Prediction of Performance and Emissions Diesel Engines Fueled-Biodiesel Using Artificial Neural Network (ANN) Resilient Backpropagation Algorithm (Rprop)

Riva Amrulloh [a,1,*], Widayat [b,2], Budi Warsito [b,3]

[a] Diponegoro University, Jl. Imam Barjo, SH No.5, Semarang and 50275, Indonesia
[b] Diponegoro University, Jl. Imam Barjo, SH No.5, Semarang and 50275, Indonesia
[1] riva.amrulloh@gmail.com*; [2] widayat@lecturer.undip.ac.id; [3] budiwarsito@lecturer.undip.ac.id
* corresponding author

ARTICLE INFO

Article history
Received 03 Aug 2024
Revised 23 Sep 2023
Accepted 10 Okt 2024

Keywords
Prediction
Biodiesel
ANN
Rprop

ABSTRACT

In order to increase energy security and improve environmental quality, the Indonesian Goverment set a target of 23% renewable energy mix in 2025, one of which is the Mandatory Bioediesel Program. A higher biodiesel blending ratio will affect the performance and emissions of diesel engines because biodiesel is chemically different from diesel oil. Research related to the prediction of diesel engine performance and emissions using *Artificial Neural Network (ANN)* has been conducted, but the author sees a research opportunity for the implementation of the ANN *Resilient Backpropagation (Rprop)* algorithm. The data used to create the ANN model prediction was secondary data from previous research. The model designed multi input and multi output (MIMO) with 4 *input variables* and 7 *output variables*. Model building done by varying the number of neurons and *hidden layers*. Model evaluation selected based on the largest coefficient of determination parameter R^2 and the smallest RMSE or MAPE. The results showed that the ANN *single layer* 4-20-7 network architecture is the best model for predicting diesel engine performance and emissions with test data R^2 , RMSE and MAPE of 0.962532, 6.699428 and 6.0% respectively, while for overall data testing has a *performance of* 0.982869, 3.908542 and 4.3%. The results also show that based on the ANN prediction results, the increasing biodiesel ratio can increase NOx emissions and decrease HC, CO and CO emissions$_2$. In terms of performance, the addition of biodiesel can increase BSFC and BP and decrease BTE. The results also show that the addition of ZnO concentration can reduce emissions while in terms of performance it will increase BTE and reduce BSFC and BP.

1. Introduction

Data from the Ministry of Energy and Mineral Resources in 2021 states that over the past decade (2011 to 2021) petroleum production has decreased by 27% (88 million barrels), while the need for fuel, especially the transportation sector, is predicted to continue to increase with an average increase of 5% per year [5]. To cover the supply and demand deficit, the Indonesian

Goverment imports fuel. BPS data shows that during the period 2011-2022, the volume of national oil and gas imports has increased by 9.18% (4,013 thousand tons). The increase in imports has made national energy security vulnerable to fluctuations in world crude oil prices and supply. In addition to energy security issues, the use of fossil-based fuels raises environmental issues, namely vehicle exhaust emissions that can trigger the greenhouse gas effect which can cause global warming. In order to increase energy security and improve environmental quality, the Government of Indonesia set an energy mix target in 2025 of 23% for the use of New Renewable Energy (EBT). One of the government's policies to increase the use of renewable energy is the *Mandatory* Bioediesel Program. Starting in 2008 Indonesia has implemented Biodiesel blending of 2.5% and in 2015 through the support of the Minister of Energy and Mineral Resources Regulation Number 12 of 2015 concerning the Provision, Utilization and Commerce of *Biofuels* as Other Fuels, the minimum blending target is 30% and by the beginning of 2023 Biodiesel blending of 35% (B35) has been implemented. In accordance with the government program, the biodiesel blending ratio will continue to be increased depending on the results of technical, economic and environmental feasibility studies. A higher biodiesel blending ratio will affect diesel engine performance because biodiesel has a low calorific value, high viscosity and density compared to diesel oil [6]. On the other hand, the high oxygen and low sulfur content in biodiesel will cause more complete combustion resulting in better exhaust emissions than diesel oil [7]. Therefore, research is needed to see the performance and emission characteristics of diesel engines with higher biodiesel blend ratios.

There have been many studies related to the prediction of diesel engine performance and emissions. Research with experiments will spend a lot of time and money. One of the *tools* that can be used to predict the performance and exhaust emissions of diesel engines is the *Artificial Neural Network (ANN)*. Researchers have used this ANN technique for applications in all fields including automotive. Reference [20] modeled the performance and emissions of a 4-stroke diesel engine fueled by *waste cooking oil (wco)* biodiesel with a mixture of ZnO *nano particles* using the ELM algorithm where the ANN model was trained with experimental data. The results found that the model for performance with the number of neurons in the *hidden layer of* 15 neurons has the best performance with MAPE 7.16% while the model for emissions with the number of neurons 25 has the best performance with MAPE 13.1%. The results also found that the addition of *nano particles* can reduce emissions and change performance. Reference [4] modeled the composition of biodiesel to get the most optimal performance and exhaust emissions using the *support vector regression machine* (SVM) approach and found that SVM can be used to model diesel engine performance and emissions in addition to ANN, where the modeling results found that BSFC performance is higher than diesel while biodiesel emissions are lower than diesel except NOx and CO_2. Reference [14], examined the prediction and emissions of WCO biodiesel-fueled diesel engines using a *single layer backpropagation* ANN algorithm with 6 inputs and 1 output. The results showed that the model with 6-5-1 design has high accuracy with R = 0.9988 and MSE = 0.0397. Reference [25] conducted multi input and multi output ANN modeling to predict the performance characteristics, stability, emissions and ignition of diesel engines fueled by diesel, biodiesel and gasoline. The results obtained a coefficient of determination of 0.9804 - 0.9998. Reference [24] conducted a literature study on production, properties and combustion characteristics, performance and emissions using various types of biodiesel generation. The results obtained a 30% reduction in CO, 50% HC and 70% Smoke while in engine performance there was a decrease due to an increase in density of 3.1%, viscosity of 89.5% and cetane 11.9%.

From all the research that has been done, there are research *gaps*, namely the use of methods/algorithms and *software tools* used. In this study, the author will predict the performance and emissions of diesel engines using the ANN algorithm *Resilient Backpropagation (Rprop)* where the model built is 1 model, Multi Input - Multi Output (MIMO) with the help of Rstudio *software* while some previous studies mostly used 2 models, namely models for performance and models for emissions and using Matlab software. Rprop algorithm is known to have advantages over the *backpropagation* method because it is efficient and accurate and easy to implement. The data source in this study is secondary data from previous research [20] with the title *Performance and emission modeling of a 4-stroke diesel engine with biodiesel extracts from waste cooking oil blends with ZnO nanoparticle using ELM*.

The objectives of this study are: 1). Build model for predicting the performance and emissions of diesel engines and measuring the accuracy of the ANN Rprop algorithm model. 2). Investigate the effect of increasing Biodiesel composition and ZnO *nano particle* to biodiesel on diesel engine performance and emissions based on the ANN Rprop prediction model.

2. Method

2.1. Research Framework

The following is the research framework that the author conducted:

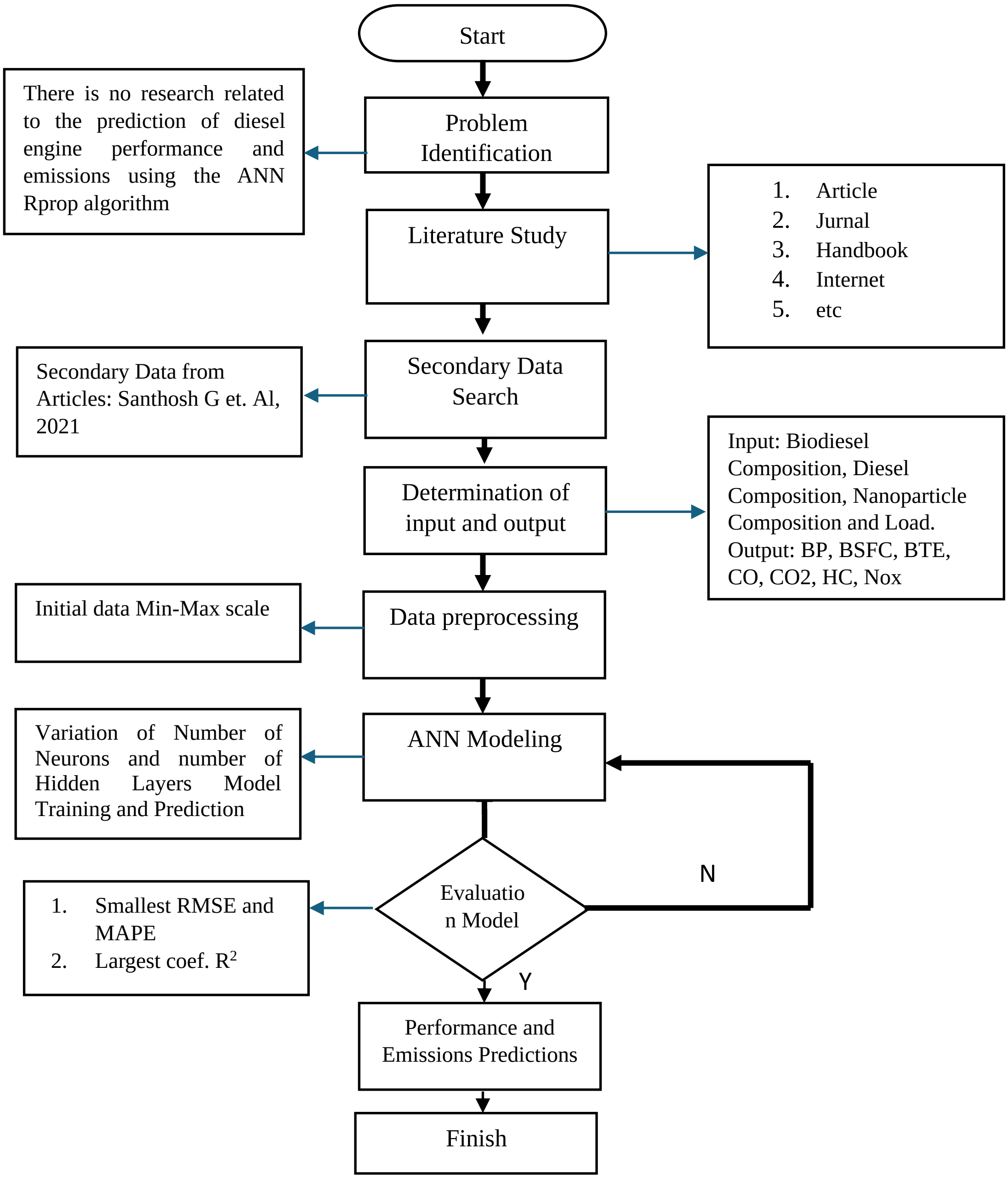


Fig. 1. Research framework

a. Problem Identification

The problem identified by the author is that there is no prediction of performance and emissions of biodiesel-fueled diesel engines based on the ANN Rprop algorithm that can be used as a reference for implementing policies to increase the biodiesel blending ratio.

b. Literature Study
Studying various literature Books, Scientific Articles and Journals that support research.

c. Secondary Data Search
Secondary data obtained from the International Journal Proceedings Category, Authors Santhosh G., Rashmi P. Shetty, Dileep Kumar M.J., Manasa G.R. with the title *Performance and emission modeling of a 4-stroke diesel engine with biodiesel extracts from waste cooking oil blends with ZnO nanoparticle using ELM*. The author's consideration for using secondary data from the Santosh et.al article is the amount of data as much as 65 data and a complete variety of input variables.

d. Determination of independent variables (input) and dependent variables (output)
At this stage the author determines the number of input variables and output variables to be modeled. The input variables consist of biodiesel composition, diesel/diesel composition, ZnO nanoparticle composition and load variation while the output variables consist of engine performance *(Brake Power/BP, Brake Thermal Efficiency/BTE and Brake Specific Fuel Consumption/BSFC)* and Emissions (NOx, CO, CO_2 and HC).

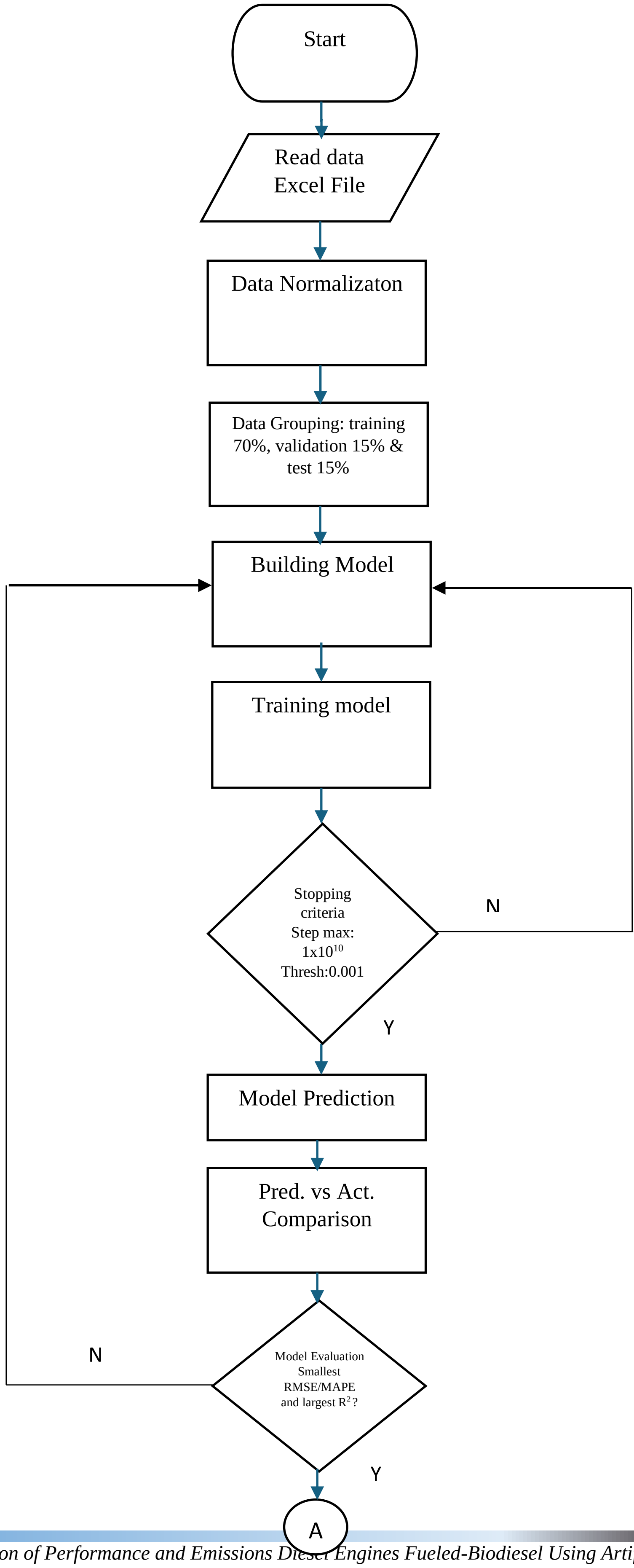
e. Data preprocessing
Furthermore, the author performs initial processing on the data, namely by normalizing / min-max scale so that the data to be processed has the same scale.

f. ANN program creation to model the normalized input and output data.
At this stage the author creates an artificial neural network model using Rstudio software. This software is an *opensource software* category so that it can be accessed by anyone. The author uses the ***neuralnet*** package that already exists in Rstudio. The following is the flow of the program creation process:

1. Reading Data in an Excel File
2. Data normalization
3. Random data grouping of 70% training data, 15% validation data and 15% test data.
4. Modeling with different number of neurons and hidden layers
5. Training model with training data
6. *Stopping criteria for* ANN training using step max=1e+10 and *threshold* 0.001, according to the features in Rstudio.
7. Model prediction with validation data
8. Comparison of predicted and actual data for each engine performance and emission variable
9. Model evaluation with the calculation of the coefficient of determination (R^2) and RMSE and MAPE
10. Repeat step 4 to find the largest coefficient of determination and the smallest RMSE and MAPE errors.
11. Creating a graph between prediction and actual in excel

g. Data retrieval
Data obtained from the results of variations in the number of neurons and hidden layers.

h. Model Evaluation
Model evaluation includes selecting the highest coefficient of determination and the lowest RMSE and MAPE errors, then on the basis of these criteria, the best model is selected to predict diesel engine performance and emissions


Start
Read data Excel File
Data Normalizaton
Data Grouping: training 70%, validation 15% & test 15%
Building Model
Training model
Stopping criteria Step max: 1x10^10 Thresh:0.001
N
Y
Model Prediction
Pred. vs Act. Comparison
Model Evaluation Smallest RMSE/MAPE and largest R^2?
N
Y
A

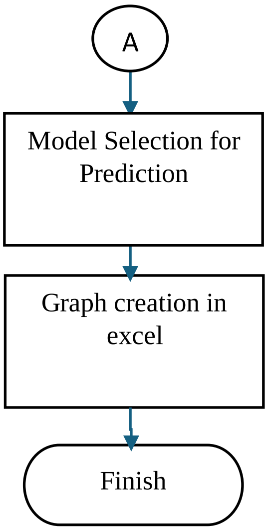


Fig 2. Programming Flowchart

i. Analysis

The best model was used to predict diesel engine performance and emissions with an independent data set of 40 and 50% biodiesel ratio variations. The model was also used to predict performance and emissions when ZnO was increased to 5 and 6 ppm.

j. Conclussion

In the last stage, the simulation results are concluded

2.2. Model Design

The model is designed with multiple inputs and multiple outputs (MIMO) with 4 input variables and 7 output variables:

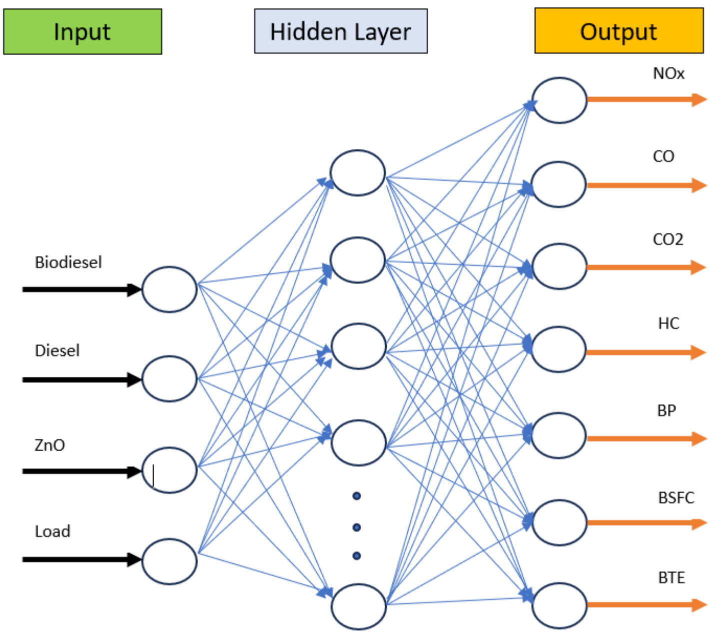


Fig. 3. MIMO Model Design

Input variables are Solar composition, Biodiesel composition, ZnO *Nanoparticle* composition and load variation, while *output* variables are diesel engine performance variables (*Break Power/BP*, *Break Specific Fuel Consumption/BSFC* and *Break Thermal Efficiency/BTE*) and emission variables (NOx, CO, CO_2 and HC).

2.3. Scope of Research

This research is limited by the following scope:

a. Simulation-based research using Rstudio software
b. Discussion of experiments in previous research is not discussed.
c. The biodiesel used is the *Waste Cooking Oil (WCO)* type.
d. The specifications of the diesel engine used are a single cylinder 4-stroke diesel engine with a compression ratio of 16.5
e. The variables studied were engine performance (BP, BSFC and BTE) and exhaust gas emissions (NOx, CO, CO_2 , HC).
f. The ANN algorithm used is *Resilient Backpropagation (Rprop)*

2.4. Data Type and Source

The data source is obtained from the International Proceeding Journal published by Elsevier Reference [20] with the title *"Performance and emission modeling of a 4-stroke diesel engine with biodiesel extracts from waste cooking oil blends with ZnO nanoparticle using ELM"*.

3. Results and Discussion

This research uses secondary data from previous research conducted by Reference [20]. Performance data and diesel engine emissions in the previous study were obtained by conducting experiments on a single cylinder 4-stroke diesel engine fueled by *Waste Cooking Oil (WCO)* with the addition of *ZnO nano-particles*. From this data, the author then created an ANN model of the *Resilient Backpropagation (Rprop)* algorithm with the help of *Rstudio* software to predict the performance and emissions of diesel engines. The following are the stages carried out by the author:

3.1. Data Preparation

3.1.1 Identification and preliminary analysis of sample data

In the first stage, the *input* variables and *output* variables were identified. The *input* variables consist of biodiesel composition, diesel composition, *nano-particle* composition and load variation while the *output* variables are engine performance parameters *(Break Power (BP)*, *Break Specific Fuel Consumption (BSFC)* and *Break Thermal Efficiency (BTE))* and emission parameters (NOx, CO, CO_2 and HC). Then to see the influence between parameters, a correlation analysis of input parameters and output parameters was conducted. Correlation analysis provides further information to understand how parameters affect each other and plays an important role in understanding the effects of several parameters [14]. The correlation calculation uses the **PEARSON** function in MS Excel. The following is the correlation matrix of the input parameters to the output parameters of the experimental results in the previous study:

Table 1. Correlation Matrix of Input Parameters to Output Parameters

Parameter	NOx	CO	CO_2	HC	BP	BSFC	BTE
Biodiesel	-0.37796	-0.56804	-0.82735	-0.57615	-0.77460	0.10924	-0.09585
Diesel	0.37796	0.56804	0.82735	0.57615	0.77460	-0.10924	0.09585
NP ZnO	-0.98480	-0.62302	-0.95163	-0.82285	0.71714	-0.83111	0.75994
Load	0.93068	-0.03604	0.86420	0.97054	1.00000	-0.82618	0.82881

Table 1 shows that the performance parameters (BP, BSFC and BTE) and emissions (NOx, CO, CO_2 and HC) are influenced by the input parameters of biodiesel composition, diesel composition, *nano particle* composition and *load* variation. Biodiesel composition has a negative correlation for NOx, CO, CO_2 , HC, BP and BTE while BSFC has a positive correlation. Diesel composition has the opposite correlation to biodiesel composition as both are blends totaling 100%. The *Nano Particle* ZnO parameter has a negative correlation for NOx, CO, CO_2 , HC and BSFC while BP and BTE have a positive correlation. The load variation parameter has a negative correlation to CO and BSFC while NOx, CO_2 , HC, BP and BTE have a positive correlation.

3.1.2 Data Grouping and Normalization

Furthermore, for the purposes of training the ANN model, a total of 65 data samples were randomly divided into 3 parts, namely 71% training data (46 data), 15% validation data (10 data) and 14% test data (9 data). Training data is used to train the model so as to get the appropriate weight and bias parameters so that the predicted value is close to the actual value, validation data is used to see the *performance of* the model as measured by 3 parameters, namely the coefficient of determination (R^2), RMSE and MAPE while test data is used to test the best ANN model from the results of testing using validation data.

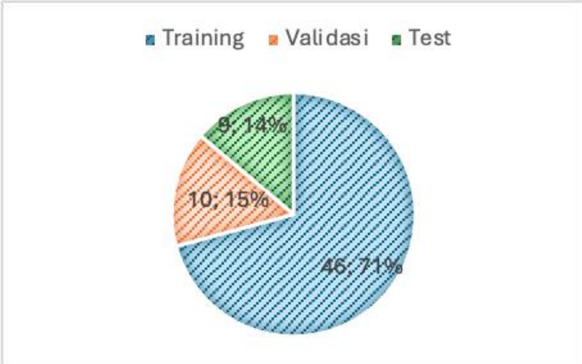


Fig 4. Distribution of Sample Data

The next step is to normalize the data to facilitate processing in the ANN network, this is done because the value / size of the sample data varies. The data normalization method used is to use a *min-max* scale where all data is converted to a scale of 0 (min) to 1 (max).

3.2. Model Building

After the initial processing of the data, the next stage is model building which can be illustrated in **Fig. 5** below:

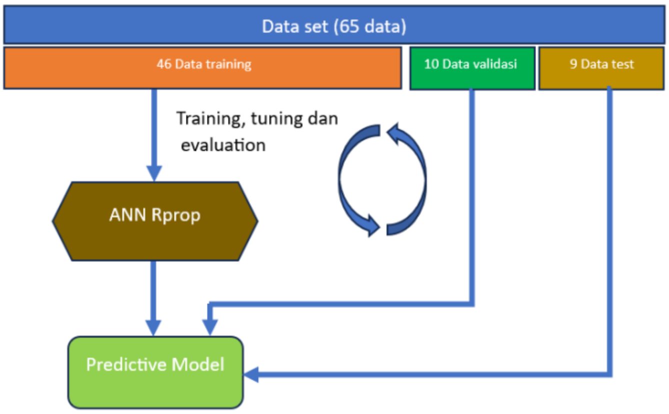


Fig. 5. Illustration of ANN model building

From **Fig. 5** above, it can be explained that the ANN model was trained with the *resilient backpropagation (Rprop)* algorithm using 46 training data and then validated using 10 validation data to measure *performance* R^2 , RMSE and MAPE. In order to get the best model based on these *performance* measures, the model architecture is *tweaked* by varying the number of neurons and the number of *hidden layers*. The determination of the number of neurons and *hidden layers* is based on a combination of *rules of thumb* [9], *trial-error* and previous research references. According to the *rules of thumb*, for 4 *inputs* and 7 *outputs*, the number of neurons can be determined as follows:

1. ***Rule 1***: The number of neurons in the *hidden layer* must be **between the size of the *input layer* and *the output layer***, so for 4 *input variables* and 7 *output variables*, the number of neurons must be between the interval {4,7};

2. ***Rule 2***: The number of neurons in the *hidden layer* should be **2/3 the size of the *input layer* plus the size of the *output layer***. This means the number of neurons in the *hidden layer* is =(2/3*4)+7 = 9.7 ~ 10 neurons;
3. ***Rule 3***: The number of neurons in the *hidden layer* **must be less than twice the size of the *input layer***. This means the number of *neurons* must be less than 8, so the number of neurons must be in the interval {1,8}.

Since {4,7} is a component of {1,8} and while 10 is outside the {1,8} component, if combined according to the *rules of thumb* above, the number of *neurons* is {1,8}. However, according to previous research by [20] that the number of neurons is in the range of 5 to 35 with a multiple of 5, the simulation of variations in the number of neurons and *hidden layers* is carried out to get the best model *performance*. Here are 18 topology architecture models that have been done in the research to get the best model.

Table 2. Variations of ANN Model Architecture

Model	Model Architecture	Description
Model 1	4-4-7	*Single Layer*
Model 2	4-5-7	*Single Layer*
Model 3	4-6-7	*Single Layer*
Model 4	4-7-7	*Single Layer*
Model 5	4-8-7	*Single Layer*
Model 6	4-9-7	*Single Layer*
Model 7	4-10-7	*Single Layer*
Model 8	4-15-7	*Single Layer*
Model 9	4-20-7	*Single Layer*
Model 10	4-25-7	*Single Layer*
Model 11	4-30-7	*Single Layer*
Model 12	4-35-7	*Single Layer*
Model 13	4-40-7	*Single Layer*
Model 14	4-20-10-7	*Multi Layer*
Model 15	4-20-15-7	*Multi Layer*
Model 16	4-25-10-7	*Multi Layer*
Model 17	4-25-15-7	*Multi Layer*
Model 18	4-20-10-8-7	*Multi Layer*

Furthermore, from the model architecture design above, each model is simulated in Rstudio software which already supports ANN programming with the Rprop algorithm using the **neuralnet** library**.** The following is an example of the ANN Rprop programming *script* in Rstudio software:

```
library(neuralnet)
annpa20a <- neuralnet(f,trainingpa20a,hidden=c(20), threshold=0.001, stepmax=1e+10, algorithm = "rprop
                      act.fct = "logistic", err.fct= "sse", linear.output=T)
```

The *script* above can be explained as follows:

The **annpa20a** variable is the ANN function of the Rprop algorithm with a *single layer* architecture of 20 *neurons* trained with **trainingpa20a training** data, the stop criteria according to the default in Rstudio using *threshold* = 0.001 and *stepmax* 1×10^{10} , the activation function uses a *logistic* function and *the error function* uses a *sum square error (sse)*, with a linear output value between 0 and 1 according to the min-max scale. The programming *script* was run for each model architecture and then recorded the R performance^2 , RMSE and MAPE for each model.

3.3. Model Selection

The best model selection is based on the highest R^2 *performance* measure and the lowest RMSE or MAPE. Simulation results using Rstudio software from 18 ANN model architectures in **Table 2** above tested with **validation data** obtained R^2 , RMSE and MAPE for each model can be depicted with a graph in **Fig. 6** as follows:

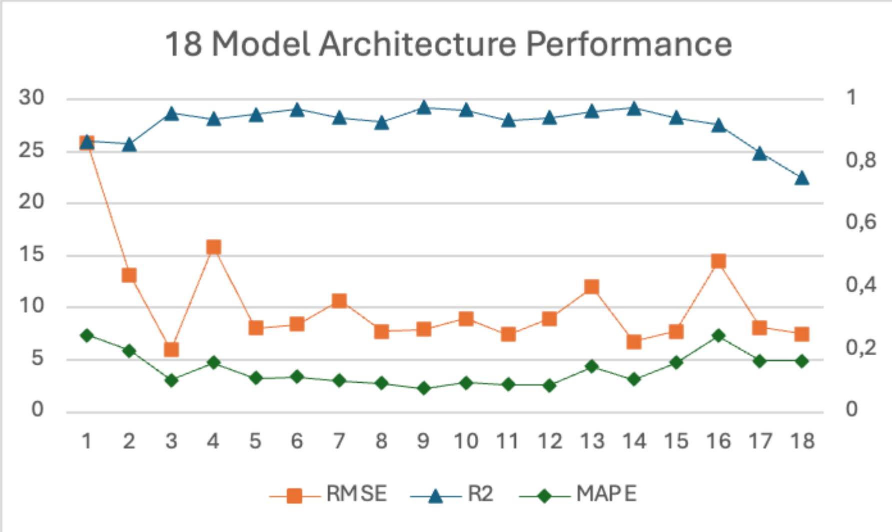


Fig 6. Performance 18 architecture model

From **Fig. 6** above, it can be seen that the 9th model with 4-20-7 architecture has the highest average performance R^2 of 0.975715 and the smallest MAPE of 7.5%. Based on the model accuracy criteria where R^2 >0.9 and MAPE <10%, it can be said that the model with 4-20-7 architecture has high accuracy. While the 18th model with 4-20-10-8-7 multilayer architecture has the lowest R^2 compared to others, namely 0.748592 and MAPE 16.4%, this is because the model is too complex so that it experiences *overfitting*. Conversely, the 1st Model with 4-4-7 architecture has the highest R^2 = 0.866097 and MAPE value of 24.5%, indicating that the model is too simple to represent the relationship between input and output variables so that one way to improve accuracy is to increase the number of neurons in the *hidden layer*. The following are the performance details of each performance and emission *output variable* for models 1, 9 and 18:

Table 3. Performance comparison of Models 1, 9 and 18

Model Architecture	Parameter	R^2 Validasi	RMSE	MAPE
4-4-7	NOx	0.937856	173.0481	100,6%
	CO	0.799382	0.004611	8.6%
	CO_2	0.580602	1.220368	32.2%
	HC	0.951254	3.392015	8.5%
	BP	0.999524	0.014142	4.0%
	BSFC	0.939209	0.033490	7.2%
	BTE	0.854854	2.936973	10.5%
	Average	**0.866098**	**25.80710**	**24.5%**
4-20-7	NOx	0.995349	52.06671	11.9%
	CO	0.899669	0.003099	7.8%
	CO_2	0.995923	0.203664	7.2%
	HC	0.994229	1.408974	4.3%
	BP	0.998991	0.022209	9.1%
	BSFC	0.975724	0.026141	6.6%
	BTE	0.970119	1.602816	5.7%
	Average	**0.975715**	**7.904802**	**7.5%**
4-20-10-8-7	NOx	0.996099	41.97189	13.9%
	CO	0.052123	0.010664	21.0%
	CO_2	0.606595	1.153220	17.7%

HC	0.929771	4.796613	17.2%
BP	0.993286	0.051592	14.6%
BSFC	0.883474	0.058622	17.3%
BTE	0.778797	4.462797	13.6%
Average	**0.748592**	**7.500771**	**16.5%**

Here's a plot of the 4-20-7 model architecture from the Rstudio simulation:

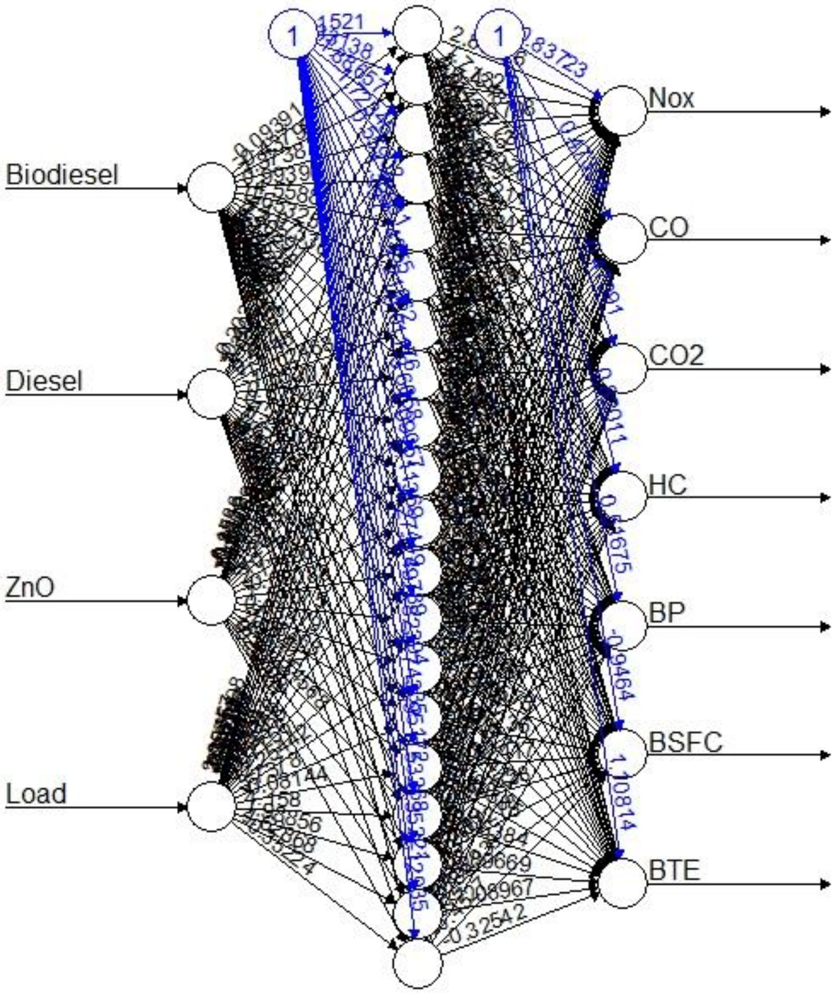


Fig. 7. Plot of 4-20-7 ANN Architecture with Rstudio

Furthermore, the 4-20-7 model was tested using test data and obtained R^2 , RMSE and MAPE as follows:

Table 4. Model *performance resume* with test data

Parameter	Sub Parameter	R^2 Test	RMSE	MAPE
Emissions	NOx	0.9992318	42.93610	8.95%
	CO	0.8619296	0.003565	6.65%
	CO_2	0.9929949	0.218514	6.74%
	HC	0.9913486	2.128333	4.08%
Performance	BP	0.9990108	0.023558	7.37%
	BSFC	0.9786160	0.014117	3.65%
	BTE	0.9145899	1.571808	4.68%

Average	**0.96253166**	**6.699428**	**6.02%**

It can be seen that the 4-20-7 model when used for predicting emissions and engine performance using **test data** obtained an average R^2 , RMSE and MAPE of **0.962532, 6.699428 and 6.0%,** respectively. This value proves that the model can be categorized as having high accuracy because R^2 >0.9 and MAPE <10%.

Furthermore, the 4-20-7 model is also used to predict the entire data set (65 data), this is done to see the accuracy of the model against the actual value of the experimental results in previous studies. The results showed that the model has an average performance of R^2 0.9828690, RMSE 3.908542 and MAPE 4.3%. The value of R^2 >0.9 and MAPE <10% illustrates that the model has high accuracy.

The following is a *resume of the* 4-20-7 Model *performance* for training, validation, test and overall data (All):

Table 5. *Performance resume of* 4-20-7 model

Parameter	Sub Parameter	R^2 Train	R^2 Val	R^2 Test	R^2 All	RMSE	MAPE
Emission	NOx	0.9995565	0.9977137	0.9992318	0.9991914	24.69549	5.4%
	CO	0.9888155	0.9111307	0.8619296	0.9427348	0.002421	3.8%
	CO2	0.9960042	0.9906891	0.9929949	0.9939452	0.142279	3.7%
	HC	0.9979193	0.9910592	0.9913486	0.9966477	1.218990	2.7%
Performance	BP	0.9997595	0.9968262	0.9990108	0.9995834	0.014631	7.2%
	BSFC	0.9895284	0.9646467	0.9786160	0.9826230	0.016810	3.7%
	BTE	0.9719732	0.9612991	0.9145899	0.9653593	1.269175	3.6%
Average		**0.9919370**	**0.9733380**	**0.9625316**	**0.9828690**	**3.908542**	**4.3%**

From **Table 5** it can be seen that between the average R^2 Train, R^2 Validation, R^2 Test and R^2 All > 0.9 and MAPE <10% this indicates that the model has high accuracy to predict all performance parameters and diesel engine emissions. From the table, it can also be seen that for all performance and emission parameters, the value of R^2 All >0.9 and MAPE <10% so that in this case the model can be used to predict the performance and emissions of diesel engines.

When compared with previous research by Reference [20], it can be summarized according to **Table 6** below:

Table 6. Comparison of research results with previous research

Parameter	Santosh et.al	Research Result
Algorithm	*Extreme Learning Machine (ELM)*	*Resilient Backpropagation (Rprop)*
Number of Model	2 models (1 performance model and 1 emission model)	1 model used for performance and emissions
Model Architecture	Performance: 4-15-4 Emissions: 4-25-4	4-20-7
MAPE	Performance: 7% Emissions: 13%	4.3%

From **Table 6** above, it can be seen that the *resilient backpropagation* algorithm with 1 multi input and multi output model has better MAPE performance than the ELM algorithm with 2 models.

3.4. Performance and Emissions Predictions

3.4.1 Prediction of performance and emissions due to the effect of increasing biodiesel composition

The model with 4-20-7 architecture is then used to predict the performance and emissions of diesel engines with simulation data outside the data set (extrapolation) by varying the biodiesel variable while the other input variables are constant. In this study, the model is only used to predict the performance and emissions of diesel engines with 40% (B40) and 50% (B50) biodiesel composition. The data selection considers that the predicted data is not too far from the range of data sets that have been used in network training where the maximum data set is 30% (B30) so that the predicted value is expected to be accurate. The following is a graph of the model prediction results:

A. Emissions

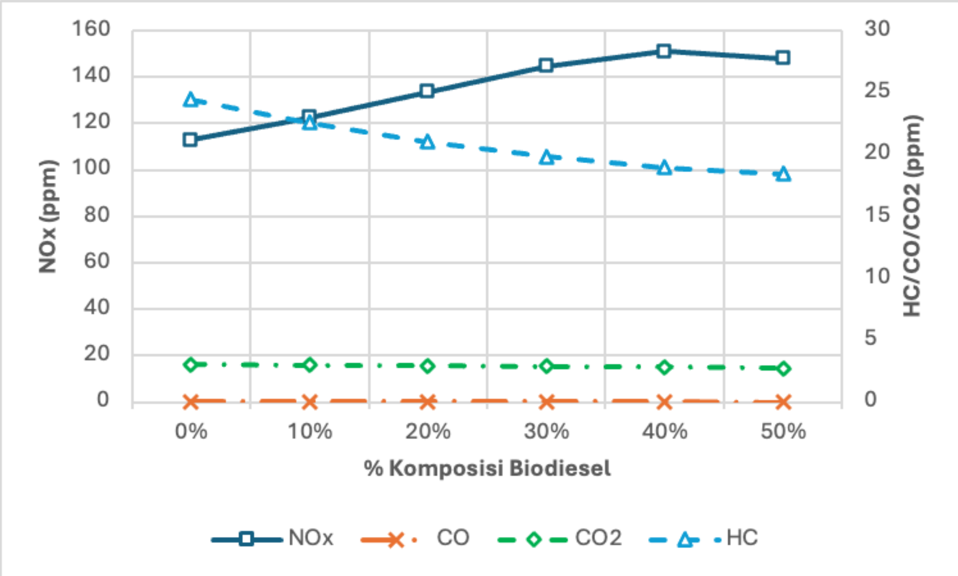


Fig. 8. Predicted Emissions due to changes in Biodiesel composition

From **Fig. 8**, it can be seen that based on the prediction results of the ANN model, adding biodiesel from 0 to 50 will increase NOx emissions by 31%. This is because the increase in oxygen content in biodiesel causes more complete combustion which increases the combustion temperature in the cylinder which can increase NOx emissions. When compared to **Table 1** where NOx has a negative correlation with Biodiesel, this is due to the lack of experimental data in the previous study (only 3 data) to represent the relationship between biodiesel and NOx. The HC and CO emissions based on the model predictions will decrease by 24.4% and 20.7% respectively. This is because the more complete combustion of biodiesel will result in smaller HC and CO emissions compared to pure diesel. **Table 1** also shows that biodiesel has a negative correlation with HC and CO where the greater the biodiesel, the smaller the HC and CO emissions. This decrease in HC and CO is also in accordance with research conducted by Reference [17] that increasing biodiesel composition will increase NOx and decrease CO and HC. Furthermore, CO emissions$_2$ based on predictions decreased by 9.7%. This is not in accordance with some literature which states that increasing biodiesel will increase CO_2 emissions because increasing oxygen in biodiesel will result in complete combustion which produces more CO_2 . Although some literature states that CO/CO emissions$_2$ are highly dependent on combustion conditions, combustion temperature and air-fuel ratio. This discrepancy could be due to the fact that the data inputted into the ANN network is outside the predefined data set (max. B30 data set), so the network does not have enough information to predict data outside the data set.

B. Performance

From **Fig. 9**, it can be seen that the addition of biodiesel from B0 to B50 will reduce BTE performance by 46.7%. This is because biodiesel has a greater viscosity than diesel oil, causing obstacles in the process of spraying fuel into the combustion chamber, besides that biodiesel also has a lower calorific value than diesel so that its thermal efficiency will decrease. This negative correlation is also consistent with what is shown in **Table 1**. In contrast to BTE, BSFC performance will increase by 23.1% due to the low calorific value and high viscosity of biodiesel, which requires more fuel to produce the same power. This positive correlation is also consistent with what is shown in **Table 1**. This prediction is also in accordance with the research conducted by Reference [17] where increasing Biodiesel will reduce BTE and increase BSFC. Furthermore, for BP performance, based on the prediction results, an increase in biodiesel composition will increase BP by 21.4%. This is not in accordance with research conducted by Reference [17] where biodiesel has a lower calorific value than diesel so that it can reduce BP performance. This negative correlation is also consistent with what is shown in **Table 1**. The discrepancy between the ANN model predictions and the literature could be because the data entered in the ANN network is outside the predetermined data set (max. B30 data set) so that the network does not have enough information to predict data outside the data set.

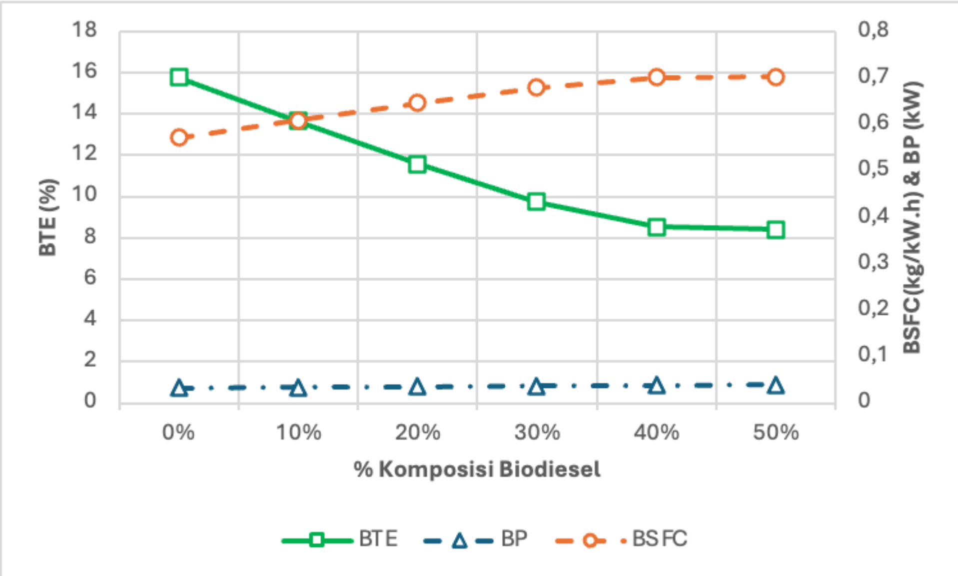


Fig. 9. Predicted performance due to changes in Biodiesel composition

3.4.2 Performance and emission prediction due to the effect of ZnO nano particle increase

This research also aims to see the effect of adding ZnO *nano particles* to biodiesel on the performance and emissions of diesel engines based on ANN model predictions. The ZnO simulation data used is outside the data set (extrapolation), namely the ZnO concentration at 5 and 6 ppm. The data selection considers that the predicted data is not too far from the range of data sets that have been used in network training, namely 4 ppm so that the predicted value is expected to be accurate. The following is a graph of the model prediction results:

A. Emissions

From **Fig. 10** it can be seen that the addition of ZnO concentration can reduce NOx, CO, CO_2, HC emissions by 11.3%, 47.7%, 19.6% and 17.4% respectively. This is because the ZnO *nano particle* can act as a catalyst that can improve fuel quality and improve the combustion process so as to reduce diesel engine combustion emissions. This prediction result is in accordance with research [10] where the addition of ZnO can reduce emissions because ZnO can increase oxidation stability, fuel viscosity and cetan number so as to improve combustion quality. This negative correlation is also shown in **Table 1** where the larger the ZnO *nano particle*, the smaller the emissions will be.

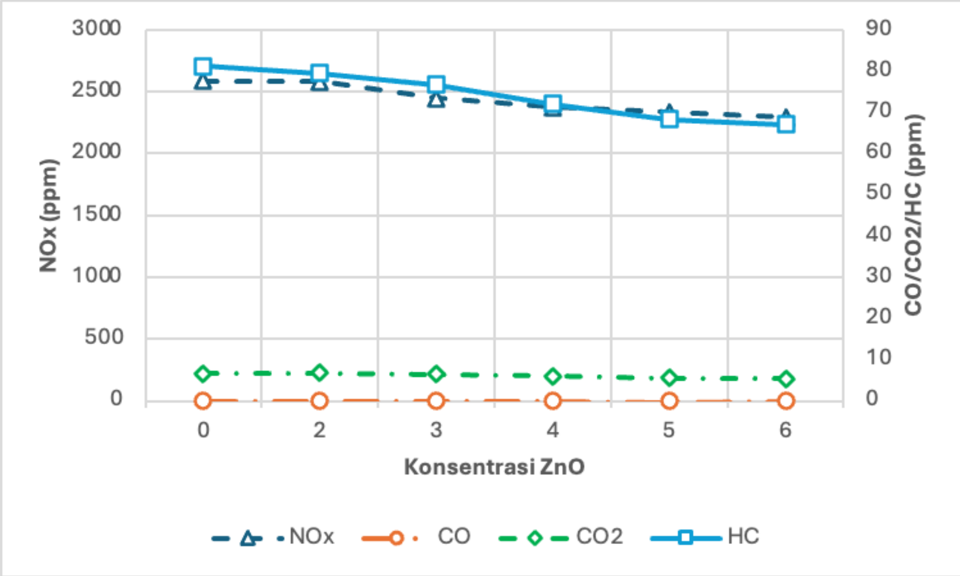


Fig. 10. Predicted emissions due to changes in ZnO concentration

B. Performance

From **Fig. 11**, it can be seen that increasing the concentration of ZnO can increase BTE, according to research conducted by Reference [12] where the addition of ZnO can improve engine performance and combustion quality because ZnO can improve fuel quality such as increasing calorific value, *flash point, cetane number* and reducing viscosity. This positive correlation is also shown in **Table 1**, the greater the concentration of nano particle ZnO, the BTE performance will increase. This increase in BTE is in contrast to BSFC, where the greater the concentration of ZnO, the BSFC will decrease, indicating that fuel consumption will be more efficient because ZnO can improve fuel quality and combustion quality. This negative correlation is also consistent with what is shown in **Table 1**. Furthermore, for BP, based on the prediction results, it will reduce BP performance by 0.5%, this is not in accordance with the research conducted by Reference [12]. This BP prediction is also not in line with the increase in BTE where if BTE increases then BP also increases. **Table 1** also shows a positive correlation between ZnO *nano particle* and BP. This difference in prediction could be due to the data entered in the ANN network is outside the predetermined data set (data set max. ZnO 4 ppm) so that the network has not enough information to predict data outside the data set.

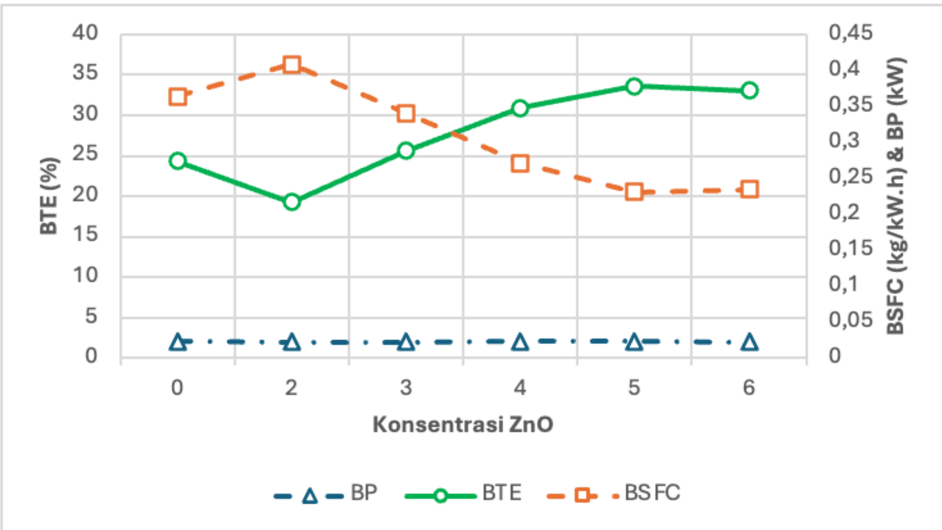


Fig. 11. Predicted performance due to changes in ZnO concentration

4. Conclusion

From the results and discussion above, the following conclusions can be drawn:

a. The best ANN model architecture with the Rprop algorithm to predict the performance and emissions of biodiesel-fueled diesel engines is a *single layer* ANN model with the number of neurons in the *hidden layer of* 20 so that the architecture is 4-20-7. The accuracy level of the 4-20-7 ANN model with overall test data has a performance of R^2 , RMSE and MAPE are NOx 0.999191, 24.69549 and 5.3%, CO 0.9427348, 0.002421 and 3.8%, CO_2 0.9939452, 0.142279 and 3.7%, HC 0.9966477, 1.21899 and 2.7%, BP 0.99958, 0.014631 and 7.2%, BSFC 0.982623, 0.0168101 and 3.7%, BTE 0.9653593, 1.269175 and 3.6%. According to the criteria if R^2 >0.9 and MAPE <10%, the ANN model has high accuracy.

b. Increasing the biodiesel composition up to B50 based on the ANN model prediction results will cause CO, CO_2 and HC emissions to decrease by 24.4%, 9.7% and 20.7%, respectively. The decrease in CO_2 is not in accordance with the literature which states that biodiesel can increase more complete combustion so that more CO_2 will be produced. The NOx emission will increase by 31%. In terms of engine performance parameters, the addition of biodiesel will increase BSFC by 23.1% and decrease BTE by 46.7%. Meanwhile, the BP performance increases by 21.4%, which is not in accordance with the literature where increasing biodiesel will decrease BP performance because biodiesel has a lower calorific value than diesel oil. The discrepancy between the BP and CO_2 predictions and the literature could be because the data entered in the ANN network is outside the predetermined data set so that the network does not have enough information to predict data outside the data set.

c. The addition of *nano particle* ZnO concentration up to 6 ppm based on the prediction results will reduce NOx, CO, CO_2 and HC emissions by 11.3%, 47.7%, 19.6% and 17.4% respectively. This is because *nano particle* ZnO can act as a catalyst that can improve fuel quality and improve combustion quality so as to reduce diesel engine combustion emissions. Furthermore, for the performance prediction results, the addition of ZnO can improve BTE performance by 35.7% and reduce BSFC by 35.5%. This is in accordance with previous research that ZnO can improve fuel quality such as increasing calorific value, *flash point, cetane number* and reducing viscosity so as to improve diesel engine performance. For the BP parameter, the ANN model prediction results show a decrease of 0.5%. This is inconsistent with some literature that increasing ZnO concentration will improve BP performance due to improved fuel quality and combustion quality. This discrepancy could be because the data entered in the ANN network is outside the predetermined data set so that the network does not have enough information to predict data outside the data set.

Acknowledgment

The author would like to thank the supervisor who has helped the author in completing this research. The author would also like to thank to the institution of where the author work for providing opportunities for its employees to continue their Masters in Energy at Diponegoro University Semarang.

Declarations

Author contribution. Riva Amrulloh: Writing-original draft, Software, Methodology, Investigation and Analysis. Widayat: Review, editing and supervision. Budi Warsito: Review, editing and Supervision
Funding statement. N/A
Conflict of interest. The authors declare that the research has no conflict of interest.

Data and Software Availability Statements

Data and Software will be made based on request.

References

[1] Abed, Morsi, Sayed, Shoib and Gad, et.al, “Effect of waste cooking-oil biodiesel on performance and exhaust emissions of a diesel engine”, Egypt. J. Pet. 27 (4) 2018 pp 985–989

[2] Anh, Nizetic, Ong, and Tarelko, “A review on application of artificial neural network (ANN) for performance and emission characteristics of diesel engine fueled with biodiesel-based fuels”. Elsevier. Sustainable Energy Technologies and Assessments 47 2021 101416. https://doi.org/10.1016/j.seta.2021.101416

[3] Addison, Miligan and Ohlig, “Comparative analysis of ZnO and CeO2 nano particle additive for the improvement of hemp biofuels as a sustainable biodiesel alternative: A comprehensive review”. Elsevier: Industrial Crops and Products, 2024.

[4] Bukkarapu and Khrisnasamy, “Support vector regression approach to optimize the biodiesel composition for improved engine performance and lower exhaust emissions”. Elsevier. Fuel 348, 2023. https://doi.org/10.1016/j.fuel.2023.128604.

[5] Dewan Energi Nasional, “ Outlook Energi Indonesia”. Jakarta: Sekertariat DEN, 2019.

[6] Fareed, Shafay, Mujtaba, Riaz and Gad, “Investigation of waste cooking and castor biodiesel blends effects on diesel engine performance, emissions, and combustion characteristics”. Elsevier. Case Studies in Thermal Engineering 60 (2024) 104721. https://doi.org/10.1016/j.csite.2024.104721

[7] Gad, Seesy, Radwan and He, “Enhancing the combustion and emission parameters of a diesel engine fueled by waste cooking oil biodiesel and gasoline additives”, Fuel 269 (2020) 117466, https://doi.org/10.1016/j.fuel.2020.117466

[8] Hassan, Rahman and Nabi. et.al, “Opportunities and challenges for the application of biodiesel as automotive fuel in the 21st century”. *Biofuels Bioprod. Biorefin,* Volume 16, 2022, p. 1353–1387

[9] Heaton, J. ”Introduction to Neural Network for C# 2nd Edition”. St. Louis: Heaton Research. Inc, 2008

[10] Lobo and Manjunath. “Influence of Zinc Oxide on Direct Injection CI engine Fuel with waste cooking oil biodiesel”. Int. Adv. Res. J.sci. eng. Technol. 6, 2019, pp 32-40

[11] Malfatti, Zampar, Galvao, Rabazza and Boiago, “Evaluating and predicting egg quality indicators through principal component analysis and artificial neural networks”. LWT, 148, Article 111720. https://doi.org/ 10.1016/j.lwt. 2021.111720

[12] Modi, Rampure, Babbar, Kumar et al, “Nano particle enhance biodiesel blends: A comprehensive review on improving engine performance and emissions”. KeAI: Material Science for Energy Technology, 2024

[13] Patalay, S. dan Bandlamudi, M.R, “Gold Price Prediction Using Machine Learning Model Trees. *Proceeding International Conference on Changing Business Paradigm (ICCBP)”*, Murshidabad, 2021

[14] Patnaik, Khatri and Rhene, ”Artificial neural networks-based performance and emission characteristics prediction of compression ignition engines powered by blends of biodiesel derived from waste cooking oil. Elsevier. Fuel 370, (2024), 131806. https://doi.org/10.1016/j.fuel.2024.131806

[15] Rawat, “Zomato Partners With BioDiesel To Produce Biodiesel From Used Cooking Oil” https://inc42.com/buzz/zomato-partners-with-biod-to-producebiodiesel-from-used-cooking-oil/, 21/10/2019.

[16] Rekhate, Prajapati, ”Production, engine performance, combustion, emission characteristics and economic feasibility of biodiesel from waste cooking oil: A review”, Environ Qual Manage. 29 (1) (2019) 7–35

[17] Roopesh and Goyal, ”A review: Effect on performance and emission characteristics of waste cooking oil Biodiesel- diesel blends on IC engine”. Elsevier. Materials Today: Proceedings 63 (2022) 643–646. https://doi.org/10.1016/j.matpr. 2022.04.447

[18] Singh, Sharma and Jhalani, “A review on feedstocks, production processes, and yield for different generations of biodiesel”, Fuel 262 (2020) 116553, https://doi.org/10.1016/j.fuel.2019.116553

[19] Sakthivel, Ramesh, Shameer and Purnachandran, “A review on the properties, performance and emission aspects of the third generation biodiesels”, Renew. Sustain. Energy Rev. 82 (3) (2018) 2970–2992, https://doi.org/10.1016/j. rser.2017.10.037.

[20] Santhosh, Shetty, Kumar and Manasa, “Performance and emission modeling of a 4-stroke diesel engine with biodiesel extracts from waste cooking oil blends with ZnO nanoparticle using ELM”. Proceedings 52 (2022) 1935–1941. https://doi.org/10.1016/j.matpr.2021.11.581

[21] Vhisnu, Praveen and John, “Performance and Emission Characteristics of a Diesel Engine Running on Biodiesel”. Ijirset. *Vol.4, Special Issue 12, 2015*

[22] Selva Rex, Edwin, Shiny and Annrose, “Power performance and emissions analysis of outboard diesel engines by use of waste cooking oil biodiesel”, Elsevier. https://doi.org/10.1016/j.matpr.2023.09.005

[23] Yahya, Wahab and Harun, “Optimization of biodiesel production from waste cooking oil using Fe-Montmorillonite K10 by response surface methodology”, Renewable Energy 157 (2020) 164–172
[24] Zhang, Zhong, Lu, “A Comprehensive Review of the Properties, Performance, Combustion, and Emissions of the Diesel Engine Fueled with Different Generations of Biodiesel”. MDPI: Processes, 2022
[25] Zandie, Kiat, Gan, Said and Cheng, “Multi input multi output machine learning predictive model for engine performance and stability, emissions, combustion and ignition characteristics of diesel-biodiesel-gasoline blends”. Elsevier: Energy, 2023