A Genetic Algorithm Optimized ANN for Prediction of Exergy and Energy Analysis Parameters of a Diesel Engine Different Fueled Blends

Ali YASAR

Department of Mechatronic Engineering, Faculty of Technology, Selçuk University, Konya, Türkiye

ARTICLE INFO

ABSTRACT

In this research, a hybrid artificial neural network (ANN) optimized by a genetic algorithm (GA) was used to estimate energy and exergy analyses parameters. This article presents an approach for estimating energy and exergy analyses parameters with optimized ANN model based on GA (GA-ANN) for different ternary blends consisting of diesel, biodiesel and bioethanol in a single-cylinder, water-cooled diesel engine. The data used in the experiments performed at twelve different engine speeds between 1000 and 3000 rpm with 200 rpm intervals for five different fuel mixtures consisting of fuel mixtures prepared by blends biodiesel, diesel and 5% bioethanol in different volumes constitute the input data of the models. Using these input data, engine torque (ET), amount of fuel consumed depending on fuels and speed (AFC), carbon monoxide emission values (CO), carbon dioxide emission values (CO₂), hydrocarbon emission values (HC), nitrogen oxides emission values (NOₓ), the amount of air consumed (AAC), exhaust gas temperatures (EGT) and engine coolant temperatures (ECT) were estimated with the GA-ANN. In examining the results obtained were examined, it was proved that diesel, biodiesel and bioethanol blends were effective in predicting all the results mentioned in engine studies performed at 200 rpm intervals in the 1000–3000 rpm range. A standard ANN model used in the literature was also proposed to measure the prediction performance of GA-ANN model. The predictive results of both models were compared using various performance indices. As a result, it was revealed that the proposed GA-ANN model reached higher accuracy in estimating the exergy and energy analyses parameters of the diesel engine compared to the standard ANN technique.

1. Introduction

Due to increasingly scarce resources and the negative impact on the environment, there has been a long-term increase in interest in clean alternative fuels instead of fossil fuels [1]. In this context, biodiesel with lower greenhouse gas emission levels can be used as an alternative to diesel from fossil fuels. [2, 3]. Due to the negative impacts such as greenhouse gas emissions caused by the increased use of diesel fuel recently and environmental pollution, review studies have been conducted for diesel fuels with additives to reduce the negative impacts [2]. Compared to fossil diesel, biodiesel is one of the most promising alternative biofuels, which has made the most progress due to its properties very similar to diesel [4, 5]. On the other hand, for reasons such as high density and poor fluidity, it may be necessary to use additives such as ethanol to facilitate the use of biodiesel in diesel engines [6, 7]. Engine studies to test engines are often difficult, time consuming and costly to do [8]. For these reasons, artificial intelligence methods can be used to model engine performance parameters to eliminate these costs, time, and complexity. Recently, artificial intelligence methods have appeared in many automotive engineering studies, such as engine performance and exhaust emissions. In recent decades, artificial intelligence algorithms have been rapidly improved and are widely used. For example, support vector machine (SVM) and random forest (RF) methods from machine learning techniques were used to predict pyrolytic gas yields and compositions by Tang et al [9]. Paul et al. in a partly changed single-cylinder direct
injection (DI) diesel engine, nitrogen oxides (NO\textsubscript{x}), brake specific energy consumption (BSEC), unburned hydrocarbon (UBHC), brake thermal efficiency (Bth), carbon monoxide (CO) and carbon dioxide (CO\textsubscript{2}) have used the ANN to estimate emission characteristics such as emissions [10]. Arcaklioğlu and Çelikten (2005) defines, using artificial neural-networks (ANNs), exhaust emissions form and the performance of a diesel engine with regard to injection pressure, throttle position and engine speed [11]. Gürген, Üner and Altun used ANN to model the cyclic variability of a diesel engine using diesel fuel and butanol-diesel fuel mixtures [12]. Aydı̇n, Uslu and Çelik, using the optimized RMS, the engine performance and emission parameters of a single cylinder diesel engine operating with biodiesel-diesel fuels were estimated by ANN [13]. Morris, Daood, and Nimmo applied and analyzed machine learning algorithms for estimating ash, K, Na, Cl, Pb, and Zn levels in commercial wood fuel mixtures used in a typical biomass power plant [14]. Zheng et al. studied various ANN techniques to estimate the viscosity of biodiesel blends and compare them with empirical correlations [15]. Uslu and Çelik conducted a study to predict exhaust emissions and engine efficiency of a diesel engine using ANN on a single-cylinder diesel engine with diethyl ether [16]. Hung at al. used ANN to estimate exhaust emissions from a water-cooled multi-cylinder diesel engine fueled by biodiesel from corn blends [17]. Grahovac et al. proposed the use of ANN for modeling and estimating bioethanol production from sugar beet processing by-products and intermediates [18]. Yılmaz et al. used the least squares support vector machine (LSSVM) and response surface methodology (RSM) to evaluate the engine performance and exhaust emission outputs of hazelnut oil methyl ester (HOME) in a turbo charged direct injection (TDI) diesel engine [19]. Yasar et al. proposed a fuzzy control system for estimating the combustion parameters, exhaust emission and engine performance parameters of bioethanol-gasoline mixtures at different engine loads [20].

In this article, the data obtained from the exergy and energy analyses parameters for a diesel engine fueled with diesel-biodiesel-bioethanol blends in a single-cylinder, water-cooled diesel engine. As dataset, all data were collected from the literature [21, 22]. Using the Genetic algorithm to optimize the initial weights in the input parameters of the ANN model will achieve faster convergence of ANN. Moreover, the ANN model is easy to achieve local optimum and overfit. It is an important task for the ANN model to reduce the problems of overfitting and local optimum, which are the biggest problems encountered in ANNs [23, 24]. The benefit of the genetic algorithm here is that it can reduce the probability of falling into the local optimum and has a strong stability [25]. For this reason, the GA-ANN model has been widely used in different fields of study recently, which makes it have higher estimation accuracy and faster convergence speed.

In the work, in addition to the simple ANN model, the GA-ANN model, which was developed for the estimation of the energy and exergy analysis parameters of the diesel engine, is also applied for the first time. Besides, the improved models including simple and hybrid techniques are compared so as to select the best models for predicting energy and exergy analysis parameters. In this study, the energy and exergy analysis parameters of the diesel engine were first estimated by creating a ANN model from these data. A 10-fold cross-validation was used to objectively measure the success of the model. Finally, the energy and exergy parameters of the diesel engine were estimated using the GA-ANN model, which was created using the genetic algorithm to optimize the starting weights so that the ANN model does not suffer from a local optimum and overfitting.

The main contributions of this study are listed below.

The proposed system is to present the GA optimized ANN model for estimating the energy and exergy parameters of the diesel engine.

Using the GA-ANN model, it contributed to the estimation of results such as engine power, instantaneous fuel consumption, engine coolant temperature, and exhaust emissions from diesel engines.

R-Squared values were calculated as 0.65-0.99 and 0.971-0.999 for ANN and GA-ANN, respectively.

The results of the estimation will help to reduce the time loss and cost of diesel engine testing.

This article has continued as follows. Chapter 2 describes the materials and methods used in this study. Chapter 3 presents the results of the estimates of the models ANN and GA-ANN. Chapter 4 addresses the results of the study.

2. MATERIALS AND METHODS

ANN architecture may differ according to the type of problem to be solved in prediction problems. Therefore, in order to fully benefit from the ANN architecture to be created according to the problem, the ANN architecture and the training process must be optimized. Generally, the ANN architecture uses trial-and-error approaches to arrive at the optimum solution. This causes ANN to be expressed as negative aspects. Researchers have proposed different algorithms for ANN architecture and training methods to eliminate these disadvantages. Some of them are; particle swarm optimization (PSO) [26], ant colony (AC) [27], simulated annealing (SA) [28], Genetic Algorithm (GA) [29] etc. [30]. ANN optimized with genetic algorithm can be used as an effective tool for estimating complex problems [31].

In this study, the ANN and GA-ANN models were used to estimate the energy and exergy analysis parameters of the diesel engine at different fuel blends and different engine speeds (rpm). The graphical model representation of this
study is given in Figure 1.

2.1. Artificial Neural Network (ANN)

ANN takes an example by from the working logic of the biological neural network in the human brain. It is a computational model inspired by information processing [32]. ANN is a machine learning method that can learn from imprecise or complex data and is inspired by the structure and functional aspects of biological neural networks [33, 34]. An ANN has three layers: input, hidden, and output. In ANN models, the number of hidden layers varies according to the number of input layers [25]. The number of features at the input and output in the designed model determines the number of neurons in the input and output layers. In contrast, there is no established method for the number of neurons in the hidden layer. However, the number of neurons in the hidden layer and the number of hidden layer(s) are determined by the complexity of the problem and trial-and-error method [35, 36]. ANN architecture showing the relationship between input variables and output variable(s) is shown in Figure 2.

Typically, when examining ANN architecture, the input layer feeds the hidden layer(s), and the hidden layer(s) feeds the output layer [37]. When the studies in the literature are examined, although there are many models of artificial neural networks, we come across with error back propagation learning algorithms, which are mostly called multi-layer feed-forward neural networks and BP networks [38]. The purpose of this algorithm is to minimize the error between the targeted output and the output produced by the ANN, and by determining the necessary weights, the error is propagated to the input layer. The biggest problem encountered in the algorithm is that it converges slowly and gets stuck at the local minimum [35].

2.2. Genetic Algorithm (GA)

Genetic algorithm (GA) is an optimization algorithm that is inspired by the biological evolution process and is frequently encountered in literature studies. It is a population-based search algorithm that uses the concept of survival of the fittest [39]. GA generates the optimal solution as an optimization algorithm from the possible solution space by using the fitness (objective) function and constraints [40]. GA is a method of optimizing the search tool for difficult problems based on the principle of genetic selection [41]. The major advantage of GA is its ability to use the accumulated knowledge about the initial unknown search space to move subsequent searches to useful areas [42]. Selection, chromosome representation, crossover, mutation, and fitness function calculation are the basic elements of GA. A simple GA algorithm is given in the figure 3 [43, 44]. GA dynamically updates the search process and reaches the optimum solution through the probabilities that occur in crossover and mutation operations. GA may
need to evaluate more than one individual in order to produce more than one optimum solution. As a result, GA has better global search capability [45].

Algorithm 1. The simple genetic algorithm proposed by Caelo and Ali

1. Initialization:
   - Generate N uniformly distributed random points (chromosomes) in \( [0,1] \) and store them to the set S.
   - Set iter = 0

2. Evaluation:
   - Evaluate the fitness value for each chromosome.

3. Selection:
   - Select \( m \) parents from S.

4. Crossover:
   - Create new points (offsprings) from the previously selected parents.

5. Replacement:
   - Replace the \( m \) worst chromosomes in the population with the previously generated offsprings.

6. Local Search:
   - Create using the local technique procedure a trial point if where is the current worst point in S, then replace by .

7. Set iter = iter + 1

   goto step 2

Figure 3. Simple genetic algorithm

2.3. Integration of GA and ANN

One of the first population-based stochastic algorithms proposed in history is the Genetic Algorithm (GA). GA has three main operators and these are selection, crossover and inheritance mutation[46]. Genetic algorithm is noted as a search process used in computing to discover precise or a nearly solution for search and optimization problems. GA make a technique for programs to evolve their parameters automatically[47]. GA dynamically modifies the search process thanks to the probabilities of inheritance mutation and crossover, reaching the optimal solution. GA can produce multiple optimum solutions. Therefore, GA has better global search ability[41]. In spite of all benefits of ANNs, they endure several problems like slow convergence and being stuck in a regional minimum. To address these issues, researchers have recently recommended a hybrid method of GA-ANN to improve the performance of ANNs and achieve the global minimum. Despite all the advantages of ANNs, they face some problems, such as slow convergence and persistence in a regional minimum. The base goal of using hybrid GA–ANN model is regulating a set of biases and weights to minimizing the purpose function[35]. Studies on the use of GA in optimization problems have proven to that they have a strong potential to achieve global optimum solutions[48]. The architecture of an optimized ANN model based on genetic algorithm is shown in Figure 4.

2.3. K-Fold Cross-Validation

K-fold cross validation is used to objectively measure the success of prediction models. Cross-validation is a technique intended to increase the security of prediction methods and eliminate the aspect of random results. In K-Fold Cross-Validation, the data to be employed in the study is divided into \( k \) different subsets. \( k-1 \) subsets are employed to train the data, with the last subset serving as test data. The obtained average error value shows the availability of the suggested model. The \( k=10 \), k-fold cross-validation method is

3. Results and Discussion

3.1. ANN modelling results

ANN model was used to estimate the energy and exergy analyses parameters values of the diesel engine at different fuel mixtures and different engine speeds (rpm) in a single-cylinder, water-cooled diesel engine. The data used in the experiments performed at twelve different engine speeds between 1000 and 3000 rpm with 200 rpm intervals for five different fuel mixtures consisting of fuel mixtures prepared by blends biodiesel, diesel and 5% bioethanol in different volumes constitute the input data of the models. Using these input data, engine torque (Nm), amount of fuel consumed depending on fuels and speed (g/h), CO emission values (%), CO\(_2\) emission values (%), HC emission values (ppm), NO emission values (ppm), the amount of air consumed (kg/h),
exhaust gas temperatures (°C) and engine coolant temperatures (°C) were estimated with the ANN. ANN model is depicting in Figure 6. The comparison of the results obtained with ANN and the results obtained as a result of the experiments data for ET, AFC, CO, CO$_2$, HC, NO, AAC, EGT and ECT is shown in Figure 7-11.
Figure 7. Comparison of prediction and experimental results for ET and AFC

Figure 8. Comparison of prediction and experimental results for CO and CO$_2$

Figure 9. Comparison of prediction and experimental results for HC and NO

Figure 10. Comparison of prediction and experimental results for AAC and EGT

Figure 11. Comparison of prediction and experimental results for ECT
The performance of the ANN model was calculated by Eqs.1-4 (R-Squared, MSE, RMSE, and MAE) and shown in Table 1.

\[
R^2 = 1 - \frac{\sum_{i=1}^{m}(X_i - Y_i)^2}{\sum_{i=1}^{m}(\text{average}(Y_i) - Y_i)^2}
\]  

Eq.1

\[
MSE = \frac{1}{m} \sum_{i=1}^{m}(X_i - Y_i)^2
\]  

Eq.2

\[
RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^{m}(X_i - Y_i)^2}
\]  

Eq.3

\[
MAE = \frac{1}{m} \sum_{i=1}^{m}|X_i - Y_i|
\]  

Eq.4

Table 1. ANN prediction Values

<table>
<thead>
<tr>
<th></th>
<th>ET</th>
<th>AFC</th>
<th>CO</th>
<th>CO₂</th>
<th>HC</th>
<th>NO</th>
<th>AAC</th>
<th>EGT</th>
<th>ECT</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-Squared</td>
<td>0.94</td>
<td>0.99</td>
<td>0.9</td>
<td>0.89</td>
<td>0.93</td>
<td>0.84</td>
<td>0.99</td>
<td>0.95</td>
<td>0.65</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.80</td>
<td>0.02</td>
<td>0.01</td>
<td>0.06</td>
<td>1.086</td>
<td>3.175</td>
<td>0.248</td>
<td>13.74</td>
<td>1.91</td>
</tr>
<tr>
<td>MSE</td>
<td>0.63</td>
<td>0.0003</td>
<td>0.0001</td>
<td>0.004</td>
<td>1.179</td>
<td>10.084</td>
<td>0.062</td>
<td>188.69</td>
<td>3.64</td>
</tr>
<tr>
<td>MAE</td>
<td>0.54</td>
<td>0.013</td>
<td>0.008</td>
<td>0.044</td>
<td>0.808</td>
<td>2.056</td>
<td>0.2</td>
<td>8.3</td>
<td>1.48</td>
</tr>
</tbody>
</table>

3.2. GA-ANN modelling results

GA-ANN model was used to estimate the energy and exergy analyses parameters values of the diesel engine at different fuel mixtures and different engine speeds (rpm) in a single-cylinder, water-cooled diesel engine. The data used in the experiments performed at twelve different engine speeds between 1000 and 3000 rpm with 200 rpm intervals for five different fuel mixtures consisting of fuel mixtures prepared by blends biodiesel, diesel and 5% bioethanol in different volumes constitute the input data of the models. Using these input data, engine torque (Nm), amount of fuel consumed depending on fuels and speed (g/h), CO emission values (%), CO₂ emission values (%), HC emission values (ppm), NO emission values (ppm), the amount of air consumed (kg/h), exhaust gas temperatures (ºC) and engine coolant temperatures (ºC) were estimated with the GA-ANN. GA-ANN model was depict in Figure 12.
The comparison of the results obtained with GA-ANN and the results obtained as a result of the experiments data for ET, AFC, CO, CO$_2$, HC, NO, AAC, EGT and ECT was shown in Figure 13-17.

**Figure 13.** Comparison of prediction and experimental results for ET and AFC

**Figure 14.** Comparison of prediction and experimental results for CO and CO$_2$

**Figure 15.** Comparison of prediction and experimental results for HC and NO

**Figure 16.** Comparison of prediction and experimental results for AAC and EGT

**Figure 17.** Comparison of prediction and experimental results for ECT

The performance of the GA-ANN model was calculated by Eqs.1-3 (R-Squared, MSE, and RMSE) and shown in Table 2.
Table 2. GA-ANN predicted performance values.

<table>
<thead>
<tr>
<th></th>
<th>ET</th>
<th>AFC</th>
<th>CO</th>
<th>CO₂</th>
<th>HC</th>
<th>NO</th>
<th>AAC</th>
<th>EGT</th>
<th>ECT</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-Squared</td>
<td>0.998</td>
<td>0.998</td>
<td>0.998</td>
<td>0.998</td>
<td>0.993</td>
<td>0.997</td>
<td>0.999</td>
<td>0.999</td>
<td>0.971</td>
</tr>
<tr>
<td>MSE</td>
<td>0.015</td>
<td>4.09 e-05</td>
<td>1.48E-06</td>
<td>7.76E-05</td>
<td>0.122</td>
<td>0.165</td>
<td>0.004</td>
<td>3.651</td>
<td>0.289</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.124</td>
<td>0.006</td>
<td>0.001</td>
<td>0.009</td>
<td>0.349</td>
<td>0.406</td>
<td>0.066</td>
<td>1.911</td>
<td>0.538</td>
</tr>
</tbody>
</table>

3.3. Comparative Analyses Based on R-Squared Measures of Each Prediction Model

In this section, the comparison of the success of the ANN and GA-ANN models’ performances is presented in a table 3.

Table 3. All Performances od R-Squared, RMSE, MSE and MAE values

<table>
<thead>
<tr>
<th>Method-Scale</th>
<th>ET</th>
<th>AFC</th>
<th>CO</th>
<th>CO₂</th>
<th>HC</th>
<th>NO</th>
<th>AAC</th>
<th>EGT</th>
<th>ECT</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN R-Squared</td>
<td>0.94</td>
<td>0.99</td>
<td>0.9</td>
<td>0.89</td>
<td>0.93</td>
<td>0.84</td>
<td>0.99</td>
<td>0.95</td>
<td>0.65</td>
</tr>
<tr>
<td>ANN RMSE</td>
<td>0.80</td>
<td>0.02</td>
<td>0.01</td>
<td>0.06</td>
<td>1.086</td>
<td>3.175</td>
<td>0.248</td>
<td>13.74</td>
<td>1.91</td>
</tr>
<tr>
<td>ANN MSE</td>
<td>0.63</td>
<td>0.0003</td>
<td>0.0001</td>
<td>0.004</td>
<td>1.179</td>
<td>10.084</td>
<td>0.062</td>
<td>188.69</td>
<td>3.64</td>
</tr>
<tr>
<td>ANN MAE</td>
<td>0.54</td>
<td>0.013</td>
<td>0.008</td>
<td>0.044</td>
<td>0.808</td>
<td>2.056</td>
<td>0.2</td>
<td>8.3</td>
<td>1.48</td>
</tr>
<tr>
<td>GA-ANN R-Squared</td>
<td>0.998</td>
<td>0.998</td>
<td>0.998</td>
<td>0.998</td>
<td>0.993</td>
<td>0.997</td>
<td>0.999</td>
<td>0.999</td>
<td>0.971</td>
</tr>
<tr>
<td>GA-ANN MSE</td>
<td>0.015</td>
<td>4.09 e-05</td>
<td>1.48E-06</td>
<td>7.76E-05</td>
<td>0.122</td>
<td>0.165</td>
<td>0.004</td>
<td>3.651</td>
<td>0.289</td>
</tr>
<tr>
<td>GA-ANN RMSE</td>
<td>0.124</td>
<td>0.006</td>
<td>0.001</td>
<td>0.009</td>
<td>0.349</td>
<td>0.406</td>
<td>0.066</td>
<td>1.911</td>
<td>0.538</td>
</tr>
</tbody>
</table>

Figure 18 shows the comparative R-Squared radar plot of the models ANN and GA-ANN.

4. Conclusion

In this study, ANN and GA-ANN were used to estimate ET, AFC, CO, CO₂, HC, NO, AAC, EGT, and ECT using ANN and GA-ANN. The model created with GA-ANN performed better than ANN. A 10-fold cross-validation was performed to determine the actual performance of the models and to rule out randomness. It can be seen that the result of the estimation with ANN was between 0.65 and 0.99, while the estimation performance of GA-ANN is between 0.971-0.999. It shows that the GA-ANN model can be successfully used to estimate energy and exergy parameters at different fuel mixtures and different engine speeds.

References

1204.


[46] Mirjalili, S., Genetic algorithm, in Evolutionary algorithms and