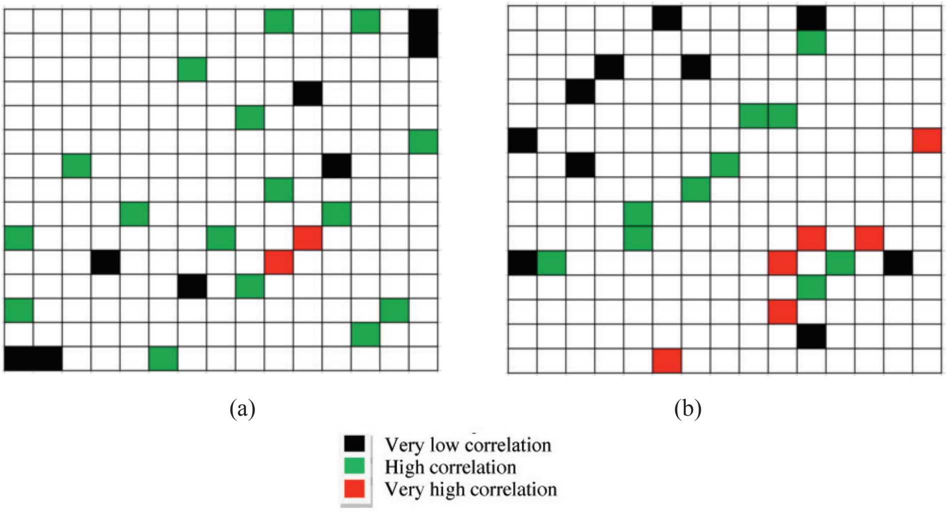
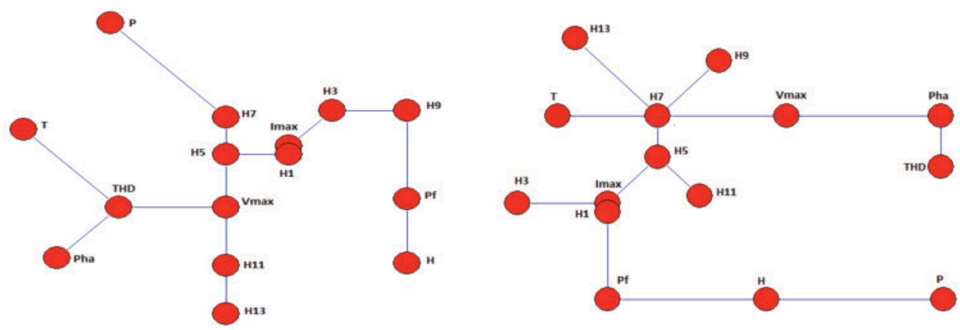


Figure 1: Dot plot matrix of clean porcelain insulator (a) and polluted one (b)

Minimum Spanning Tree

To elaborate the above finding more clearly, we present the corresponding MST and we come up with Fig. 2(a) for clean insulator and Fig. 2(b) for polluted insulator. These figures show the most important relationship among all variables in terms of MST.



(a) (b)

Figure 2 MST of clean porcelain insulator (a) and polluted one (b)

From these figures, we learn that all the first five variables in tables 1 and 2 except *Vmax* (applied voltage amplitude) and *Imax* (leakage current amplitude) are not separated from the remaining variables. This is true for clean as well as polluted insulators. This means that both variables are strongly influenced by the remaining variables.

Degree Centrality

Degree centrality indicates the connectivity of variables (nodes). It provides information on how many number of edges incident upon a given node. It can be used to measure the importance of any particular nodes. The application of degree centrality can be found in many areas of research. See, for example, Borgatti and Everett (2006) its application in social network and Xu et.al (2009) in E-Commerce. This measure is defined by (Park and Yilmaz (2010)),

$$C_{Degree}(N_i) = \sum_{j=1}^{p} a_{ij} \tag{2}$$

where α_{ij} is the element in *i*-th row and *j*-th column of an adjacent matrix N_i and is the *i*-th node.

In Table 3, we present the degree centrality of each variables for clean as well as polluted insulators. Based on this table, a more attractive MST can be contructed. This is showed in Fig. 3 where the size of each node corresponds to its degree centrality.

Table 3: Degree centrality

No.	Variable	Clean	Polluted
1	*Vmax*	0.214286	0.142857
2	*T*	0.071429	0.071429
3	*H*	0.071429	0.142857
4	*P*	0.071429	0.071429
5	*Imax*	0.142857	0.214286
6	*Pha*	0.071429	0.142857
7	*Pf*	0.142857	0.142857
8	*H1*	0.142857	0.142857
9	*H3*	0.142857	0.071429
10	*H5*	0.214286	0.214286
11	*H7*	0.142857	0.357143
12	*H9*	0.142857	0.071429
13	*H11*	0.142857	0.071429
14	*H13*	0.071429	0.071429
15	*THD*	0.214286	0.071429

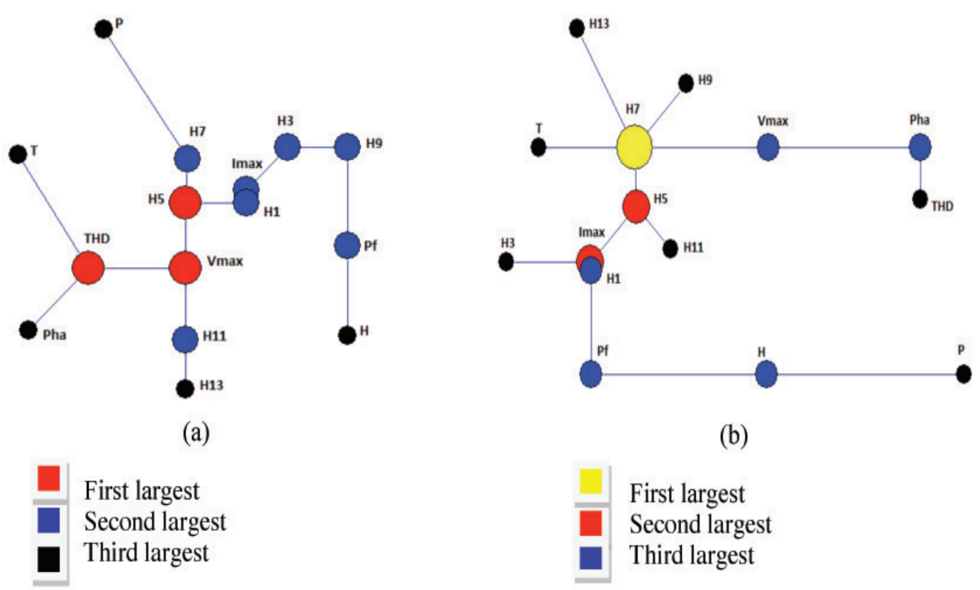


Figure 3 Degree centrality for clean porcelain insulator (a) and polluted one (b)

From Table 3 and Fig. 3, for clean insulator, nodes 1 (*Vmax*), 10 (*H5*) and 15 (*THD*) have the highest number of connections (3) in network. The rests are of 1 and 2 connections only. While, for polluted insulator, node 11 (*H7*) has the highest number of connections (5) in network. Followed by nodes 5 (*Imax*) and 10 (*H5*) have 3 connections, respectively. The rests are of 1 and 2 connections only. The higher the number of connections the more influential of a particular variable.

CONCLUSION

Dot plot matrices in Fig. 1, show that the correlation between *Pf* and *H* is shifted from high correlation to very low correlation. We also learn that the variables for clean insulator are more dispersed than the variables for polluted insulator. This means that geothermal environment strongly influences the performance of porcelain insulator.

From the MST in Fig. 2, we learn that *Vmax* (applied voltage amplitude) and *Imax* (leakage current amplitude) are not separated from the remaining variables either in clean or polluted insulators. This supports the claim in Forrest (1941) that leakeage current is an important indicator of the performance of an insulator. Furthermore, the variables *T*, *H* and *P* are well separated from the remaining variables either in clean insulator or polluted one. However, *T* and *P* have different roles in both types of insulator. In clean insulator, *T* and *P* is directly related with *THD* and *H7*, respectively, and in the polluted insulator, they relate with *H7* and *Pf*, respectively. This agrees with the result in Rivera (2007) that temperature is among the parameters that can seriously affect the lifetime and reliability of insulator.

According to degree centrality, the number of variables directly related to each of the following variables *H*, *Imax*, *Pha* and *H7* increases from clean to polluted insulators while those that relate with *Vmax*, *H3*, *H9*, *H11* and *THD* decrease. These variables are responsible for the inequality of clean and polluted insulators in terms of degree centrality.

RECOMMENDATION

The first five variables in Tables 1 and 2, i.e., Applied voltage amplitude (*Vmax*), temperature (*T*), humidity (*H*), pressure (*P*), and leakage current amplitude (*Imax*), which can be considered as independent variables, should be given special attention when porcelain insulator is used in geothermal environment.

ACKNOWLEDGEMENT

We acknowledge financial support from the Ministry of Higher Education, via FRGS vote numbers 78484 and 4F013. The first author would like to thank Universiti Teknologi Malaysia, Universiti Malaysia Pahang and Ministry of Higher Education for the opportunity to do this research. Special thanks go to the referees for their comment and suggestion.

REFERENCES

Abbasi, A. and Altmann, J. (2010). On the Correlation Between Research Performance and Social Network Analysis Measures Applied to Research Collaboration Networks. TEMEP Discussion Paper, No.2010:66, *Technology Management, Economics and Policy Papers*, Seoul National University

Borgatti, S.P. and Everett, M.G. (2006). A graph-theoretic perspective on centrality. *Social Networks*, 28: 466-484.

Bw'Obuya, N.M. (2002). The Socio-Economic and Environmental Impact of Geothermal Energy on the Rural Poor in Kenya. *AFREPREN Theme Group on Special Studies of Strategic Significance.*

Cid-Fernandez, J. A. and Araujo, P. A. (2007). Ves Characterization of a Geothermal Area in the Nw of Spain. *Electronic Journal of Environmental, Agricultural and Food Chemistry*, 6 (7): 2173-2177.

Dixit, P., and Gopal, H.G. (2007). Letter to the Editor: Variation of Leakage Current with SDD on an Artificially Polluted Porcelain Pin Insulator- An Experimental Study. *Electrical Power Components and Systems*, 35 (3): 359-365.

Forrest, J. S. (1941). The Characteristics and Performance in Service of High-Voltage Porcelain Insulators. *Central Electricity Board*, 89: 60-80.

Georgsson, L. S. and Karlsdottir, R. (2010). Resistivity Methods-DC and TEM with Examples and Comparison from the Reykjanes Peninsula and Oxarfjordur, Iceland. Presented at the *Short Course V on Exploration for Geothermal Resources, organized by UNU-GTP, GDC and Kengen*, Kenya.

Hunt, T. M. (2000). Geothermal and the Environment. *Geothermal Training Programme*, *United Nations University*, 1: 9-22.

Jennrich, R. I. (1970). Test for the Equality of Two Correlation Matrices. *Journal of the American Statistical Association*, 65 (330): 904-912.

Kruskal Jr, J.B. (1956). On the Shortest Spanning Subtree and the Travelling Salesman Problem. *Proceedings of the American Mathematical Society*, 7 (1): 48-50.

Li, J. and Ma, J. (2008). Complexity and Optimization Analysis of Spatial Network. *Proceedings of the 2008 International Seminar on Future Information Technology and Management Engineering*: 206-209.

Mantegna, R.N. (1999). Information and Hierarchical Structure in Financial Markets. *Computer Physics Communications, 121-122*: 153-156.

Ohta, J. (2006). Visualization of Metabolic Networks as Networks of Atoms by Pajek: An Application of Connectivity Matrix Method. *17th International Conference on Genome Informatics*, Japan.

Park, K. and Yilmaz, A. (2010). A Social Network Analysis Approach to Analyze Road Networks. *ASPRS Annual Conference*. San Diego, CA.

Rivera, M. A. (2007). Design Considerations for Reliable Electrical, Control and Instrumentation System in Geothermal Power Plants with Emphasis on Hydrogen Sulphide Related Problems. *Geothermal Training Programme*, *United Nations University*, 20: 461-489.

Reddy, B.S., and Nagabhushana, G.R. (2003). Study of Leakage Current Behaviour on Artificially Polluted Surface of Ceramic Insulator. *Plasma Science and Technology*, 5(4).

Schindler, F. (2009). Correlation Structure of Real Estate Markets Over Time. *Journal of Property Investment & Finance*, 27 (6): 579-592.

Sogut, A. R., Guzel, A., Zedef, V. and Bayram, A. F. (2010). Some Geological and Hydrogeochemical Characteristics of Geothermal Fields of Turkey. *Scientific Research and Essays*, 5 (20): 3147-3151.

Tang, G. Y. N. (1998). The Intertemporal Stability of the Covariance and Correlation Matrices of Hong Kong Stock Returns. *Applied Financial Economics*, 8: 359-365.

Woluyo, Sinisuka, N. I., Suwarno and Djauhari, M. A. (2010). Leakage Current and Pollutant Properties of Porcelain Insulators from the Geothermal Area. *ECTI Transactions on Electrical Eng. Electronics and Communications*, 8 (1): 126-145.

Xu, Y., Ma, J., Sun, Y., Hao, J., Sun, Y. and Zhao, D. (2009). Using Social Network Analysis as a Strategy for E-Commerce Recommendation. *Pasific Asia Conference on Information System (PACIS)*, India.