ADCHEM 2006

International Symposium on Advanced Control of Chemical Processes
Gramado, Brazil – April 2-5, 2006

A COMPARATIVE STUDY OF PREDICTION OF ELEMENTAL COMPOSITION OF COAL USING EMPIRICAL MODELLING

A. Saptoro, H.B. Vuthaluru and M.O. Tadé*

Department of Chemical Engineering, Curtin University of Technology
GPO BOX U1987, Perth, Western Australia 6845, Australia

Abstract: This paper presents empirical modelling approach in predicting elemental composition of coal. The model is developed to estimate carbon, hydrogen and oxygen content of coal. In the present work, several methods are applied to formulate the model including multiple regression (MR), principal component regression (PCR), partial least squares (PLS) and back propagation neural networks (BP-ANN). The use of BP-ANN shows the best result among the tested methods and appears to be a promising tool for predicting elemental composition of coal because it gave the least root mean square of error (RMSE) and the highest correlation coefficient (R^2).

Keywords: BP-ANN, elemental composition of coal, empirical modelling, multiple regression, PCR, PLS, ultimate and proximate analysis of coal

1. INTRODUCTION

Coal properties have many significant impacts on boiler operation and performance during coal combustion. Burning an unfamiliar fuel can reduce the efficiency of a power plant, increase pollutant emissions and, in some cases, actually damage the boiler or other system components. This can seriously affect the profitability and safety of a power plant. Power plant operators need to be confident that they can adequately know the coal properties and predict the consequences of using off-design or unfamiliar coals before they are fed into the boiler. Consequently, it is very important to provide coal properties data for power plant operators.

Unfortunately, there are some limitations of existing assessment of coal properties especially by using both conventional laboratory procedures and current on-line analysers. In laboratory procedure, chemical analysis was done on samples taken to the laboratory. Proximate analysis is relatively easy and quick to perform because it can be obtained using common laboratory equipment and is useful in practical application, however, it does not present detail information about the actual composition of coal. On the other hand, elemental analysis of coal requires highly trained analyst compared with proximate analysis, which only requires standard laboratory equipment and can be run by any competent scientist or even a skilled operator. Additionally, elemental analysis can take a day to obtain the results and in process control point of view, this traditional laboratory analysis of coal samples does not allow real time control if some adjustments need to be made to the system, for example controlling air to fuel ratio. Meanwhile, current on-line analysers of coal such as prompt gamma neutron activation analysis (PGNAA) are expensive. On-line analyser can cost from about £ 30,000 to £ 150,000 for a single parameter unit, with prices rising to as much as £ 400,000 for a prompt gamma neutron activation analysis (PGNAA) unit (Carpenter, 2002). Moreover, PGNAA still cannot be placed anywhere at the interface between the mine and the power plant (Yao et al., 2005). For these reasons, efforts must be directed to develop fast, reliable and inexpensive coal analyser which has capability for real time measurement / prediction.

One of the approaches is to establish suitable correlations to enable prediction of coal properties as a function of other available and easily obtainable coal properties. The mathematical models are developed using lab- and full-scale plant data as inputs to assess and predict other variables as on-line system outputs. Several variables exist in full scale and experimental data documented in power plant database and these data are used in the present work to develop process models. For determining elemental composition, proximate analysis can be used mainly due to its availability in power plant data base. However, due to great variability of coal properties, it is difficult to propose a suitable model, which can

* Corresponding author. Tel.: +61 8 92667581; fax: +61 8 92662681.
Email address: M.Tade@exchange.curtin.edu.au (M.O. Tadé)

represent the correlation between properties of coal. Coals are complex materials and can vary in qualities even from one mine to another or from one seam to another one. Therefore, extreme care must be taken in formulating a suitable model for representing the relationship between elemental composition and proximate analysis data.

To date, there is very limited work in the literature relevant to the elemental prediction of coal using proximate analysis data. To the best of our knowledge, only Yao et al. (2005) has developed a model for predicting hydrogen content and demonstrated the potential use of BP-ANN to tackle the difficulties in predicting elemental composition. This paper will present a comparative study of empirical modelling to predict carbon (C), hydrogen (H) and oxygen (O) content in coal using proximate analysis. Several methods are applied to formulate the model including multiple regression (MR), principal component regression (PCR), partial least square (PLS) and back propagation neural networks (BP-ANN).

2. THEORETICAL BACKGROUND

Several mathematical tools are considered to arrive at an empirical model which can predict elemental composition from proximate analysis. The description of several mathematical techniques and associated algorithms is presented below.

2.1 Multiple Regression (MR)

In simple linear regression, a dependent variable (y) is predicted from a single independent variable (x). In multiple regressions, a dependent variable is predicted from several independent variables. For predicting C, H and O content using proximate analysis (ash content (ash), volatile matter (VM), moisture content (MC) and fixed carbon (FC)), the models are formulated as follows:

$$C = \alpha_{0C} + \alpha_{1C} ash + \alpha_{2C} VM + \alpha_{3C} MC + \alpha_{4C} FC + \varepsilon \quad (1)$$

$$H = \alpha_{0C} + \alpha_{1C} ash + \alpha_{2C} VM + \alpha_{3C} MC + \alpha_{4C} FC + \varepsilon \quad (2)$$

$$O = \alpha_{0O} + \alpha_{1O} ash + \alpha_{2O} VM + \alpha_{3O} MC + \alpha_{4O} FC + \varepsilon \quad (3)$$

where $\alpha_o, \alpha_1, \alpha_2, \alpha_3 \; and \; \alpha_4$ are the model parameters and ε is the error term. As a note, sometimes, multiple regression models are developed involving interaction terms among independent variables to improve its prediction performance. Based on linear regression, one can estimate $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ and α_4 with reasonable accuracy. The estimates of $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ and α_4, will be denoted as $a_o, a_1, a_2, a_3, and \; a_4$. The predicted values of C, H and O using these estimates, will be further denoted as $\hat{C}$, $\hat{H}$ and $\hat{O}$ so that

$$\hat{C} = a_0 + a_{1C} ash + a_{2C} VM + a_{3C} MC + a_{4C} FC \quad (4)$$

$$\hat{H} = a_{0H} + a_{1H} ash + a_{2H} VM + a_{3H} MC + a_{4H} FC \quad (5)$$

$$\hat{O} = a_{0O} + a_{1O} ash + a_{2O} VM + a_{3O} MC + a_{4O} FC \quad (6)$$

To get estimates for $a_o, a_1, a_2, a_3, and \; a_4$, use the values of $\alpha_o, \alpha_1, \alpha_2, \alpha_3 \; and \; \alpha_4$ that result in minimum values of the sum of squared errors (SSE). In other words, if Ci, Hi and Oi is an observed value of C, H, and O, the values of $a_o, a_1, a_2, a_3, and \; a_4$ are obtained so that the following parameters are as small as possible (Berk and Carey, 2004; Draper and Smith, 1998; Brereton, 2003)

$$SSE_C = \sum_{i=1}^{N} (C_i - \hat{C}_i)^2 \quad (7)$$

$$SSE_H = \sum_{i=1}^{N} (H_i - \hat{H}_i)^2 \quad (8)$$

$$SSE_O = \sum_{i=1}^{N} (O_i - \hat{O})^2 \quad (9)$$

2.2 Principal Component Regression (PCR)

If **X** is the matrix of predictor / independent variables (proximate analysis: ash, VM, MC and FC) and **Y** is the matrix of response / dependent variables (elemental composition: C, H or O), principal components of **X** are constructed through principal component analysis (PCA) which can be expressed as follow

$$\mathbf{X} = \mathbf{U}\mathbf{V}^{\mathbf{T}} + \mathbf{E} \quad (10)$$

where **U** is **X**-scores ,**V** is **X**-loadings and **E** is **X**-residuals. Principal component regression (PCR) takes the scores from the decomposed **X** matrix and regresses them on the dependent data set, **Y** (Beebe et al., 1998; Martin et al, 1995; Brereton, 2003).

2.3 Partial Least Square (PLS)

If **X** is the matrix of predictor / independent variables (proximate analysis: ash, VM, MC and FC) and **Y** is the matrix of response / dependent variables (elemental composition: C, H or O), the correlation

between **Y** as function of **X** usually can be described as follows

$$\mathbf{Y} = \mathbf{X}\,\mathbf{b} \quad (11)$$

where vector b contains the model coefficient. The PLS model has the form

$$\mathbf{X} = \mathbf{U}\mathbf{V}^{\mathbf{T}} + \mathbf{E} \quad (12)$$

$$\mathbf{Y} = \mathbf{W}\mathbf{Z}^{\mathbf{T}} + \mathbf{F} \quad (13)$$

The matrices on the right-hand side of these models are defined by **U**=**X**-scores, **W**=**Y**-scores, **V**=**X**-loadings **Z**=**Y**-loadings, **E**=**X**-residuals, and **F**=**Y**-residuals. The final PLS prediction model can be re-expressed as

$$\hat{\mathbf{Y}} = \mathbf{X}\boldsymbol{\beta}^{\mathbf{T}} \quad (14)$$

$$\mathbf{B}^{\mathbf{T}} = \mathbf{P}(\mathbf{P}^{\mathbf{T}}\mathbf{P})^{-1}\mathbf{Q}^{\mathbf{T}} \quad (15)$$

where $\hat{\mathbf{Y}}$ are the predictions of **Y** and **β** are the regression coefficient vectors (Ramadhan, 2005; Martin et al., 1995; Brereton, 2003).

2.4 Back Propagation Artificial Neural Networks (BP-ANN)

Let **X** be a set of n input neurons, **Y** a set of m output neurons, ξ_i real input potential and y_i real output of neuron i. The neuron's output for this type of network is defined by the equation:

$$y_i = \sigma(\xi_i) \quad (16)$$

where the activation function, σ, maybe linear, threshold, sigmoid, hyperbolic tangent or radial basis function.

The network error E(w) related to a training set is defined as a sum of square errors $E_k(w)$ of network concerning each training example and depends on the configuration of network w.

$$E(w) = \sum_{k=1}^{p} E_k(w) \quad (17)$$

where $E_k(w) = 0.5\sum_{j\in Y}\left(y_j(w, x_k) - d_{kj}\right)^2$ (18)

Partial network error $E_k(w)$ related to the k-th Training example is directly proportional to a sum of square of difference between the real network value and the desired output where $y_j(w, x_k) - d_{kj}$ is the j-th output error for the k-th training sample. An error of zero would indicate that all the outputs pattern computed by the neural network perfectly match the expected values and the network is well trained (Haykin, 1999, Bulnová and Kostúr, 2003, Pham, 1995).

3. RESULTS AND DISCUSSION

Four methods are used to develop elemental composition predictor of C, H and O content in coal using proximate analysis. The total data set sizes are 167 x 4 for **X** and 167 x 1 for **Y**. The data from different countries and mines (Pisupati et a.l, 1992; Furimsky et al., 1990; Artos and Scaroni, 1993; Peralta et al., 2001; Coimbra et al., 1993; Fan et al., 1999; Visona and Stanmore, 1997; Armesto et al., 2003; Bailey et al., 1990; Peralta et al, 2002; Lockwood et al., 1998; Brewster et al., 1995; Su, 1999, Carlson, 1996; McLennan et al., 2000; Guo et al., 1997; Charland et al., 2003) were considered for developing and testing the models. 75 % of the data is used to develop the models and a quarter of the total data is used as independent data for testing the models. The range of data used for training and testing is presented in Table 1.

Table 1 Proximate analysis and elemental composition of coals (%)

	Average	Range
Proximate analysis		
Ash	11.33	0.5 - 40.4
VM	30.34	2.6 - 54.9
MC	6.76	0.1 - 36.8
FC	52.72	23.6 - 87.6
Elemental Composition		
C	74.07	40.6 - 94.6
H	4.45	0.4 - 6.7
O	10.61	0.2 - 38

3.1 Multiple Regression (MR)

The multiple regression model predicts C, H and O as a linear function of ash, VM, MC and FC. All the results, root mean square of error (RMSE) and coefficient correlation of the linear fit of measured and predicted (R^2) data are summarized in Tables 2, 3 and 4.

3.2 Principal Component Regression (PCR)

PCR is done for each possibility where we generate matrices **U** and **V** into one, two, three and four components. The result for each number of components (RMSE and R^2) is listed in Tables 2, 3 and 4.

3.3 Partial Least Square (PLS)

PLS is done for each possibility where we generate matrices **U**, **V**, **W** and **Z** into one, two, three and four components. The result for each number of components (RMSE and R^2) is shown in Tables 2, 3 and 4.

3.4 Back Propagation Artificial Neural Networks (BP – ANN)

Training process is done to develop suitable models for predicting C, H and O content. In this case,

Levenberg-Marquardt algorithm is used for the learning process and constructing the topology of BP-ANN containing input layer with four nodes (ash, VM, MC, and FC), one hidden layer and one output layer with one node (C, H, or O). To find the optimum number of neurons in the hidden layer, the number of neurons was changed from 2 up to 50 and arrived at optimum results which give minimum value of RMSE and maximum value of R^2. The 4-22-1 network for predicting C, 4-10-1 network for predicting H and 4-9-1 for predicting O give the most accurate prediction for this present study. The prediction performances are summarized in Tables 2, 3 and 4.

Table 2 Summary of Prediction Performance of C

Model		Training		Test
		RMSE	R^2	R^2
MR		6.259187	0.6915	0.6762
PCR	1 comp	6.71869	0.6568	0.7866
	2 comp	6.56509	0.6672	0.776
	3 comp	6.51491	0.6741	0.7678
	4 comp	6.40771	0.6874	0.6737
PLS	1 comp	6.69081	0.6546	0.7853
	2 comp	6.48453	0.6757	0.7494
	3 comp	6.3821	0.6875	0.6776
	4 comp	6.40771	0.6874	0.6737
BP-ANN		4.99097	0.879	0.91

Table 3 Summary of Prediction Performance of H

Model		Training		Test
		RMSE	R^2	R^2
MR		0.56654	0.705	0.3969
PCR	1 comp	1.24138	0.0063	0.0055
	2 comp	0.6664	0.6628	0.4262
	3 comp	0.61319	0.6628	0.4563
	4 comp	0.58019	0.7006	0.3942
PLS	1 comp	1.18171	5E-05	0.0142
	2 comp	0.63392	0.6436	0.4288
	3 comp	0.57805	0.7005	0.3998
	4 comp	0.58019	0.7005	0.3942
BP-ANN		0.44272	0.899	0.888

3.5 Discussion

From the results of each methods used in the present work, it is obvious that BP – ANN based predictor of C, H, and O content of coal show the best result compared with other methods (the prediction performance of BP-ANN is the best). Overall, it appears that BP-ANN based model is a valuable tool to assess the coal properties for any coal-fired power plant. The residual plots for prediction of C for each method are as shown in Figures 1, 2, 3 and 4. It can be seen that the Figure 4 for BP-ANN show the most random residual plot indicating the good fit of this model compared with the other three methods.

Table 4 Summary of Prediction Performance of O

Model		Training		Test
		RMSE	R^2	R^2
MR		4.273568	0.5902	0.5118
PCR	1 comp	7.59932	0.3684	0.5101
	2 comp	4.69999	0.5137	0.4032
	3 comp	4.47801	0.5608	0.5516
	4 comp	4.34347	0.5902	0.5115
PLS	1 comp	7.28019	0.5902	0.4428
	2 comp	4.55755	0.542	0.4474
	3 comp	4.32669	0.59	0.5186
	4 comp	4.34347	0.5902	0.5115
BP-ANN		2.33773	0.939	0.894

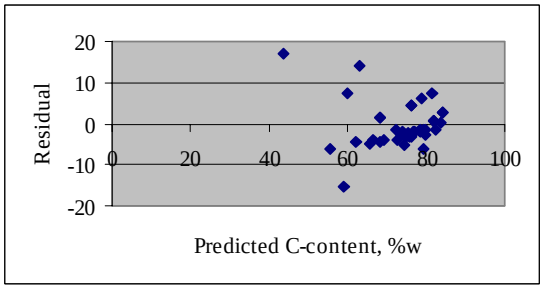


Fig. 1. Residual Plot for Prediction of C using MR

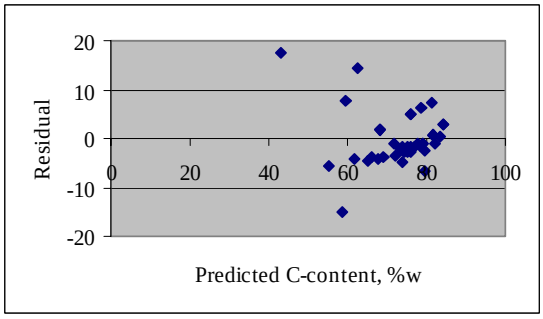


Fig. 2. Residual Plot for Prediction of C using PCR

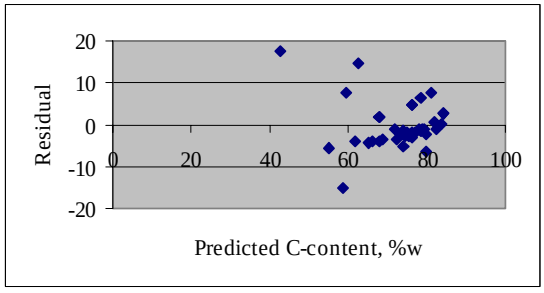


Fig. 3. Residual Plot for Prediction of C using PLS

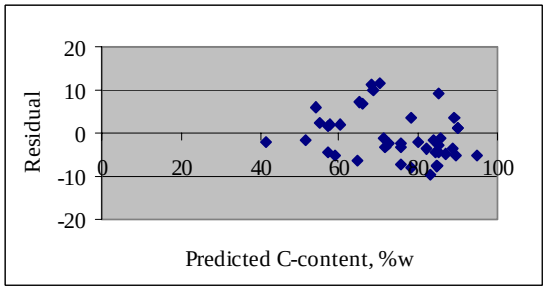
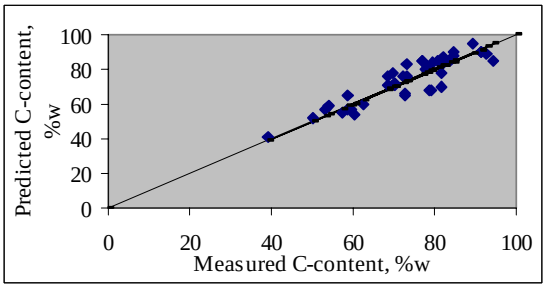


Fig. 4. Residual Plot for Prediction of C using BP-ANN

The linear fit of measured values of C, H, and O and their predicted values using BP-ANN based model are shown in Figures 5, 6 and 7. Predicted values are within ± 10 % from the measured elemental compositions.



.

Fig.5. Measured and Predicted Values of Carbon

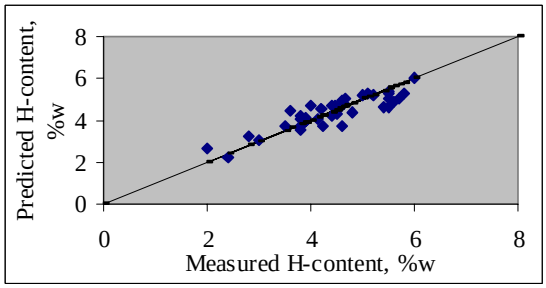


Fig.6. Measured and Predicted Values of Hydrogen

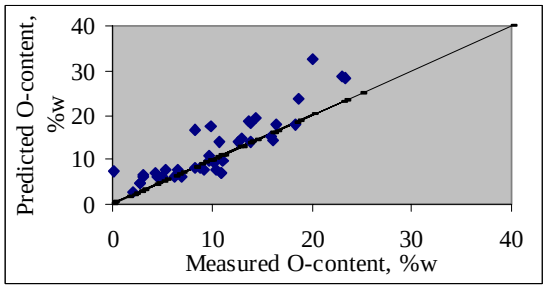


Fig. 7. Measured and Predicted Values of Oxygen

5. PRACTICAL IMPLICATIONS

Most of the coal-fired power stations receive coal from different mining operations. The coal, generally vary in quality even from the same coal mine. Moreover, due to the changing nature of the coal, off-line ultimate analysis may not be so accurate when the coal reaches the mill inlet. The challenge is to monitor the coal through the process and the quality of coal at the mill inlet is known so that the combustion can be appropriately controlled. Empirical model such as BP-ANN model would be a useful tool in this regard to provide the on-line information of elemental composition of coal, which can be used to determine the stoichiometric air requirement for various coal samples. Thus, the empirical model can provide fast and reliable prediction of elemental composition of coal to enhance performance of the combustion control system for power utilities.

6. CONCLUSION

In this paper, four empirical modelling approaches were applied to predict C, H, and O content in coal based on the proximate analysis data. The methods included multiple regression (MR), principal component regression (PCR), partial least square (PLS) and back propagation neural networks (BP-ANN). The use of BP-ANN gave the best result among the tested methods and appears to be a promising tool as elemental composition predictor. However, further improvements are needed for BP-ANN by utilizing additional data for training the model and using other learning algorithms. Also it is important to find the optimum value of number of epoch and learning rate of the networks. Furthermore, development of a good estimator for predicting complete elemental compositions of coal (C, H, N, S, and O) is also challenging which could potentially provide useful and valuable data for power plant operators.

ACKNOWLEDGEMENTS

Authors wish to acknowledge Prof Richard Brereton and Centre for Chemometrics, School of Chemistry, University of Bristol, UK, for the permission to use multivariate analysis add-in software.

REFERENCES

Armesto, L., H. Boerrigter, A. Bahillo and J. Otero (2003). N_2O emissions from fluidised bed combustion: The effect of fuel characteristics and operating conditions. *Fuel*, **82**, 1845-1850.

Artos, V. and A.W. Scaroni (1993). T.g.a. and drop-tube reactor studies of the combustion of coal blends. *Fuel*, **72**, 927-933.

Bailey, J.G., A. Tate, C.F.K. Diessel and T.F. Wall (1990). A char morphology system with applications to coal combustion. *Fuel*, **69**, 225-239.

Beebe, K.R., R.J. Pell and M.B. Seasholtz (1998). *Chemometrics A Practical Guide*, pp. 250-335. John Wiley & Sons, Inc., New York.

Berk, K.N. and P. Carey (2004). *Data Analysis with Microsoft Excel*, pp. 294-367. Thomson Brooks/Cole, Toronto.

Brereton, R.G. (2003). *Chemometrics data analysis for the laboratory and chemical plant*, pp. 271-338. John Wiley & Sons, Chicester.

Brewster, B.S., L.D. Smoot and S.H. Barthelson (1995). Model comparison with drop tube combustion data for various devolatilization submodels. *Energy & Fuels*, **9**, 870-879.

Bulnová, A. and K. Kostúr (2003). Development of control system by Studgard neural network simulator. *Acta Montanistica Slovaca*, **8**, 217-219.

Carlson, K.E. (1996). Fossil Fuels. In: *Power Plant Engineering* (L.F. Drbal, P.G. Boston and K.L. Westra, Ed.), Chap. 4, pp. 71-123. Chapman & Hall, New York.

Carpenter, A.M. (2002). *Coal Quality Assessment – The Validity of Empirical Test*, pp. 5-21. IEA Clean Coal Centre, London.

Charland, J.P., J.A. MacPhee, L. Giroux, J.T. Price and M.A. Khan (2003). Application of TG-FTIR to the determination of oxygen content of coals. *Fuel Processing Technology*, **81**, 211-221.

Coimbra, C.F.M., J.L.T. Azevedo and M.G. Carvalho (1994). 3 – D numerical model for predicting NO_x emissions from an industrial pulverized coal combustor. *Fuel*, **73**, 1128-1134.

Draper, N.R. and H. Smith (1998). *Applied Regression Analysis*, pp. 217-231. John Wiley & Sons, Inc., New York.

Fan, J.R., P. Sun, Y.Q. Zheng, Y.L. Ma and K.F. Cen (1999). Numerical and experimental investigation on the reduction of NO_x emission in a 600 MW utility furnace by using OFA. *Fuel*, **78**, 1387-1394.

Furimsky E., A.D. Palmer, W.D. Kalkreuth, A.R. Cameron, G. Kovacik (1990). Prediction of coal reactivity during combustion and gasification by using petrographic data. *Fuel Processing Technology*, **25**, 135-151.

Guo, B., Y. Shen, D. Li and F. Zhu (1997). Modeling coal gasification with a hybrid neural network. *Fuel*, **76**, 1159-1164.

Haykin, S. (1999). *Neural Network A Comprehensive Foundation*, pp. 156-175. Prentice Hall, New Jersey.

Lockwood, F.C., T. Mahmud and M.A. Yehia (1998). Simulation of pulverised coal test furnace performance. *Fuel*, **77**, 1329-1337.

Martin, E.B., A.J. Morris and J. Zhang (1995). Artificial neural networks and multivariate statistics. In: *Neural Networks for Chemical Engineers* (A.B. Bulsari, Ed.), Vol. 6, Chap. 26, pp. 627-658. Elsevier, Amsterdam.

McLennan, A.R., G.W. Bryant, B.R. Stanmore and T.F. Wall (2000). Ash formation mechanism during pf combustion in reducing conditions. *Energy & Fuels*, **14**, 150-159.

Peralta, D., N.P. Paterson, D.R. Dugwell and R. Kandiyoti (2001). Coal blend performance during pulverised-fuel combustion: estimation of relative reactivities by a bomb-calorimeter test. *Fuel*, **80**, 1623-1634.

Peralta, D., N.P. Paterson, D.R. Dugwell and R. Kandiyoti (2002). Development of a reactivity test for coal-blend combustion: The laboratory-scale suspension firing reactor. *Energy & Fuels*, **16**, 404-411.

Pham, D.T. (1995). An introduction to artificial neural networks. In: *Neural Networks for Chemical Engineers* (A.B. Bulsari, Ed.), Vol. 6, Chap. 1, pp. 1-19. Elsevier, Amsterdam.

Pisupati, S.V., B.G. Miller and A. Scaroni (1992). Effect of blending low-grade anthracite products with bituminous coals on combustion characteristics in a bench-scale stoker simulator. *Fuel Processing Technology*, **32**, 159-179.

Ramadan, Z., P.K. Hopke, M.J. Johnson and K.M. Scow (2005). Application of PLS and Back-Propagation Neural Networks for the estimation of soil properties. *Chemometrics and intelligent laboratory systems*, **75**, 23-30.

Su, S (1999). Combustion behaviour and ash deposition of blended coals, *PhD Thesis*, p. 10. Department of Chemical Engineering, The University of Queensland.

Visona, S.P. and B.R. Stanmore (1997). Modelling NO formation in a swirling pulverized coal flame. *Chemical Engineering Science*, **53**, 2013-2027.

Wu, H., G. Bryant and T. Wall (2000). The effect of pressure on ash formation during pulverized coal combustion. *Energy & Fuels*, **14**, 745-750.

Yao, H.M., H.B. Vuthaluru, M.O. Tadé and D. Djukanovic (2005). Artificial neural network-based prediction of hydrogen content of coal in power station boilers. *Fuel*, **84**, 1535-1542.