103. Fuel Optimization in Power Plant Based on Its Parameters Using Artificial Neural Network (ANN)

Syed Subhan Ahsen[a*], Muhammad Nauman Mustafa[a], Abdullah Munir[a], Syed Muhammad Zain[a] , Hassaan Siddiqui[a]

[a]*NED University of Engineering and Technology, Electrical Engineering Department,University Road, 75270, Karachi, Pakistan*
**Email Address: subhan.ahsen2010@gmail.com*

Abstract

Natural gas is one of the most employed fuels in the world and especially in Pakistan, where is it consumed in residential, commercial and power sector. Power sector contribution in natural gas usage is rising because of lower carbon footprints and also economical as compared to other conventional sources. It accounts for more than 42% in thermal energy mix, as per National Electric Power Regulatory Authority (NEPRA) state of industry report 2015. The increased consumption is causing the gas reserves to deplete, so it becomes indispensable to optimize natural gas in order to save its reserves for future use in diverse areas. This paper addresses technique based on neural network to optimize the natural gas usage in power plants by handling plant parameters correctly. Neural Network is an effective pattern recognition tool used to predict the output with the help of input parameters.

Process variables from various regions in power plant were taken as inputs and fuel flow as output. To improve generalization of the network, variable reduction techniques are applied after preprocessing the data. To find the degree of importance on the set of reduced variables, sensitivity analysis is done. By bringing variation in critical parameters isolated from sensitivity within the operational constraints on the trained neural network, change in output i.e. fuel flow was noted. It was observed that significant amount of fuel can be saved if the parameters are manipulated in the right dimension.

Keywords: *Process Variables; Artificial Neural Network; Data Preprocessing; Variable Reduction; Sensitivity Analysis*

1. Introduction

Sustainable development and economic stature of any nation depends on the indigenous energy reserves and its utilization. Owing to the ample resource and reliable supply, the consumption of natural gas is increasing worldwide at the most rapid rate amongst conventional sources. Natural gas not only occupy an enormous percentage of usage in domestic users but it also accounts for significant utilization in power plants and industries. Newly commissioned power plants eye natural gas as primary source of fuel for the tag of being more environmental friendly than its competitors, leading to lower carbon foot prints globally [1].

Among the developing countries like Pakistan, the economic balance immensely depends on the health of energy sector. It also affects the social development and future planning for the utilization of available domestic resources. Since the energy crisis is still unresolved, the gap between current energy demand and indigenous supplies is widening. Consequently, the dependency on the foreign oil market is increasing thus weakening the economic back bone of the country [2]. So natural gas utilization in Pakistan is increasing in all the different sectors which includes commercial, domestic and especially power, which is acting as an affordable and eco-friendly alternative to oil. The power sector is heavily dependent on natural gas, it accounts for more than 42% in thermal energy mix since 2009, as per NEPRA industry report 2015 statistics, shown in Fig. 1 [3].

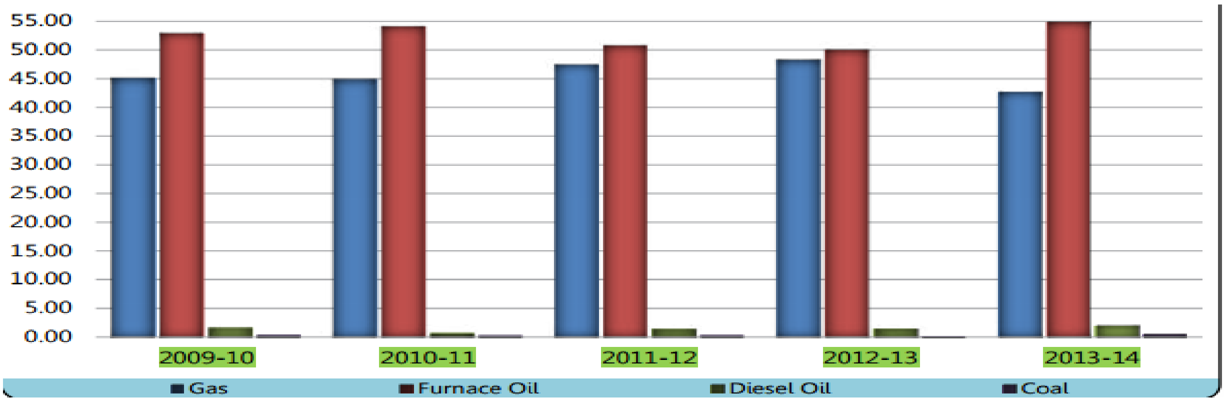

Fig. 1- Thermal Energy Mix [3]

The increasing energy consumption in various sectors in affecting our gas network by widening the demand and supply gap as shown in Fig. 2 [4].

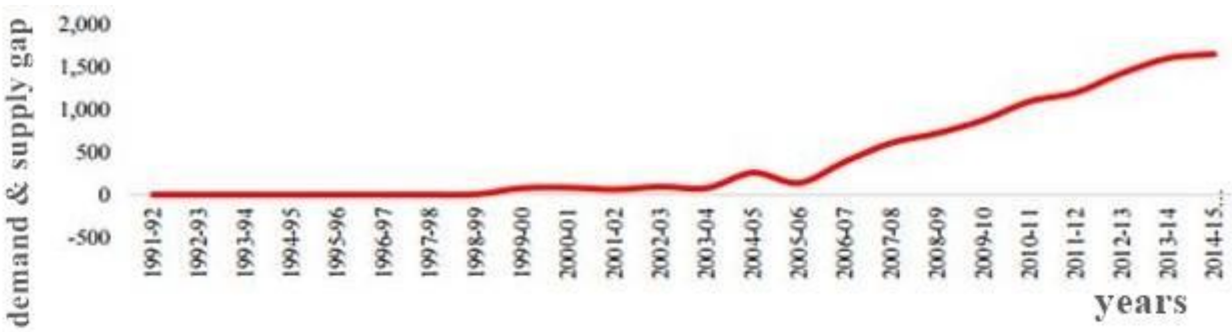

Fig. 2-Natural Gas Shortfall Data [4]

Perpetual usage is thereby affecting the reserves, so it becomes indispensable to address this issue and concrete steps are required for its optimization. Since power sector is utilizing natural gas more, so researchers are keen to work on its optimization in power plants. The trend is shifting towards artificial intelligence in order to solve highly complex issues like the one involved in power plant regarding its thermodynamics. Hence Artificial Neural Network (ANN) of the thermodynamics of a power plant can be used to determine the influence of changes in different process variables (inputs) upon the fuel flow (output). This information can be used to provide guidance to the plant operators and engineers as to where they should expand their efforts in order to regulate the fuel flow. The ANN is a technique with a flexible mathematical structure that can identify obscure patterns between input and output parameters, especially in case like power plant thermodynamics [5]. This paper presents step by step organized approach to optimize plant parameters using neural network. To capture important elements, variable reduction techniques like principal component analysis (PCA) and classification and regression trees (CRT) are used. To rank the inputs, sensitivity analysis is done. Raw data can harm the model and training process so appropriate preprocessing methods have been utilized.

2. Process Variables

Process or control variables are integral part of analyzing any power plant's performance. Their measurements are very important in controlling plant's performance indicators such as fuel flow. Real time analysis of process variables and maintaining them near to design constraints can improve fuel flow. The power plant's performance analysis is dependent upon innumerous complex process variables which are non-linearly related to each other and to performance indicators. Thus conventional methods fail to propose a universal relationship amongst them [6]. Process variable mainly comprise of temperature, pressure, level and flow variables. Temperature is one of the widely measured variables and holds immense importance from operation and maintenance aspect as its instability and non-linearity can harm equipment. Pressure measurement is needed for instance to ensure the safe operation of the boiler. For efficient quantity and quality of fuel burned, air flow measurement becomes necessary. While incorrect level measurement in vessels whether high or low than their optimum values can cause vessels to harm equipment or overflow leading to hazardous conditions [7].

3. Artificial Neural Network (ANN)

Owing to the complex non-linearity between the process variables and fuel flow ANN is used. ANN creates a predictive model which can formulate values of output (fuel flow) for different values of input (process variables). ANN is designed to model the organization and operation capabilities of the human brain thus ANN as compared to different predictive modeling techniques, is more adaptive towards highly varying scenarios posed while training for various nonlinear systems [8]. ANN comprises of many small units for processing called nodes or neurons, these nodes are arranged in layers named input, hidden and output layers also nodes are interconnected via links known as weights. Along with functionality of hidden layer and the activation function assigned to each node, input is mapped upon output [9]. Remarkable privilege that ANN provides lies in the fact that ANN generates a predictive model while being curtained to fundamental function equation between input and output, overshadowing the logistic regression and numerical methods which require vivid correspondence between input and output [10].

4. Methodology

The methodology is composed of four stages, each having its unique significance,

4.1 Power Plant Data Acquisition

This is the basic step which involves data preparation from power plant. The data was extracted by instrumentation and control (I&C) department and was provided to us for our analysis. Data is hourly averaged and is comprised of six months i.e. February, June, July, August, October and November which makes the tally of the sample size to 4320. The total number of variables are 30, includes process variables from various sections of power plant i.e. Gas Turbine (G.T), Steam Turbine (S.T) and Heat recovery steam generator (H.R.S.G). Refer table 1 for complete list of variables provided to us.

Table 1 - Plant Parameters

Plant Parameters(Unit)		
Active Power G.T (MW)	Compressor Discharge Pressure (PSIA)	Vacuum (bar)
PF	Compressor Pressure Ratio	Condenser Cooling Water Flow (tons/hr)
LP Shaft Speed (RPM)	STG Speed (RPM)	Sea Water Temperature (C)
Compressor Inlet Air Temperature(F)	Active Power S.T (MW)	HRSG Outlet Temperature (C)
Compressor Inlet Pressure (mbar)	HP Steam Pressure (barg)	GT -03 Exhaust (C)
Inlet Duct Pressure Loss (inch WC)	HP Steam Temperature (C)	GT -04 Exhaust (C)
Fuel Lower Heating Value (BTU/SCF)	HP Steam Flow (tons/hr)	HRSG Exhaust (C)
Water Injection Flow (lbs/hr)	LP Steam Pressure (barg)	HP DRUM CBD (T/HR)
Exhaust Duct Loss (inch WC)	LP Steam Temperature (C)	LP DRUM CBD (T/HR)
Fuel Flow (lb/hr)	LP Steam Flow (tons/hr)	Condensate Flow (T/HR)

4.2 Data Preprocessing

Data preprocessing plays a vital role in data mining, eliminating redundancy and noisy information thereby improving an analysis at the cost of time consumption. Data at first is prepared, cleaned, normalized, and transformed with appropriate data reduction techniques [11].

4.2.1 Data Cleaning

It includes operations that correct bad data, filter some incorrect data out of the data set and reduce the unnecessary detail of data [11]. The variables having constant data were removed because it can reduce neural network performance and increase training times. Inlet Duct Pressure Loss, Exhaust Duct Loss, HP Drum CB, power factor (approximated to be constant), Fuel Lower Heating Value, LP Steam Pressure and LP Drum CBD were removed as their values were constant throughout. The variables were reduced to 23 by data cleaning.

4.2.2 Data Transformation

In this preprocessing step, the data is converted or consolidated so that the mining process result could be applied or may be more efficient [11]. Compressor inlet pressure and discharge pressure can be transformed to compressor ratio, so instead using all the three mentioned, only compressor ratio can be selected, which makes our variables down to 21.

These were just inspection methods and based on their apparent condition, the variables were reduced. Now to further reduce the variables, proper reduction techniques based on certain principles will be applied after pre-processing the data.

4.2.3 Data Normalization

Data normalization is usually performed to achieve uniformity by assigning all parameters equal weights. New parameters named as modeling variables or analytic variables are obtained by varying original data parameters [11]. Decimal normalization technique is applied, which simply transformed the data into one significant Fig. before decimal point.

4.3 Variable Reduction

Variable reduction aims to augment the generalization of the network on datasets provided to it resulting in fast learning of the algorithm used for training. [11]. To achieve this, we've used two methods, one based on dimensionality reduction called Principal Component Analysis (PCA) and the other relies on regression and classification principle called Classification and Regression Trees (CRT).

4.3.1 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a dimensionality reduction approach which tries to reduce the dimension of input correlated variables, by transforming them into new set of variables known as principal components and which are not interrelated. PCA methods accomplishes this by keeping maximum variability from the data set [12]. This method is performed on MATLAB®. The first step is to find the minimum number of components which are able to explain maximum variability in the dataset. The first component generated always have highest variability, which tends to fall in the next components. We can see in table 2 that if we select starting five components, they explain approximately 96% of variability, which is obviously better.

Table 2 -Variability Explained

Components	Variability Explained (%)
C1	71.7
C2	10.95
C3	7.25
C4	3.79
C5	2.23
C6	1.74
C7	1.14
C8	0.28
C9	0.27
C10	0.25
C11	0.12
C12	0.11
C13	0.09
C14	0.04
C15	0.02
C16	0.01
C17	0.01
C18	0
C19	0
C20	0

This method is basically used for feature transformation, but in this paper we've used it as feature selection by analysing the coefficients of principal components. PCA assumes that the components are the linear combination of the variables present in the actual dataset, thus if coefficient of any variable has significant value, it will contribute more to the component and component itself explains greatly variability. Table 3 elucidates the coefficients of principal components.

Table 3-Coefficients of Principal Components

Variables	Components				
	C1	C2	C3	C4	C5
Active Power G.T	0.00028	0.160805	0.040841	0.035915	0.008074
LP Shaft Speed	0.000905	0.004677	0.002265	0.005167	0.000146
Compressor Inlet Air Temperature	0.13688	0.955645	0.083182	0.145961	0.044292
Water Injection Flow (lbs. /hr.)	0.034251	0.04525	0.037904	0.076094	0.042038
Compressor Ratio	0.005207	0.077501	0.012181	0.010055	0.000704
STG Speed	0.052569	0.077514	0.410187	0.113369	0.070926
Active Power S.T	0.26056	0.0446	0.017495	0.023694	0.026219
HP Steam Pressure	0.034187	0.046999	0.317633	0.028388	0.029368
HP Steam Temperature	0.010284	0.013975	0.08677	0.042839	0.015121
HP Steam Flow	0.9132	0.114939	0.081309	0.03224	0.07294
LP Steam Temperature	0.00589	0.011579	0.067171	0.037132	0.013392
LP Steam Flow	0.234712	0.056132	0.028633	0.026723	0.028348
Vacuum (abs)	0.020507	0.107088	0.736739	0.240499	0.088846
Condenser Cooling Water Flow	0.072479	0.10179	0.354393	0.882053	0.104901
Sea Water Temperature	0.015521	0.062237	0.152835	0.308146	0.003684
HRSG Outlet Temperature	0.014378	0.006518	0.005934	0.000871	0.024826
GT -03 Exhaust	0.014152	0.005152	0.092124	0.109233	0.515703
GT -04 Exhaust	0.064126	0.008127	0.024346	0.095769	0.787566
HRSG Exhaust	0.014723	0.005162	0.006702	0.004913	0.075068
Condensate Flow	0.091785	0.006884	0.002018	0.018972	0.267807

Now to select the variables, we have simply ranked the coefficients of each component and established a *shortlisting* criteria that if a coefficient in first three components ranks less than six, the corresponding variable of that component is selected which is shown in table 4. The highlighted cells explains this criteria. The top three components also explains almost 90% variability. The approach is adopted from [13].

Table 4-Ranking of Coefficients

Variables	Ranking of Coefficients				
Active Power G.T	20	2	11	11	17
LP Shaft Speed	19	20	19	18	20
Compressor Inlet Air Temperature	4	1	8	4	9
Water Injection Flow (lbs. /hr.)	9	11	12	8	10
Compressor Ratio	18	7	16	17	19
STG Speed	8	6	2	5	8
Active Power S.T	2	12	15	15	13
HP Steam Pressure	10	10	4	13	11
HP Steam Temperature	16	13	7	9	15
HP Steam Flow	1	3	9	12	7
LP Steam Temperature	17	14	10	10	16
LP Steam Flow	3	9	13	14	12
Vacuum (abs)	11	4	1	3	5
Condenser Cooling Water Flow	6	5	3	1	4
Sea Water Temperature	12	8	5	2	18
HRSG Outlet Temperature	14	17	18	20	14
GT -03 Exhaust	15	19	6	6	2
GT -04 Exhaust	7	15	14	7	1
HRSG Exhaust	13	18	17	19	6
Condensate Flow	5	16	20	16	3

4.3.2 Classification and Regression Trees

This regression based method is extensively being used by the researchers for feature identification. This task is accomplished by using IBM SPSS Statistics software. This method helps to identify groups and the relations present among them, further information is explained in [14]. Though classification is not a priority in our case, but we can use results from IBM SPSS generated report which will help us to identify crucial variables which makes up the decision tree. A very useful information that the software provides is the normalized importance of each variable, as shown in table 5

Table 5- CRT Inputs

Independent Variable Importance		
Independent Variable	**Importance**	**Normalized Importance**
Active Power G.T	**.002**	**100.0%**
Compressor Inlet Air Temperature	**.001**	**54.3%**
Compressor Ratio	**.001**	**50.3%**
Water Injection Flow (lbs. /hr.)	**.001**	**49.7%**
HP Steam Flow	**.001**	**25.7%**
Active Power S.T	**.001**	**22.4%**
LP Steam Flow	**.001**	**22.2%**
GT -03 Exhaust	**.001**	**21.4%**
Sea Water Temperature	**.000**	**14.2%**
Condenser Cooling Water Flow	**.000**	**13.9%**
Condensate Flow	**.000**	**7.7%**
Vacuum (abs)	**.000**	**7.0%**
HRSG Exhaust	**.000**	**6.7%**
HP Steam Pressure	**.000**	**5.9%**
HP Steam Temperature	**.000**	**5.8%**
LP Steam Temperature	**.000**	**5.3%**
HRSG Outlet Temperature	**.000**	**4.0%**
GT -04 Exhaust	**8.503E-5**	**3.5%**
LP Shaft Speed	**5.153E-5**	**2.1%**
STG Speed	**4.718E-5**	**1.9%**
Growing Method: CRT		
Dependent Variable: Fuel Flow		

From the table it is clear that the starting variables defines more about the target variable (fuel flow) as compared with the others. Thus, we selected top 10 variables for further analysis. Since this method works on regression principle, it takes into account the output variable i.e. fuel flow as well unlike PCA method.

4.4 Neural Network Implementation

After having preprocessed data and reduction in variables, we will test the performance of both the models used in variable reduction i.e. PCA and CRT. Levenberg–Marquardt training algorithm was chosen and data was divided as 70% training and remaining 30% was split equally among testing and validation. The table compares the results of both the models with the one having all the variables present. The model having overall performance closest or having least mean squared error (mse) to that when all the inputs are used, will be selected. Table 6 provides all the information about performance of models used.

Table 6- ANN results

Mean Squared Error	**All Inputs**	**PCA Inputs**	**CRT Inputs**
Training	4.51E-04	7.79E-04	5.87E-04
Testing	5.51E-04	8.16E-04	6.22E-04
Validation	4.63E-04	7.41E-04	6.29E-04
Average	4.68E-04	7.79E-04	5.99E-04
Correlation	94.14%	90.64%	92.44%

CRT models provides better result as compared to PCA. Ten inputs almost providing the same results as if all were selected, thus a better model is achieved in which output is dependent on lesser inputs though providing same accuracy. If we observe closely, the variable 'Vacuum' was in PCA, and also ranks just below 10 in CRT model. If this parameter is added in our existing model, it becomes more accurate. Let this model be called 'SUB model' which contains CRT inputs plus one extra added variable. Table 7 explains final results and Fig. 3-5 shows their respective regression models.

Table 7- ANN Final Results

Mean Squared Error	**All Inputs**	**PCA Inputs**	**CRT Inputs**	**SUB Inputs**
Training	4.51E-04	7.79E-04	5.87E-04	4.70E-04
Testing	5.51E-04	8.16E-04	6.22E-04	5.29E-04
Validation	4.63E-04	7.41E-04	6.29E-04	5.27E-04
Average	4.68E-04	7.79E-04	5.99E-04	4.88E-04
Correlation	94.14%	90.034%	92.44%	93.88%

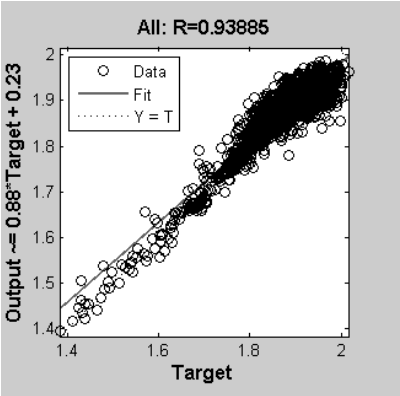

Fig. 3-SUB Model

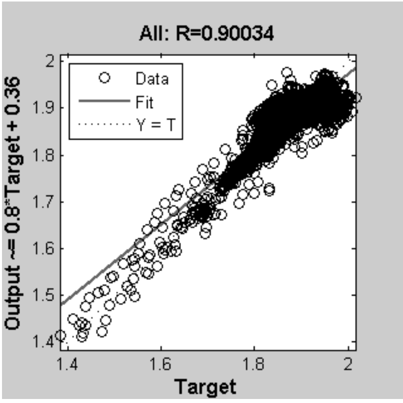

Fig. 4-CRT Model

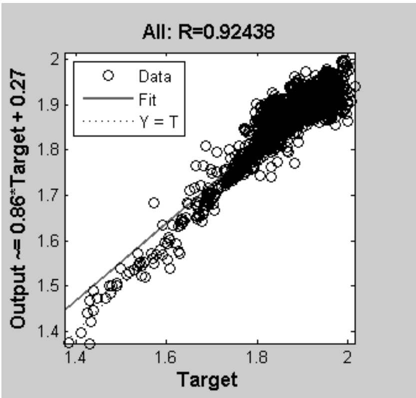

Fig. 5-PCA Model

This SUB model is achieving nearly the same performance when all variables are selected. Thus we've managed to reduce the variables from 20 to 11 after apply rigorous reduction techniques. The mse of neural network changes every time whenever the training is restarted, because of different initial conditions. In order to create homogeneity in all the models, we've assigned an initial value of '1' so that its performance doesn't vary.

4.5 Sensitivity Analysis

We've gathered all the inputs which have greater importance, now they must be ranked under sensitivity analysis. The method used is called Change of mean square method (COM)

According to the COM method, firstly all the input parameter were used for neural network training and mse is noted then in order to identify the importance of each parameter neural network is retrained by removing one parameter at a time, the alteration in mse indicates the relative importance of that parameter. If mse increased drastically the parameter removed holds critical importance, whereas slight variations in mse undermine the importance of respective parameter [15]. Table 8 shows most sensitive top 5 variables that have profound impact on fuel flow.

Table 8- Sensitivity Analysis

Sensitive Parameters	M.S.E Change
Vacuum (abs)	0.000409273
Active power G.T	0.00029079
Compressor Inlet Air Temperature	0.000270867
Sea Water Temperature	0.000269752
Water Injection Flow (lbs/hr)	0.000263957
Compressor Ratio	0.000229618

Here apart from active power, all the other variables can be manipulated (directly or indirectly) and the corresponding change in fuel flow can be observed. Active power neglected because it is itself a dependent variable, and its inclusion was only helpful for neural network to learn the relation between input and output.

5. Results

Among the variables isolated from sensitivity, vacuum and water injection flow can be easily varied within the limits in power plant. Parameters like Inlet air temperature or sea water temperature, needs to be separately addressed. Now we take a sample from our data and test the change in fuel requirement,

It is seen from data that at 07/08/2015 15:00,
Actual Fuel Flow = 18728.55 lb./hr.
Predicted using ANN = 18778.23589 lb. /hr.

Now by using our two crucial variables obtained which are Vacuum and Water injection Flow on ANN, the following results were observed.

By lowering Condenser Vacuum from 0.1 bar to 0.09 bar
Fuel Flow = 18706.02469 lb/hr
By increasing Water Injection Flow from 13.33 lb/hr to 13.46 lb/hr
Fuel Flow = 18677.07918 lb/hr
Deviation in fuel flow 'β'= 18728-18677
= 51 lb/hr.
The net saving for whole month would be,
$$B = 51 \times 24 \times 30$$
= 36720 lb

Here it is evident that by manipulating parameters in right dimension and within operational constraints, significant fuel can be conserved.

6. Conclusion

A simple approach is addressed in this paper on how to manage and reduce excessive fuel usage in power plants by manipulating the parameters correctly. Plant data is first prepared and after applying variable reduction techniques, the final model was created for testing. Variables most dependent on fuel flow were isolated using sensitivity study and were varied. It is seen that up-to 37000 lb. of fuel can be saved if the crucial variables are operated on values provided by neural network. This analysis also provides operators influential parameters where they can maximize their efforts to work on, resulting in fuel (natural gas in our case) conservation. This approach can be used in other plants such as coal or oil fired to optimize their fuel sources as well.

Acknowledgements

We would like to express our sincere gratitude to Mr. Javed Mateen (DGM K-electric) for providing plant data. The efforts of Mr. Ali baig (Lecturer NEDUET,EED) cannot be neglected as well who made significant contributions in this study.

References

[1] "International energy outlook 2016-Natural gas - energy information administration," 2016. [Online]. Available: http://www.eia.gov/forecasts/ieo/nat_gas.cfm. Accessed: Sep. 12, 2016.
[2] [Online]. Available: https://www.usip.org/sites/default/files/SR375-Pakistans-Power-Crisis-The-Way-Forward.pdf. Accessed: Sep. 23, 2016.
[3] Available: http://www.nepra.org.pk/Publications/State%20of%20Industry%20Reports/State%20of%20Industry %20Report%202015.pdf. Accessed: Sep. 23, 2016.
[4] Available: http://www.finance.gov.pk/survey/chapters_15/14_Energy.pdf. Accessed: Sep. 23, 2016
[5] P. Deshpande, N. Warke, P. Khandare, and V. Deshpande, "Thermal power plant analysis using artificial neural network," 2012 Nirma University International Conference on Engineering (NUiCONE), Dec. 2012.
[6] B. Saha, V. Patel, and K. Chaterjee, "Assessment of process parameter to improving power plant performance," 2014 Innovative Applications of Computational Intelligence on Power, Energy and Controls with their impact on Humanity (CIPECH), Nov. 2014.
[7] S. Basu and A. Debnath, Power plant instrumentation and control handbook: A guide to thermal power plants. United States: Academic Press, 2014.
[8] L. Zhang, J.-H. Jiang, P. Liu, Y.-Z. Liang, and R.-Q. Yu, "Multivariate nonlinear modelling of fluorescence data by neural network with hidden node pruning algorithm," Analytica Chimica Acta, vol. 344, no. 1-2, pp. 29–39, May 1997.
[9] O. A. Abdalla, M. N. Zakaria, S. Sulaiman, and W. F. W. Ahmad, "A comparison of feed-forward back-propagation and radial basis artificial neural networks: A Monte Carlo study," 2010 International Symposium on Information Technology, Jun. 2010
[10] J. V. Tu, "Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes," Journal of Clinical Epidemiology, vol. 49, no. 11, pp. 1225–1231, Nov. 1996.
[11] S. García, J. Luengo, and F. Herrera, Data Preprocessing in data mining in Intelligent Systems Reference Library. Springer International Publishing, 2015.
[12] I. T. Jolliffe, Principal component analysis in Springer Series in Statistics. Springer New York, 1986.
[13] R. Platon, V. R. Dehkordi, and J. Martel, "Hourly prediction of a building's electricity consumption using case-based reasoning, artificial neural networks and principal component analysis," Energy and Buildings, vol. 92, pp. 10–18, Apr. 2015.
[14] "SPSS decision trees - SPSS decision trees," 2016. [Online]. Available: http://www-03.ibm.com/software/products/en/spss-decision-trees. Accessed: Jul. 20, 2016.
[15] A. H. Sung, "Ranking importance of input parameters of neural networks," Expert Systems with Applications, vol. 15, no. 3-4, pp. 405–411, Oct. 1998.