GRANT SPONSORED BY YAYASAN FELDA

USM PROJECT CODE NUMBER A0567

COMPREHENSIVE REPORT

Project Title: Controlling the Emission of Pollution from the Boiler of Palm Oil Mill by Using Neural Network and Genetic Algorithm

"Pengawalan Pencemar Pencemar Udara Dari Dandang Kelapa Sawit Dengan Menggunakan Rangkaian Neural Dan Algoritma Genetik"

Project Leader: Assoc. Prof. Dr Ishak Hj Abdul Azid/ Prof. K.N.Seetharamu
School of Mechanical Engineering
Universiti Sains Malaysia
Engineering Campus
14300 Nibong Tebal
Penang

Co-Researcher: Prof Abdul Latif Ahmad*, Prof. K.N.Seetharamu^, Assoc. Prof. Dr Zaidi Mohd Ripin^
School of Chemical Engineering*,
School of Mechanical Engineering^,
Universiti Sains Malaysia
Engineering Campus
14300 Nibong Tebal
Penang

Duration of Project: 1 July 1999 – 1 July 2000

Contents:

APPENDIX A

Extract from MSc Thesis entitled: Predicting Smoke Emission from Palm Oil – Ahmad Razlan Yusof (March, 2003).

1. INTRODUCTION

For the palm oil mills in Malaysia, the palm oil waste materials, i.e. fibre and shell are used as fuel as a way to save the biomass material being exhausted. In other words, the palm oil mill can independently utilize the waste material to produce electrical energy. However, the use of these biomass materials greatly releases emission from the chimney to the atmosphere. From the combustion process, the pollutants released mainly consists of Carbon Monoxide (CO), Particulate Matter (PM), Nitrogen Oxide (NO_x) and Sulphur Dioxside (SO_2). According to a survey in 1999, only 76% of the 309 palm oil mills in operation meet the allowable limit of emission [1]. According to Leong [2], for every 1 tonne fresh fruit bunches (FFB) burnt, 1.5 g/Nm^3 particulate level is produced. For a typical 500 000 tonnes of crude palm oil generally produced from a mill in a year, it may require the burning of 17 500 000 tones of FFB. This may result the release of the pollutant of 1 to 1.5 tonnes particulate matter to atmosphere per mill. For an extended period, this level of pollution will certainly threaten the lives of human, animals, plants and buildings.

The smoke emission factors are something that are complex, and are influenced by many processes. The factors such as combustion, boiler process and mill process are related to each other in producing emission from steam power plant generation. The combustion of fibre and shell is recognized as the main factors for boiler emission due to the fluctuating fuel flowrate entering into the furnace. The variable fuel flow affects combustion, steam generation, electric energy and finally the emission [3]. Despite the fuel transfer system being driven constantly by motor to supply the fuel, the problem exists as fibre and shell inlet capacity is not stable in the trolley. It will also cause fluctuating calorific value and moisture, or fuel quality interrupting the combustion rate and its temperature. Briefly, not only fuel capacity causes the emission, but also the fuel types and its quality play a role in the emission production and cannot easily be controlled.

Steam generation from palm waste material combustion depends on the power load required. Li and Priddy [4] said the processes and power needed influence fuel inlet capacity, air and water. When power is unstable, it influences other factors to change their parameters automatically. High power electricity requires low pressure and high temperature steam [3]. Therefore, the electric energy generated from turbine is proportional to the mill operation which involves many processes. The ventilation system is automatically controlled by the fuel inlet and inlet water capacity. The air supply will increase when excess of water in boiler drum and fuel flows in the furnace exist. Although excess air supply is to achieve complete combustion, it also causes low combustion temperature and in return producing NO_x gas. However, less air supply causes CO emission with attendant less steam generation. It will decrease boiler efficiency, power turbine generation and interrupt optimum heat transfer. As a result, the air capacity, temperature and excess air are so complex to control and optimise in order to minimize smoke production. These processes are related to each other and are highly non-linear. They cannot be illustrated by direct equations to forecast the emission.

As discussed, the complex factors and contributors to many processes in the palm oil mill have no direct or simple relationship in predicting the emission. The operations in the palm oil mill, the processes and operations involve furnace, boiler and turbine which are closely related to each other. Figure 1 illustrates the factors complexity to emission such as steam pressure (Sp), steam capacity (Sc), feed water (Fw), steam temperature (St), furnace combustion (Fc), boiler outlet (Bo), water level (Wl), ambient temperature (At), flue gas temperature (Ft), excess air (Ea), fibre flowrate (Ff), shell flowrate (Sf), power output (Po), oxygen (O_2) and carbon dioxide (CO_2). The output emissions include carbon monoxide (CO), nitrogen oxide (NO_x), sulphur dioxide (SO_2) and particulate matter (PM). This Figure also simplifies process of input and output variables. Controlling emission can be done if the process of input factors can be controlled as the input indirectly is responsible for the outputs or pollutants. As these parameters are highly complex and non-linear, any activity happening inside the process can be regarded as a black box. The main concern from this process is to monitor and control the release of emission. As the whole process involves the cause and effect or input and output, ANN is found to be feasible method to model the boiler emission from palm oil mill.

In this paper, ANN is applied to monitor the release of the emission from the palm oil mill. ANN is briefly discussed and later the method of selecting the major and minor input variables is explained.

2. ARTIFICIAL NEURAL NETWORK AND ITS APPLICATION IN THE PREDICTION OF EMISSION

ANN is adapted from human biological neuron that reacts to give result after receiving the information [5]. Actually, this ANN is the integrated combination between signal processing using mathematical formula that function as a human nerve. Neural networks must learn how to process input before they can be utilized in an application. The process of training a neural network involves adjusting the input weights on each neuron such that the output of the network is consistent with the desired output. This involves the development of a training file, which consists of data for each input node and the correct or desired response for each of the network's output nodes. Once the network is trained, only the input data are provided to the network, which then recalls the response of the network learned during training.

Information received by neuron determines the command signal product either to the back or forward or to lateral. Maren *et al.* [6] stated the characteristics of the ANN structure can be divided into three basic models; i.e. feed forward network, feed backward network and lateral network. Feed forward network is indicated by the neurons supplying the output to the next neuron layer only while feed backward network enables the neurons to supply the output to the next or previous neuron.

Lateral network is a network when the output becomes the lateral input neuron. The network layer can be classified into single layer, bilayer and multiple layer. Bilayer networks have feedforward or feedbackward without lateral network. Multilayer network consists of one or more hidden layers. Current study uses feedforward structure and backpropagation algorithm as they are suitable for modelling boiler emission. Feedforward network is a fast trained stable network which is suitable for nonlinearity. The utilization of backpropagation in the project is selected due to its acceptability to learn the problem efficiently. The structure and algorithm is very widely used due to its successes in several applications [7-9], especially in boiler emission model [7-8]. Backpropagation ANN is a versatile tool and has been successfully applied to a wide range of complex science and engineering problems.

2.1 Review of Published Works in Predicting Boiler Emission

Grott [7] has reported successful boiler emission modelling in Rheinbran Corplent and Health Project in German and Ostrulak power plant in Poland. Other projects are petroleum refinery factory in Anaco, Texas city, Westinghouse power plant, Wisconcin electric plant, Florida and Pavilion power plant in USA. In Britain, the same successful projects are also developed by BCURA (British Coal Utilization Research Association), GNOCIS (Generic NO_x Control Intelligent System) and DTI (Department Trade and Industry) [7,11] Interested readers can refer to paper based reviews [8,12-15] to find more success study of application of ANN in boiler emission controlling.

Table 1 summarizes several papers related to current study of predicting boiler emission by using ANN. Several of them used ANN to control boiler emission as a substitution in using CEMs at coal power plant. However, [16,17] implemented NN at paper mill boiler, [18] used it at incinerator and [19] used NN to predict biomass gasification. Therefore, most of the papers in Table 1 describe about the utilization of NN at coal power plant and only [19] focuses on biomass emission modeling. No one yet applies on palm oil mill boiler emission that uses biomass as fuel.

Table 1 also shows the input and output variables, as obtained by the 18 researchers in ANN boiler emission modeling. Based on the papers, some researchers applied different variables from other researchers even though to predict the same pollutants. For example, Radl and Tronci [20,21] focused on predicting CO and NO_x, but they used different input variables. However, the use of different input variables does not greatly affect the prediction accuracy.

Based on the papers listed in Table 1, the number of times certain input and output variables used in the prediction can be counted. From the total of the last row in Table 1, fuel capacity and excess air are commonly used as input variables (total 10). The second commonly used input variables are air capacity, air fuel ratio, and combustion temperature. These variables are chosen depending on the processes and factors that contribute to the pollutants. Based on their investigation and experimental results, some researchers employed different variables as input parameters to predict the output. For example, to model NO_x as an emission from coal boiler, Rowland [22] used air capacity, fuel capacity, excess air, ambient temperature and firing rate as input variables. However, Grott [7] found hooter flame temperature in combustion process caused NO_x released, while [17] justified that NO_x emission was due to the pneumatic conveyer transportation system. As a result, Grott [7] used air capacity, air fuel ratio, combustion temperature, firing rate and boiler efficiency, while Euhos and Blanc [17] employed air capacity, air temperature, fuel capacity, combustion temperature and excess air as input variables. Thus, the input variables cannot easily be determined and they depend on the processes and conditions contribute to the pollutants.

So far to predict 8 output pollutants, the researchers have used 20 input variables as listed in Table 1. Radli and Troci *et. al.* [20,21] focused on predicting CO and NO_x only, while other researchers [18,19,23,24] carried out work to predict various pollutants. The production of NO_x, CO and PM increase is due to the maximum combustion under unstable load contributed to fluctuating of air fuel ratio [24]. So far from the literature, nobody has compared the simulation boiler emission with 2 or more different mills as reported by EPA USA. The successful Pavilion project installation of 200 PEM in USA used their own power plant data collection to validate RATA (Relative Accuracy Test Audit).

The input and output variables need to be measured during the data collection. There are online or offline data collection that can be carried out, but most researchers used online data collection as it can give data in real-time sequence. For input data collection, Laux [25] used electric charge transfer system (ECT); Hou *et.al.* [26] applied digital computer signal (DCS) and other people employed sensor and gage that interface with software to collect the data through online. Most of the researchers used the available continous emission monitoring system (CEMs) in order to determine the pollutants amount. Androscoggin [18] used source sampling method and Kamal [24] used electric static precipitator to collect PM. In this paper, the data collection from steam plant palm oil mill is taken at five locations (turbine, boiler, fuel inlet, exhaust and stack chimney). These locations are chosen as the sources of the input and output variables, totalling to 15 input variables such as fibre, shell, steam generated etc as shown in Figure 2. Output variables of emission (CO, NO_x, SO_2 and PM) are collected from the stack chimney. The data at boiler and steam turbine are directly taken from their reading display. However, at fuel inlet capacity, the fibre and shell capacity have to be weighted manually since no automatic measurement is available. Gas analyser and isokinetic sources sampling are utilized at the stack chimney to measure the pollutants released.

2.2 Use of Sensitivity Analysis via Neural Network (SANN) to find Major and Minor Input Variables

Each output is caused by several different input variables. These input variables influence and contribute to output variables. Some certain variables might have big contribution or influence to certain specific output variables only, and some might give small influence to those specific output *variables. The input variables which have great impact on a particular output variable are called major input variables whereas which give small impact on that particular output variables are called minor input variables. Major input variables reduce simulation time, increase performance and make the model more generalized. All 15 input parameters taken can be divided into major and minor variables. In order to reduce and eliminate minor input variables several methods have been proposed by many researchers through statistics and neural network simulation. In neural network approaches, several methods can be used such as sequential zeroing weights (SZW), system variation variables (SVV), sensitivity analysis (SA), fuzzy curves, changing mean square error and analysis of effectiveness. In this research, the sensitivity analysis is carried out to identify minor and major input variables used in boiler emission modelling. Most resérchers such as [27-29] suggested using partial differential sensitivity analysis. All types of SANN can rank and select the major and input variables through its analysis. Present study will used SANN using partial diffrential approach as it was successfully used by other researchers. Sensitivity analysis with partial differential is based on calculation of input, weights and output variables from the ANN simulation [27]. All neural network models for sensitivity analysis use the training data to perform sensitivity value. By referring to sensitivity value, the input variables can be ranked for their contribution to the output. The results with a value more than 1 represent major factors and the value less than 1 represents minor factor.*

Prior to ANN application, an approach is required to make the network structure optimum in order to achieve generalized and high accuracy network with less simulation time. During training, neural network will learn from input data and adjust the weight smoothly to give desired output. According to Maren *et.al.* [6], the number of hidden layers depends on the data, which can be determined through training. More number of hidden layers cause long time for training and will produce unstable network [30] and overtrain [9]; but it produces better results [31]. The suitability of time and accuracy must be considered. Trial and error method is usually applied to get the optimum nodes and layer numbers [6]. High number of hidden layer/nodes causes a slow time for training, but usually gives better prediction. In contrast to that, less number layer/node gives incorrect prediction. As a result, to find an optimum structure through the trial and error method, time for training, number of iterations and square error achieved by network with different hidden node as indicated by 'h' in Table 2 should be considered. In this research, only single layer is employed referring to Kolmogorov's theorem [32]. Briefly stated, Kolmogrov claimed that a single hidden layer or three layers backpropagation network is enough to map any function to very high precision and better prediction.

After obtaining the optimum structure for all 5 model of simulation (4 for each output and 1 for all output), the SANN are carried out. A ThinkPro software version 1.05 is utilized for this purpose. The software uses neural network setting parameters as shown in Table 2. In this study, the Levenberg-Marquart algorithm is used as suggested and used by many researchers [8,30,33-35]. During this neural network process, the weights and bias are initialed in randomized order. For a present task, to minimize the time taken for simulation, and also to follow a common practice regarded to be successful, the learning rate value of 0.01 and momentum 0.95 are utilized. Both values are applied as initial value for a network that can be adjusted during training. The weights are automatically adjusted depending on the learning rate and momentum to obtain minimum comparison value between actual and predicted output using backpropagation method. The simulation stops after reaching 1000 epoch or square error of 0.001 between the actual and predicted value.

2.3 Parameters Settings of Neural Network

After obtaining the optimum structure of the network and before ANN simulation is carried out, the ANN parameters are needed to be optimized in order to obtain high accuracy, less simulation time and low square error networks. These are carried out similar to the SANN approaches, but the parameter optimization is focused on transfer function and scaling method. The transfer functions utilized are logsig, purelin and tansig for the connection between the input, hidden and output layers. These three parameters are interchanged with each other such as logsig purelin, tansig purelin, logsig logsig, tansig tansig, logsig tansig and tansig logsig as combination of transfer functions. A technique based on the trial and error method is employed as successfully done by other researchers such as [5,36,37]. For scaling method, many different approaches can be applied such as conventional method, standard deviation & mean method and minimum & maximum method. These three scaling methods and transfer functions are exchanged together to find the optimum parameter settings based on the time requirement, accuracy and error of goal achieved by the network.

A full ANN models are then constructed based on similar concept to SANN, but it requires prolonged simulation process. The boiler emission model simulation process uses training, validation and testing in order to obtain a generalized model [38]. Training period is an initial approach to create randomized weights that will be adjusted to achieve a certain square error. During validation stage, the final weights from training period will be updated using simulation for new sets of training data obtained from lower square error than training period. This is also known as generalization phase in which a wider range of data value is used than in the training period. In order to verify the network either successful or not, the testing period data is needed. In this testing period, input data that are excluded from both training and validation periods are used. The training and validation periods are needed to increase the network capability, accuracy and will solve the overfitting problem that is always faced by feed forward back propagation. This problem is the error shown during training is small, but large error occurs during testing section [5,39]. A low error between actual and predicted output indicates the accuracy of the model.

3. RESULTS AND DISCUSSION

3.1 Modelling Boiler Emission for Mill No. 1

3.1.1 Sensitivity *Analysis* Neural Network (SANN)

Prior to the sensitivity analysis, trial and error method as discussed is employed to optimize the NN structure. For example, Figure 3 shows the results of trial and error method for CO. From the graph, the optimum structure can be determined from the minimum mean square error (mse) of 0.0089239 at the value of 27 (number of hidden neuron). So, the optimum structure for CO model is [15,27,1] as illustrated by Figure 4. The first column refers to the input neuron number, the second column for the number of hidden nodes and the third column for the number of output nodes. Similar procedure is carried out for other output variables to find optimum structure of NN and the results are as follow:

NO_x	[15,27,1]
SO_2	[15,28,1]
PM	[15,26,1]

With the same trial and error method, the optimum structure of [15,30,4] is found for multiple input and output model. The minimum mse for multiple output is slightly bigger than that for single output because of more than one output neurons used in the output layer make the network difficult for prediction.

3.1.2 Major and Minor Input Variables Selection

The SANN is carried out to find the major and minor input variables to the emission released based on the above optimum model structures. The analysis is carried out by 'ThinkPro' software using all sets of data and the results are shown in Table 3. Based on the sensitivity analysis, the ranking of all input parameters for each pollutant is established. Minor input variables are represented by the value less than 1, while major input variables are indicated by the value more than 1. The SANN values with the bold show the major input variables for the corresponding output. As shown in Table 3, about 40 % from all input variables are major input variables to CO, while about 47 % to NO_x. However, the percentage of major input variables increase for both PM and SO_2, i.e. about 60 %. For multiple output model, about 60 % from all input variables are major input variables. These major input variables are slightly different from the single model. Different emission outputs demonstrate different major input variables from all 15 input variables.

In general, the main contributors to the emission come from fuel, turbine, air and boiler parameter. From SANN and CC analysis, fibre and shell appear as major contributors to all pollutants. For multiple output models, only 9 out of 15 input variables are found to be major input variables.

3.1.3 Optimization of Structure and Parameters

The scheme to find optimum structure are carried out here similar to the procedure in Section 3.1.1, but only the major input variables are utilized to find suitable number of hidden nodes. This approach depends on the results determined from Section 3.1.2 in which inputs are the major input variables. The major input variables for each model are different from each other. The CO model employs 6 input variables, NO_x 7 major inputs, SO_2 9 major inputs and PM 9 major inputs. For multiple output model, only 6 major variables are used. From the structure optimization analysis using trial and error method, the optimum structures for the models are as follows:

CO	*[6,10,1]*
NO_x	*[6,101]*
SO_2	*[9,18,1]*
PM	*[9,17,1]*
Multiple	*[9,18,4]*

After the optimum structures for the neurons are achieved, the investigation is further carried out to find the optimal parameter for each network. To define the optimal parameter setup, the considerations should refer to the accuracy, simulation time and number of epochs or iterations taken by the models. The parametric study of the parameters used for ANN is carried out and the results are as follows:

CO	*minmax-scaling with logsig purelin transfer functions*
NO_x	*minmax scaling with logsig purelin transfer functions*
SO_2	*stdmn scaling with tansig purelin transfer functions*
PM	*minmax scaling with logsig purelin transfer functions*

Multiple stdmn scaling with tansig purelin transfer functions

3.1.4 *Simulation with Single Output Models*

Single output model with all input variables and with only major input variables for each pollutant are developed. The regression analysis between actual and predicted output is applied to demonstrate the accuracy of the models as shown in Table 4. It shows the r-value for training, validation and testing period for both models. Both models show accurate prediction during training and validation and a slight deviation for r-value in the testing segment. In Table 4, number 1 refers to all input variables and number 2 refer to major input variables only. The model for major input variables gives better accuracy than all input variables model as demonstrated by high r-values. Therefore, the results from major input variables model is used to make comparison between actual and NN prediction for the emission.

Figures 5.a, b, c, and d show the comparison values between actual and predicted values by NN for CO, NO_x, SO_2 and PM respectively. The predicted values of CO by NN are almost at the same values with the actual value collected from the mill. High accuracy is obtained where the average error is only 0.1 %. The maximum percentage error in testing part is about 5.4 % at data point 110. For NO_x, the same pattern for both the actual and predicted values is obtained as shown in Figure 5.b. The average error is 0.5 % and the maximum error of 5.8 % occurs at data point number 97. High accuracy of prediction is also obtained for SO_2 and PM as shown in Figure 5.c and 5.d respectively. The average error for both predictions is less than 1.0 % and the maximum percentage error for both predictions is less than 8.5 %.

Based on these results, NN can accurately predict the value for all the emission. The accuracy of the prediction is very high as the average error for all the prediction is less than 2 % and the maximum percentage error is less than 8.5 %.

3.1.5 *Simulation with Multiple Output Model*

Figures 6.a, b, c and d show the comparison between predicted and actual outputs for multiple outputs for CO, NO_x, SO_2 and PM respectively. The simulation took 113.59 seconds computational time to get 2.16E-15 mse for training and 9.42E-15 mse for validation. The multiple model shows the lowest mse due to a high number of hidden nodes to predict the output. On the whole, the figures show that NN with multiple output can predict accurately. The average error between the predicted value and experimental data for all pollutants is within 6 % with maximum error within 16 % occurs during the testing period. The details of the average error and maximum error for each pollutant are as follows:

CO average error 0.2 % and maximum error of 12.9 % at data number 97

NO_x average error 2.0 % and maximum error of 8.3 % at data number 112

SO_2 average error 5.1 % and maximum error of 11.8 % at data number 112

PM average error 2.7 % and maximum error of 15.8 % at data number 110

3.1.6 *Comparison Between Single and Multiple Output NN Modelling*

The comparison between the single and multiple model of NN can be made by referring to the iterations taken by the network to converge, mean square error (mse), maximum error achieved and the time required during the whole simulation. Both amount of epoch and time used by multiple model is better than those of single model. The mse of multiple model also gives lower square error than single models. However, the single model give better accuracy where the maximum error is less than 8.5 % compared to the multiple model which is 16 %. The single model also gives better accuracy where the average percentage error is less than 1.5 % compared to the multiple model which is 2.5 %. Thus, single model can predict better than the multiple output model. However, both models can be employed for the emission modeling as the average error for prediction is very small, i.e. 2.5%.

3.2 Modelling Boiler Emission for Mill No. 2

By using the previous procedures and equipments, the 65 sets of data are collected from another palm oil mill at Selama, Perak. This palm oil mill used similar boiler's specification as the first mill, but this mill No. 2 uses newer steam power boiler than the mill No. 1.

3.2.1 Major and Minor Input Variables

The analyses of new data collected from mill No. 2 use the methods similar to mill no.1. Based on SANN, the major and minor input variables for each emission are the same as previously obtained from mill No. 1. The utilization of neural network sensitivity analysis is accomplished well to both data samplings. Therefore, the previous optimum structure of major input variables will be applied in NN simulation.

3.2.2 ANN Model

In this simulation, similar procedure as in the first mill is employed to mill No. 2 data. The optimum model and parameter settings of all input variables for the previous simulation of mill No. 1 are utilized to update the weights of the new data collected from the second mill. All input variables and only major input variables are carried out to model NN where all input variables are simulated in order to make comparison with MLR which also uses all input variables.

Figures 5.7.a, b, c and d show the comparison between predicted and actual outputs for single output for CO, NO_x, SO_2 and PM respectively. The simulation took 44.00 seconds computational time to get 3.11E-7 mse for validation. On the whole, the figures show that

NN with multiple output can predict accurately. The average percentage error between the predicted value and experimental data is within 6 % with maximum error within 12 % is obtained. The details of the errors for each pollutant are as follow:

CO average error of 0.6 % and maximum error of 7.8 % at data number 60

NO_x average error 1.9 % and maximum error of 8.9 % at data number 63

SO_2 average error 3.5 % and maximum error of 11.8 % at data number 112

PM average 3.0 % and maximum error of 7.9 % at data number 57

As determined from mill No. 1, multiple model output NN model with major input variables gives better prediction than multiple model with all input variables, therefore multiple output NN with major input variables model is employed to mill No. 2. Figures 5.8.a, b, c and d show the comparison of values between predicted and actual values for multiple output NN model. It converges at 144 epochs with 3.584 sse and using 221.62 seconds simulation time. Those figures demonstrate similar results for both actual and predicted values during early validation phase. The average percentage error of 1.0 %, 2.5 %, 4.5 % and 3.3 % is achieved for CO, NO_x, SO_2 and PM respectively. The maximum percentage error for CO model shows 10.5 %, NO_x model gives 12.5 %, SO_2 model produces 13.1 % and PM model gives 14.5 %.

The comparison between the single and multiple model of NN can be made. Both amount of epoch and time used by multiple model is better than that of single model. The mse of multiple model also gives higher square error than single models. The accuracy of the single models are better with less than 10 percent for maximum percentage error whereas multiple model shows around 15 percent for maximum percentage error, while the average error of single model is also less than multiple model of major input variables. Thus, the results for the single major variable model are better than the multiple output in terms of accuracy. The error predicted increases during simulation period for most models. It is due to the data used to update weights are different to the previous mill. Either single output or multiple models can be employed to predict the CO, PM, NO_x and SO_2 output for mill No. 2.

4. CONCLUSION

In this paper, the following conclusions can be made:

a) The major and minor input variables are identified and they turn out to be same for both palm oil mills

b) ANN modeling with major input variables gives better accuracy and speed

c) Single and multiple output models can be used in ANN modelling of palm oil mill emission

d) ANN predicts well and agrees with the actual data taken from the palm oil mills.

References:

[1] Berita Harian, (1999). *Amaran JAS Kilang Kelapa Sawit,* Isnin, 1 Februari, 1999.

[2] Leong, L.W. (1986). *Air Pollution and Control in Palm Oil Mill.* Technical report from PORIM, Kuala Lumpur.

[3] Weisman, J., Eckart, R.,(1985) *Modern Power Plant Engineering.* pg. 103-155. London: Prentice Hall Inc.

[4] Li, K. W. & Priddy, A.P. (1985). *Power Plant System Design,* John Wiles & Sons, pp. 138-181.

[5] Demuth, H. & Beale, M. (1998). *Neural Network Toolbox for Use with MATLAB.* The MathWorth, Inc. USA.

[6] Maren, A. J., Harston, C. T. and Pap, R. M., (1990). *Hankbook Of Neural Computing Applications,* pp. 1-250, Academic Press Inc.

[7] Grott, J. J, (2001). Power Tool, Westinghouse Process Control. *Magazine of Environmental Protection.* **12**(4), 46-47. Feb 2001. Steven Publication Corp.

[8] Lennox, B., Montague, G.A., Frith A.M., Gent, C and Beven, V., (2001). Industrial Application of Neural Networks – An Investigation. *Journal of Process Control.* **11**, 497-507.

[9] Yue, H.H, Valle, S and Qin, S.J.(1998). Comparison of Several Methods of Multivariate Soft Sensor For Emission Monitoring. *Proc of Environmental Engineering*. 23-28.

[10] Gibson, R.A. (1999). Emission Monitoring: CEMS or PEMS ?. *Magazine of Winter International*, 10-13. Dis 5, 1999.

[11] Lisboa, P.J.G, (2001a). A Review of Evidence of Health Benefict From Artificial Neural Networks in Medical Intervention. *Neural Networks*, 1-29. London: Addison-Wesley Publication.

[12]. Keller, P. E., Kouzes, R.T., Kangas, L.J. and Hashemi, S. (1994). Neural Network Based Sensor Systems for Manufacturing Application. *Advanced Information systems & Technology Conference*, 28-30, March 1994.

[13] Tuma, A., Haasis, H.D. and Rentz , O., (1996). A Comparison of Fuzzy Expert Systems, Neural Networks and Neuro-Fuzzy Approaches Controlling Energy and Material Flows. *Ecological Modelling*. **85**, 93-98.

[14] Kalogirou, S.A.(1999). Applications of Artificial Neural Networks in Energy System A review, *Energy Conversion & Management*. **40**, 1073-1087.

[15] Lisboa, P. J. G. (2001b). Industrial Use of Safety-related Artificial Neural Networks. *Contract Research Report*. Health & Safety Executive Liverpool John Moores University.

[16] Briggs, S., Stapper, B and Crow, W (1997). A Predictive Emission Monitoring System (PEMS) for A Paper Mill Power Boiler. *TAPPI Proceeding*, 599-602.

[17] Euhus L., & Blanc, L. (2000). *Process Control News (for the Pulp and Paper Industries)*, v20, 11, Nov, 2000, pp. 2

[18] Androscoggin,M., (1998). *Industrial Report for International Paper Project XL Application*, Jay, Maine 15 September 98.

[19] Guo, B, Li, D, Cheng, C, Lu, Z and Shan, Y, (2001). Simulation of Biomass Gassification With A Hybrid Neural Network Model. *Bioresource Technology*. **76**, 77-83.

[20] Radl, B. J., (1999). Neural Networks improve performance of Coal-fired Boiler. *Pegasus Newsletter*. **1.** 5-8 , Jan 1999 CADDET Energy Efficiency.

[21] Tronci, S., Barrati, R and Serrida, A., (2002). Monitoring Pollutant Emissions In A 4.8 MW Power Plant Through Neural Network. *Journal Neurocomputing*. **43**, 3-15.

[22] Rowland W. Jr (1999). Neural Network Improve Performance of Coal-fired Boilers. *Pegasus Technologies Ltd Technical Report*, 1999.

[23] Booth, K.C. & Roland, W.B. Jr, (1998). Pegasus NN Based Combustion Opt Reduces Nox Emission While Simultaneously Improving Thermal Effeciency. *Pegasus Technologies Ltd. Magazine of KFx Net Power Sollution*, 1-5, 15 May 1998.

[24] Kamal, R.,(2000). *Neural Modelling & Control of Coal-fired Boiler Plant (Neuromon).* Dept. of Trade & Industry, Project Profile, BCURA, Apr. 2000.

[25] Laux, S(1999). On-Line Coal From Measurement for Improved Boiler Performance. *Industrial Furnaces Magazine.* **1,** 10-13, Jan 1999.

[26] Hou, Z, Kefa, C and Jianbo, M., (2001). Combining Neural Network and Genetic Algorithms to Optimize Low NO_x Pulverized Coal Combustion. *Fuel,* 80, 2163-2169

[27] Sun, H, (1998). Rangking inportance of input parameters of neural networks. *Expert Systems with Application.* **15** , 405-411

[28] Howes, P & Crook, N. (1999). Using Input Parameter Influences to Support the Decisions of Feedforward Neural Network. *Neurocomputing* **24** ,191-206

[29] Ricotti, M. E. & Zio, E, (1999). Neural Network approach to sensitivity and uncertainly analysis. *Reliability Engineering and System Safety.* **64,** 59-77.

[30] Radhakrishnan, V.R. & Mohamed, A.R (2000). Neural Networks for identification and control of blast furnace hot metal quality. *Journal of Process Control.* **10**, 509-524.

[31] Crowther, W. J. & Cooper, J. E. (2001). Flight Test Flutter Prediction Using Neural Network. *Proc. Insth Mech Eng.* **215**, Part G ImechE 2001, 37-47.

[32] Syu, M.J. & Liu J. Y. (1998). Neural Network Sensitivity Analysis of the Detected Signal From An SO_2 electrode. *Sensor and Actuators* B. **50**, 1-8.

[33] Ender, L & Filho, R. M. (2000). Design of Multivariable Controller Based on Neural Network. *Computer & Chemical Engineering.* **24**, 937-943

[34] Hung, S.L., Kao, C.Y. , Lee. J.L., (2000), Active Pulse Structural Control Using ANN, *Journal Engineering Mechanic.* **8,** 839-849.

[35] Mu, H. H. ,Kahad, Y. P. and Sherlock, B.G. (2001). Design of Control Systems Using Neural Network. *CARS & FOF 2001*, Durban.

[36] Martines, F. G. & Ceolho, M.A.N. (2000). Application of Feedforward ANN to Improve Process Control of PID (proportional, integral, deravative). *Computer & Chemical Engineering.* **24**, 853-858.

[37] Shemseldin, A., Black and B. J. O'Brien, E. (2000). Multiple-Sensor Weight-In Motion of Road Vehicle Using A Neural Network Algorithm, *Computer Engineering Using Methapors from nature,* pp. 45-51, Civil Computer Press.

[38] Baines, G.H., Hayes, R.L. and Stabell, J.A. (1997). Predicting Boiler Emission With Neural Networks, *TAPPI Proceeding*, 57-61.

[39] Babavic, C. V., Canazares, R, Jensen, H. R. and Klinting, A (2001). Neural Networks As Routine For Error Updating Of Numerical Models. *Journal of Hydraulic Engineering.* **3**, 181-193.

NO	Researchers	INPUT PARAMETERS																				OUTPUT PARAMTERS							
		Power load	*Feed water*	*Air capacity*	*Air temperature*	*Fuel quality*	*Fuel capacity*	*Air fuel ratio*	*Steam capacity*	*Steam pressure*	*Steam temperature*	*Combustion temp.*	*Excess air*	*Stack pressure*	*Ambient temp.*	*Stack temp.*	*Firing/heat rate*	*Boiler efficiency*	*Slag/soot*	CO_2	O_2	CO	NO_x	PM	SO_2	CO_2	H_2	O_2	CH_4
1	Briggs *et.al*										●		●				●						●					●	
2	Bainess *et.al*						●	●				●											●						
3	Fisher	●						●			●						●	●	●			●	●		●				
4	Noble & Mayhew			●			●			●	●		●										●						
5	Booth & Roland	●		●	●	●	●	●		●			●		●				●				●		●	●			
6	Yu *et.al*		●	●	●				●				●	●	●	●							●						
7	Laux			●		●	●								●								●					●	
8	Radl					●	●	●					●				●					●	●						
9	Rowland			●			●						●		●		●						●		●	●			
10	Androscoggin				●	●	●	●			●												●	●	●				
11	Euhos & Blan			●	●		●					●	●										●						
12	Ikonen *et.al*						●					●									●		●						
13	Kamal	●						●	●				●		●					●		●	●	●	●	●		●	
14	Kersch *et.al*								●				●	●		●	●						●						
15	Grott			●				●				●					●	●					●						
16	Guo *et. al*											●				●						●				●	●		
17	Hou *et.al*					●						●			●								●						
18	Tronci *et.al*	●					●			●		●	●	●								●	●						
	Total	4	1	7	4	5	10	7	3	3	4	7	10	3	6	3	6	2	2	1	1	5	17	2	5	4	15	3	

Table 1: The researchers with their input and output in modeling the boiler emission

Table 2: Typical neural network setup for sensitivity analysis

Structure	Feedforward
Algorithm	Backpropagation
Type of training	Trainlm
Network structure	[15, h, 1/4]
Transfer function	Sigmoid and linear
Number of iterations	1000 (max epoch)
Performance function	Mean square error (mse)=1×10^{-3}
Data scaling	Standard deviation and mean
Learning rate, η	0.01
Momentum, α	0.95

Table 3: The results for sensitivity analysis and their ranking of input variables contribution to the emission for mill No. 1

Variables	CO	Rank	NO_x	Rank	SO_2	Rank	PM	Rank	Multiple	Rank
Sp	0.86684	12	**1.55302**	1	**1.07971**	5	0.90396	12	**1.20349**	1
Sc	**1.06865**	4	0.99319	8	**1.03235**	8	0.99237	11	0.83873	14
Fw	0.96484	10	0.92563	11	**1.30689**	1	**1.03918**	8	**1.14398**	2
St	**1.05742**	5	0.98979	9	**1.10477**	3	0.90155	13	**1.06845**	5
Fc	0.9989	7	**1.00994**	6	**1.0873**	4	**1.22868**	1	**1.00167**	9
Bo	0.99307	8	0.73308	14	0.71661	15	**1.09274**	4	**1.02611**	7
Wl	0.99151	9	0.71747	15	0.72699	14	**1.05451**	6	0.8017	15
At	0.82782	13	0.86459	13	0.95746	10	**1.00005**	9	0.94437	10
Ft	0.73764	15	0.97334	10	**1.07918**	6	**1.05155**	7	0.88853	13
Ea	0.96355	11	**1.00618**	7	0.93311	12	0.77389	15	0.928	11
Ff	**1.0067**	6	**1.07738**	5	**1.03274**	7	**1.06324**	3	**1.03693**	6
Sf	**1.14213**	3	**1.0813**	4	**1.01125**	9	**1.10966**	2	**1.01912**	8
Po	**1.17815**	2	**1.45547**	2	**1.18367**	2	0.99977	10	**1.08743**	3
O_2	**1.58434**	1	0.91663	12	0.95364	11	**1.05697**	5	**1.08454**	4
CO_2	0.81222	14	**1.10301**	3	0.89433	13	0.83187	14	0.92696	12

Table 4: Regression analysis between predicted and actual output for all single models using all input variables and major input variables for mill No. 1 by Neural Network

	r-value			
	CO	NO_x	SO_2	PM
Training	1	1	1	1
Validation	1	1	1	1
Testing (all)	0.9964	0.9946	0.9968	0.9527
Testing (major)	0.9983	0.9991	0.9994	0.9701

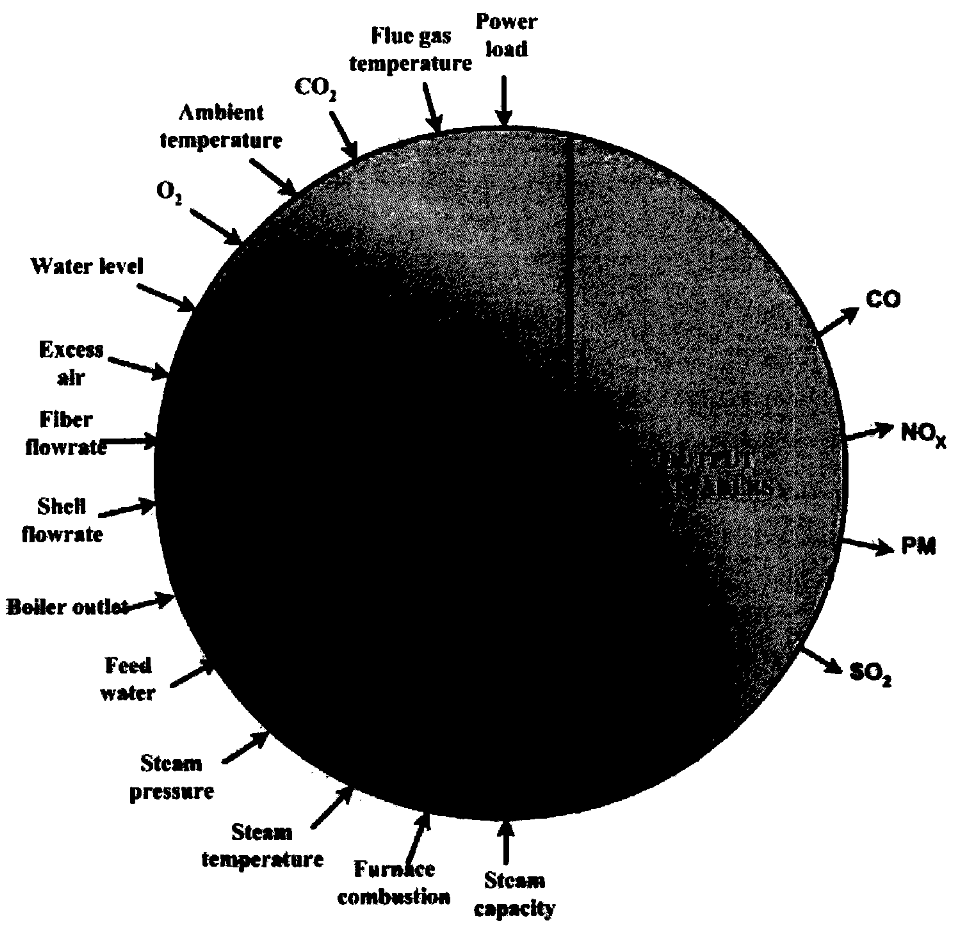


Figure 1: Input and output parameters for a typical boiler emission

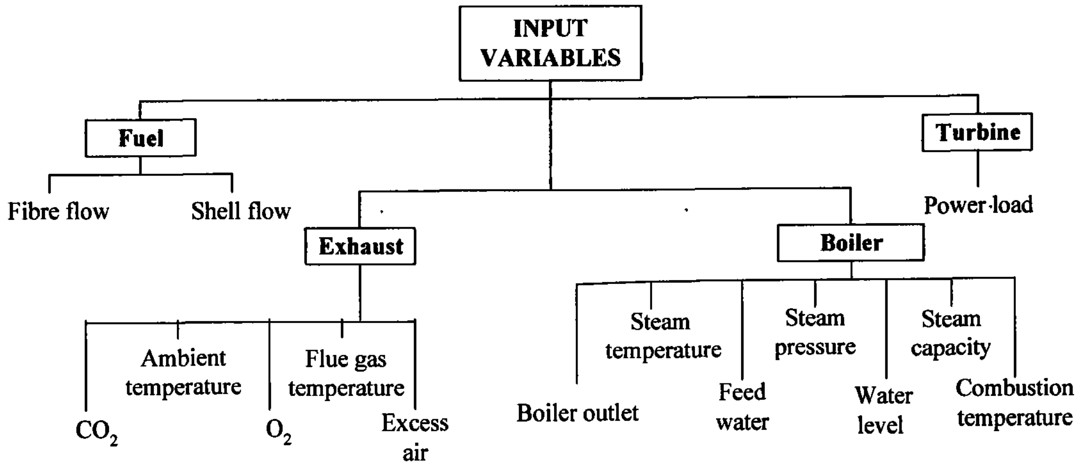


Figure 2: Selected input variables for fuel, exhaust, turbine and boiler parameters

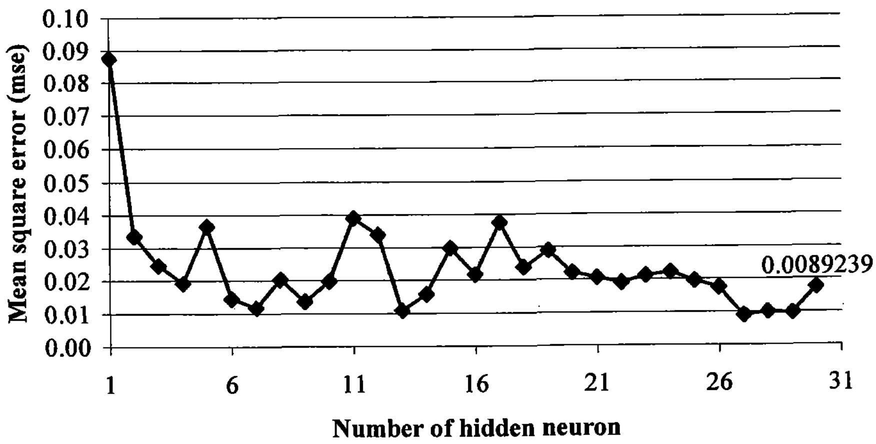


Figure 3: Finding the best structure of CO neural network for NN by sensitivity analysis

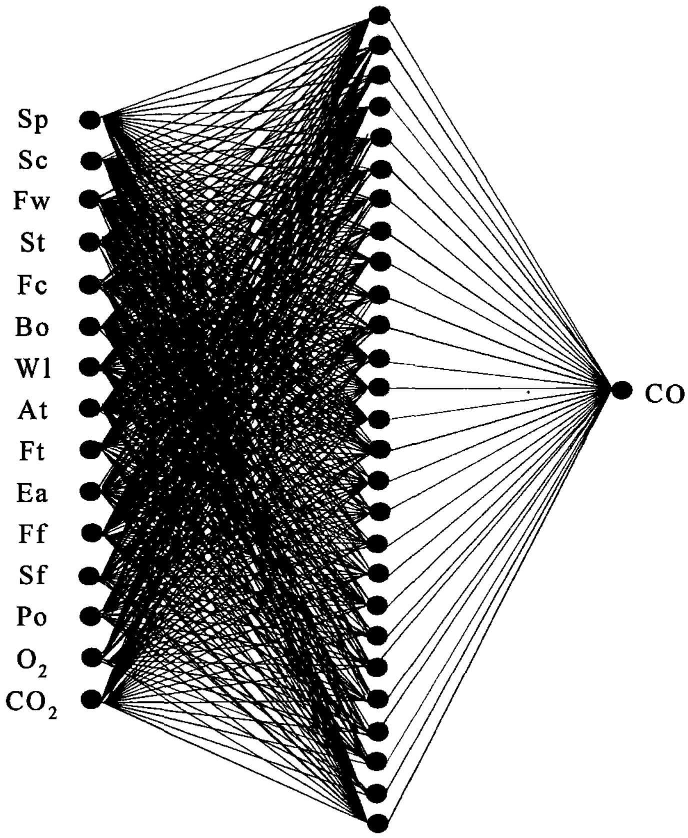
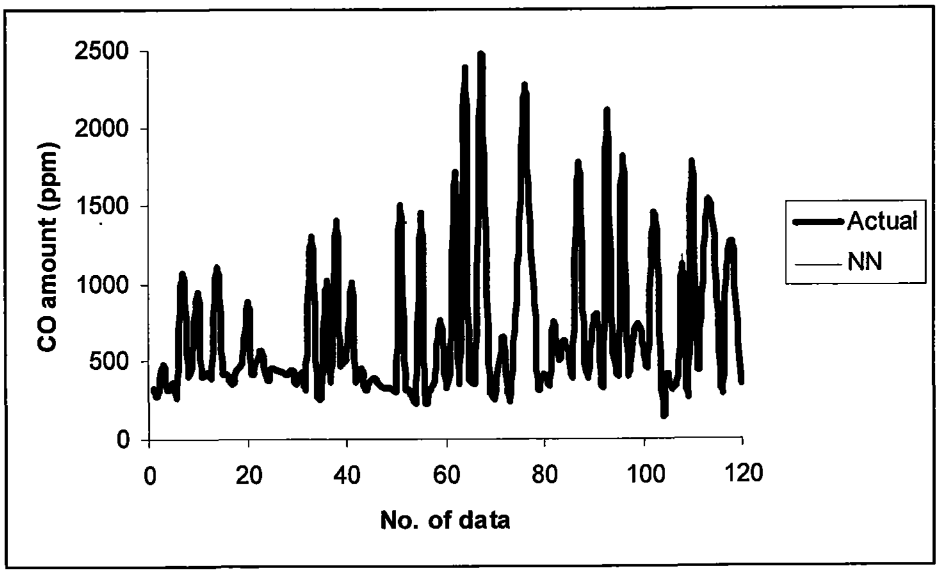
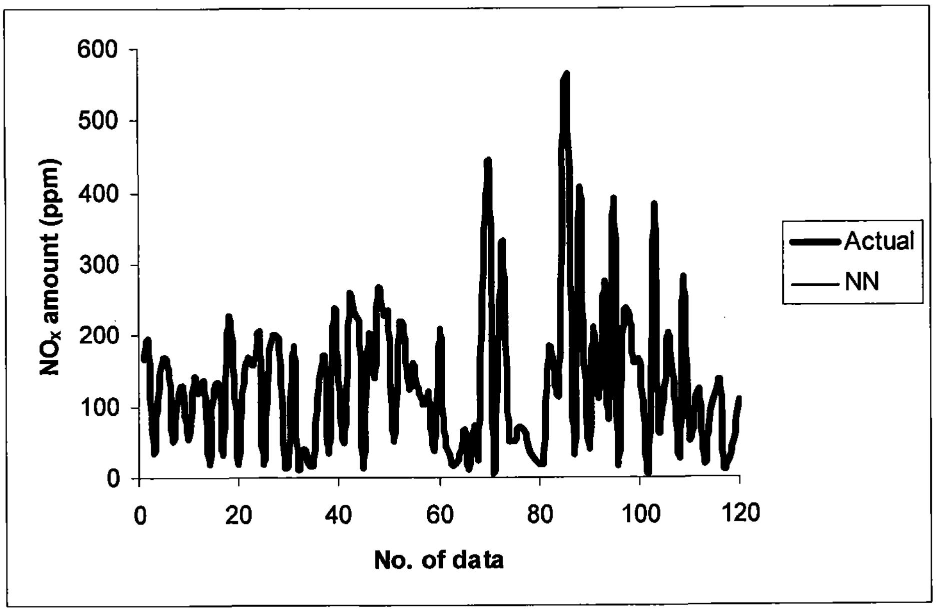


Figure 4- CO optimum neural network structure for sensitivity analysis [15,27,1]

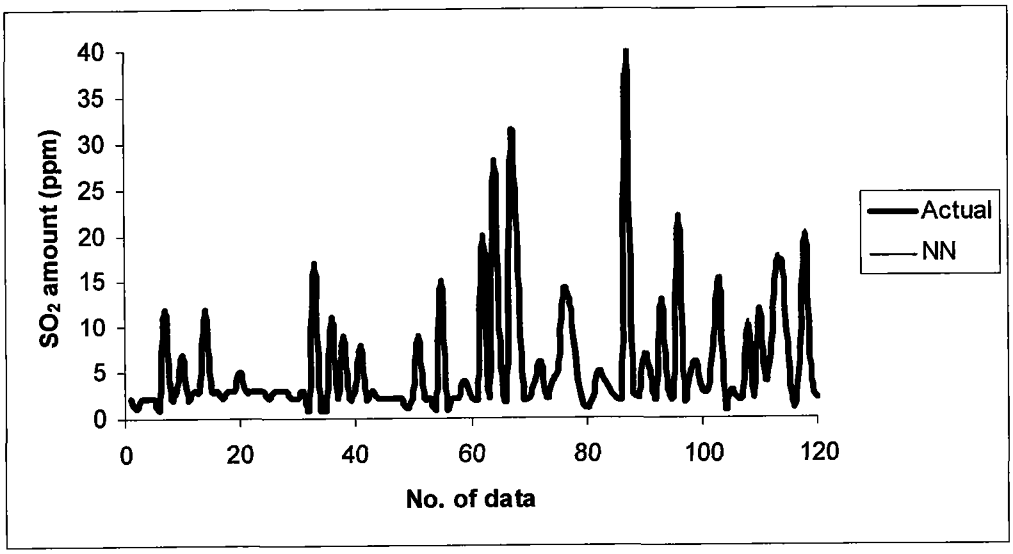

2500
2000
1500
1000
500
0
CO amount (ppm)
0
20
40
60
80
100
120
No. of data
Actual
NN

a)

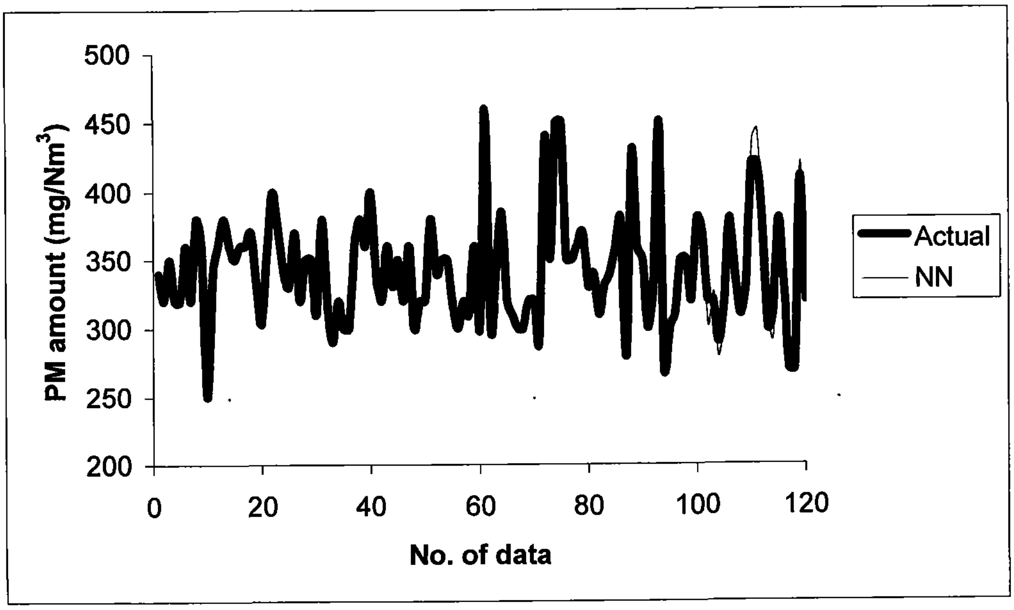

600
500
400
300
200
100
0
NOx amount (ppm)
0
20
40
60
80
100
120
No. of data
Actual
NN

b)

Figure 5: Results comparison between actual and single model of NN for a) CO, b) NO_x, c) SO_2 and d) PM for mill No. 1

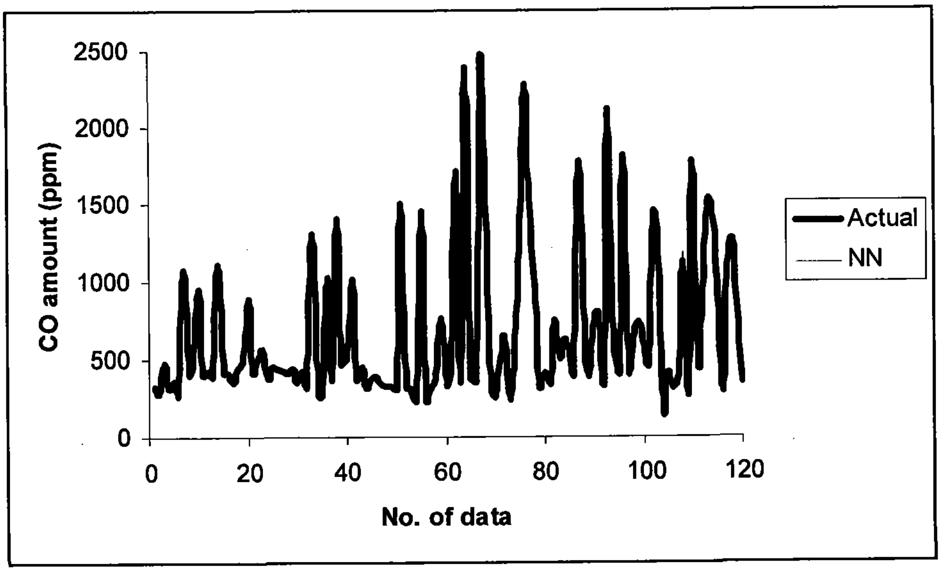


c)

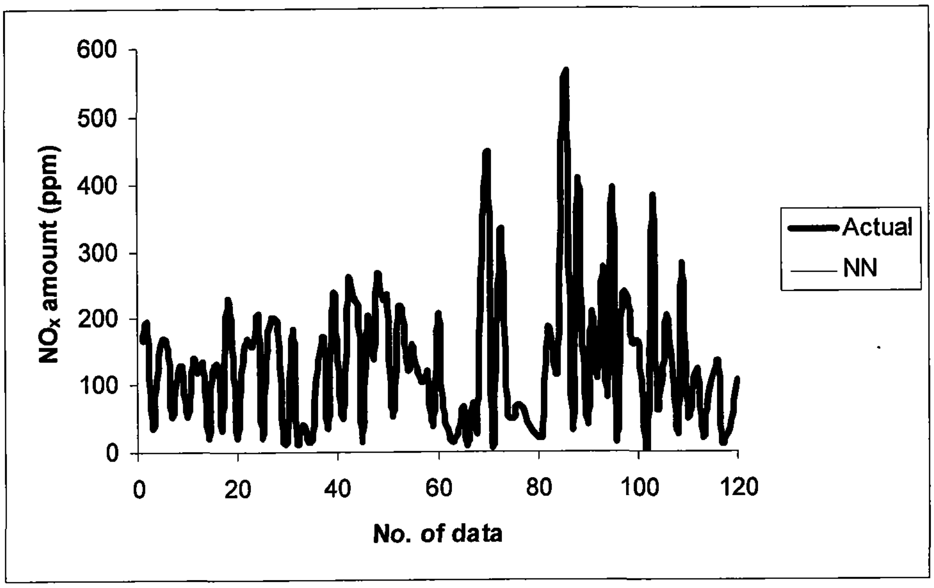


d)

Figure 5, Continued

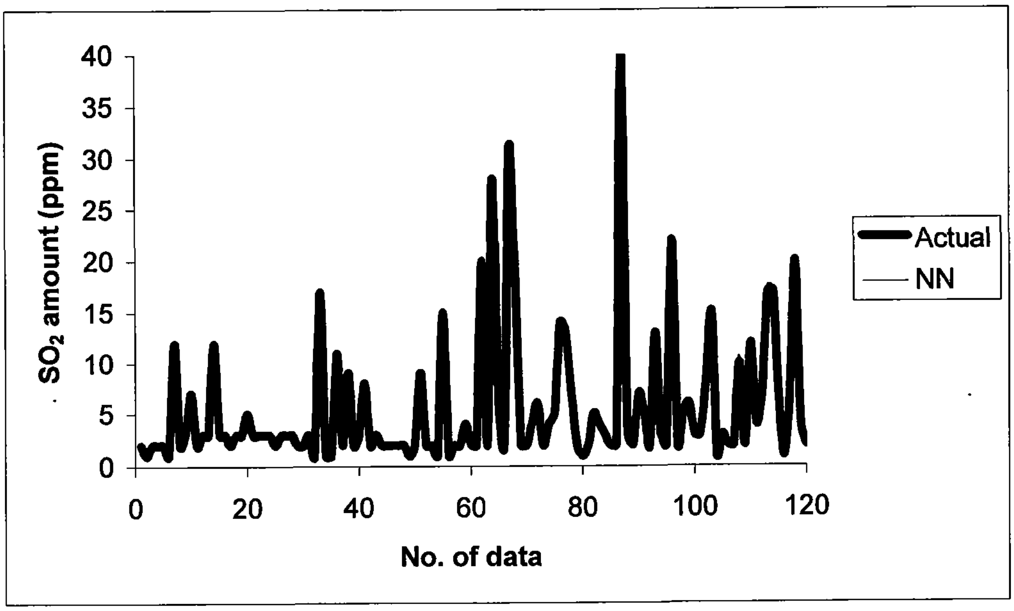


a)

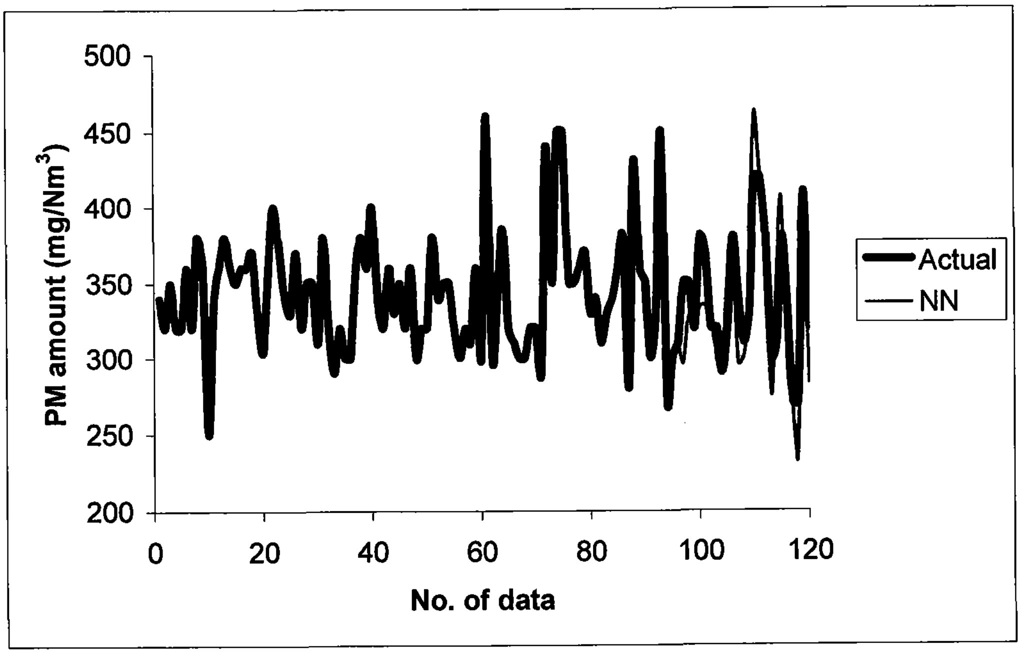


b)

Figure 6: Results comparison between actual and multiple model of NN for a) CO, b) NO_x, c) SO_2 and d) PM for mill No. 1

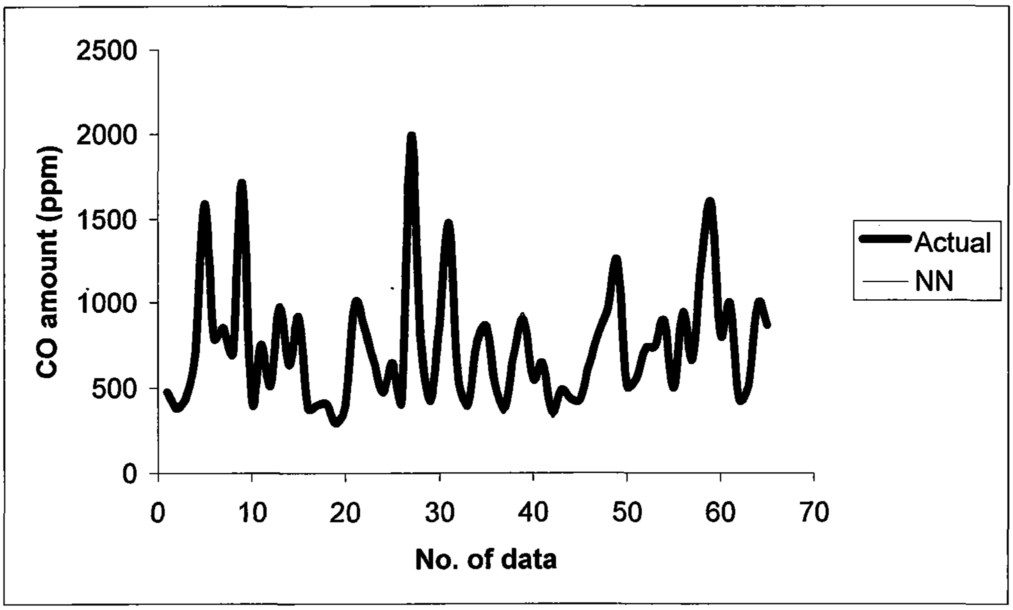


c)

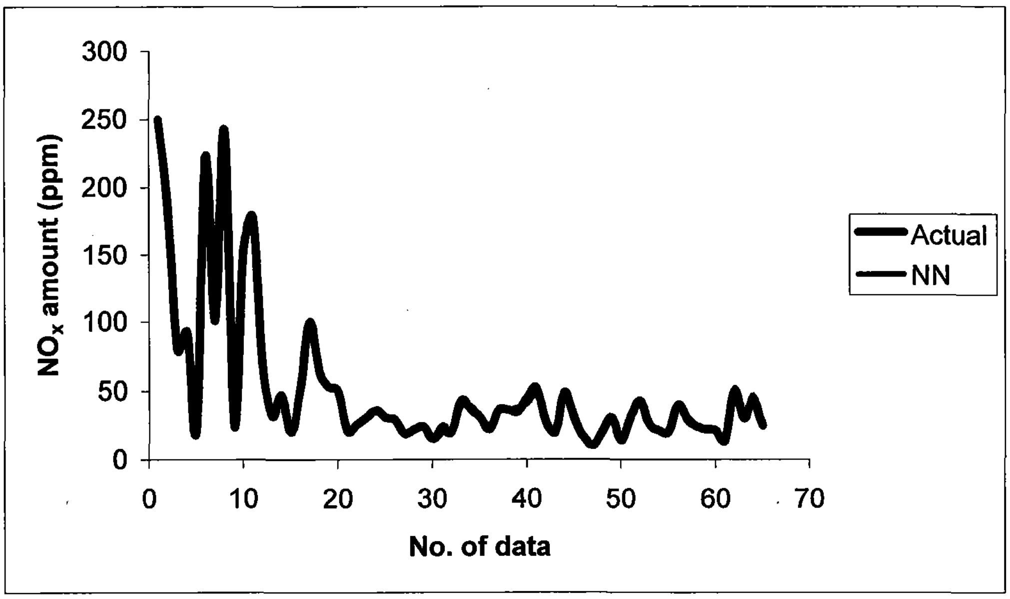


d)

Figure 6, Continued

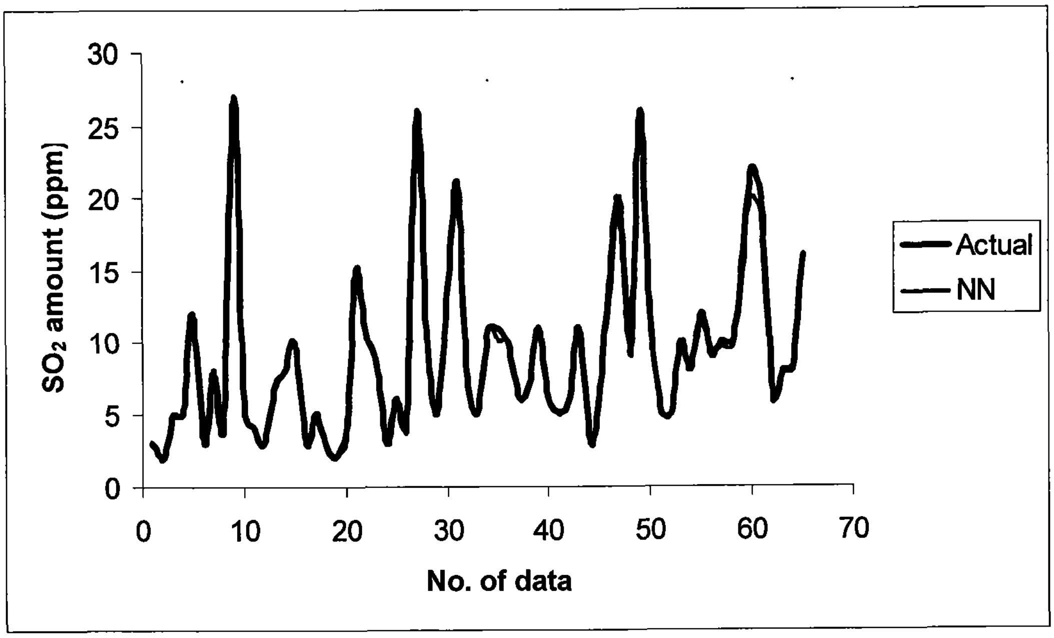

2500
2000
1500
1000
500
0
0
10
20
30
40
50
60
70
CO amount (ppm)
No. of data
Actual
NN

a)

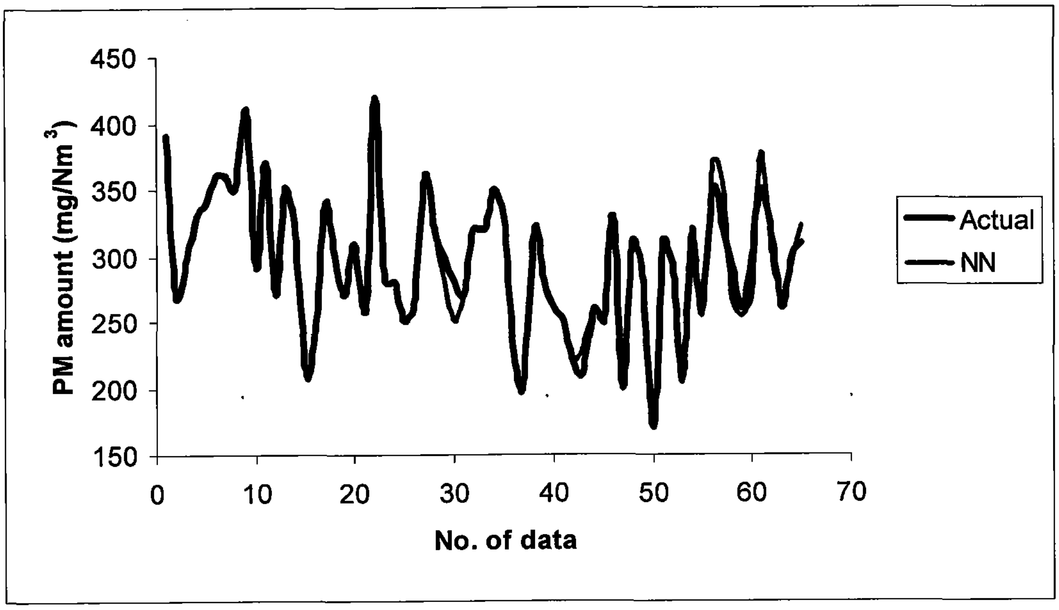

300
250
200
150
100
50
0
0
10
20
30
40
50
60
70
NO_x amount (ppm)
No. of data
Actual
NN

b)

Figure 5.7: Results comparison between actual and single model of NN for a) CO, b) NO_x, c) SO_2 and d) PM for mill No. 2

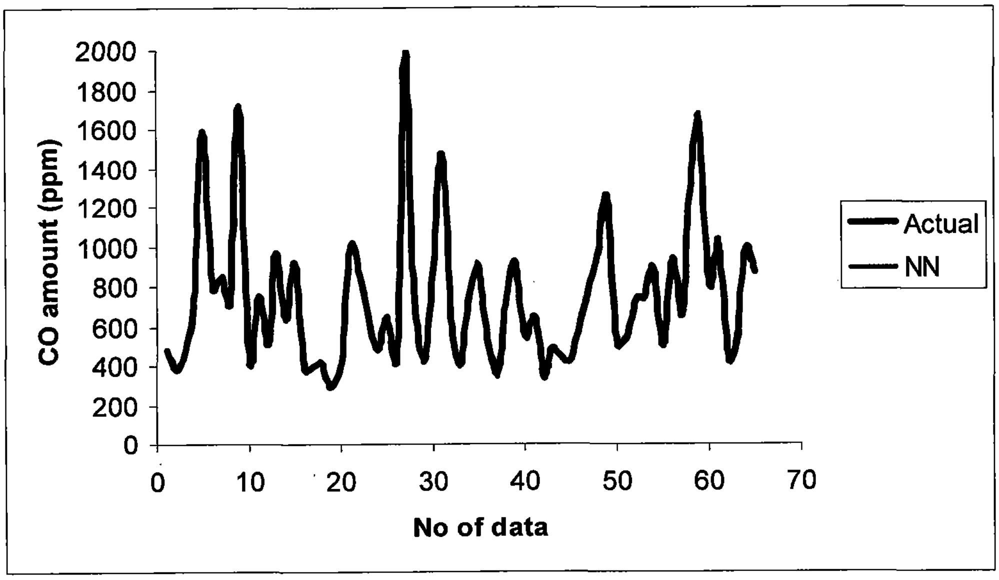


c)

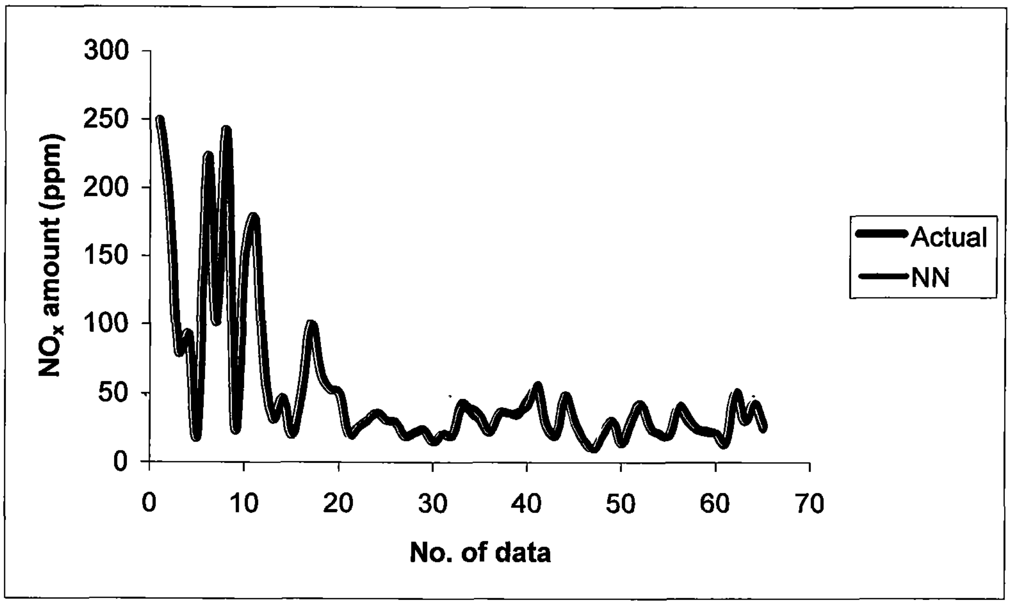


d)

Figure 5.7, Continued

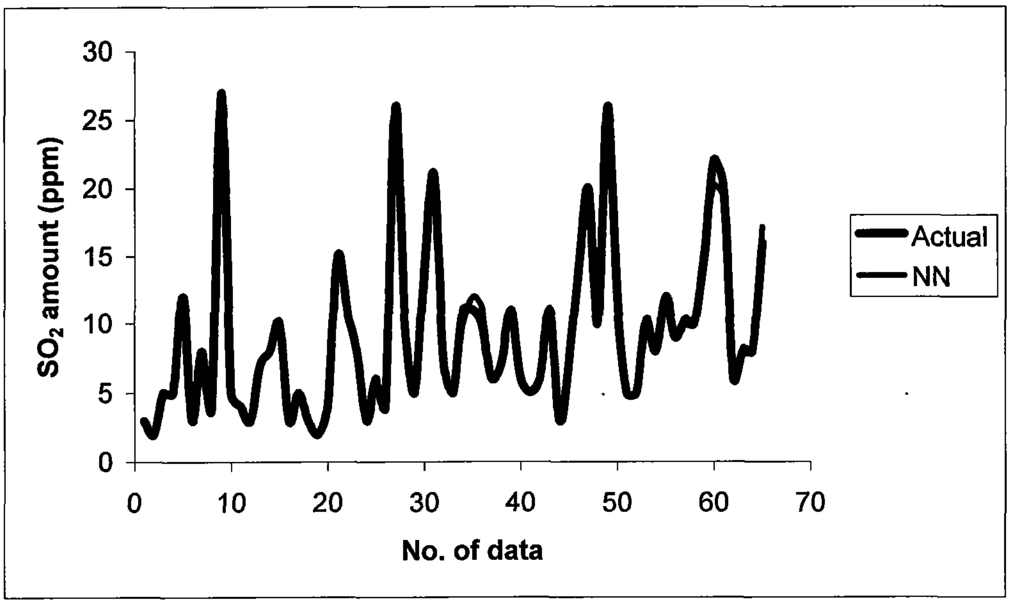


a)

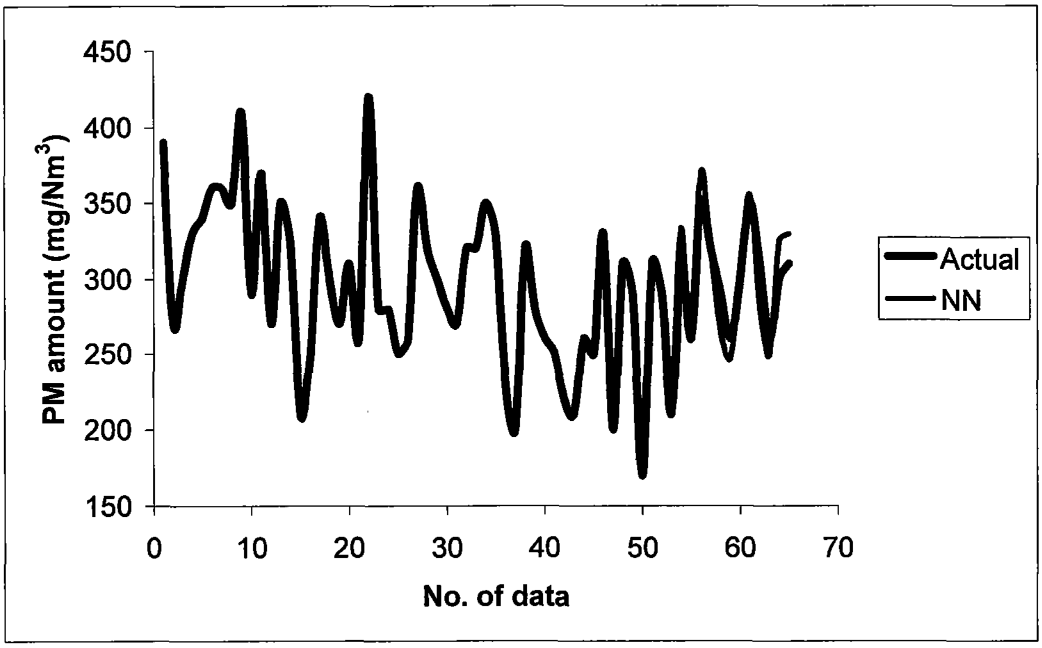


b)

Figure 5.8: Results comparison between actual and multiple model of NN for a) CO, b) NO_x, c)SO_2 and d) PM for mill No. 2



c)



d)

Figure 5.8, Continued

Appendix B

A.L. Ahmad, I.A. Azid, A.R. Yusof and K.N. Seetharamu, “Emission control in palm oil mills using artificial neural network and genetic algorithm”, Computers and Chemical Engineering 28 (2004) 2709–2715

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

ELSEVIER

Computers and Chemical Engineering 28 (2004) 2709–2715

Computers & Chemical Engineering

www.elsevier.com/locate/compchemeng

Emission control in palm oil mills using artificial neural network and genetic algorithm

A.L. Ahmad[a,*], I.A. Azid[b], A.R. Yusof[b], K.N. Seetharamu[b]

[a] *School of Chemical Engineering, Engineering Campus, Universiti Sains Malaysia, Seri Ampangan, 14300 Nibong Tebal, Pulau Pinang, Malaysia*

[b] *School of Mechanical Engineering, Engineering Campus, Universiti Sains Malaysia, Seri Ampangan, 14300 Nibong Tebal, Pulau Pinang, Malaysia*

Received 6 April 2003; received in revised form 6 April 2004; accepted 20 July 2004
Available online 25 September 2004

Abstract

The present study utilized a combination of artificial neural network (ANN) and genetic algorithms (GA) to optimize the release of emission from the palm oil mill. A model based on ANN is developed from the actual data taken from the palm oil mill. The predicted data agree well with the actual data taken. GA is then employed to find the optimal operating conditions so that the overlimit release of emission is reduced to the allowable limit.

Keywords: Emission control; Neural networks; Optimization; Genetic algorithms

1. Introduction

The emission released from the palm oil mill in Malaysia has quite significant environmental impact. The use of biomass material from the palm oil byproduct, i.e. fiber and shell as fuel is identified as the main cause of the emission released. Economically, the usage of palm waste material as fuel is seen as productive as this material will not be wasted. However, the combustion process in the furnace of the boiler releases the emission such as particulate matters (PM), carbon monoxide (CO), nitrogen oxide (NO_x) and sulphur dioxide (SO_2). The monitoring and control of these pollutants from the palm oil mill is a great concern to the community. Currently, the limit is set by the Department of Environment (DOE), Malaysia, as shown in **Table 1**.

It is well known that the emission released depends on many complex factors which are either directly or indirectly related to each other and are influenced by many variables. These variables can appear from the combustion process, boiler process, etc. The unsteady fuel flow affects combustion and the emission released (**de Nevers, 1999**). Despite the fuel transfer system being driven constantly by motor to supply the fuel, the problem exists as fibre and shell inlet capacity is not stable in the conveyor. It will also cause fluctuating calorific value and moisture, or fuel quality interrupting the combustion rate and its temperature. Not only fuel capacity causes the emission, but also the fuel types and its quality play a role in the emission production and cannot easily be controlled. These processes are related to each other and are highly non-linear. When a particular operation changes, the new optimum parameters cannot be determined immediately.

The current work deals with the monitoring and control of the emission being released from the palm oil mill. An ANN is developed to model the pollutants released under various operating conditions. When the pollutants released are beyond the allowable limits, GA is then applied to find the optimal input parameters that release the allowable pollutant level.

* Corresponding author. Tel.: +60 5593 7788; fax: +60 4594 1013.
E-mail address: chlatif@eng.usm.my (A.L. Ahmad).

0098-1354/$ – see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compchemeng.2004.07.034

Table 1
Standardization of exhaust/chimney separation (**de Nevers, 1999** and Malaysian Ministry of Industry's 2nd Mandate, 1993, DOE, EPA-USA)

Rules	Pollutants			
	PM	NO_x	SO_2	CO
DOE Malaysia (mg/N m^3)	400	3500	1700	
Tenaga Nasional Berhad (TNB) (mg/N m^3)	40	980–1480	615–1640	
Typical North America (mg/N m^3)	50	750	650	
World Bank (mg/N m^3)	400	750	650	
Typical ASEAN (mg/N m^3)	100	850–1700	940	
Typical Europe (mg/N m^3)	100	1430	720	
Typical ASIA (mg/N m^3)	100	1430	720	
Ministry of Industry (mg/N m^3 or ppm)	400/351	470/250	1300/500	1000/870

2. Modeling and optimization method by using genetic algorithm and artificial neural network (GAANN)

Due to its ability to model complex and nonlinear problems, the widely accepted method, artificial neural network (ANN) is chosen to model the complex behavior between input and output in the process of releasing emission from the palm oil mill. It has been used extensively in many fields such as forecasting, pattern recognition, robotics, etc. It is originally based on the human thoughts of receiving and transferring the information in making decision. A simple model of ANN consists of an input layer, a hidden layer and an output layer. With sets of input–output patterns stored in input and output layers, the hidden layer interconnects different strength of information from the input to the output layers, through so-called weights. The weights are adjusted in the learning process in which all the patterns of input–output are presented in the learning phase repeatedly. There are many learning algorithms available and the most popular and successful learning algorithm used to train multilayer network is the back propagation scheme. Any output point can be obtained after this learning phase, and good results can be achieved. In this study, a feed forward back propagation with one hidden layer is employed for ANN modeling. The number of neurons for input and hidden layer can be determined depending on the optimization process using sensitivity analysis by neural network (SANN) and coefficient correlation analysis (CC). Commercial software of MATLAB Version 5.3 and Neural Network Toolbox from Mathworks Inc. (**MATLAB User's Guides, 1998**) are used to generate neural network modeling. The stopping criteria used in the current study was set at 1000 maximum epoch number, and the characteristics of the training set was train multiplayer (Trainlm). Whereas the validation set and the testing set was set once the difference between sum square error of the actual and predicted values is $\leq 1 \times 10^{-3}$.

Genetic algorithms (GA) is an optimization algorithm based on the principle of survival of the fittest during the evolution. In nature, individuals must adapt to the frequent changing environment in order to survive. **Holland's proposed GA (1975)** is based on this process where GA mimics the evolutionary processes of the nature. GA is one of the strategic randomized search techniques, which are well known for its robustness in finding the optimal or near-optimal solution since it does not depend on gradient information in its walk of life to find the best solution. The approach uses and manipulates a population of potential solutions to find the optimal solution. A generation is complete after each individual in the population has performed these genetic operators. The individuals in the population will be better adapted to the environment, as they have to in order to survive in the subsequent generations. In GA, environment is defined as the objective function or fitness function for a specific optimization problem. By favoring the optimal and near-optimal individuals in each generation, the average fitness level in the population is expected to improve over time. The interested readers can look into more comprehensive texts such as **Goldberg (1989)** for a complete explanation of GA.

Several researchers employed genetic algorithm and artificial neural network (GAANN) in their applications to get optimal operating conditions or parameter set up for the input variables such as in a fuzzy controller for fruit storage (**Morimoto, Suzuki, & Hashimoto, 1997**), in control system (**Sette, Boullart, & Langauhore, 1998**), in stereo lithography process (**Cho, Par, Chi, & Lew, 2000**); in fluidized controlling cracking catalytic (**Zhao, Chen, & Hu, 2000**), in chiller system (**Chow, Zhang, Lin, & Song, 2002**) and in coal combustion (**Hou, Kefa, & Jianbo, 2001**).

In this paper, ANN is combined with GA to get the optimum value. To search for the optimum, GA requires the predicted value from ANN. ANN then employs GA to generate parameters as a new input to predict the new output value. Consequently, both GA and ANN programs should be linked-up and exchanged data with each other. In current study, the procedure adopted is as follows. First GA writes the selected input parameters that are written in the text file. The text file is then read by the ANN and received as a new input parameter. Then, ANN will predict the output value. The output generated from this prediction is compared with the pollutant limit. If the value exceeds the limit, GA generates new input parameters from the GA operator, i.e. mutation and crossover. These steps are repeated until the optimal input values of fuel are found. This is an iterative process at the end of which the GA arrives at the optimum set of fuel parameters which produce emission within acceptable limit.

Table 2
Regression analysis between predicted and actual output by neural network

	r-value			
	CO	NO_x	SO_2	PM
Training	1	1	1	1
Validation	1	1	1	1
Testing	0.9983	0.9991	0.9994	0.9701

A program written in C^{++} is modified to suit the problem for GA in emission controlling in which the objective function used is the concentration of the emission.

Before GAANN is applied, GA parameters such as population size, mutation rate etc. need to be determined. As the parameter settings are right, so do the effects of crossover rates, population size and mutation rates on the GAANN performance. The first study is carried out for different crossover rate (i.e. 0%, 5%, 10% and 100%) while other parameters are kept constant. Similar studies were also carried out to find the mutation rate and population size. From this parametric study, it is found that the best rate for crossover and mutation is 88.67% and 4.33%, respectively. It is also found that the best number of size for population is 100.

3. Experimental procedure and data analysis

In this study, gas analyzer is used for collecting CO, NO_x and SO_2, and isokinetic source sampling equipment of method 5 is used for collecting PM. There are a variety amount of data collected by other researchers to model boiler emission. **Yue, Valle, and Qin (1998)** gathered about 350 sets of data and **Fisher–Rosemount (1997)** collected the data in each minute for 2 and 4 weeks for boiler emission modeling. However, for current study, by considering the equipment provided and the operation time for palm oil mill, the data collected were 120 sets for the first mill and 65 sets for the second mill in 15 min interval time for several days. The data taken are then divided into two major parts: simulation (training and validation) and testing. The division of the data follows the rule set by Environmental Protective Agency of United States of America (EPAUSA), which requires 20% of the data must be taken for testing and the rest are for the training and validation. The training of ANN is needed to update the weights according to new data sets. The testing segment is then used to evaluate the methods using unseen data (**Demuth & Beale, 2000**; **Martines & Ceolho, 2000**; **Freeman & Skapura, 1991**. **Table 2** shows the regression analysis between predicted data and actual data. During training and validation, the regression 'r' value of 1 is obtained which means that the predicted and actual points coincide. During testing, the predicted and actual values almost coincide in which the regression 'r' value of almost 1 is obtained. It shows that the prediction by ANN is almost accurate, and the result is reliable.

From the coefficient correlation analysis (**Baines, Hayes, & Stabell, 1997**; **Noble & Mayhew, 1997**; **Yue, Valle, & Qin, 1998**) and sensitivity analysis by neural network (SANN) (**Ricotti & Zio, 1999**; **Sun, 1999**), major and minor input variables to a particular emission released for ANN simulation is determined. By using these analyses, out of 15 input parameters selected based on the contributing factors to palm oil mill emission, major input variables for CO are four (steam capacity, steam temperature, fibre flow rate and shell flowrate). The major input variables for NO_x are also four with the same major input variables as CO except steam temperature is replaced by furnace combustion. Finally for particulate matters (PM), the major input variables are feed water, furnace combustion, boiler outlet, water level, fibre flow rate and shell flow rate. In general, major input variables have great impact on a particular output variable whereas minor input variables give small impact on a particular output variable.

4. Result and discussion

The data above the permissible limits of the pollutants released could be determined by referring to the standardization of exhaust separation in **Table 1**. Each pollutant has its own allowable limit according to DOE of Malaysia. However, for CO and NO_x, the limit of 870 ppm and 250 ppm as laid down by Ministry of Industry (MI) of USA, is taken for this study.

Fig. 1 shows the graph of CO produced and the fuel flow rate during 120 data collection. The broken line indicates the

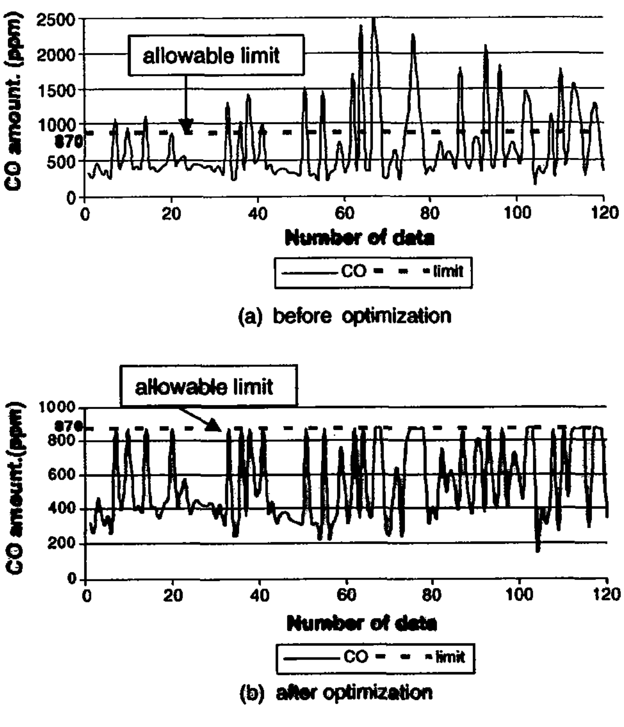


Fig. 1. Optimized conditions of palm oil mill showing the control of CO emission.

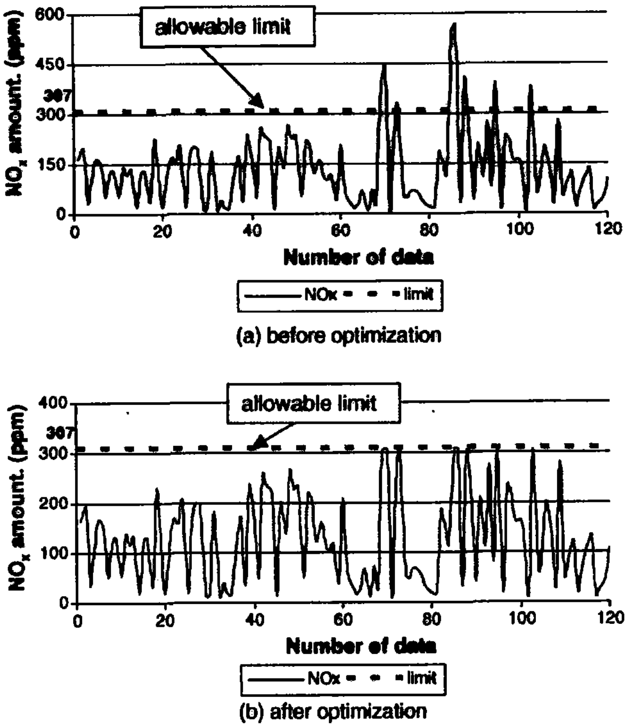


Fig. 2. Optimized conditions of palm oil mill showing the control of NO_x emission.

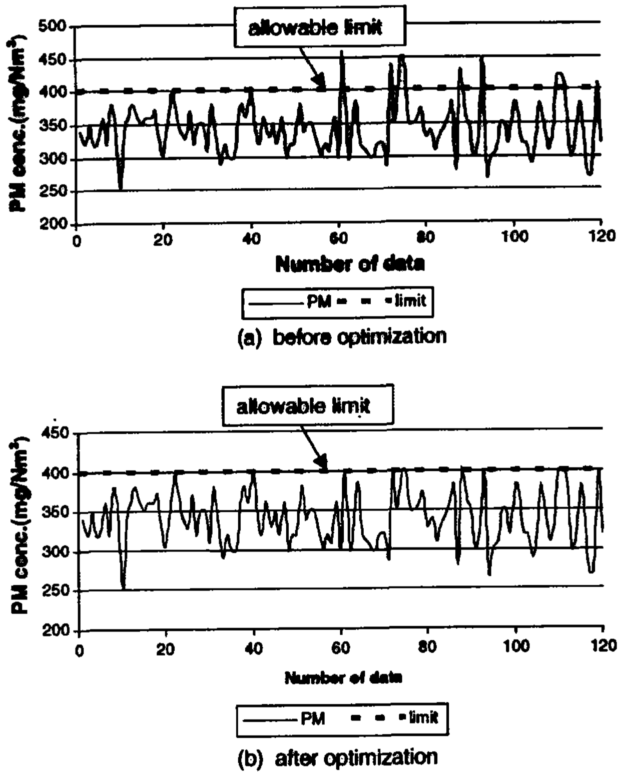


Fig. 3. Optimized conditions of palm oil mill showing the control of PM emission.

allowable limit of 870 ppm. Many points of CO released are over the permitted value as shown in **Fig. 1**(a). As can be seen from the graph, these 30 over limit points need to be within the allowable limit, and the processes of releasing CO needs to be optimized. **Figs. 2(a) and 3(a)** show the models for NO_x and PM output, respectively, with the broken lines represent the allowable limit. Those figures show 12 points for NO_x and 11 points for PM model are above the allowable limit. There is no overlimit for SO_2 emission. Hence, only three pollutants are used for the optimization process with GAANN for that particular mill.

Single major input variables of ANN models of those three pollutant models are employed and linked with GA to find the optimal configuration that emits allowable emission. The output predicted by ANN is based on the new input values given by GA. **Table 3** shows the optimization results of CO emission from the over limit input variables. Four major input variables (steam capacity, steam temperature, fibre flow rate and shell flow rate), which contribute to the CO emission, are considered for optimization. All 30 set of data which are beyond the allowable limit are used in the optimization so that the generated CO is lower than 870 ppm. The values in bracket represent the original values before the optimization. The percentage of reduction between actual and GAANN results are illustrated in which the maximum percentage reduction is 64.36% at data point 67. With fibre flowrate of 1474.11 kg/h and shell flowrate of 532.13 kg/h, steam capacity of 15.10 tonne/h and steam temperature of 286.01 °C, the CO released is 868.90 ppm, which is 19.10% reduction from the original data of 1074 ppm. Other 29 points are optimized accordingly with constant power load and the results are presented in **Table 3**. **Fig. 1**(b) shows the optimized conditions of the emission released. The optimized points are clearly below the allowable limit.

Similar procedure is carried out to NO_x and PM. **Table 4** shows the optimization results of NO_x emission from the over limit input variables. Four major input variables (steam pressure, furnace combustion, fibre flowrate and shell flowrate), which contribute to the NO_x emission, are considered for optimization. The values in bracket represent the original values before the optimization. The percentage of reduction between actual and GAANN results are illustrated in which the maximum percentage reduction is 55.55% at data point 86. With the same power load, the NO_x released is now within the allowable limit of 250 ppm. With fibre flowrate of 1905.12 kg/h and shell flowrate of 660.33 kg/h, steam pressure of 19.30 kg/cm^2 and furnace combustion of 0.30 mm/h, the NO_x released is 249.56 ppm, which is 2.89% reduction from the original data of 257 ppm. All other points are optimized accordingly with constant power load and the results are presented in **Table 4**. **Fig. 2**(b) shows the optimized condition of NO_x. **Table 5** shows the optimization results of PM pollutant from the over limit input variables of 11 data. Six major input variables (feed water, furnace combustion, boiler outlet, water level, fibre flowrate and steam temperature), which contribute to the PM released, are considered

Table 3
GANN optimization of CO emission

Data number	Input parameters generated by GANN				CO output		
	Steam capacity (tonne/h)	Steam temperature (°C)	Fibre flowrate (kg/h)	Shell flowrate (kg/h)	Actual (ppm)	GANN (ppm)	Reduction (%)
7	15.20 (15.1)	286.01 (304)	1474.11 (1584)	532.13 (720)	1074	868.9	19.097
10	14.22 (14.5)	326.23 (331)	1874.95 (19839)	575.89 (519)	942	869.11	7.738
14	15.01 (14.2)	311.19 (327)	1450.06 (1800)	446.88 (576)	1113	869.8	21.851
20	14.53 (15.2)	343.05 (354)	1305.22 (1008)	453.09 (1008)	881	868.98	1.364
33	15.52 (14.6)	306.03 (346)	1179.42 (720)	420.36 (540)	1301	869.15	33.194
36	14.51 (14.1)	339.42 (334)	1654.44 (1980)	737.23 (1980)	1019	868.88	14.732
38	14.02 (14.3)	320.21 (369)	1445.09 (1548)	733.43 (1548)	1408	868.9	38.288
41	15.23 (14.5)	323.51 (343)	1544.03 (1440)	544.20 (972)	1007	869.45	13.659
51	15.22 (15.2)	383.44 (383)	1415.23 (1404)	576.33 (576)	1503	869.05	42.179
55	14.52 (14)	326.33 (366)	1243.02 (1830)	456.22 (395)	1459	868.98	40.440
62	14.50 (14.2)	351.02 (363)	1650.32 (1548)	621.11 (720)	1711	869.16	49.202
64	15.53 (13.4)	315.45 (376)	1131.23 (1404)	405.20 (540)	2387	869.5	63.574
67	15.82 (15.7)	257.33 (348)	1114.02 (1260)	360.80 (360)	2441	869.31	64.387
68	15.52 (14.1)	333.13 (383)	1606.42 (900)	724.28 (864)	1659	869.89	47.565
75	15.53 (15.1)	342.11 (382)	1214.34 (1098)	604.06 (769)	1287	869.02	32.477
76	15.57 (14.5)	266.61 (347)	1119.52 (576)	735.19 (936)	2273	868.99	61.769
77	15.26 (14.5)	290.88 (337)	1205.46 (1512)	524.33 (576)	1536	869.56	43.388
78	14.55 (14.3)	341.54 (343)	1721.11 (1620)	712.36 (828)	944	867.98	8.053
87	14.48 (14.4)	248.42 (348)	1422.55 (1476)	412.89 (216)	1779	869.72	51.112
93	15.40 (14.2)	279.02 (328)	1587.33 (1756)	415.36 (952)	2116	868.9	58.937
96	15.03 (14)	299.43 (364)	1555.23 (1440)	623.31 (720)	1813	869.25	52.055
102	15.52 (14.5)	303.63 (325)	1411.35 (1512)	722.36 (900)	1437	869.97	39.459
103	15.17 (14.1)	303.15 (352)	1422.28 (1757)	425.72 (439)	1269	869.17	31.507
108	15.11 (14.6)	333.35 (341)	1601.23 (1800)	440.56 (540)	1125	869.45	22.716
110	15.52 (14.2)	247.63 (270)	1022.55 (1548)	303.57 (828)	1773	868.76	51.001
113	15.03 (14.1)	304.53 (334)	1389.06 (1692)	499.52 (468)	1522	869.06	42.900
114	14.50 (14.1)	311.26 (342)	1523.43 (1620)	520.36 (720)	1382	867.9	37.200
115	15.53 (15.2)	313.66 (333)	1315.20 (1464)	450.66 (366)	978	869.24	11.121
117	14.87 (15.1)	310.56 (325)	1666.54 (1281)	452.87 (732)	1095	869.94	20.553
118	15.36 (14.9)	311.45 (353)	1355.77 (1584)	543.62 (720)	1273	869.51	31.696

Table 4
GANN optimization of NO_x emission

Data number	Input parameters generated by GANN				NO_x output		
	Steam pressure (kg/cm^2)	Furnace combustion (mm/h)	Fibre flowrate (kg/h)	Shell flowrate (kg/h)	Actual (ppm)	GANN (ppm)	Reduction (%)
42	19.30 (20.1)	0.30 (0.2)	1905.12 (2229)	660.33 (882)	257	249.56	2.89
48	21.15 (20.5)	0.55 (0.25)	1651.11 (1898)	824.07 (1056)	264	248.9	5.72
69	20.06 (20.1)	0.25 (0.2)	1131.53 (1814)	356.02 (936)	324	249.11	23.11
70	18.44 (20.2)	0.22 (0.25)	1720.33 (1556))	690.43 (930)	435	248.87	42.79
73	19.43 (20.1)	0.25 (0.2)	1456.87 (1670)	525.18 (828)	328	249.56	23.91
85	17.54 (20.4)	0.33 (0.4)	1536.42 (1784)	530.42 (1018)	547	249.69	54.35
86	17.26 (20.3)	0.41 (0.3)	1444.23 (1327)	526.34 (637)	562	249.39	55.62
88	19.56 (20.3)	0.52 (0.3)	1660.58 (1804)	333.56 (801)	405	249.89	38.30
93	17.12 (17.5)	0.50 (0.44)	1541.45 (1757)	509.36 (952)	276	249.55	9.58
95	18.53 (20.3)	0.42 (0.3)	1404.56 (2578)	527.30 (856)	392	249.09	36.46
103	17.56 (18)	0.52 (0.38)	1422.76 (1757)	425.37 (439)	382	248.49	34.95
109	20.06 (20.2)	0.34 (0.25)	1211.43 (1800)	355.43 (540)	280	249.68	10.83

for optimization. The values in bracket represent the original values before the optimization. The percentage of reduction between actual and GAANN results are illustrated in which the maximum percentage reduction is 13.48% at data point 61. For example, at that data point with the same power load, the PM released is now within the allowable limit of 400 mg/N m^3. With feed water of 13.9 tonne/h, furnace combustion of 0.35 mm/h, boiler outlet of 0.65 mm/h, water level of 51.36%, fibre flowrate of 1668.41 kg/h and shell flowrate of 558.23 kg/h, the PM released is 397.97 mg/N m^3, which is 13.48% reduction from the original data of 460 mg/N m^3. Other points are optimized accordingly with constant power load and the results are presented in **Table 5**. The conditions of the release of PM after optimization are shown in **Fig. 3**(b).

Table 5
GANN optimization of PM pollutant

Data no.	Input parameters generated by GANN						PM output		
	Feed water (tonne/h)	Furnace combustion (mm/h)	Boiler outlet (mm/h)	Water level (%)	Fibre flowrate (kg/h)	Shell flowrate (kg/h)	Actual (ppm)	GANN (ppm)	Reduction (%)
22	15.42 (15.2)	0.34 (0.25)	0.75 (0.8)	51.03 (49.1)	1843.22 (1965)	854.10 (960)	400	395.89	1.03
40	15.21 (14.9)	0.35 (0.25)	0.75 (0.6)	49.12 (52.8)	1806.21 (1807)	733.26 (1062)	400	396.54	0.86
61	13.19 (9.9)	0.35 (0.44)	0.70 (0.65)	51.36 (49)	1668.41 (1720)	558.23 (732)	460	397.97	13.48
72	14.51 (14)	0.40 (0.44)	0.81 (0.9)	49.00 (49.1)	1732.09 (1464)	740.07 (659)	440	397.91	9.57
74	11.52 (10.5)	0.44 (0.42)	0.72 (0.65)	48.23 (48)	1405.36 (1537)	405.66 (695)	450	396.56	11.88
75	10.50 (9.8)	0.35 (0.4)	0.71 (0.65)	50.22 (51.2)	1101.33 (1098)	359.23 (769)	450	398.97	11.34
88	14.58 (13.8)	0.23 (0.3)	0.51 (0.9)	51.36 (49.1)	1660.87 (1804)	333.59 (901)	430	399.45	7.10
93	15.5 (13)	0.51 (0.44)	0.62 (0.5)	49.51 (50)	1541.01 (1757)	509.12 (952)	450	398.8	11.38
110	11.28 (10.4)	0.36 (0.4)	0.62 (0.7)	51.36 (49.1)	1441.52 (1548)	435.89 (828)	420	398.39	5.15
111	12.51 (13.3)	0.37 (0.4)	0.59 (0.9)	.52.0 (49.2)	1450.01 (1391)	. 505.01 (732)	420	397.70	5.48
119	9.95 (10.1)	0.33 (0.4)	0.69 (0.65)	50.52 (49.7)	1556.23 (1647)	433.16 (585.6)	410	397.87	2.96

Similar procedure is carried out to a different mill. Fig. 4(a) shows the release of CO during 65 data collection. From the graph, many points of CO released are beyond the allowable limit as indicated by the broken lines. These 17 points, which are beyond the allowable limit, are required to be within the allowable limit and the processes of releasing it need to be optimized. Fig. 5 shows the output of PM in which two points are above the allowable limit. All single major input variables of ANN models of those two pollutant models are employed and linked with GA to find the optimal configuration that emits allowable emission. Figs. 4(b) and 5(b) show the output of CO and PM after the implementation of the optimization process. The optimal points are clearly within the allowable limits for both emission.

In the above GAANN process, it is worth to note that the data presented in Tables 3–5 are only the data which were beyond the permissible limit of particular emission. The optimal points obtained in all of the pollutants are within the limit of minimum and maximum values used in the training phase for all the data sets; for instance, for furnace combus-

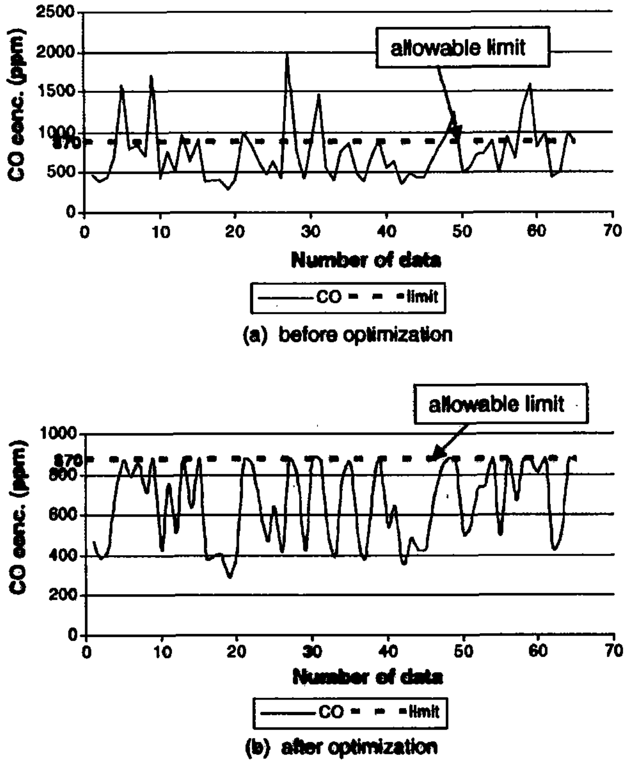


Fig. 4. Optimized conditions of palm oil mill showing the control of CO emission for a different mill.

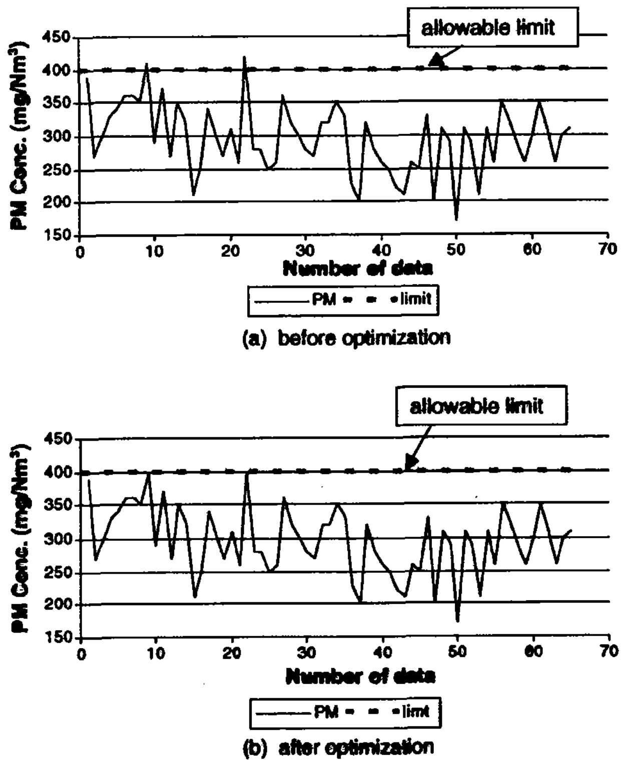


Fig. 5. Optimized conditions of palm oil mill showing the control of PM emission for a different mill.

tion, the minimum value is 0.2 and the maximum value is 0.6.

From the above GAANN process, all optimum input parameters can be included in new optimized data as a guided manual for boilerman or operator at the palm oil mill to operate the steam plant with minimum pollutant released. A table of the optimum input values can be tabulated and listed next to the operator. The smoke production from the stack gas will be maintained and controlled under the permitted level while satisfying the power load capacity required. As the input variables can be focused on certain pollutant emission, so this GAANN procedure can be applied in the case where only certain pollutant is mainly considered. In Malaysia, as black smoke which consists of particularly PM is put under much consideration, GAANN can be used to only consider the PM released. To ensure the black smoke, which is consisting of PM is controlled under DOE limit, an optimized value for input parameters combinations can be obtained. In summary, the combination of ANN and GA has achieved the objective of the study which is to find the optimal operating condition in boiler emission processes.

5. Conclusion

The optimization of operating condition in palm oil mill is successfully carried out with GAANN. This optimal operating condition will ensure the emission released is under the pollutant limit for both mills. Any data can be analyzed and predicted by ANN, and the operating condition can be optimized by GAANN. The CO, NO_x and PM optimized input parameters are useful as a guideline for mill operator in order to control the pollutants released from boiler. The tool developed from this study can be utilized to predict and control any pollution.

Acknowledgement

The authors would like to gratefully acknowledge Yayasan FELDA, Malaysia, for their financial support for this research.

References

Baines, G. H., Hayes, R. L., & Stabell, J. A. (1997). Predicting boiler emission with neural networks. *TAPPI Journal, 80*(5), 57–61.

Cho, A. S., Par, W. S., Chi, B. W., & Lew, M. C. (2000). Determining optimal parameters for stereolithography process via genetic algorithm. *Journal of Manufacturing System, 9*, 18–27.

Chow, T. T., Zhang, G. Q., Lin, Z., & Song, C. L. (2002). Global optimization of absorption chiller system by genetic algorithm and neural network. *Energy and Buildings, 34*, 103–109.

de Nevers, N. (1999). Air pollution. In *Control engineering*. USA: Prentice Hall.

Demuth, H., & Beale, M. (1998). *Neural network toolbox for use with MATLAB*. USA: The Math Works Inc.

Fisher-Rosemount Company. (1997). Industrial report for Westinghouse process control. *Smart process (Plant opt. software Westinghouse)*.

Freeman, J. A., & Skapura, J. M. (1991). *Neural network algorithm applications and programming techniques*. California: Addison-Wesley Pub.

Goldberg, D. E. (1989). *Genetic algorithms in search, optimisation, and machine learning*. Reading, Massachusetts: Addison-Wesley.

Holland, J. H. (1975). *Adaptation in natural and artificial systems*. Ann Arbor: University of Michigan Press.

Hou, Z., Kefa, C., & Jianbo, M. (2001). Combining neural network and genetic algorithms to optimize low NO_x pulverized coal combustion. *Fuel, 80*, 2163–2169.

Martines, F. G., & Ceolho, M. A. N. (2000). Application of feedforward ann to improve process control of PID (proportional, integral, derivative). *Computer & Chemical Engineering, 24*, 853–858.

MATLAB User's Guide. (1998). *Neural network toolbox* (Version 5.3). The Math Works Inc.

Morimoto, T., Suzuki, J., & Hashimoto, Y. (1997). Optimization of a fuzzy controller for fruit storage using neural networks and genetic algorithm. *Engineering Application Artificial Intelligence, 10*(5), 453–461.

Noble, T. H., & Mayhew, J. W. (1997). Keys to successful power boiler predictive emission monitoring system installation (PEMS). *TAPPI Journal*, 593–597.

Ricotti, M. E., & Zio, E. (1999). Neural network approach to sensitivity and uncertainty analysis. *Reliability Engineering and System Safety, 64*, 59–77.

Sette, S., Boullart, L., & Langauhore, L. V. (1998). Using genetic algorithms to design a control strategy of an industrial process. *Control Engineering Practise, 6*, 523–527.

Sun, H. (1999). Ranking importance of input parameters of neural networks. *Expert Systems with Application, 15*, 405–411.

Yue, H. H., Valle, S., & Qin, S. J. (1998). Comparison of several methods of multivariate soft sensor for emission monitoring. *Journal of Environment Engineering, 3*(6), 23–28.

Zhao, W., Chen, D., & Hu, S. (2000). Optimizing operating conditions based on ANN and modified gas. *Computer and Chemical Engineering, 24*, 61–65.