ANN and MLR Model of Specific Fuel Consumption for Pyrolysis Oil Blended with Diesel used in a Single Cylinder Diesel Engine: A Comparative Study

Saumil C Patel
Research Scholar,
PAHER University, Udaipur, Rajasthan, India

Pragnesh K Brahmbhatt, Ph.D.
Associate Professor, Government Engineering
College, Modasa, Gujarat, India

ABSTRACT

The main objective of the research is to compare the accurateness of Artificial Neural Networks (ANN) and Multiple Linear Regressions (MLR) model for Specific Fuel Consumption for pyrolysis oil blended with diesel used in a single cylinder diesel engine. In this study, parameters i.e. Injection Timing, Injection Pressure, Compression Ratio, and Load are taken. Artificial Neural Networks (ANN) and Multiple Linear Regressions (MLR) models were prepared using the results of Experiments to predict Specific Fuel Consumption for pyrolysis oil blended with diesel used in a single cylinder diesel engine. The results show that ANN prediction is more accurate than MLR prediction.

Keywords
C.I. Engine, Pyrolysis Oil, SFC, ANN, MLR

1. INTRODUCTION

Developing renewable energy has become an important part of worldwide energy policy to reduce greenhouse gas emissions caused by fossil fuel. Alternative transport fuels such as hydrogen, natural gas and biofuels are seen as an option to help the transport sector in decreasing its dependency on oil and reducing its environmental impact. Alternative fuels for the diesel engines are becoming increasingly important due to the diminishing petroleum reserves and environmental consequences of the exhaust gases from petroleum fuelled engines. Biofuel sources, particularly pyrolysis oil have attracted much attention as an alternative energy source. It is renewable, available everywhere and has proved to be a cleaner fuel and more environment friendly than the fossil fuels. However long term engine test results showed durability problems with pyrolysis oil because of deposit formation, carbon build up and lubricating oil contamination. Blending and transesterification may overcome these problems.

2. LITERATURE REVIEW

Surface soil shear strength could be a useful dynamic index for soil erodibility and thus a measure of soil resistance to water erosion. In this study, researcher evaluated the predictive capabilities of artificial neural networks (ANNs) and an adaptive neuro-fuzzy inference system (ANFIS) in estimating soil shear strength from measured particle size distribution (clay and fine sand), calcium carbonate equivalent (CCE), soil organic matter (SOM), and normalized difference vegetation index (NDVI). The results showed that the ANN model was more feasible in predicting the soil shear strength than the ANFIS model. The root mean square error (RMSE), mean estimation error (MEE), and correlation coefficient (R) between the measured soil shear strength and the estimated values using the ANN model were 0.05, 0.01, and 0.86, respectively. In ANFIS analysis, the RMSE was 0.08 and a lower correlation coefficient of 0.60 was obtained in comparison with the ANN model. Furthermore, the ANN and ANFIS models were more accurate in predicting the soil shear strength than was the conventional regression model. Results indicate that the ANN model might be superior in determining the relationships between index properties and soil shear strength [1].

MLR were used to predict the somatic production/biomass ratio of animal population from empirical data of population parameters and environmental variables. Based on data from 899 benthic invertebrate population, compared the prediction of P/B by MLR models and by ANN. The latter showed a slightly but significantly better performance. The accuracy of both approches was low at population level, but both may be used to estimate production and productivity of large population [2].

Quantitative structure–activity relationship (QSAR) models had developed to predict for CCR5 binding affinity of substituted 1-(3, 3-diphenylpropyl)-piperidinyl amides and ureas using multiple linear regression (MLR) and artificial neural network (ANN) techniques. A model with four descriptors, including Hydrogen-bonding donors HBD(R7), the partition coefficient between n-octanol and water logP and logP(R1) and Molecular weight MW(R7), showed good statistics both in the regression and artificial neural network with a configuration of (4-3-1) by using Bayesian and Leven- berg-Marquardt Methods. Researcher had Compared the descriptor's contribution obtained in MLR and ANN analysis shows that the contribution of some of the descriptors to activity may be non-linear [3].

In a machining operation, the productivity depends on the work-tool combination, speed, feed and depth of cut etc. Among the properties of the work-tool materials combination, hardness plays a crucial role in machining. To select the appropriate parameters including the tool material to meet the above objectives, the hardness of the workpiece was to be known. In small and medium scale machining industries the job orders will be of different materials with varying hardness values. This demands an online hardness measuring system. The methodology proposed was to measure the feed motor current using a current sensor and relate it with the hardness of the material to be machined. This was achieved using a tool with high hardness as a pilot machining tool with standard speed, feed and depth of cut. An artificial neural network (ANN) model had been built to predict hardness using spindle motor current. The ANN had been optimized to predict best possible values and the result of the same has been compared with that of regression analysis [4].

Structural optimization using computational tools has become a most important research field in recent years. Methods commonly used in structural analysis and optimization may

demand considerable computational cost, depending on the problem complexity. Among these ANN and MLR may be combined with classical analysis, to reduce the computational effort without affecting the ultimate result quality. In present study comparison of different predicting tool was explained. Researcher was compare the accurateness of ANN and MLR model for shear stress of EICHER 11.10 CHASSIS FRAME. The chassis frame was made of two side members joined with a series of cross members. The design variables were Number of cross members, their locations, cross-section and the sizes of the side. The chassis frame model was developed in Solid works and analyzed using Ansys. Then FEA was carried out on those models. ANN and MLR models were trained using the results of FEA to predict shear stress onthe chassis frame. The results indicate that ANN prediction was more accurate than MLR prediction [5].

Air pollution was a important issue that had been influencing human health, agricultural crops, forest and ecosystem. Forecasts was usually depended on statistical relationships between weather conditions and ambient air pollution concentrations. Multiple linear regression models have been widely used for this reason, and well-specified regressions can provide reasonable results. The aim of this study was to determine the best technique between Multiple Linear Regression (MLR) and Feedforward Backpropagation Artificial Neural Network (ANN) models for predicting concentration in Pulau Pinang. Multiple regression models and neural networks were examined for Seberang Jaya, Pulau Pinang with the same independent variables, enabling a comparative study of the two approaches [6].

3. THEORETICAL BACKGROUND

3.1 Multiple Linear Regression Analysis Model (MLR)

A multiple regression equation is used to describe linear relationships involving more than two variables. A multiple linear regression equation expresses a linear relationship between a response variable y and two or more predictors variable $(X_1, X_2, ..., X_K)$. The general form of a multiple regression equation is [5]:

$$\hat{Y} = b_0 + b_1X_1 + b_2X_2 + ... + b_kX_k \qquad \text{(I)}$$

A multiple linear regression equation identifies the plane that gives the best fit to the data,

$$\hat{Y} = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 \qquad \text{(II)}$$

Where,

$\hat{Y}$: predicted value of Specific Fuel Consumption

X_1 : Injection Pressure

X_2 : Injection Timing

X_3 : Compression Ratio

X_4: Load

b_0 : estimate value of y-intercept

b_1, b_2, b_3, b_4 : estimate value of the independent variable coefficient .

3.2 Artificial Neural Network (ANN)

Artificial Neural Network (ANN) captures the domain knowledge. The ANN can handle continuous as well as discrete data and have good generalization capability as with fuzzy expert systems. An ANN is a computational model of the brain. They assume that the computation is distributed over several simple units called neurons, which are interconnected and operate in parallel thus known as parallel distributed processing systems. Implicit knowledge is built into a neural network by training it. Several types of ANN structures and training algorithms have been proposed. Fig. 1 shows a flowchart for comparison of ANN and MLR model. Both modes are compared on the basis of error [5,7].
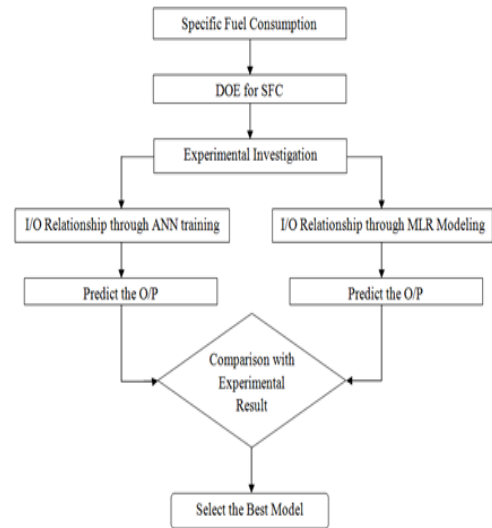

Fig. 1 Flowchart for comparison of ANN and MLR model

4. RESULT AND DISCUSSION

4.1 MLR Model

Linear regression is an approach to modeling the relationship between a scalar dependent variable y and one or more explanatory variables denoted X. The case of one explanatory variable is called simple linear regression. For more than one explanatory variable, it is called multiple linear regression [5].

A_j = 0.01 - 0.053 IT + 0.112 CR - 0.00030 INJ. PRE. + 0.252 LOAD   (III)

$$\hat{Y} = 0.01 - 0.053X_1 + 0.112X_2 - 0.00030X_3 + 0.252X_4 \qquad \text{(IV)}$$

From experimental data in Table 1 the Specific Fuel Consumption estimation formula (Equation.III) was calculated by using MLR. There are three chosen independent variables with 25 cases. Where the potentials independent variables are X_1 = Injection Timing, X_2 = Compression Ratio, X_3 = Injection Pressure, X_4 = Load and the dependent variable Y = Specific Fuel Consumption

Table 1. Multiple Linear Regression (MLR) Analyzed Data For Specific Fuel Consumption

Predictor	Coef	SE Coef	T	P
Constant	0.010	4.073	0.00	0.998
I.T.	-0.0531	0.1266	-0.42	0.680
C.R.	0.1120	0.1540	0.73	0.476
I.P.	-0.000300	0.007702	-0.04	0.969
Load	0.2520	0.1540	1.64	0.118
S = 1.08926 R-Sq = 14.5% R-Sq(adj) = 0.0%				

Table 1 shows the highest R Square (0.145) and Adjusted R Square (0.00) values. Hence, as it is found that formula for the Specific Fuel Consumption is ideal. From the calculation we can conclude that the Specific Fuel Consumption estimation formula using Multiple Linear Regression is as shown in Equation (III) [5].

Factors to be taken into consideration to choose best Equations:

➢ Use common sense and practical considerations to include or exclude variables.

➢ Consider the equation with high values of adjusted R2 and try including only a few variables.

➢ Consider the P-value (the measure of the overall significance of multiple regression equation-significance F value) displayed in computer output.

➢ The smaller P-value is the better.

Find the linear correlation coefficient r for each pair of variables being considered. If 2 predictor values have a very high r, there is no need to include them both. Exclude the variable with the lower value of r. Analysis of variance is given in Table 2 [5,9].

Table 2. Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	4.012	1.003	0.85	0.513
Residual Error	20	23.730	1.186		
Total	24	27.742			

4.2 ANN Model

ANN is a massive parallel-distributed information proc-essing system that has certain performance characteristics. Literature review shows that ANN models have better prediction capability than the regression models. So ANN models are also created for shear stress prediction. This section describes pre processes, model design and training, model simulation and post processes in the generation of ANN prediction models. All 25 experimental data sets are divided for training, validation and testing. Using GUI in Neural Network Toolbox in MATLAB, different network configuration with different number of hidden neurons is trained and their performance is checked. There are 17 data sets are used for training, 4 data sets for validation and 4 data sets for testing. It is clear that more data sets in training reduces processing time in ANN learning and improves the generalization capability of models, so large number of data sets are used to train the models. Attempts have been made to study the network performance with a different number of hidden neurons. A network is constructed each of them is trained separately, and the best network is selected based on the accuracy of the predictions in the testing phase [7,9].
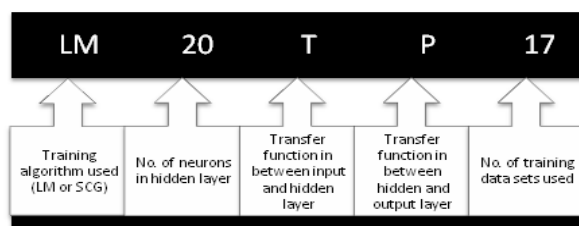

Fig. 2 ANN Model designation

Fig. 2 suggests how this model is designated. This designation covers various properties of the ANN model created. It covers types of training algorithm used, number of neurons in the hidden layer, transfer function used in between input and hidden layer, and in between hidden and output layer.A feed-forward neural network with back propagation is used. As shown in Fig. 3 the network consists of three layers. The first layer, which is the input layer, is triggered using the sigmoid activation function whereas the second layer is hidden layer and third layer is the output layer which is triggered using the linear activation function as shown in Fig. 4. A network of two transfer function, where the first transfer function is tansig and the second transfer function is purelin, can be trained to approximate any function [7,10].
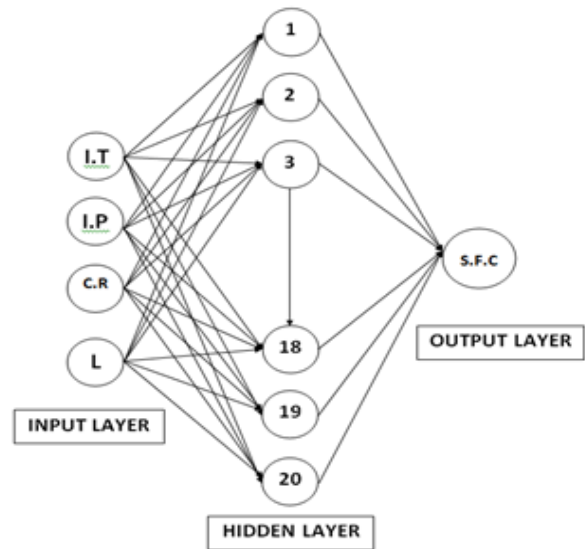

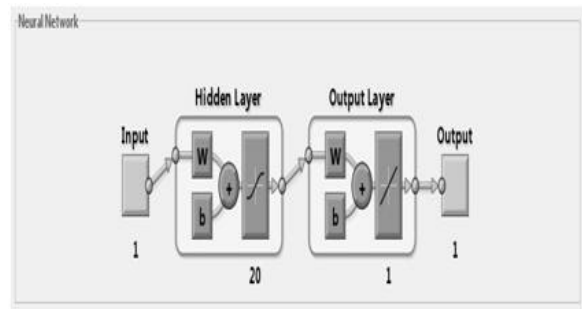**Fig. 3 General view of LM20TP Model view with three layers**

Fig. 4 Abbreviated view of LM20TP Model in MATLAB window

The network is trained using Levenberg-Marquardt algorithm. In the case of supervised learning, the network is presented with both the input data and the target data called the training set. The network is adjusted based on comparison of the output and target values until the outputs match the targets.After the data have been normalized, input data files and targets data files are created for training purpose. These input data files include file for training, validation and testing which contains input data sets in random order. Target data files include targets (normalized measured shear stress values respectively of input data sets) for training, validation and testing data sets. The work in this paper included a function approximation or prediction problem that required the final error to be reduced to a very small value [7,8].
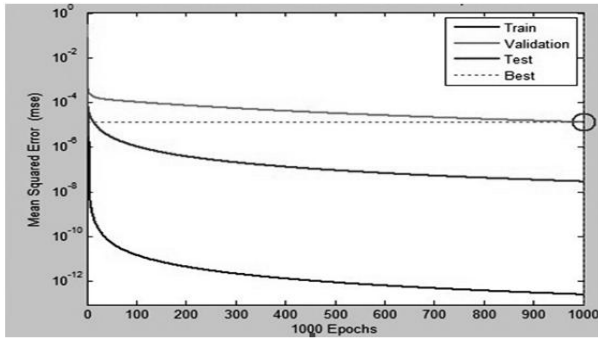
Fig. 5 LM20TP Model Training Performance Graph (retrained)

Fig. 5 shows retrained performance (MSE) graph of LM20TP model, created during its training. The training stopped after 1000 epochs because the validation error increased. It is a useful diagnostic tool to plot the training, validation, and test errors to check the progress of training. The result here is reasonable, because the test set error and the validation set error have similar characteristics, and it doesn't appear that any significant over fitting has occurred. After initial training of LM20TP model, it was retrained for epochs and performance MSE was obtained 2.53e-13 in training which took about 10 minutes on the test laptop [7,9].

Table 3. Experimental Results Table

| Sr. No	IT (° BTDC)	CR	INJ. PRE. (bar)	LOAD (kg)	Experimental SFC Kg/Kwh A_j	ANN Predicted SFC Kg/Kwh Pa_j	MLR Predicted SFC Kg/Kwh Pm_j	ANN Error= A_j-Pa_j	MLR Error= A_j-Pm_j
1	21	14	140	1	1.64	1.641	0.728	-0.001	0.912
2	21	15	160	2	0.87	0.869	1.086	0.001	-0.216
3	21	16	180	3	0.69	0.69	1.444	0	-0.754
4	21	17	200	4	0.61	0.609	1.802	0.001	-1.192
5	21	18	220	5	1.9	1.903	2.16	-0.003	-0.26
6	22	14	160	3	1.21	1.209	1.12	0.001	0.09
7	22	15	180	4	0.61	0.609	1.478	0.001	-0.868
8	22	16	200	5	4.56	4.554	1.836	0.006	2.724
9	22	17	220	1	1.06	1.059	0.934	0.001	0.126
10	22	18	140	2	0.7	0.699	1.322	0.001	-0.622
11	23	14	180	5	0.53	0.53	1.565	0	-1.035
12	23	15	200	1	1.34	1.339	0.663	0.001	0.677
13	23	16	220	2	0.73	0.73	1.021	0	-0.291
14	23	17	140	3	1.25	1.249	1.409	0.001	-0.159
15	23	18	160	4	4.73	4.729	1.767	0.001	2.963
16	24	14	200	2	1.12	1.119	0.75	0.001	0.37
17	24	15	220	3	0.72	0.719	1.108	0.001	-0.388
18	24	16	140	4	1.38	1.392	1.496	-0.012	-0.116
19	24	17	160	5	1	0.998	1.854	0.002	-0.854
20	24	18	180	1	0.79	0.79	0.952	0	-0.162
21	25	14	220	4	1.59	1.615	1.195	-0.025	0.395
22	25	15	140	5	1.26	1.259	1.583	0.001	-0.323
23	25	16	160	1	0.61	0.61	0.681	0	-0.071
24	25	17	180	2	0.52	0.52	1.039	0	-0.519
25	25	18	200	3	0.95	0.948	1.397	0.002	-0.447
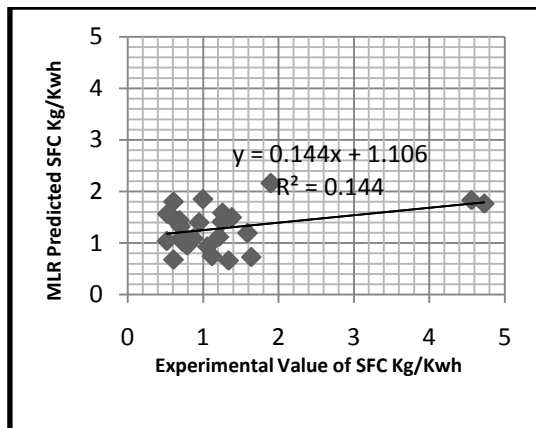
Fig. 6 Regression Model of MLR
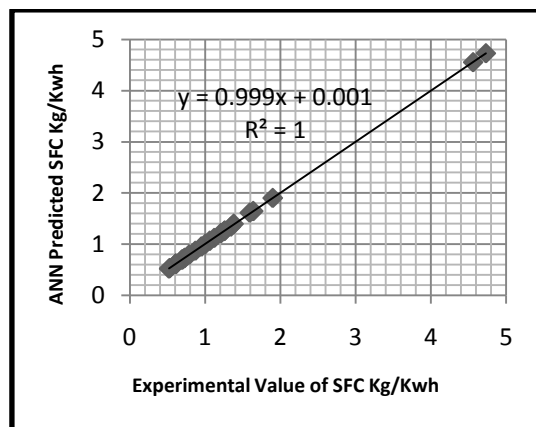
Fig. 7 Regression Model of ANN

4.3 Comparison of MLR and ANN

Fig. 6 shows a regression model of MLR and Fig. 7 shows a regression model of ANN. They show that ANN technique is more feasible in predicting the shear strength than the MLR technique. This might be due to the large amount of data required for developing a sustainable regression model, while the neural network could recognize the relationships with less data for distributed and parallel computing natures. A second reason is the effect of the predictors on the dependent variable, which may not be linear in nature. In other words, the ANN model could probably predict shear stress with a better performance owing to their greater flexibility and capability to model nonlinear relationships. Therefore, in the case of data sets with a limited number of observations in which regression models fail to capture reliably, advanced soft computing approaches like ANN may be preferred [5,7].

5. CONCLUSIONS

The present investigation aimed at the comparison of ANN and MLR model for Specific Fuel Consumption prediction. This optimization is carried out by developing Specific Fuel Consumption models based on L_{25} orthogonal array. An ANN model and MLR model are developed to predict Specific Fuel Consumption for pyrolysis oil blended with diesel used in a single cylinder diesel engine. The comparative study of MLR model and the ANN model for Specific Fuel Consumption prediction draws the following conclusions.

- The results obtained during preliminary test suggest that the ANN approach is a promising tool for accurately estimating Specific Fuel Consumption compare to MLR model.

- The result shows that the ANN techniques better than MLR method.

- Comparison of the goodness-of-fit statistics of the MLR models with those of the ANN models indicated that ANN gives better prediction results.

6. FUTURE SCOPE

- Results of ANN can be compared with other mathematical model to check the credibility of ANN.

- Different mathematical models can be compared with each other to findout the best prediction model.

7. REFERENCES

[1] Besalatpour, A., Hajabbasi, M. A., Ayoubi, S., Afyuni, M., Jalalian, A., & Schulin, R. (2012). Soil shear strength prediction using intelligent systems: artificial neural networks and an adaptive neuro-fuzzy inference system. *Soil Science and Plant Nutrition*, 58(2), 149-160.

[2] Brey, T., Jarre-Teichmann, A., & Borlich, O. (1996). Artificial neural network versus multiple linear regression predicting P/B ratios from empirical data.*Marine ecology-progress series*, 140, 251-256.

[3] Mouhibi, R., Zahouily, M., El Akri, K., & Hanafi, N. (2013). Using Multiple Linear Regression and Artificial Neural Network Techniques for Predicting CCR5 Binding Affinity of Substituted 1-(3, 3-Diphenylpropyl)-Piperidinyl Amides and Ureas. *Open Journal of Medicinal Chemistry*, 3, 7.

[4] Balamuruga, G. (2010) PREDICTION OF WORK PIECE HARDNESS USING ARTIFICIAL NEURAL NETWORK. *International Journal of Design and Manufacturing Technology (IJDMT)*, pp. 29-44

[5] Tushar, M. P., & Bhatt, N. (2013). ANN and MLR Model for Shear Stress Prediction of Eicher 11.10 Chassis Frame: A Comparative Study. *International Journal of Mechanical Engineering & Technology (IJMET)*, 4(5), 216 - 223.

[6] Ul-Saufie, A. Z., Yahya, A. S., Ramli, N. A., & Hamid, H. A. (2011). Comparison Between Multiple Linear Regression And Feed forward Back propagation Neural Network Models For Predicting PM10 Concentration Level Based On GaseousAndMeteorological arameters. *International Journal of Applied*, 1(4).

[7] Patel, T. M., & Bhatt, N. M. (2013). Shear Stress Prediction Using FEA-ANN Hybrid Modeling Of Eicher 11.10 Chassis Frame. *IOSR Journal of Mechanical and Civil Engineering*. PP 22-32.

[8] Zahouily, M., Rhihil, A., Bazoui, H., Sebti, S., & Zakarya, D. (2002). Structure-toxicity relationships study of a series of organophosphorus insecticides.*Molecular modeling annual*, 8(5), 168-172.

[9] Sahoo, S., & Jha, M. K. (2013). Groundwater-level prediction using multiple linear regression and artificial neural network techniques: a comparative assessment. *Hydrogeology Journal*, 21(8), 1865-1887.

[10] Gevrey, M., Lek, S., & Oberdorff, T. (2006). Utility of sensitivity analysis by artificial neural network models to study patterns of endemic fish species. In*Ecological Informatics* (pp. 293-306). Springer Berlin Heidelberg.