ARTIFICIAL NEURAL NETWORK BASED PREDICTION OF ENGINE-OUT RESPONSES FROM A BIODIESEL FUELLED COMPRESSION IGNITION ENGINE

by

Cheikh KEZRANE^a, S. Houcine HABIB^b, Mustafa BAYRAM^{c*}, Sultan ALQAHTANI^d, Sultan ALSHEHERY^d, Omolayo M. IKUMAPAYI^{e,f}, Esther T. AKINLABI^g, Stephen A. AKINLABI^g, Khaled LOUBAR^h and Younes MENNI^{i,j}

^a LDMM Laboratory, University of Djelfa, Djelfa, Algeria
 ^b Department of Mechanical Engineering, University of M'hamed Bougara, Boumerdes, Algeria
 ^c Department of Computer Engineering, Biruni University, Istanbul, Turkey
 ^d Mechanical Engineering Department, College of Engineering, King Khalid University,

Abha, Saudi Arabia

^e Department of Mechanical and Mechatronics Engineering, Afe Babalola University, Ado Ekiti, Nigeria

^f Department of Mechanical Engineering Science, University of Johannesburg, Johannesburg, South Africa

^g Department of Mechanical and Construction Engineering,

Faculty of Engineering and Environment, Northumbria University, Newcastle, UK ^h GEPEA, UMR 6144, DSEE-IMT Nantes La chantrerie 04 rue Alfred Kastler,

Nantes Cedex 3, France

ⁱ Department of Technology, University Center Salhi Ahmed Naama (Ctr. Univ. Naama), Naama, Algeria

^j National University of Science and Technology, Dhi Qar, Iraq

Original scientific paper https://doi.org/10.2298/TSCI2304433K

Numerical simulations, based on relatively complex physical models developed for CFD, can accurately predict engine-out responses, but they require huge memory space and/or computation time. In terms of resources and computer time, artificial intelligence methodologies are more cost-effective. In this work, we used an ANN to predict the performance and exhaust emissions of a single-cylinder Diesel engine running on fossil diesel, biodiesel, and their blends under various speed and load regimes. To perform the modeling, we employed multilayer perceptrons and a back-propagation gradient algorithm with momentum to train the network weights. The modification of the network weights was done using the second-order method of Levenberg-Marquardt, and the technique of early termination was utilized to avoid overtraining the model. The study involved using 70% of the complete experimental data to train the neural network, allocating 15% for network validation, and reserving the remaining 15% to evaluate the trained network effectiveness. The ANN model that was created demonstrated remarkable accuracy in predicting both engine performance and emissions. This is evident from the strong correlation coefficients observed, which ranged from 0.987 to 0.999, as well as the low mean squared errors ranging from $7.44 \cdot 10^{-4}$ to $2.49 \cdot 10^{-3}$.

Key words: neural network, multilayer perceptron, back-propagation, diesel, biodiesel, learning

^{*} Corresponding author, e-mail: mbayram34@gmail.com

Introduction

Around 25% of the world power is produced by internal combustion (IC) engines using fossil fuel oil. It is obvious that RES for power generation and entirely electric transportation will not happen for many years, if ever [1, 2]. Future IC engine development, however, should focus on techniques to lessen reliance on fossil fuels and research to increase efficiency with respect for the environment [2]. Biofuels have been used as suitable alternative fuels instead of current fossil fuels for the sake of reducing exhaust emissions' effect on the environment and preserving worldwide petroleum reserves from excessive consumption [3, 4]. In theory, these fuels hold promise as alternative options for compression ignition engines. However, to operate such engines, appropriate modifications are required, either on the engine, the fuel, or both [5]. These modifications require studying and modeling the engine with the suitable alternative fuels in different ways [6-9]. The experimental bench test is a method with high accuracy. On the other way, physical modelling (models based on an understanding and deep knowledge of the phenomena observed within the systems to be modelled), can miss accuracy. The simulations produced by these models agree very well with the results of bench tests, but this approach is still limited and lacks the capacity and precision required to model complex phenomena such as the formation of pollutants such as NO_x . Pollutant formation and oxidation mechanisms are not always well understood. This difficulty, as well as the dependence of the physical models on the geometrical characteristics of the system, pushed us to find a less complex solution. Among them, the black box models of the neural networks can be used.

The ANN hold significant importance as diagnostic tools within the realm of machine learning, enabling the simulation of system performance. It is based on using a system of numerous processing units connected to each other. The information is transmitted throughout the units, starting with entering the inputs and ending with delivering the outputs via predefined functions. The ANN models can be used as diagnostic, modeling, control, and optimization tools, as well as a computational model of analysis in engineering for predictive investigations [10-12]. As an advantage, the ANN does not need all the information in the system. Moreover, the information between the input variables is obtained as in the non-linear regression method, *i.e.*, by a preliminary study of recorded data. Hafner et al. [13] demonstrated how the engine control design was done using the quick neural network models. They later incorporated such neuro-models into more advanced emission optimization methods. The use of such ANN models in this context comprised pollution formation, virtual monitoring, and engine management optimization [14, 15]. The study conducted by Karthickeyan et al. [16] employed the implementation of an ANN model for the purpose of engine performance prediction. This encompassed various metrics including BTE, as well as emission characteristics such as CO, HC, and NO_x. The predictions were made under varying loads and different compression ratios (CR). In another independent study, see ref. [17], researchers developed an ANN model by employing the backpropagation learning algorithm. They found that the ANN model they developed showed good accuracy when tested with real-time data obtained from experiments. Hoang et al. [18] stated that there are numerous engine parameters and fuel qualities as input data, and there are numerous output factors that require prediction. However, adding too much detail to the input layer and forecasting too many output variables could make it difficult for the ANN algorithm to learn and build itself. As a result, the ANN outcomes could not be accurate. Given this, careful consideration must be paid to the choice of suitable data for input, layers that are hidden, and output variables.

3434

Kezrane, C., et al.: Artificial Neural Network Based Prediction of	
THERMAL SCIENCE: Year 2023, Vol. 27, No. 4B, pp. 3433-3443	3435

To expand upon the current body of research regarding the application of ANN in predicting engine performance and emissions, this study aims to investigate how different input parameters affect the accuracy of ANN models. While previous studies have employed a diverse range of input parameters, the focus of this investigation lies in utilizing fundamental fuel and engine data to assess whether simplified input parameters can still generate accurate predictions of engine performance and emissions. This study aims to identify a more efficient and cost-effective approach to modeling engine responses through the utilization of ANN. Moreover, this research provides valuable insights into the recommended practices for selecting input parameters in ANN models to ensure precise and dependable predictions of engine performance and emissions.

Experimental investigation

The data for this study was collected from a test bench that is composed mainly of a direct injection Diesel engine (7.5 kW at 2500 rpm), a dynamometer, a bay for analysis of the exhaust gas, an analyzer of nitric oxides, and a particle analyzer. The test specifications are listed on tab. 1 [4].

Table 1. Test engine specifications

Peak power	7.5 kW at 2500 rpm	Length of the connecting rod	165.3 mm
Refrigeration mechanism	Air cooled	Volume swept by the piston	630 cc
Cylinder count	1	Pressure of fuel injection	250 bar
Displacement volume	95.3 mm × 85.5 mm	Timing of fuel injection	13 CA bTDC

Figure 1 presents visual representations of the experimental set-ups, showcasing the diverse components and procurement systems employed [19, 20].

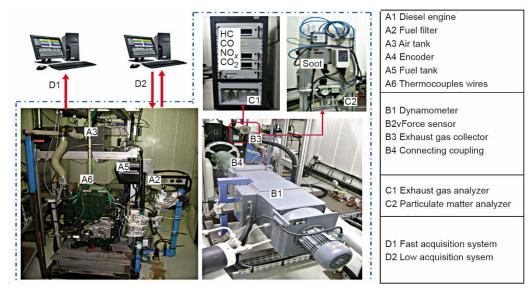


Figure 1. Experimental set-up scheme

Process of biodiesel preparation

3436

The production of biodiesel from waste cooking oil (WCO) was achieved through an alkali-catalyzed reaction, utilizing methanol as the alcohol and caustic potash as the catalyst. Various properties were assessed to analyze the fuels, including diesel fuel, WCO biodiesel, and their blends.

The experimental trials encompassed various fuel variants, namely diesel fuel B0, B100 (pure WCO biodiesel), as well as their respective blends: B25, denoting a 25% biodiesel content in diesel fuel, and B50, indicating a 50% biodiesel content in diesel fuel.

Experimental set-up and test procedure

Four speeds were used in this work: 1500 rpm, 1800 rpm, 2200 rpm, and 2500 rpm. Besides, 4 load parts were considered at each engine speed, varying from 25% to 100% as a maximum limit. Therefore, the present test can give a clear idea about the combustion characteristics, engine emissions, and performance since the complete interval of useful load and speed are taken into account.

The ANN design

The combustion and emission formation in the engine are established as non-linear processes, which makes the existing mathematical models unable to treat this property. As a real-time method, ANN is well known by its capability to handle non-linearities in the system. This characteristic renders it a viable alternative method for engineering analysis and predictions. Operating akin to a *black box* model, it directly utilizes the provided data as an input to generate the required information, thanks to its distinctive learning capabilities. Therefore, ANN can learn from the non-linear data of complex problems and calculate the unknown values with good accuracy. The foundational principle of neural networks is grounded in the concept of:

- the organization of neurons into layers,
- selecting an appropriate activation functions,
- the arrangement of connections among the hidden layers, and
- identifying the ideal number of hidden neurons within the hidden layers.

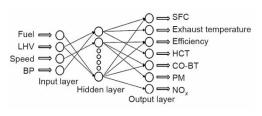


Figure 2. Structure of the ANN

According to experimental data, operating variables as the fuel type, its lower heating value (LHV), speed of engine and its brake power (BP) at the rated speed are the inputs of our network. The model considers the outputs to be the specific fuel consumption (SFC), exhaust gas temperature, thermal efficiency, and engine emissions. Here, the chosen multilayer perceptron is consisted of three layers as defour networks the second (hidden) will contain

picted in fig. 2. The first one (inputs) contains four neurons, the second (hidden) will contain a number of neurons and the third (output) consists of seven neurons.

The objective is to identify the optimal network structure, specifically determining the ideal number of hidden neurons that enable the most accurate approximation of the model. However, before that, we should normalize the data set. In practice, we put between zero and one using the following mathematical function:

$$y = \frac{(y_{\max} - y_{\min})(x_{\max} - x_{\min})}{x_{\max} - x_{\min}} + y_{\min}$$
(1)

where x is the vector that needs to be normalized and y – the normalized value corresponding to x.

Determining the optimal structure of network

The optimal structure of the model is determined in an iterative manner. To do this, the data are randomly divided into three sets: a learning data set consisting of 70% of the sample therefore 44 samples; a validation data set comprising 15% of the sample and 10 samples; and a test data set containing 15% of the sample therefore 10 samples.

Choice of network parameter

The non-linearity into neural networks is considered by means of special functions, named activation functions, which have to be continuous and differentiable, such as Log-sigmoid function. This one has been used in the hidden neurons of the feed-forward network. The log-sigmoid activation function has self-limiting property that can ensure automatic control in order to maintain a suitable output avoiding infinite values. There is also another suitable activation function (logistic sigmoid transfer) that can be used due to its properties [21]. The activation function is given by:

$$f(v) = \frac{1}{1 + \exp(-v)}$$
 (2)

where

$$v = w_0 + \sum_{i=1}^{n-1} w_1 x_i$$
(3)

Due to its advantageous optimization properties, the loss function chosen to be minimized was the mean square error (MSE). Its equation can be written as a sum of *n* terms:

$$MSE = \frac{1}{n} \sum_{j=1}^{n} (t_1 - o_j)$$
(4)

where n is the number of data sets, t – the target value, and o – the output value. Therefore, the prediction accuracy can be optimized by minimizing the MSE of the network using back propagation learning algorithms especially gradient descent algorithm. This approach is typically used for multilayer perception.

The Levenberg-Marquardt learning algorithm is commonly employed in back-propagation neural networks to handle the additional second derivative of error and enhance computational efficiency [22, 23].

Selecting the optimal ANN structure

In order to determine the optimal structure of the ANN, a range of 1 to 10 neurons in the hidden layer was explored, conducting multiple tests for each number of hidden neurons. The number of these neurons, which represent the units of non-linear adoption, can be increased to get a network with high storage feature that allow using more complex patterns [24].

Many networks of different structures have been tested before selecting the optimal ANN that gives the smallest possible test error. Figure 3 shows a sample of these tests, MSE versus number of hidden neurons [25].

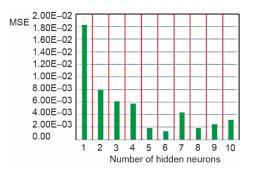


Figure 3. Analysis of MSE as a function of the

After analyzing the obtained results, we notice the important effect of the initialization of the parameters. This suggests that by adjusting the initialization of weights and biases, each number of neurons in the hidden layer can produce distinct outcomes. The criterion of choice between these different models is of course the MSE on the test set, fig. 3.

Based on this criterion, it is observed that the network comprising six neurons in the hidden layer exhibits the lowest test error MSE of $1.32 \cdot 10^{-3}$, accompanied by favorable MSE values on the training set $(1.23 \cdot 10^{-3})$ and the validation set $(2.22 \cdot 10^{-3})$.

Hence, the optimal pattern of network is (4-6-7). Comprehensive information regarding the network parameters can be found in tab. 2.

Optimal ANN			
Topology	The configuration consists of 4 inputs, a hidden layer comprising 6 neurons, and 7 outputs, denoted as 4-6-7.		
Data	Training 70% (44 samples); validation 15% (10 samples); test 15% (10 samples)		
The activation function	Function log-sigmod		
The training algorithm	Lenvenberg-Marquardt		
The loss function criteria	The smallest achievable value of the MSE		
The stopping criteria	Terminate the network training process when the validation error shows a noticeable rise, indicating a deviation from the desired trend.		

Table 2. The network parameters

hidden neurons

Results and discussions

The model output values were compared to the actual engine-out responses to assess their accuracy. The comparison between experimental data and the values calculated by the ANN for SFC, BTE, and exhaust gas temperature is illustrated in figs. 4-6. To evaluate the accuracy of the ANN predictions, they were compared to an ideal prediction represented by a linear trend line. The results clearly demonstrate the close proximity of all predicted values to the trend lines.

The results indicate that the predictions for the SFC are highly accurate, as evidenced by a MSE of $8.967 \cdot 10^{-4}$ and a correlation coefficient, *R*, of 0.97653. These values suggest that the predicted SFC values closely align with the experimental values.

The low MSE indicates minimal deviation between the predicted and experimental SFC, while the high correlation coefficient suggests a strong linear relationship, fig. 4.

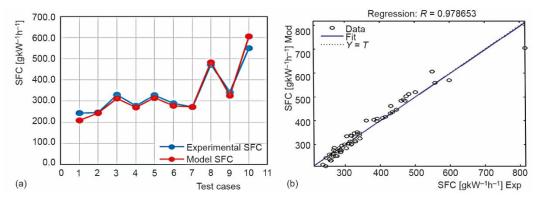


Figure 4. The side-by-side evaluation of the predicted and measured values for SFC; (a) SFC for different tests and (b) correlation coefficient

Similarly, the brake thermal efficiency predictions also show excellent concordance with the experimental values. The points are closely aligned with the trend line, as indicated by a correlation coefficient of 0.98799 and an MSE of $9.24 \cdot 10^{-4}$. This implies that the predicted values for brake thermal efficiency closely match the observed values, with a minimal margin of error, fig. 5.

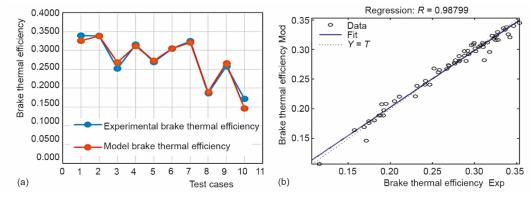


Figure 5. Comprehensive comparison between the predicted values and the corresponding measured values of thermal efficiency; (a) brake thermal efficiency for different tests and (b) correlation coefficient

Furthermore, the exhaust gas temperature predictions demonstrate a high level of accuracy, with an MSE of $7.44 \cdot 10^{-4}$ and a correlation coefficient of 0.99329. These values indicate a strong agreement between the predicted exhaust gas temperatures and the actual measurements. The low MSE suggests minimal deviation between the predicted and observed values, while the high correlation coefficient indicates a close linear relationship, fig. 6.

Figures 7-10 show the comparison between experimental results and the obtained values for hydrocarbon emission HCT, CO, NO_x , and PM, respectively. The ideal predictions are represented by straight trend lines in order to evaluate the accuracy of the ANN predictions.

The results indicate that the predictions for HCT, CO, NO_x, and PM exhibit a high level of accuracy. The MSE values for HCT, CO, NO_x, and PM are $7.444 \cdot 10^{-4}$, $1.4 \cdot 10^{-3}$, $2.49 \cdot 10^{-3}$, and $2.49 \cdot 10^{-3}$, respectively. These low MSE values suggest that the predicted values closely align with the experimental measurements, with minimal deviation. Additionally, the correlation coefficients for HCT, CO, NO_x, and PM are 0.96107, 0.95911, 0.97887, and

3439

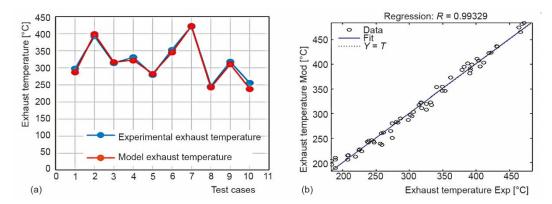


Figure 6. The side-by-side comparison of predicted and measured values for exhaust *T*; (a) exhaust temperature for different tests and (b) correlation coefficient

0.96698, respectively. These high correlation coefficients indicate a strong linear relationship between the predicted and experimental values. The closer the correlation coefficient is to 1, the better the concordance between the predicted and observed data.

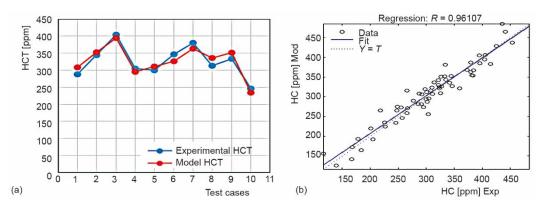


Figure 7. Comprehensive comparison between the predicted values and the corresponding measured values of HCT; (a) HCT for different tests and (b) correlation coefficient

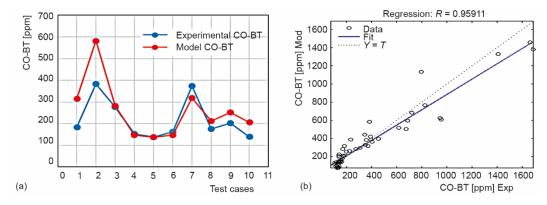


Figure 8. Detailed comparison between predicted values and the corresponding measured values of CO-BT; (a) CO-BT for different tests and (b) correlation coefficient

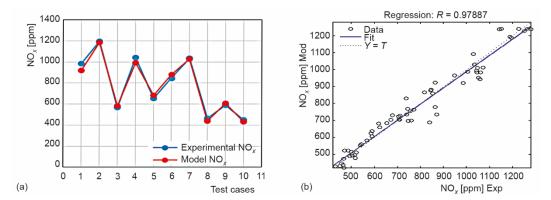


Figure 9. Comparative analysis of predicted and measured NO_x values; (a) NO_x for different tests and (b) correlation coefficient

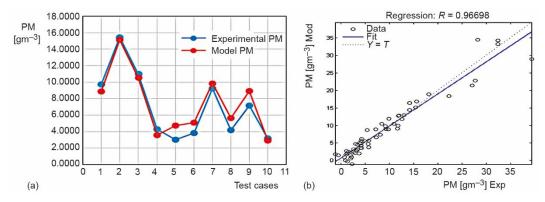


Figure 10. Comprehensive comparison between the predicted values and the corresponding measured values of PM; (a) PM for different tests and (b) correlation coefficient

Furthermore, the interpretation of the results highlights that the predicted values closely follow the experimental trend across all ranges of load and speed. This observation suggests that the predictive model captures the underlying patterns and variations in the data accurately.

The close alignment between the predicted and experimental values for all the mentioned pollutants implies that the model is reliable and capable of estimating pollutant emissions with high precision. In summary, the enhanced interpretation of the results confirms the accuracy of the predictions for HCT, CO, NO_x , and PM. The low MSE values and high correlation coefficients demonstrate the close agreement between the predicted and experimental values. The fact that the predicted values align well with the experimental trend across different load and speed ranges further validates the reliability of the predictive model.

Conclusions

This research utilizes an extensive collection of empirical data encompassing the complete speed/load range of an engine running on fossil diesel (B0), pure WCO biodiesel (B100), as well as their blends (B25) and (B50). The gathered experimental data were then employed for the training, validation, and testing of the ANN model. In order to describe the

engine-out responses, seven MLP were used to estimate the seven engine characteristics (SFC, exhaust gas temperature, BTE, HCT, CO, NO_x , and PM) as a function of the four engine control parameters (fuel, LHV, speed, BP). It is demonstrated that the seven trained MLP are capable of accurately predicting each of the seven engine reactions and, as a result, accurately capturing the engine's characteristics under a variety of operational conditions.

The MLP were trained, validated, and tested by experimental results for different fuels at different engine operational conditions, which demonstrates the utility of the ANN as a precise and effective modelling technique on the one hand, and on the other hand, certifies the robustness of the control parameters chosen to be learned by the model.

The significance of employing this ANN model and its ability to accurately predict engine parameters is emphasized by the correlation coefficient values, which range from 0.9765 to 0.9923 for performance and 0.9591 to 0.97887 for emission. This underscores the reliability and effectiveness of the ANN in making accurate predictions.

For all scenarios examined, the mean square error values fell within the range of 0.0007-0.0009 for performance and 0.0007-0.0014 for emission. These results serve as compelling evidence for the efficacy of the selected inputs incorporated into the model.

While the model has demonstrated its capability to forecast engine responses in a compression ignition engine fueled by biodiesel, there is still a lack of refinement in accounting for the inherent discrepancies resulting from the diverse physicochemical properties of various fuel blends. Therefore, the authors intend to further study the capability to incorporate physico-chemical property models into the ANN for future research.

Acknowledgment

The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through large group Research Project under grant number RGP2/290/44.

References

- [1] Reitz, R. D., *et al.*, IJER editorial: The Future of the Internal Combustion Engine, 2020, SAGE Publications Sage UK: London, England, pp. 3-10
- [2] Naima, K., et al., Experimental and Numerical Investigation of Combustion Behaviour in Diesel Engine Fuelled with Waste Polyethylene Oil? *Journal of Engineering Science and Technology*, 13 (2018), 10, pp. 3204-3219
- [3] Bui, T. T., et al., Characteristics of PM and soot emissions of internal combustion engines running on biomass-derived DMF biofuel: a review, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 44 (2022) 24, pp. 8335-8356
- [4] No, S. Y., *Application of Liquid Biofuels to Internal Combustion Engines*. Springer Nature, New York, USA, 2020
- [5] Naima, K., et al., Effect of EGR on Performances and Emissions of DI Diesel Engine Fueled with Waste Plastic Oil: CFD Approach, In Annales de Chimie-Science des Matéeriaux, 45 (2021), 3, pp. 217-223
- [6] Devarajan, Y., et al., A Comprehensive Study on Emission and Performance Characteristics of a Diesel Engine Fueled with Nanoparticle-Blended Biodiesel, *Environmental Science and Pollution Research*, 26 (2019), Feb., pp. 10662-10672
- [7] How, H. G., *et al.*, Influence of Injection Timing and Split Injection Strategies on Performance, Emissions, and Combustion Characteristics of Diesel Engine Fueled with Biodiesel Blended Fuels, *Fuel*, 213 (2018), Feb., pp. 106-114
- [8] Hoang, A. T., Combustion Behavior, Performance and Emission Characteristics of Diesel Engine Fuelled with Biodiesel Containing Cerium Oxide Nanoparticles: A Review, *Fuel Processing Technolo*gy, 218 (2021), July, 106840
- [9] Chen, H., *et al.*, NO_x Emission of Biodiesel Compared to Diesel: Higher or Lower?, *Applied Thermal Engineering*, *137* (2018), June, pp. 584-593

- [10] Odesola, I. F., et al., An ANN Approach for Estimation of Thermal Comfort and Sick Building Syndrome, Rev. d'Intelligence Artif., 33 (2019), 2, pp. 151-158
- [11] Ghritlahre, H. K., Prasad, R. K., Prediction of Exergetic Efficiency of Arc Shaped Wire Roughened Solar Air Heater Using ANN Model, *International Journal of Heat and Technology*, 36 (2018), 3, pp. 1107-1115
- [12] Chen, J., Randall, R. B., Improved Automated Diagnosis of Misfire in Internal Combustion Engines Based on Simulation Models, *Mechanical Systems and Signal Processing*, 64 (2015), Dec., pp. 58-83
- [13] Hafner, M., et al., Fast Neural Networks for Diesel Engine Control Design, Control Engineering Practice, 8 (2000), 11, pp. 1211-1221
- [14] Arsie, I., et al., Neural Network Models for Virtual Sensing of NO_x Emissions in Automotive Diesel Engines with Least Square-Based Adaptation, Control Engineering Practice, 61 (2017), Apr., pp. 11-20
- [15] Kakati, D., et al., Development of an Artificial Neural Network Based Virtual Sensing Platform for the Simultaneous Prediction of Emission-Performance-Stability Parameters of a Diesel Engine Operating in Dual Fuel Mode with Port Injected Methanol, Energy Conversion and Management, 184 (2019), Mar., pp. 488-509
- [16] Karthickeyan, V., et al., Developing of ANN Model for Prediction of Performance and Emission Characteristics of VCR Engine with Orange Oil Biodiesel Blends, Journal of the Brazilian Society of Mechanical Sciences and Engineering, 39 (2017), Mar., pp. 2877-2888
- [17] Musthafa, M., Performance Characteristics of a Variable Compression Ratio CI Engine Simulation Using Artificial Neural Network, *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 44 (2022), 1, pp. 1912-1922
- [18] Hoang, A. T., et al., A Review on Application of Artificial Neural Network (ANN) for Performance and Emission Characteristics of Diesel Engine Fueled with Biodiesel-Based Fuels, Sustainable Energy Technologies and Assessments, 47 (2021), Oct., 101416
- [19] Cheikh, K., et al., Experimental Assessment of Performance and Emissions Maps for Biodiesel Fueled Compression Ignition Engine, Applied energy, 161 (2016), Jan., pp. 320-329
- [20] Naima, K., et al., A Comparative Assessment of Combustion Behavior and Emissions Characteristics of DI Diesel Engine Fueled with Waste Plastic Oil and Eucalyptus Biofuel for Sustainable Development Applications, *International Journal of Low-Carbon Technologies*, 17 (2022), Dec., pp. 1399-1405
- [21] Watrous, R. L., (1988), Learning Algorithms for Connectionist Networks: Applied Gradient Methods of Non-Linear Optimization, University of Pennsylvania Department of Computer and Information Science, Technical Report No. MS-CIS-88-62, Philadelphia, Penn., USA, date Posted: 18 September 2007
- [22] Hagan, M. T., *et al.*, An Introduction to the Use of Neural Networks in Control Systems, *International Journal of Robust and Non-linear Control: IFAC-Affiliated Journal*, *12* (2002), 11, pp. 959-985
- [23] Nelles, O., Non-Linear System Identification: from Classical Approaches to Neural Networks, Fuzzy Models, and Gaussian Processes, Springer Nature, New York, USA, 2020
- [24] Blayo, F., Verleysen, M., Les Reseaux de Neurones Artificiels, Universitaires de France, Paris, France, 1996
- [25] Dreyfus, G., Apprentissage statistique, Editions Eyrolles, Paris, France, 2008