World Academy of Science, Engineering and Technology
International Journal of Chemical and Molecular Engineering
Vol:4, No:9, 2010

Artificial Neural Network Prediction for Coke Strength after Reaction and Data Analysis

Sulata Maharana*, B Biswas, Adity Ganguly, Ashok Kumar

***Abstract*—**In this paper, the requirement for Coke quality prediction, its role in Blast furnaces, and the model output is explained. By applying method of Artificial Neural Networking (ANN) using back propagation (BP) algorithm, prediction model has been developed to predict CSR. Important blast furnace functions such as permeability, heat exchanging, melting, and reducing capacity are mostly connected to coke quality. Coke quality is further dependent upon coal characterization and coke making process parameters. The ANN model developed is a useful tool for process experts to adjust the control parameters in case of coke quality deviations. The model also makes it possible to predict CSR for new coal blends which are yet to be used in Coke Plant. Input data to the model was structured into 3 modules, for tenure of past 2 years and the incremental models thus developed assists in identifying the group causing the deviation of CSR.

***Keywords*—**Artificial Neural Networks, backpropagation, Coke Strength after Reaction, Multilayer Perceptron.

I. INTRODUCTION

THE quality of coke has a strong effect on the performance of a Blast furnace. Coke fed to the iron blast furnace serves three purposes, as a fuel, as a reducing agent, and as a structural material supporting the deep bed of iron oxide/fluxes that makes up much of the furnace volume. Research by major coke producers shows that the operational efficiency of Blast furnaces depends on the post reactive strength CSR and reactivity CRI of the coke. Hence it is important to optimize the initial coking conditions so as to obtain coke with the best CSR[1].Each steel company employs models adapted to its operational characterization which change as a function of the coals available in the market. Due to the great diversity of the coals used for coke making, there is universally no applicable model, as some of the most relevant methods applied in coke quality prediction have been reviewed [2].

In the present work, an attempt has been made to predict and diagnose CSR with coal characterization and coke making process data as inputs and subsequently correlation with the technological value of the blast furnace coke produced is being done. In Module 1, ANN Prediction model for coal characterization has been developed. Module 2 consists of ANN Prediction model based on Coke making process and Module 3 consists of Prediction model consisting of both Module 1 and Module 2.The aim was to identify the group or the combination or groups that is causing the deviation of CSR.

The paper is organized as follows. Section II briefly describes the ANN, its theoretical background and role in prediction. Section III describes the development of ANN model. Section IV describes the Back propagation algorithm. Section V describes methodology used in the model. Section VI provides the experimental results. Section VII summarizes and concludes the paper. Section VIII provides Future scope.

II. ARTIFICIAL NEURAL NETWORKS

Artificial neural Networks (ANNs) are most often chosen for its ability to generalize results from unseen data, specially for dynamic systems on real time basis. ANN's are parallel computational models comprised of densely interconnected adaptive processing units. These networks are fine grained parallel implementations of dynamic systems [3].

ANNs can identify and learn correlated patterns between input data sets and corresponding actual target values. ANNs are networks of highly interconnected neural computing elements that have the ability to respond to input stimuli and to learn to adapt to the environment. ANN includes two working phases, the phase of learning and that of recall. During the learning phase, known data sets are commonly used as a training signal in input and output layers. The recall phase is performed by one pass using the weight obtained in the learning phase. A single-layer network has single input/output units, a multi-layer network has one or more hidden units between input and output layers [4]. An important feature of these networks is their adaptive nature, where ''learning by example'' replaces ''programming'' in solving problems [3]. The application domains are hence innumerous where training data is already available.

III. DEVELOPMENT OF ANN MODEL

ANN's are basically generated with layers of units and are termed as Multilayer ANN's. These multilayer networks are capable of any linear or non-linear computation, and is well suited for approximation functions. MLP basically consist of 3 layers namely input layer, hidden layer and an output layer.

Sulata Maharana, *Corresponding Author, Raw Material and Iron Making Technology Group(RMIMTG),Tata Steel, Jamshedpur- 831001,India.
E-mail:sulata.maharana@tatasteel.com,fax:91 657 2756548,ph:9204058106.

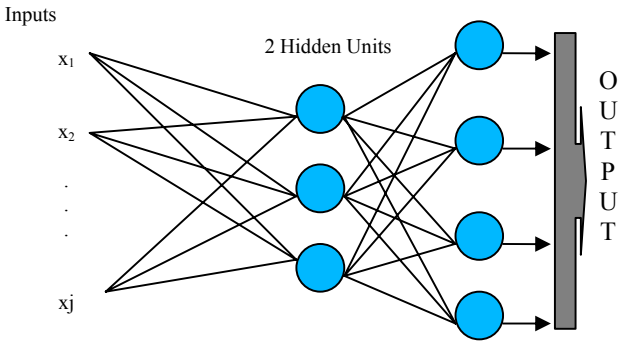

Fig. 1 A simple representation of Neural Network Structure

The role of neurons in the input layer is to transmit or distribute the scalar input to the next stage i.e. the hidden layer.

The hidden layer adds up the weighted inputs as received from the input nodes, and further adds a bias before passing the result to a non-linear transfer function.

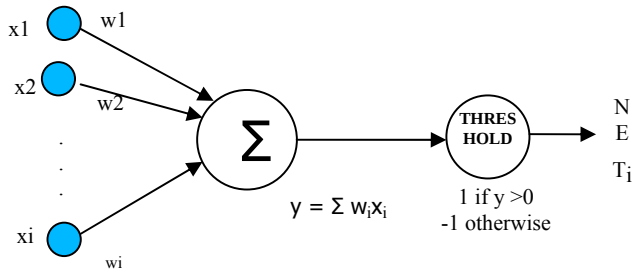

Fig. 2 represents a Single Perceptron Rule

The output node net_i is the result of inputs x_i weighted by w_{ij} which are the weights connecting neuron j to i. μ_i is the threshold for neuron i.

The output neti is a function of the sum, viz.,

$$net_i = \sum w_{ij} x_i + \mu_i \quad (1)$$

Neural network is trained by adjusting its weights by mean of training algorithms. The aim is to minimize the sum of squared differences between actual and desired values from output neurons, where

$$Y=1/2 \sum f (net_{pi}-net_i)^2 \quad (2)$$

where net_i is actual output of neuron and net_{pi} is desired output for neuron i.

Weight w_{ij} has to be adjusted which again depends on training algorithm utilized so as to reduce Y to a minimum.

The activation functions, learning algorithms, learning rates, and number of neurons in the hidden layer play a vital role in creating a network which gives optimum result.

The learning algorithm used is scaled conjugate gradient algorithm and is based on conjugate directions. It is designed to avoid time consuming line search which makes it distinct from other conjugate gradient algorithms. [5]

Training stops if any of the following conditions are met:

- The maximum number of epochs (repetitions) is reached.
- Performance has been minimized to the goal.
- The performance gradient falls below mingrad.
- The maximum amount of time has been exceeded.

Input variable selection is very important for modeling and/or forecasting so as to categorize the main parameters affecting the output variable. The data set is categorically divided into training, testing and validation sets.

Neural Network architecture describes the structure which consists of details regarding the number of hidden layers, number of output nodes, activation function etc.

Multiple feed forward neural network or multiple layer perceptron (MLP), learned by back propagation algorithm is usually based on supervised procedure where network is constructed by training input to the model with known output based on the generalized delta rule [6]

IV. BACKPROPAGATION ALGORITHM

The simplest implementation of backpropagation learning updates the network weights and biases in the direction in which the performance function decreases most rapidly - the negative of the gradient [7].

One iteration of this algorithm is as shown below:

$$N_{h+1} = N_h - \alpha_h K_h$$

where N_h is a vector of current weights and biases, K_h is the current gradient, and α_h is the learning rate.

A. Training Set

Comprises of input and output patterns which are trained in a batch. In batch mode, the weights and biases of the network are updated only after the entire training set has been applied to the network [6].

In this model, the training is done by Batch Gradient Descent. The weights and biases are updated in the direction of the negative gradient of the performance function. Only one training function is associated with a given network [6].

Training consists of

- Assemble the training data
- Create the network object
- Train the network
- Simulate the network response to new inputs.

Learning rate must be aligned corresponding to the desired result. If the learning rate is too large, the algorithm becomes unstable and if the learning rate is too small, then the algorithm may take longer time to converge.

B. Testing Set

This comprises of input and output patterns which assess the network performance. The network error and its standard deviation from actual outputs are taken into consideration while selecting networks.

C. Error Calculation

The network error is determined by Mean Square Error (MSE) and Root Mean Square Error (RMSE). In spite of the complexity of input elements, properly trained backpropagation networks with appropriate number of hidden layers tend to give accurate results. Typically, a new input leads to an output similar to the correct output for input vectors used in training that are similar to the new input being presented. This generalization property makes it possible to train a network on a representative set of input/target pairs and get good results without training the network on all possible input/output pairs.

V. METHODOLOGY

The application of neural network technique is divided into three modules. The first module discuses ANN model developed for Coal inherent characterization data, the second module discusses ANN model for process data and the third module explains the methodology for developing ANN model for the final data set comprising of both process data and coal characterization data sets. While preparing data sets for each module, the corresponding time lag for each data parameter with the output CSR has been taken into account. Coal characterization data which is tested on a weekly basis is repeated for the whole week. Coke Process inputs are directly aligned with output CSR and there is no lag except for some parameters like stamping time and bulk density which have 1 day lag each. This has been done to link each module data to its corresponding output CSR value for better accuracy in NN prediction. This module approach is done to analyze the main reasons for deviation for the predicted output value is due to which group or combination of groups as shown in Fig. 3. Earlier no such method was available which could link the deviation in CSR to particular group or groups.

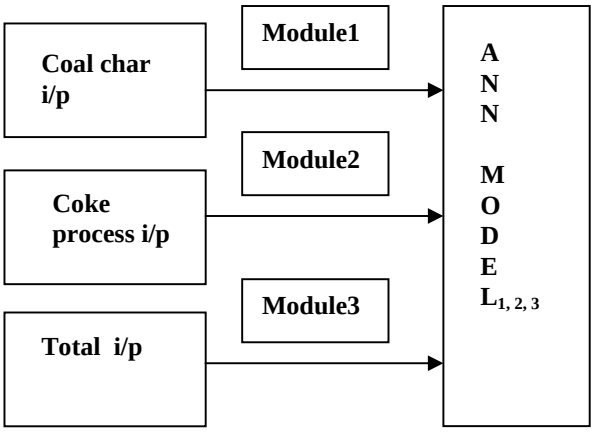

Fig. 3 Brief Sketch of ANN Developed Modules

During NN training phase, a large number of input variables can provide an accurate description of the problem being considered, but yields over-parameterized model and requires more computational processing time and often more data for an effective understanding of the relationship between inputs and outputs.

ANN in this study was trained and simulated using Matlab 7.1 developed by The Math Works Inc, Natick, Massachusetts.

VI. EXPERIMENTAL RESULTS

All the procedures of ANN Modeling have been coded by Matlab programming. Each category of input-output patterns have distinct Prediction model as discussed in Section V.

In this ANN Modeling and simulation, the steps involved are Task Identification and Design, Feasibility, Data Coding, Data Collection, Data checking, Network Design, Training and Testing, Error analysis, Network Analysis, System Implementation.

A brief flow diagram showing the steps followed in ANN model development is given in Fig. 4

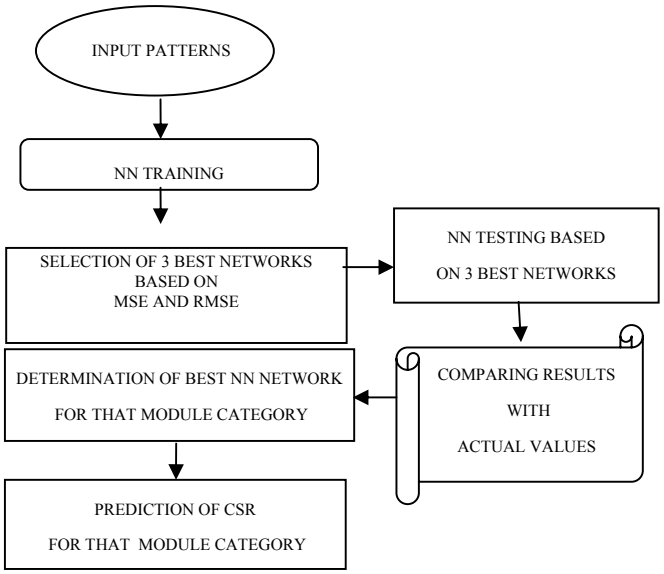

Fig. 4 Selected ANN Model

A. ANN Model for Coal Characterization Data

The development of this model included identifying characterization data which could affect the CSR. Feasibility check was carried out, noisy data was eliminated which also included plant test trials data. The network design was initiated with 1 and 2 layers. Based on the performance, network design was decided upon 2 layer networks.

In order to increase the performance of the network, the number of neurons was increased step by step as shown in the TABLE I.

TABLE I
INFLUENCE OF NUMBER OF NEURONS IN THE HIDDEN LAYER

Neurons in 1st layer	Neurons in 2nd layer	STD Deviation	RMSE
50	30	0.0723	0.00518
50	40	0.0860	0.00541
50	50	0.0360	0.00271
50	60	0.0367	0.00282

Simulation for characterization data was attempted with 50 neurons in the first layer and 30 to 60 neurons in the second layer. The results show that ANN fed with 50 neurons in the 2nd hidden layer provided the best results.

It is also experimentally seen that by increasing the number of epochs during training further enhances the network performance. The results are as shown in Fig. 5

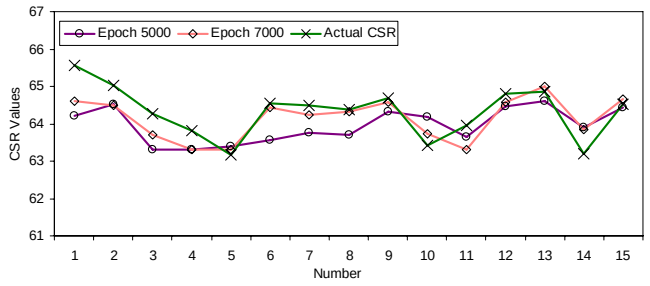

Fig. 5 NN Performance with same network and different epochs

The network thus developed was programmed for prediction of plant process data on a real time basis. The outputs of the ANN coal characterization model for a tenure of 18 days is as shown in FIG. 6

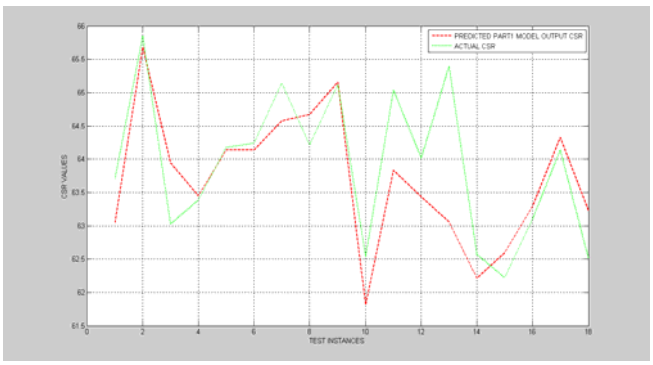

Fig. 6 NN output For Coal Characterization Prediction Model

B. ANN Model For Coke Process Data

Similar to Section VI A prediction model for Coke Process data was also attempted. The aim is to diagnose the main parameters affecting the overall CSR value during a wide operating regime.

The network selection was based on the RMSE and standard deviation results, wherein the number of neurons was increased step by step as shown in the TABLE II.

TABLE II
INFLUENCE OF NUMBER OF NEURONS IN THE HIDDEN LAYER

Neurons in 1st layer	Neurons in 2nd layer	STD Deviation	RMSE
50	30	0.0698	0.00483
50	40	0.0647	0.00352
50	50	0.0436	0.00319
50	60	0.0461	0.00323

From the results of TABLE II, it is clearly evident that the best network chosen for this model is 50 neurons in the 2nd hidden layer.

The outputs of the ANN coke process model is as shown in Fig. 7

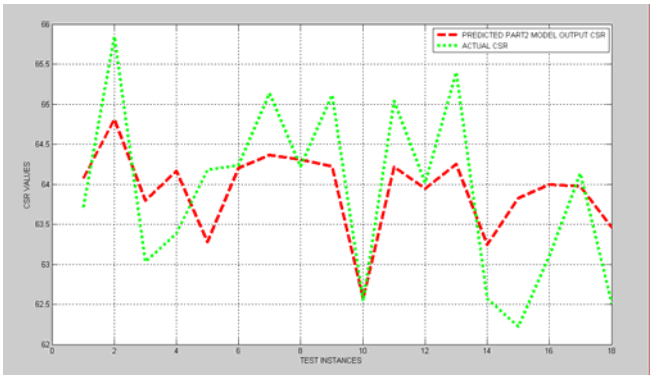

Fig. 7 NN Output For Coke Process Prediction Model

C. ANN Model For Both Coal Characterization and Coke Process Data

The input data in this model is an additive of both Module1 and Module2 models. Here the overall effect of all the input parameters on the coke quality i.e. CSR output has been studied. The network is decided as per performance criteria of least standard deviation and RMSE as shown in the TABLE III.

TABLE III
INFLUENCE OF NUMBER OF NEURONS IN THE HIDDEN LAYER

Neurons in 1st layer	Neurons in 2nd layer	STD Deviation	RMSE
50	30	0.0574	0.00375
50	40	0.0439	0.00346
50	50	0.0338	0.00265
50	60	0.0316	0.00202

After Network Design and programming for the overall prediction set, the final output of the model is as shown in FIG. 8

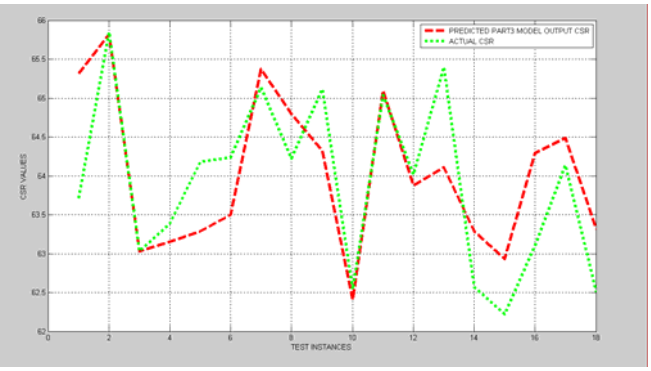

Fig. 8 NN Output For Total Coal Coke Prediction Model

The predicted output and its variation with actual CSR corresponded to changes in various characterization and process data.

After comparing the ANN output of each module i.e. FIG. 6, FIG. 7 and FIG. 8 say for any particular test instances, we can very well say that the deviation in CSR is due to Module 1 for test instances numbered 8,12. Similarly test instance numbered 5,9 is due to Module 2 and test instance numbered 7,11 is due to Module 3.

VII. CONCLUSION

Results from this study prove that ANN models developed are reliable to estimate Coke Strength after reaction so that process visibility and detailed diagnosis is possible. This approach for predicting CSR would be useful in industrial practice where database of coke quality data is already available. The significance of this work lies with the detection of the parameters which have more impact on the overall CSR value. Improvement in the predictions came by increased number of inputs, adopting multilayer hidden strategy and greater number of epochs for training data. Earlier no method was used for prediction of Process parameters. This experimental model has helped coke plant personnel to diagnose, analyze and understand the coke making process in greater detail.

VIII. FUTURE SCOPE

Further improvements which can be done:

- To predict CSR for a blend which are yet not been used in the plant process, by including the input data to the NN training set.
- Quick Diagnostic Tool for day to day process visibility to the Coke plant personnel to take alternative actions to tune the process.

REFERENCES

[1] G. R. Gainieva, L. D. Nikitin, M. M. Naimark, N. N. Nazarov, and G. P. Tkachenko, Influence of Batch Composition and Clinkering Properties on the Hot Strength of Coke and Blast-Furnace Operation, ISSN 1068-364X, Coke and Chemistry, 2008, Vol. 51, No. 10, pp. 390–393. © Allerton Press, Inc., 2008.

[2] M.A. Díez, R. Álvarez and C. Barriocanal, Coal for metallurgical coke production: Predictions of coke quality and future requirements for cokemaking, Int J Coal Geol **50** (2002), pp. 389–412.

[3] Girish Kumar Jha,''Artificial Neural Networks'',Indian Agricultural Research Institute.

[4] Mustafa Taskin, Halil Dikbas and Ugur Caligulu, ''Artificial neural network (ANN) approach to prediction of Diffusion bonding behavior (shear strength) of Ni-Ti alloys manufactured by powder metallurgy method'',*Mathematical and Computational Applications*, Vol. 13, No. 3, pp. 183-191, 2008.

[5] http://www.mathworks.com

[6] D.E Rummelhart, G.E.Hinton, and R.J.Williams,''Learning internal representations by error backpropagation'', *Parallel Distributed Processing*,vol. 1 ,Cambridge,Mass:MIT Press,1986.

[7] H. Demuth, and M. Beale, *Neural Network Toolbox For Use with MATLAB*. Natick, MA: The Mathworks Inc., 2001.