Prediction of LAM Coke Price Using ANN and ANFIS Model

Sanjita Jaipuria

Assistant Professor, Operations and Quantitative Techniques Indian Institute of Management Mayurbhanj Complex, Meghalaya, India

ARTICLE INFO	ABSTRACT
Keywords: *Artificial Neural Network (ANN)* *Adaptive neuro-fuzzy inference system (ANFIS)* *Wavelet analysis* *Discrete wavelet transformation*	In the procurement decision, the future price of the raw material plays an important role as it affects the budget and procurement plan. Inaccuracy in raw material price prediction affects the performance of procurement activity and budget plan. Low Ash Metallurgical (LAM) coke is one of the important raw materials used by various alloy manufacturing company. In India, this coke generally imported from the foreign market and followed a long-term contract with the supplier. The LAM coke price is highly volatile in nature. Hence, accurate forecasting of the LAM coke price is very important for any alloy manufacturing company to manage its budget for the procurement. The intelligent models such as Artificial Neural Network (ANN) and Adaptive Neuro-fuzzy Inference System (ANFIS) can learn the past pattern through their self-learning and adapting capability. In this study, the discrete wavelet transformation has been used to extract the past LAM coke price pattern and resultant series applied to train the ANN and ANFIS model to improve the prediction accuracy. From the study, it has found that the prediction error for both the model is less than 5%, hence those models can be used by the alloy industry to predict the LAM coke price.

1. Introduction

Forecasting is the basic and inescapable job of the buyer. Hence, the buyer should have some idea of the future quantity requirement and price of the item for better planning and budgeting. The accuracy of the price forecast depends on the reliability of raw data and the methods used (Gopalakrishnan, 2010). Generally, raw material procurement cost contributes more than fifty percent in the total production cost. Therefore, accurate forecasting of the raw material price is essential in procurement decisions. Coal and coke are the major raw materials used in alloy industries. Coke is a special form of coal contain a high percentage of carbon. In the year, 2010 the worlds' crude steel production was 1.4 billion and around 721million tons of coking coal was used in the manufacturing of steel. (http://www.worldcoal.org/coal/uses-of-coal/coal-steel/). The alloy industries in India, raw material coal and coke are primarily imported from the international market. These raw materials are associated with high landed cost and procurement is done through a long-term contract. There is a high fluctuation in coal prices. Therefore, the coal price prediction for the alloy industries is an important issue. Many researchers have highlighted this issue and proposed different models and methods to overcome the problem. Dong et al. (2010) have studied

_____

the impact of an accurate forecast of coal price employing a time-series approach. Zhang et.al (2011) have proposed partial least-square regression to predict the coal price index. Hu et al. (2010) examined the influence of fluctuation in coal prices on the financial risk of 19 listed companies in the Chinese coal industry. They concluded that a rise in coal price results in increased debt and bad investments. Ellerman (1995) has reported that credited increased in the trade of coal over the past twenty years for the development of a unified coal world market in which price movement is synchronized. Accurate forecasting of coal and coke price is one of the major issues in the procurement decision of steel industries. The artificial intelligence model, Artificial Neural Network (ANN) and Adaptive Neuro-Fuzzy Inference System (ANFIS) have the property to self-learn and self-adapt the data pattern and hence applied to predict the data series possess stationary, non-stationary or non-linear pattern. The ANN model used by many authors in different areas for forecasting (Yao & Tan, 2000; Zhang, 2001; Khashei & Bijari, 2010). The ANFIS model is applied in various areas of research for prediction (Amjady & Keynia, 2008; Khayamian & Esteki, 2004). The prediction accuracy of the intelligence models ANN and ANFIS can be enhanced when these models are properly trained with a given data pattern. The mathematical tool, wavelet analysis can extract the data pattern concerning time and frequency. Therefore, used in various areas of research to extract the data pattern (Sengur, 2008; Boyacioglu & Avci, 2010). Previously, an integrated approach of wavelet analysis and intelligent model has not been used to predict the LAM coke price. Hence, in this study, an integrated approach of wavelet analysis and ANN and ANFIS has been used to predict the LAM coke price and verified that these models are suitable models to predict the coal price.

2. The Predictive Model

In this study, an integrated approach of discrete wavelet transformation and ANN and ANFIS model has been used to predict the LAM coke price. Next paragraph, brief the discrete wavelet transformation, ANN and ANFIS model.

2.1. Discrete Wavelet Transformation

The tool, wavelet analysis popularly used to extract the pattern concerning time and frequency. Wavelet analysis is of two types such as continuous wavelet transformation and discrete wavelet transformation (DWT) according to the type of data it deals with. In wavelet analysis, the original data, f(t) decomposed into approximation coefficient (cAn) at level-n provides background information on the original data and the detail coefficients (cD1, cD2,cD3,…,cDn) at different levels of 1,2,…,n contains the detail information such as period, break and jump. Hence, the original signal (f (t)) can be represented as:

$$f(t) = cA_n(t) + \sum cD_n(t) \qquad (1)$$

The Daubechies wavelet family represents as 'dbN' where, db is the 'surname', and N is the order of wavelet i.e. number of decomposition levels (Wei et al., 2012). Further, the resultant decomposed data series i.e. approximations and details data series are taken as input to train the ANN and ANFIS model and predict the future price of LAM coke. The proposed model denoted as DWT-ANN and DWT-ANFIS.

2.2. Artificial Neural Network (ANN)

The feed-forward multilayer (FFML) perceptron network is one of the most commonly used ANN models to predict future value. The structure of an ANN model comprises three different layers these are input, hidden and output layer and each layer has interconnected nodes. The number of neurons in the input layer equals the external inputs applied to the network. The number of neurons in the hidden layer is optimized through experimentation and the number of neurons at the output layer depends on the anticipated number of output. (Kumar, 2011). The ANN prediction process follows two steps one is training and the other is testing. For this, the available data series divided into two data sets i.e. training and testing data set. In the training process, the training data set applied as input to train the ANN model using the gradient descent with momentum parameter backpropagation training function. In this process, the inputs and outputs are mapped and the weight value of the neurons present in each layer of the network is regulated.

Let $I_p = (I_{p1}, I_{p2} \dots I_{pl})$, $p = 1, 2 \dots$ N is the pth input among the N inputs. The output from the neurons present in the input layer (O_{pi}) is shown in Equation (2) whereas the output from the hidden layer (O_{pj}) and output layer (O_{pk}) is shown in Equation (3) and Equation (4) respectively.

$$O_{pi} = I_{pi}, i = 1,2,...,l \tag{2}$$

$$O_{pj} = f\left(\sum_{i=0}^{l} W_{ji} O_{pi}\right), j = 1,2,...,m \tag{3}$$

$$O_{pk} = f\left(\sum_{j=0}^{m} W_{kj} O_{pj}\right), k = 1,2,...,n \tag{4}$$

where, i symbolizes the ith neuron and l is the total number of neurons in the input layer. Here it is equal to the number of data series obtained after decomposing the LAM coke price using the wavelet analysis. j represents j^{th} neurons in the hidden layer and m is the total neurons that are optimized through experimentation. Similarly, k^{th} is the number of neurons in the output layer and n is the total number of neurons, here it is one that is the predicted value. The parameters W_{ji} denotes the connecting weights between i^{th} input neuron to j^{th} hidden neuron. Similarly, W_{kj} is the connecting weight between j^{th} hidden neuron to k^{th} output neuron respectively.

The input data set is adjusted through the connecting weights and generates a sum of modified value (x). Further, the sigmoidal transfer function f(x) as shown in Equation (5) applied for a modification. The forecasted output is compared with the desired output during the training process by calculating the prediction error i.e. mean square error (E_p) using Equation (6). If E_p is higher than a defined targeted value (here taken 0.10), then it is backpropagated from output to input layer and weights are further adjusted using Equation (7) till the error or number of iteration reach to the defined target value.

$$f(x) = \frac{1}{1 + e^x} \tag{5}$$

$$E_p = \sum_{i-1}^{n} \frac{1}{2}\left(D_{pi} - O_{pi}\right)^2 \tag{6}$$

$$\Delta W(t) = -\eta E_p(t) + \alpha \times \Delta W(t-1) \tag{7}$$

η is the learning rate and α is the momentum coefficient, both the parameters take a value between 0 and 1. t represents the number of iteration (epoch).

2.3.The Adaptive Neuro-Fuzzy Inference System (ANFIS) model
Like ANN, ANFIS is also an intelligent model. It has also self-adapting and self-learning capability for the data pattern. The learning procedure of past data patterns in the case of ANN model is treated as a black box, whereas in the ANFIS model it is based on if-then rules. Jang (1993) introduced this intelligent model, an adaptive neuro-fuzzy inference system (ANFIS). The input membership function and associated parameters are used by the model to maps the input and output membership function and output. The fuzzy if-then rules and membership functions are used to describe the input-output behavior of the complex system. In ANFIS, the five-layer creates the inference system. Each layer comprises several nodes and is described with the node function. The presentation layer takes the output signals from the nodes present in the previous layer as its input signal. The general structure of the ANFIS model is shown in Figure 1.
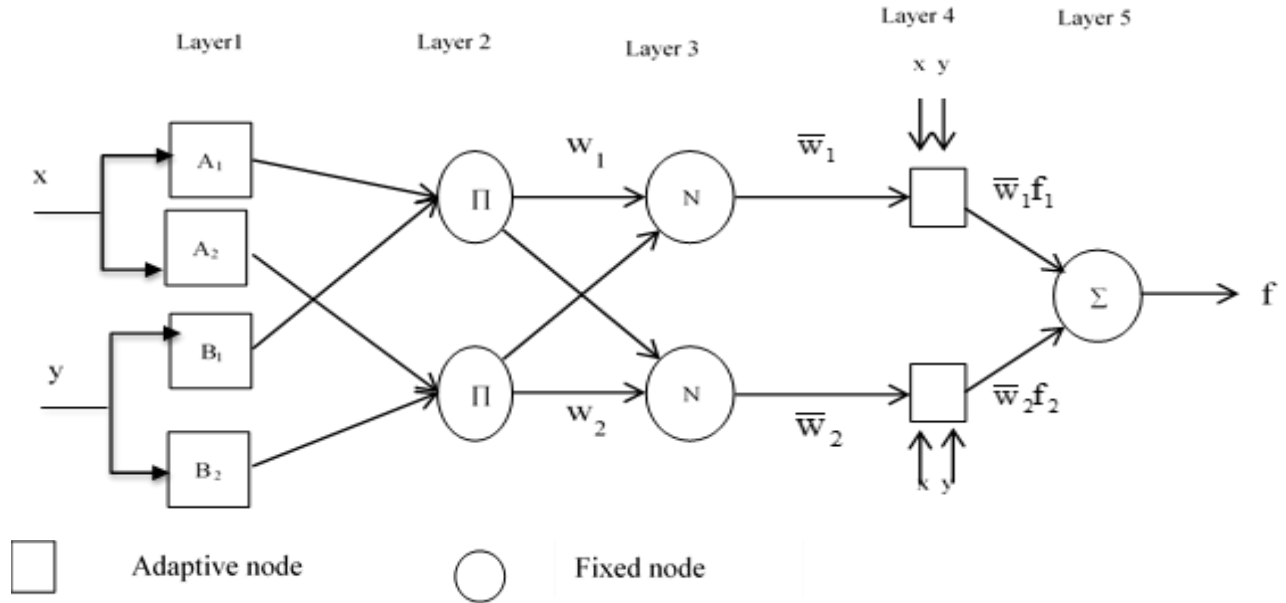

Figure 1. ANFIS model structure

One degree of Sugeno's function used to represent the rule used in the ANFIS model can be defined as (Jang, 1993):

Rule 1: if x is A_1 and y is B_1 then $f = p_1 x + q_1 y + r_1$

Rule 2: if x is A_2 and y is B_2 then $f = p_2 x + q_2 y + r_2$

From Figure 1, x and y are two inputs variables and f is the output of the fuzzy inference system. The output of each layer is represented by O_i^L where L=1,2,…,5 layers. The output from the layer 1 is:

$$O_i^1 = \mu_{Ai}(x) \qquad i=1,2 \qquad\qquad (8)$$
$$O_i^1 = \mu_{Bi-2}(y) \qquad i=3,4 \qquad\qquad (9)$$

where, A_i and B_{i-2} is the linguistic variable. The membership function (as shown in Equation 10) usually adopts bell-shape with a maximum value equal to 1 and minimum value equals to 0.

$$\mu_{Ai}(x) = \frac{1}{1 + \left\{((x - c_i)/a_i)^2\right\}^{bi}}$$

(10)

a_i, b_i, c_i represents parameter set called premise parameter.

The layer 2 output is represented in Equation (11)

$$O_i^2 = w_i = \mu_{Ai}(x)\mu_{Bi}(y) \qquad i=1,2$$

(11)

The layer 3 output is represented in Equation (12)

$$O_i^3 = \overline{w}_i = \frac{w_i}{w_1 + w_2} \qquad i=1,2$$

(12)

The layer 4 output is represented as

$$O_i^4 = \overline{w}_i f_i = \overline{w}_i (p_i x + q_i y + r_i) \qquad i=1,2$$

(13)

The layer 5 output is represented as

$$O_i^5 = \sum_{i=1}^{2} \overline{w}_i f_i = \frac{\left(\sum_{i=1}^{2} w_i f_i\right)}{w_1 + w_2}$$

(14)

The final output of the inference system can be represented as

$$f = \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2$$

(15)

The least-square method can be used to determine the optimal value setting for the parameters $p_1, q_1, r_1, p_2, q_2, r_2$. While the premise parameters are not fixed, the search space becomes larger and convergence of training becomes slower. Consequently, the hybrid-learning algorithm that combines the backpropagation gradient descent and least squares method has applied here for optimal parameter setting of the ANFIS model. Based on the above description the node outputs move in the forward direction up to layer 4 and the resulting parameters are determined by the least-square method through forward learning process. Whereas, the error signal moves in backward direction till layer 1. To obtain the optimal parameter settings for ANFIS model, the hybrid learning process that combines the backpropagation gradient descent and least squares method has been used in this study.

3. Results and Discussion
The raw material LAM coke is a carbon material produced by the "distillation" of bituminous coal. This coke used in different operational area such as 1) Iron foundries 2) Steel foundries 3) Non-ferrous metal castings 4) Lead and Zinc smelters 4) Chemical Plants, and 5) Pig iron producers.

Its price is highly volatile in nature and associated with high landed cost. According to the specific consumption rate, 550-570 Kg of LAM coke is required to produce one ton of Hot metal, depending upon the installed capacity of the plant. Similarly, to produce 1MT of Ferro Alloy 600Kg of LAM coke is used as reducing agent. The weekly LAM coke price data, from April-2013 to March-2014 (48 weeks) in rupees is collected for the study. Table 1 describes the statistics of the LAM coke price.

Table 1.
Descriptive statistics value for LAM coke price

Variable	Q1	Median	Q3	Range	IQR	Skewness	Kurtosis
	17300	17500	17658	7200	358	2.67	6.41
	N	Mean	StDev	Variance	CoefVar		
LAM coke price	48	17893	1667	2778966	9.32		

From the estimated statistical values (Table 1), it can be observed that LAM coke price is highly positively skewed in nature as its value is greater than 1. The Kurtosis value says that LAM coke price is following Leptokurtic distribution as value is greater than 3, this means data are heavy-tailed or profusion of outliers. This can be observed in Figure 2 and Figure 3. From both the figure outliers can be observed, thus indicating those four data points are higher price at four different periods as compared to the rest of the data points. The estimated values of variance, coefficient of variance, range and interquartile rate represent the dispersion in LAM coke price, which is very high.
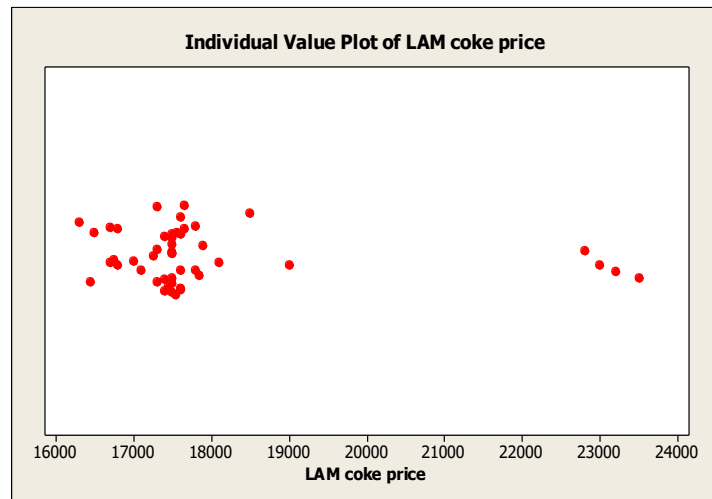
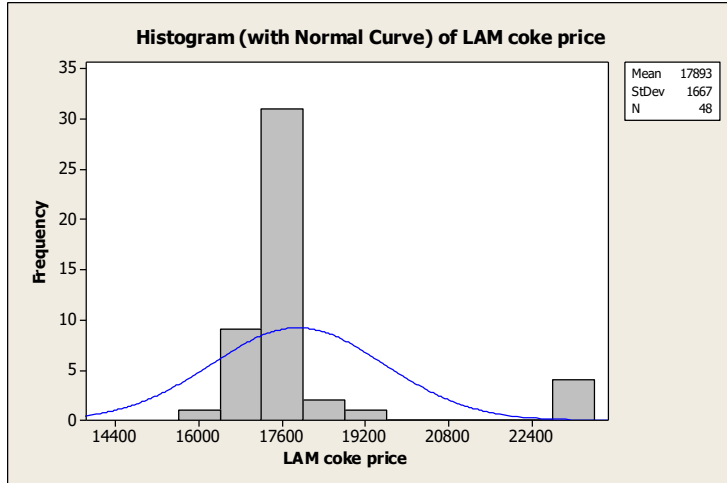

Figure 2. Individual value plot for LAM coke price

Figure 3. Histogram with normal probability plot

The time series plot of LAM coke price is shown in Figure 4. This figure signifies that the data series is non-stationary in nature. Further, Augmented Dickey-Fuller (ADF) t-test is conducted at a significance level of 0.05 to check the existence of stationary in the LAM coke price. The test resulted in H=0 and p-Value = 0.1145 (fail to reject the null hypothesis) and hence proof that LAM coke price is non-stationary in nature. The db4 wavelet family has applied to the price series to decompose the series into five levels. The decomposed price series after wavelet analysis has generated an approximation series (A5) and five detail series (D1, D2…, D5) as shown in Figure 5. The whole data points in the subseries are divided into two sets these are training dataset i.e. 90% of the data points (43weeks) and rest 10% of the data points (5-weeks) as testing dataset. The training set is used as input to train the ANN and ANFIS model. The rest 5weeks price data has used for prediction.
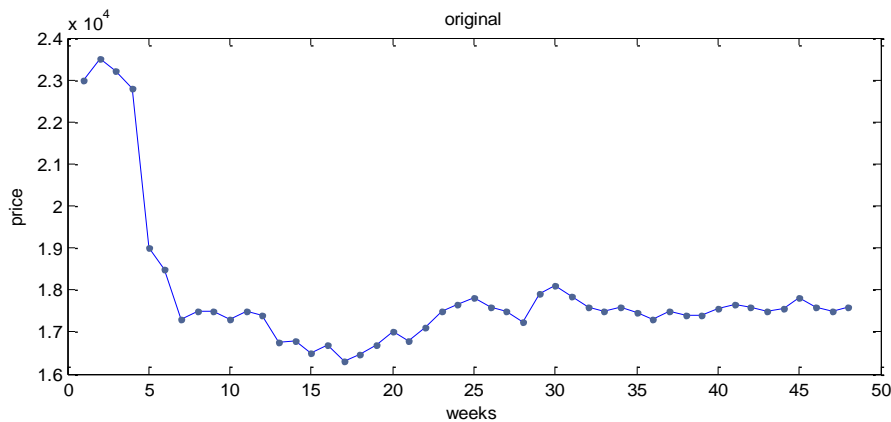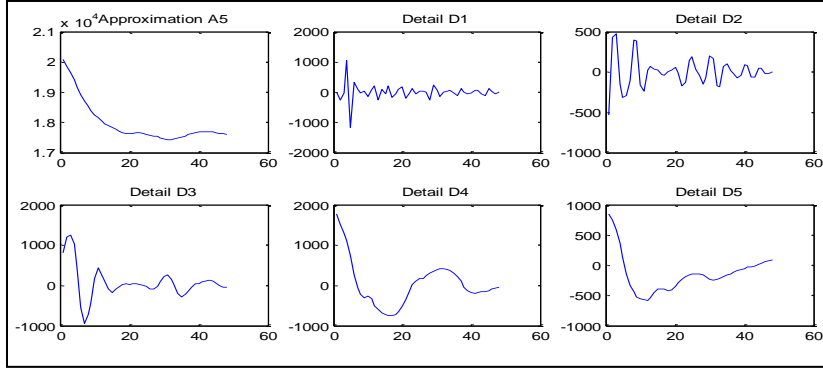

Figure 4. Time series plot of LAM coke price

Figure 5. Decomposed LAM coke price data

After the successful training of the ANN and ANFIS model, optimal model parameters setting are obtained as listed in Table 2. Using these optimal parameter settings, the testing data set is applied to ANN and ANFIS model for prediction. The Figure 6 summaries the training performance and Figure 7 summaries the testing performance of the ANN model.The testing output summary of ANFIS model is shown in Figure 8. The forecasting accuracy of the proposed models has been measured in terms of the Mean Absolute Percentage Error (MAPE) using the Equation (16). The estimated forecast error MAPE for the ANN model is 0.48 whereas in case of ANFIS it is 0.50 which is less than 5% error signifies the high accuracy of the forecasting model.

$$\text{Mean absolute percentage error (MAPE)} = \frac{1}{n} \sum_{t=1}^{n} \left| \left(\frac{Y_t - F_t}{Y_t} \right) \times 100 \right| \qquad (16)$$

Table 2.
Model parameter settings

| Analysis	Model parameter	Model parameter settings
DWT analysis	Wavelet family	db4
	Decomposition levels	5
ANN	Learning parameter	0.01
	Momentum parameter	0.1
	Epoch	20000
	Goal	0.001
	Number of neurons in input layer	6
	Number of neurons in hidden layer	8
	Number of neurons in output layer	1
ANFIS	Membership function assign to each input value	3 3 3 3 3 3
	Input membership function type	gauss2mf
	Output membership function type	Linear
	ANFIS training Optimization method	Hybrid
	Error tolerance	0.00001
	Number of epochs	3
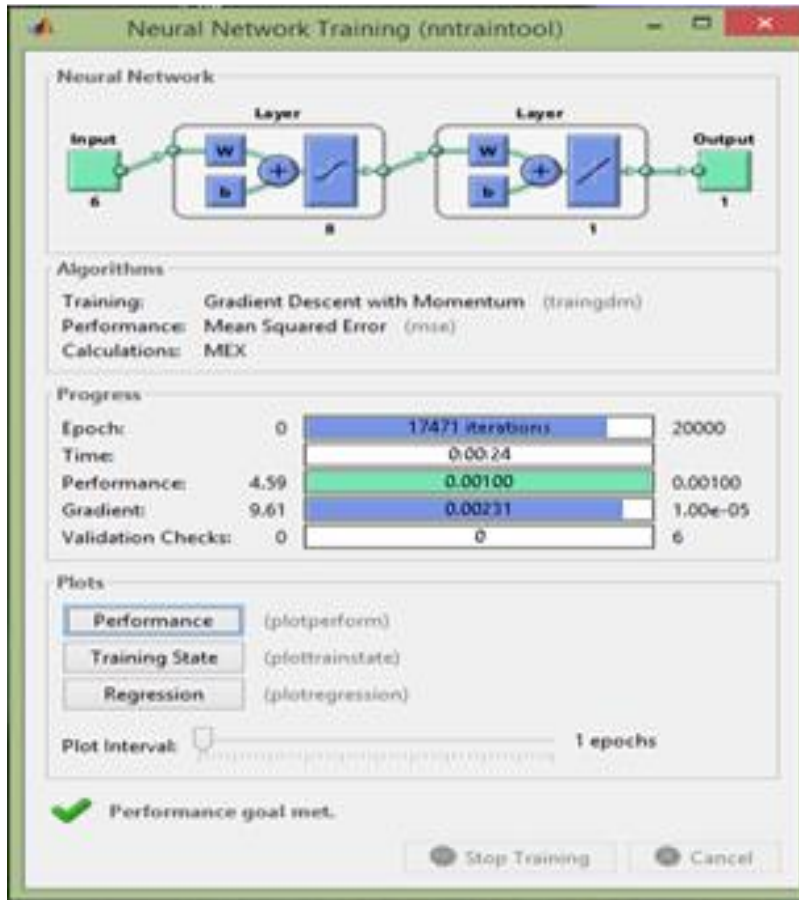
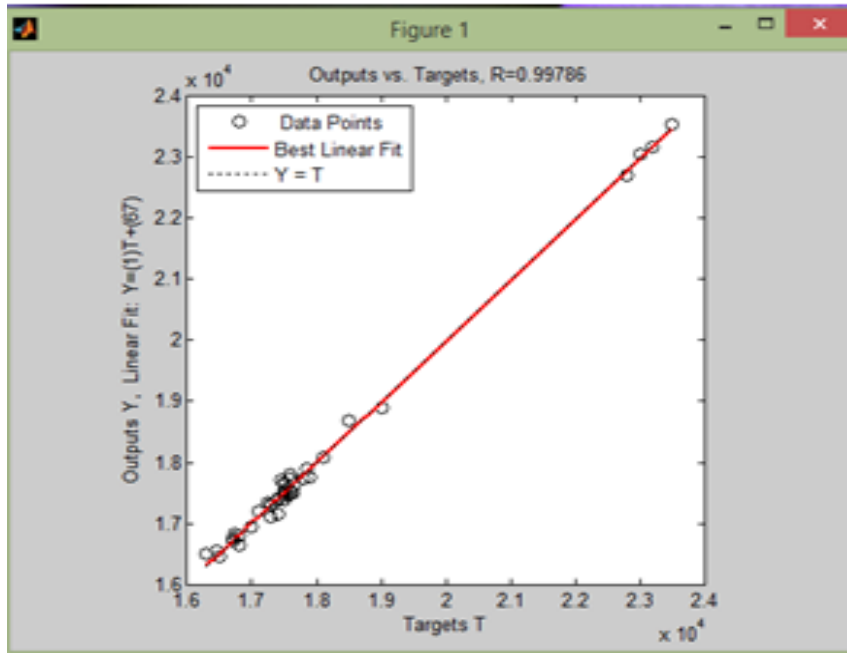*Figure 6.* Training performance of ANN model
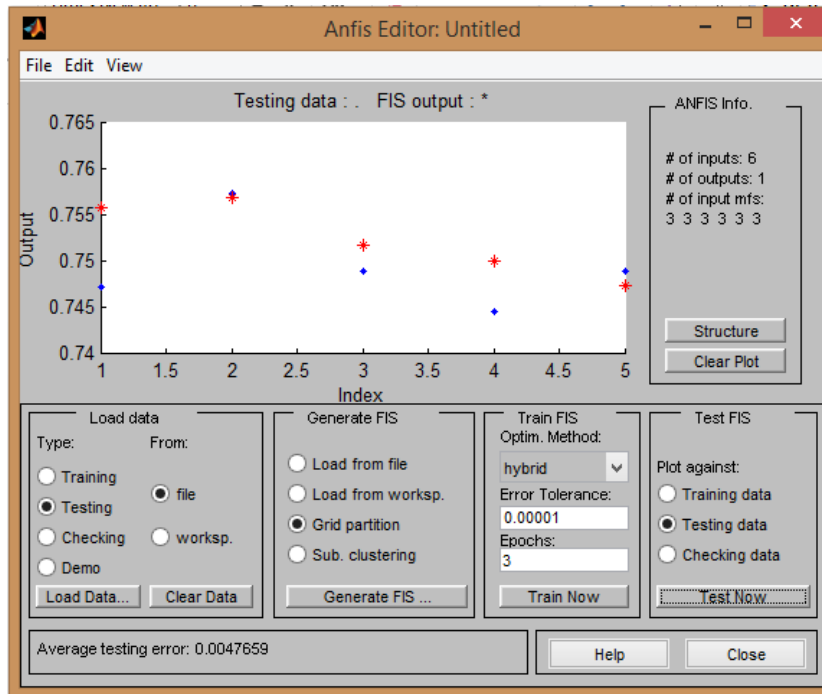

Figure 7. Testing performance of ANN model

Figure 8. Prediction performance of ANFIS model

4. Conclusion

In this study, a hybrid approach of discrete wavelet transformation and ANN and ANFIS denoted as DWT-ANN and DWT-ANFIS is adopted for accurate forecasting of LAM coke price. The DWT-ANN and DWT-ANFIS model can be used as a suitable model to predict the price of LAM coke as the forecast error MAPE estimated is less than 5%. These predictive models can help an alloy manufacturing organization to predict the future price of raw material. Through predicting the future price of the raw material, a procurement executive can schedule the procurement activity and can make a proper decision of quantity to procure. The proposed model can be extended with other different variants of the intelligence models like support vector machine, least-square support vector machine, etc.

References

Amjady, N., & Keynia, F. (2008). Day ahead price forecasting of electricity markets by a mixed data model and hybrid forecast method. International Journal of Electrical Power & Energy Systems, 30(9), 533-546.

Boyacioglu, M. A., & Avci, D. (2010). An adaptive network-based fuzzy inference system (ANFIS) for the prediction of stock market return: the case of the Istanbul stock exchange. Expert Systems with Applications, 37(12), 7908-7912.

Dong, B., Li, X., & Lin, B. (2010). Forecasting Long-Run Coal Price in China: A Shifting Trend Time-Series Approach. Review of Development Economics,14(3), 499-519.

Ellerman, A. D. (1995). The world price of coal. Energy Policy, 23(6), 499-506.

Gopalakrishnan, P. (2001). Purchasing and materials management. Tata McGraw-Hill Education, New Delhi, India.

Hu, S., Bing, W., & Liming, M. (2010, March). The impact of fluctuations in coal prices on financial risk of China's coal listed companies. In Power and Energy Engineering Conference (APPEEC), 2010 Asia-Pacific (pp. 1-4). IEEE.

Jang, J. S. (1993). ANFIS: adaptive-network-based fuzzy inference system. IEEE transactions on systems, man, and cybernetics, 23(3), 665-685.

Khashei, M., & Bijari, M. (2010). An artificial neural network (p, d, q) model for timeseries forecasting. Expert Systems with applications, 37(1), 479-489.

Khayamian, T., & Esteki, M. (2004). Prediction of solubility for polycyclic aromatic hydrocarbons in supercritical carbon dioxide using wavelet neural networks in quantitative structure property relationship. The Journal of supercritical fluids, 32(1), 73-78.

Kumar, S. (2011). Neural networks - a classroom approach. Tata McGraw-Hill Education Private Limited, New Delhi, India.

Sengur, A. (2008). Wavelet transform and adaptive neuro-fuzzy inference system for color texture classification. Expert Systems with Applications,34(3), 2120-2128.

Wei, S., Song, J. and Khan, N. I. 2012. 'Simulating and predicting river discharge time series using a wavelet-neural network hybrid modelling approach', Hydrological process, 26(2):281-296.

Yao, J., & Tan, C. L. (2000). A case study on using neural networks to perform technical forecasting of forex. Neurocomputing, 34(1-4), 79-98.

Zhang, B., & Ma, J. (2011). Coal price index forecast by a new partial least-squares regression. Procedia Engineering, 15, 5025-5029.

Zhang, G. P., Patuwo, B. E., & Hu, M. Y. (2001). A simulation study of artificial neural networks for nonlinear time-series forecasting. Computers & Operations Research, 28(4), 381-396.