Islamic Azad University
Mashhad Branch

A. K. Verma et al./ Iranian Journal of Earth Sciences 2 (2010) / 32-38

Intelligent prediction of heating value of coal

A. K. Verma*[1], T. N. Singh[2] and M. Monjezi[3]

1. Center for Research on Energy Security, The Energy and Resources Institute, IHC Complex, Lodhi Road, New Delhi - 110 003, India
2. Department of Earth Science, Indian Institute of Technology, Powai, Bombay-76, India
3. Department of Mining Engineering, Tarbiat Modares University, Iran
Received 25 November 2009; accepted 25 February 2010

Abstract

The gross calorific value (GCV) or heating value of a sample of fuel is one of the important properties which defines the energy of the fuel. Many researchers have proposed empirical formulas for estimating GCV value of coal. There are some known methods like Bomb Calorimeter for determining the GCV in the laboratory. But these methods are cumbersome, costly and time consuming. In this paper, multivariate regression analysis and Co-active neuro-fuzzy inference system (CANFIS) backed by genetic algorithm technique is used for the prediction of GCV, taking all the major constituents of the proximate and ultimate analyses properties as input parameters and the suitability of one technique over the other has been proposed based on the results.
Correlations have been developed using multivariate regression analysis that are simple to use based on the proximate and ultimate analysis of data sets from 25 different states of USA because a very through study has been done and the data available is less variable. Also, CANFIS backed by genetic algorithm model is designed to predict the GCV of 4540 US coal samples from the abovementioned datasets. Optimization of the network architecture is done using a systematic approach (genetic algorithm). The network was trained with 4371, cross validation with 100, predicted with rest 69 datasets and the predicted results were compared with the observed values. The mean average percentage error in prediction is found to be negligible (0.2913%) and the generalization capability of the model was established to be excellent. A useful concept of sensitivity analysis is adopted to set the hierarchy of influence of input factors. The results of the present investigation provide functional and vital information for prediction of GCV of any type of coal in USA.

Keywords: *GCV, CANFIS, Neural networks, Genetic algorithm, Epoch, Hidden layer, Sensitivity analysis*

1. Introduction

Recently, the energy demands of the whole world is increasing and are mostly compensated from fossil based fuels such as fuel-oil, natural gas and coal. Among the fossil fuels, coal is one of crucial energy sources for many countries, which is converted to heat and electrical power by different technologies for our daily life requirements. Therefore, predicting coal quality is an important task and depends on the knowledge of its physical and chemical constitution. In order to determine the chemical composition of coals, proximate and ultimate analyses are usually used. With proximate analysis, the substances of moisture, ash, volatile matter and fixed carbon in the coal content are determined as weight percent. On the other hand, ultimate analysis specifies the elements of carbon, hydrogen, nitrogen, oxygen and sulphur [1].
Accurate measurement of coal gross calorific value are needed to define its energy content in relation to different requirements, e.g. defining its energy potential, finding efficiently usage area, exact determination of its price and properly design and operation of thermal systems [2].
When a data stream is analyzed using a neural network, it is possible to detect important predictive patterns that were not previously apparent to a non-expert. Thus, the neural network can act as an expert. A particular network can be defined using three fundamental components: transfer function, network architecture and learning law [3]. One has to define these components depending upon the problem to be solved. A network first needs to be trained before interpreting new information. Several different algorithms are available for training of neural networks, but the back-propagation algorithm is the most versatile and robust technique for it provides the most efficient learning procedure for multilayer neural networks. Also, the fact that back propagation algorithms are especially capable to solve problems of prediction makes them highly popular [4].
During training of the network, data are processed through the network until they reach the output layer (forward pass). In this layer, the output is compared to the measured values. The difference or error between the two is processed back through the network (backward pass) updating the individual weights of the

----
*Corresponding author.
E-mail address (es): amit.verma@teri.res.in

connections and the biases of the individual neurons. The input and output data are mostly represented as vectors called training pairs. The process as mentioned above is repeated for all the training pairs in the data set, until the network error has converged to a threshold minimum defined by a corresponding cost function, usually the root mean squared error (RMS) or summed squared error (SSE). A simple ANN network is shown in Fig.1.

2. Survey of GCV correlations and need for CANFIS-based models

Given et al. [5] used theoretical physical constants to develop a relation to determine the gross calorific value from elemental composition based on data from US coals (expressed in SI units):

$Q = 0.3278C + 1.419H + 0.09257S - 0.1379O + 0.637$ (MJ/kg) --- (1)

where, C, H, S, and O are on a dry, mineral-matter free basis, mineral matter is from the modified Parr formula, O is by difference, C is adjusted to a carbonate-free basis, and H is adjusted to exclude hydrogen in bound water present in clay minerals [5].

Regression analysis has been done on the data from 775 USA coals (with less than 30% dry ash) to develop an empirical equation by Mason and Gandhi [6] that estimates the calorific value of coal from C, H, S and ash (all on dry basis); expressed in SI unit their equation is:

$Q = 0.472C + 1.48H + 0.193S + 0.107A - 12.29$ (MJ/kg) --------- (2)

Cordero et al. [7] have given the correlation between HHV, NM and ASH as follows:

$HHV = 35.43 - 0.1835VM - 0.3543\ ASH$ (MJ/kg) ------ (3)

Akkaya [2] has given the following relationship using multiple non-linear regression analysis:

$HHV = 0.561\ M^{-6.137}\ VM^{0.381}\ FC^{0.666}$ (MJ/kg) ------------ (4)

Huang et al. [8] have given the following relationship using multiple linear regression analysis:

$HHV = -2737 - 160.43\ M + 266.76\ VM$ (MJ/kg) ------ (5)

According to S. Mesroghli et al. [9]:

$HHV = 37.77 - 0.647\ M - 0.387\ A - 0.089VM$ (MJ/kg) ----------- (6)

Majumder et al. [10]: have developed a new proximate analysis based correlation:

$HHV = -0.03\ A - 0.11M + 0.33\ V_M + 0.35\ F_C$ (MJ/kg) ---------- (7)

Patel et al. [1] and S. Mesroghli et al. [9] have used Artificial Neural Network (ANN) to predict GCV values and obtained satisfactory predictions. But it has its own limitation like 'black box' nature and local minimization. So, there is a need to develop a technique which can also include the uncertainties of ultimate and proximate analysis of coal. During the 1980s a new technique of genetic algorithms (GAs) based on stochastic optimization techniques have been applied to a number of engineering problems [11]. The application of optimization procedures and probabilistic methods to mining problems is discussed by a number of researchers including (Anderson et al. [12], Verma and Singh [13]; Singh et al. [14], Singh et al. [15, 16]).CANFIS stands for coactive neuro-fuzzy inference systems. The CANFIS model integrates fuzzy inputs with a neural network to speedily solve poorly defined problems.

3. Approach

Data used in this paper to predict GCV have been taken from U.S. Geological Survey Coal Quality (COALQUAL) database, open file report 97-134 [17]. The samples with more than 50% ash as well as the samples with proximate and/or ultimate analysis different from 100 were excluded from the database. A total of 4540 set of coal sample analysis were used. The number of samples and range of GCV for different states of USA are shown in Table1. Proximate and ultimate analysis of the coal samples from different states have been shown in Table 2.

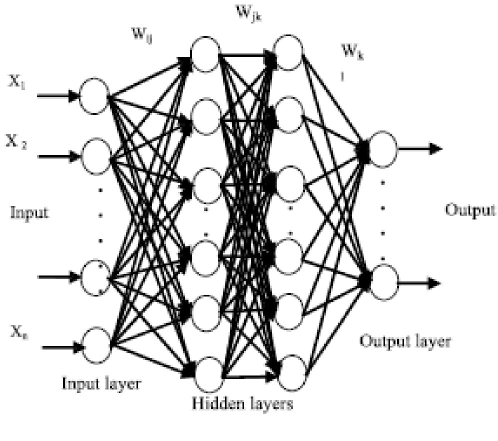

Fig. 1. A typical ANN model

Table 1. Ranges of GCV for different states of USA

Sr. No	State	Range of GCV (MJ/kg)
1	Alabama	6.05–34.80
2	Alaska	8.65–27.42
3	Arizona	18.54–24.36
4	Arkansas	5.57–34.68
5	Colorado	7.24–33.81
6	Georgia	24.03–34.85
7	Indiana	19.23–28.96
8	Iowa	16.03–26.59
9	Kansas	20.78–28.86
10	Kentucky	18.68–34.03
11	Maryland	23.04–33.48
12	Missouri	22.83–28.63
13	Montana	5.55–20.63
14	New Mexico	8.81–32.15
15	North Dakota	4.85–13.61
16	Ohio	16.43–31.14
17	Oklahoma	23.89–33.31
18	Pennsylvania	13.58–33.10
19	Tennessee	24.61–33.48
20	Texas	9.54–27.74
21	Utah	4.82–30.14
22	Virginia	19.49–34.80
23	Washington	13.14–27.45
24	West Virginia	14.29–34.75
25	Wyoming	6.27–34.23

Table 2. Ranges of proximate and ultimate analysis of coal samples

Variable	Minimum	Maximum	Mean	Std. Deviation
Moisture	0.4	49.6	8.09	9.9
Volatile matter	3.8	55.7	32.3	6.32
Ash	0.9	32.9	10.84	5.97
Hydrogen	1.7	8.1	5.27	0.69
Carbon	24.1	89.6	65.72	12.02
Nitrogen	0.2	2.41	1.29	0.33
Oxygen	0.9	54.7	14.86	11.27
Sulfur	0.07	17.3	1.9	1.73
Hydrogen exclusive	0.19	5.86	4.36	0.79
Oxygen exclusive	0.09	22.14	7.5	3.27

3-1. Network Architecture for Co-active Neuro-Fuzzy Model

For Co-active Neuro-fuzzy model data from 4540 set of coal samples were used. The network was trained with 4371, validated with 100 and predicted with rest 69 datasets.

3-2. Co-active Neuro-Fuzzy adaptive network for Gross Calorific Value:

1. Number of input Processing Elements = 11 (genetically optimized)
2. Number of Membership functions for each input = 3
3. Type of membership functions for each input = Bell
4. Type of membership functions for each output = Bell
5. Number of output Processing Elements = 1
6. Number of output membership functions = 3
7. Number of hidden layers = 0

3-3. Parameter of Output Layer

8. Transfer function of output layer =Axon
9. Learning rule = Momentum
10. Learning rate

Step size = 0.5(Genetically optimized)
Momentum rate = 0.01(Genetically optimized).\
Supervised learning control

11. Termination criteria = MSE
12. Maximum epochs = 100 for each Run
13. Number of Runs = 10
14. Type of Fuzzy model = TSK
15. No. of training epochs: = 100

Number of training datasets: = 4371
Number of cross validation datasets: = 100
Number of testing datasets: = 69

4. Results and Discussion

4-1. Multivariable relationships of GCV with ultimate and proximate analysis parameters

By a least square mathematical method, the correlation coefficients of Moisture, Volatile matter, Fixed Carbon, Standard Ash, Hydrogen, Hydrogen exclusive, Carbon, Nitrogen, Oxygen, Oxygen exclusive and Sulfur with GCV, were determined to be -0.9236,0.1790, +0.8543,-.2330, -0.5064, -0.5064, +0.8504, +0.9773, +0.72269,-0.9273,-0.7130 and +0.1137respectively.

Fig.2 shows the regression analysis of proximate and ultimate analysis of coal samples. The R^2 value of 0.9952 shows that there is an overall good fit of data with 95% confidence level. Hence, the best-correlated multivariable equations, between the various mentioned parameters and GCV can be presented as following equations:

GCV (MJ/kg) = 91.4621 -0.0556 M+0.02800 V-0.9039 A-0.5687 C-0.6972 N-1.1252 O-0.8775 S R^2=0.995

4-2. CANFIS-based models for GCV estimation

The network was trained with 500 training epochs, 25 numbers of chromosomes and 50 numbers of generations. Average of minimum Mean absolute percentage errors (MSE) for training was found to be 0.000729582 while for cross validation it was found to be 0.000551403. Fig.3 shows the performance of the network during testing process and Fig.4 shows performance of network output with number of generations. Sensitivity analysis of each input has been done and shown in Fig.5.

4-3. Comparison of obtained CANFIS results with ANN

Some researchers have used Artificial Neural Network (ANN) as a predictive tool to determine GCV values [1,9].Although their results are satisfactory but it has been found that neurofuzzy modeling is more reliable and accurate than ANN because the former takes into account uncertainty on the system which is lacking in the later [16]. Table3 shows the comparison of results obtained by different researchers in the past. It can be concluded that CANFIS has superior predictive capability than any other intelligent models like ANN, MVRA etc

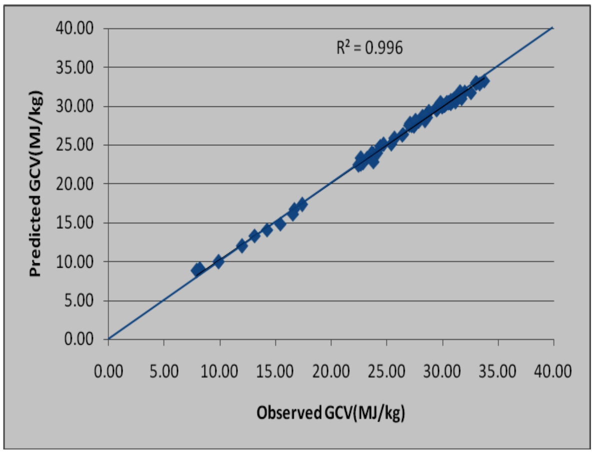

Fig. 2. Correlation coefficient between observed and predicted values of GCV (MJ/kg)

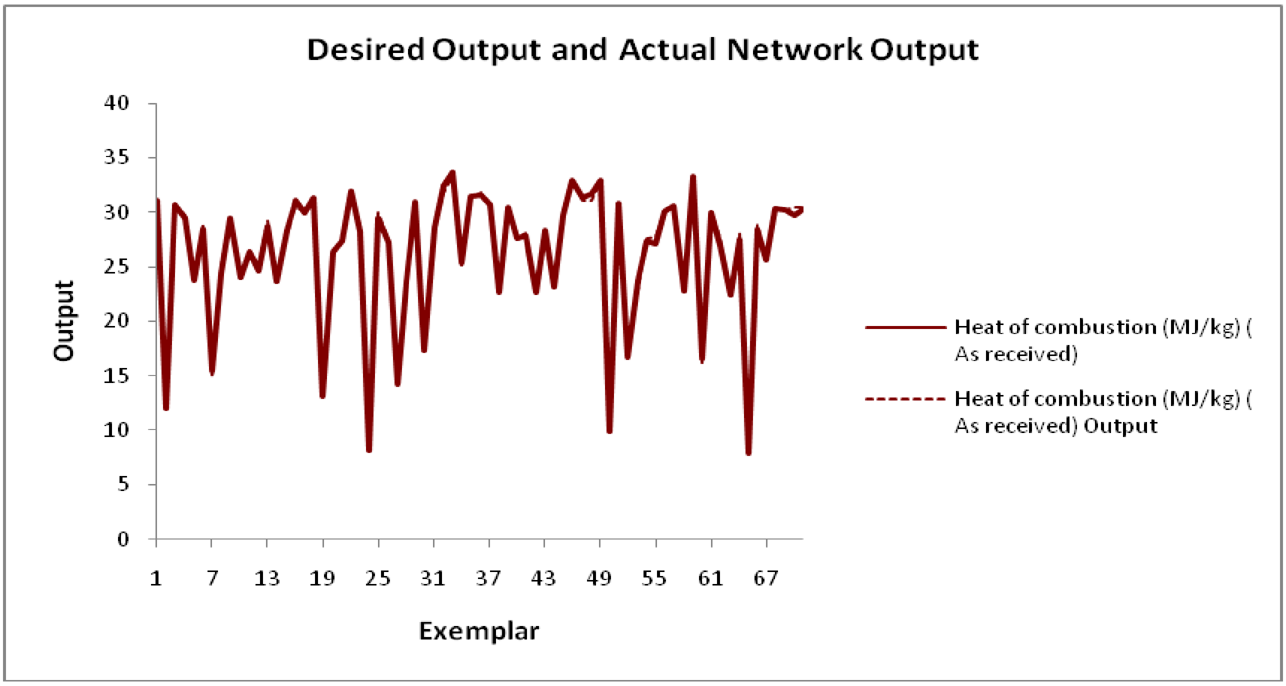

Fig. 3. Desired output and actual network output

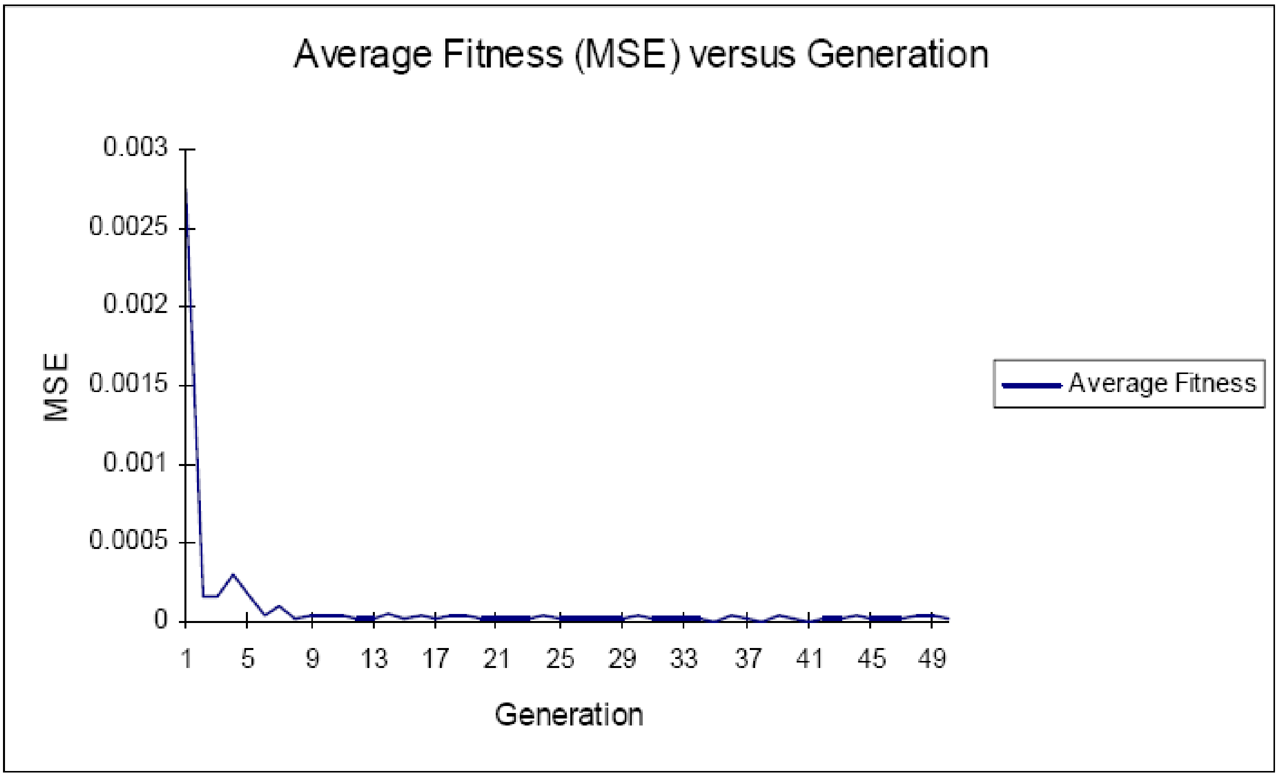

Fig. 4. Performance graph of mean square error versus no. of generations

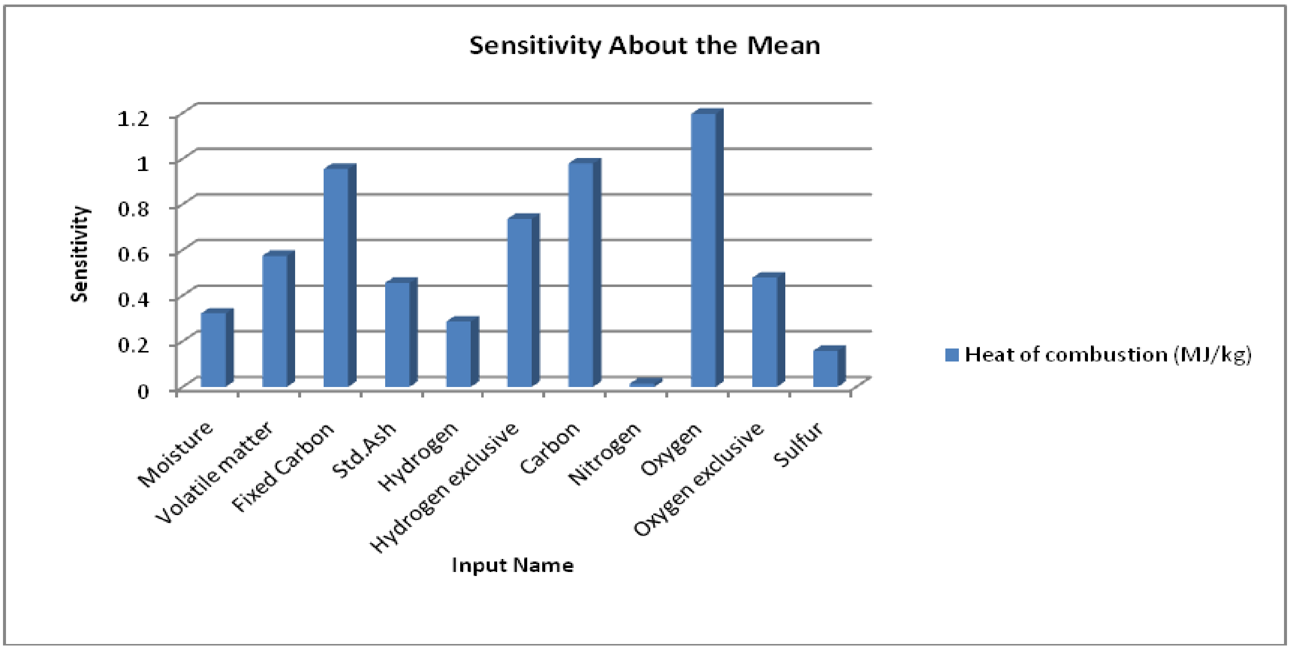

Fig. 5. Measure of the relative importance among the inputs of the neural model

Table 3. Comparison of the results with the previous researchers

Researcher	Type of Model	Coefficient of correlation
S.U. Patel et al. [1]	**ANN Model**	**0.987**
S. Mesroghli et al.[9]	**ANN Model**	**0.995**
Current Work	**CANFIS Model**	**0.996**

5. Conclusions

• A Co-active neuro-fuzzy inference system (CANFIS) optimized by genetic algorithm was used to estimate GCV (MJ/kg) of coals from ultimate and proximate analysis. The network employed simulates a fuzzy inference system. Hence, unlike ordinary neural networks there is nothing vague or unclear as to what happens inside the network. Based on the data from U.S. Geological Survey Coal Quality (COALQUAL), an interrelationship between the basic ultimate and proximate results and the heat content (GCV) could be acquired using CANFIS. It is also possible to understand the degree of influence of the input parameters used in this model on the output by conducting a sensitivity analysis. The method of genetic algorithms provides a powerful computational tool for any optimization problems that can be coded in string form.

• Sensitivity analysis about mean was performed using trained model in this study to determine hierarchy of influence of the input parameters over the output (GCV). Although, all the selected inputs have strong effect over the output but according to the hierarchy setup by sensitivity analysis oxygen content of coal had the maximum influence and nitrogen content had the minimum influence over GCV (MJ/kg).This also verifies the importance of oxygen to help in the combustion of coal. Also, nitrogen contributes very little because of its inactive nature. Genetic algorithm coupled with ANN has been found to be a better tool to predict the heat content of coal. The identification of GCV for coals by this method demonstrated that CANFIS optimized by genetic algorithm is the better alternative. This can be seen by the good coefficient of correlation between the observed and the predicted values. In virtue of this work, it is expected that the problem of determination of Gross Calorific Value (GCV) for different coals can be solved and this nonlinear modeling can be easily adopted in existing methods by simple modifications to cope up with time-consuming and cumbersome contemporary methods. The model presented in this paper shows a good potential to model complex, non-linear and multivariate problems; hence this approach should be extended to various other critical engineering problems, especially in coal engineering.

References

[1] Patel, S.U., Jeevan, K.B., Yogesh, P.B., Sharma B.K., Saha S. and Biswas, S., 2007, Estimation of gross calorific value of coals using artificial neural networks, Fuel, v. 86, p.334–344.

[2] Akkaya, E. and Demir, A., 2009, Energy content of Municipal Solid Waste by Multiple Regression Analysis. Proc. 5th International Advanced Technologies Symp., (IATS'09) - Karabuk, Turkey.

[3] Simpson, P.K., 1990, Artificial neural system-foundation, paradigm, application and implementation; New York: Pergamon Press.

[4] Maulenkamp F. and Grima, M. A. 1999, Application of neural networks for the prediction of the unconfined compressive strength (UCS) from Equotip hardness, J. Rock Mech. Min. Sci. v.36, p.29-39.

[5] Given, P.H., Weldon, D., Zoeller, J.H., 1986, Calculation of calorific values of coals from ultimate analyses: theoretical basis and geochemical implications. Fuel, v. 65, p.849–854.

[6] Mason, D.M., and Gandhi,K.N.,1983, Formulas for calculating the calorific value of coal and coal chars: Development, tests, and uses, Fuel Processing Technology, v.7, 1, p. 11-22

[7] Cordero, T., Marquez, F., Rodriquez-Mirasol, J., Rodriguez, J.J., 2001, Predicting heating values of lignocellulosic and carbonaceous materials from proximate analysis, Fuel, v.80, p.1567–1571.

[8] Huang C., Han, L., Liu. X., Yang, Z., 2008, Models Predicting Calorific Value of Straw from the Ash Content , International Journal of Green Energy, v. 5, (6) p.533 - 539

[9] Mesroghli, Sh., Jorjani, E., and Chehreh Chelgani, S.,2009, Estimation of gross calorific value based on coal analysis using regression and artificial neural networks, Int. J. of Coal Geology v.79, p.49–54.

[10] Majumder, A.K., Jain, R., Banerjee, J.P., Barnwal, J.P., 2008, Development of a new proximate analysis based correlation to predict calorific value of coal, Fuel, v.87, p.3077–3081.

[11] Goldberg, D.E., 1989, Genetic Algorithms in Search Optimization and Machine Learning. Addison-Wesley, p.412-430.

[12] Anderson, J., Shapiro, A. M. and Bear, J., 1984, A stochastic model of a fractured rock Conditioned by measured information. Water Resources Research, v. 20 (1) p. 79-88.

[13] Verma, A.K. and Singh, T.N., 2009, A Neuro-Genetic approach for prediction of compressional wave velocity of rock and its sensitivity analysis, Int. J. of Earth Sci. and Engg., v. 2, no.2, p.81-94.

[14] Singh T.N., Verma A.K, and Sharma P.K., 2007, A Neuro-Genetic approach for prediction of time dependent deformational characteristic of rock and its sensitivity analysis, Int. J. of Geotechnical and Geological Engg., v. 25, p.395-407.

[15] Singh, T.N., Kanchan R., Verma A.K. and Saigal K., 2005a, A comparative study of ANN and Neuro-fuzzy for the prediction of dynamic constant of rockmass", J. Earth Syst. Sci., v.114(1), p.75-86.

[16] Singh, T.N., Verma, A.K, Singh, Singh, V. and Sahu, A., 2005b, A neurofuzzy approach for the prediction of Slake Durability Index of some coal measure rocks, Int. J. of Environmental Geology, Springer Publication, v.47, No.2, p. 246-253.

[17] U.S.GEOLOGICAL SURVEY OPEN FILE REPORT 97-134; USGS COAL QUALITY (COALQUAL) DATABASE: VERSION 2.0