Energy Proceedings
Vol 31, 2023
ISSN 2004-2965

Application of Machine Learning Methods to Predict NOx Emissions for Next Generation Zero-Carbon Ammonia/Hydrogen Fueled Engines

Ruomiao Yang[1], Xiao Zheng[1], Zhentao Liu [1*], Yu Zhang[2], Jiahong Fu[2]

1 Power Machinery and Vehicular Engineering Institute, College of Energy Engineering, Zhejiang University, Hangzhou 310027, PR China;

2 Mechanical Engineering Department, Zhejiang University City College, Hangzhou 310015, PR China;

ABSTRACT

After industrialization and informatization, the world economy is moving toward "decarbonization". As two carbon-free fuels, hydrogen and ammonia are attracting increasingly widespread interest around the world. And ammonia and hydrogen are promising and practical alternative energy sources for internal combustion (IC) engines, which may be used to power the next generation of engines due to their zero-carbon footprint. However, the combustion of ammonia produces large amounts of nitrogen oxides (NOx), which pollute the environment. Therefore, it is very significant to control nitrogen oxides (NOx) emissions from spark-ignition (SI) engines. Machine learning (ML) approaches are an alternative analytical tool to three-dimensional (3D) simulations, in-depth experiments and empirical phenomenon models that can accelerate the development of IC engines. The objective of this study was to assess the applicability of ML models in predicting NOx emissions. A calibrated spark-ignition engine fueled with gasoline operating under different conditions was used to provide sufficient data for model training, validation, and testing. The results indicated that the artificial neural network (ANN) and support vector regression (SVR) have good prediction performance and high accuracy, and the prediction accuracy of the RF model is acceptable. In general, ANN and SVR have comparable performance and both models are recommended to predict NOx emissions from ammonia/hydrogen fueled engines. And the prediction performance of the RF model will be less accurate compared to the other two ML models.

Keywords: machine learning, NOx, ammonia/hydrogen, artificial neural network, support vector regression, random forest.

NONMENCLATURE

Abbreviations

1D	One dimensional
3D	Three dimensional
ANN	Artificial neural network
ATDC	After top dead center
CA	Crank angle
CAD	Crank angle degree
CFD	Computational fluid dynamics
CR	Compression ratio
IC	Internal combustion
MBT	Minimum advance for best torque
ML	Machine learning
NOx	Nitrogen oxides
R^2	Coefficient of determination
RF	Random forest
RMSE	Root mean squared error
RPM	Revolutions per minute
SI	Spark ignition
ST	Spark timing
SVR	Support vector regression

1. INTRODUCTION

With the massive development and use of fossil energy, it has brought three serious challenges: energy crisis, environmental pollution and climate change. And carbon dioxide from fossil energy accounts for 57% of global greenhouse gas emissions and is the main cause of climate change. Therefore, many countries have invested a lot of effort in studying the low-carbon transition of their energy system, which is an important condition for achieving the goals of carbon peaking and

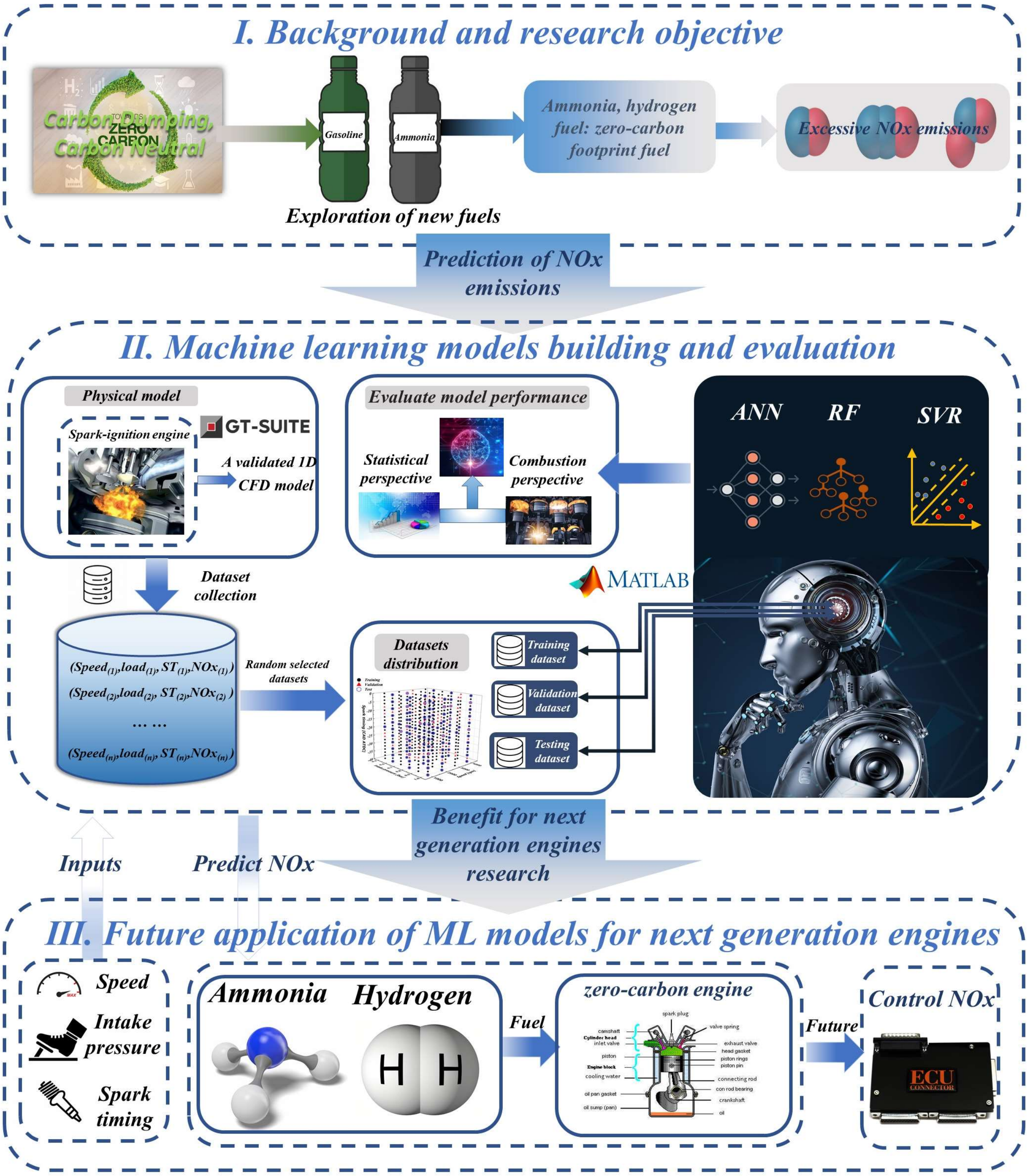

Fig. 1. Application of ML models to predict NOx processes and future application scenarios

carbon neutrality [1]. Many researchers are working on finding suitable alternative fuels such as natural gas, methanol, biodiesel, etc. Since decarbonization is currently the most important strategic goal, carbon-free fuels are receiving particular attention. Both ammonia and hydrogen do not produce carbon dioxide during combustion, which meets the requirements as a future energy source: no acceleration of global warming [2]. The processes of manufacturing, storing, and transporting ammonia are also quite simple. Moreover, as countries increase their restrictions on greenhouse gas emissions, ammonia/hydrogen, a fuel without a carbon base, will emerge as an advantage [3]. However, the ammonia/hydrogen engine has one major disadvantage: it produces a large amount of NOx during combustion, which is very harmful to the environment.

Therefore, in the context of decarbonization, it is of great interest to control NOx emissions from zero-carbon ammonia/hydrogen spark-ignition engines.

Most of the nitrogen oxides (NOx) emissions from IC engines are products of fuel combustion in the cylinder, such as ammonia, natural gas, gasoline, etc. The amount of NOx emissions generated is mainly determined by the design parameters and operating conditions of the engine. Currently, data-driven approaches have shown great potential in predicting some key parameters and learning the internal laws of physical models in various fields, such as energy, architecture, medicine, etc. The machine learning algorithm can be utilized as a surrogate model to replace the multidimensional simulations, bench tests and physical models with sufficient high-quality data. As a result, it can assist in engine design and shorten the development time for the next generation of engine updates and iterations [4]. The existing literature suggests that machine learning methods can be used to predict engine-out emissions with acceptable ability. However, only a limited amount of literature has investigated the effectiveness of different ML models for NOx emission prediction from spark ignition engines. And both gasoline and ammonia/hydrogen engines will produce NOx emissions. Therefore, this study evaluates the performance of three ML models (ANN, SVR, and RF) in predicting NOx from SI engines fueled with gasoline from both statistical and combustion perspectives, respectively, which will provide a suitable choice of machine learning models for the evaluation of NOx emissions from next-generation zero-carbon ammonia/hydrogen engines. This is because the ML methods used to predict NOx emissions are the same for gasoline and ammonia/hydrogen engines. Meanwhile, the background and research objectives, the specific process of applying machine learning for SI engine NOx prediction, and the role of future machine learning models for next-generation zero-carbon engine development are shown in Fig. 1.

2. DATA COLLECTION AND ML MODELS

A validated 1D CFD model of a four-stroke SI engine fueled with gasoline, was operated at different operating conditions including various spark timings, intake pressures and engine speeds to provide a large number of training, validation and testing datasets for the statistical models. The compression ratio (CR) of the engine is 9.5, the stroke and bore are about 86mm, and the connecting rod length is 175mm. Reference [5] provides more detailed information about the CFD model validation.

The different operating conditions of SI engine considered the changes of key parameters including ST, engine speed and load. In detail, the spark timing ranges from -40 to 0 CA ATDC with an interval of 2 degrees, in order to include data in the MBT condition at constant speed and load. For the engine speed, it operates in the range of 1000-4000rpm, with an increment of 1000rpm. As for the intake pressure, it ranges from 0.5 to 1 bar with an interval of 0.1 bar, which can be used to control the engine load. The rest of the operating conditions such as intake air temperature, equivalence ratio, and cooling system are kept constant. About 500 groups of steady-state points were collected by the validated CFD model for the regression model.

The machine learning algorithm is used to model the mapping between the three input variables (ST, intake pressure, speed) and NOx emissions. The obtained dataset is randomly divided into two different parts, where 80% of the dataset is used as the training dataset, which can be used to establish the intrinsic connection between the model inputs and outputs metrics. The remaining 20% is used as the validation dataset, which will be used to test the validity of the model since the validation dataset is unseen data for the trained ML models. For the statistical indicators, the coefficient of determination (R^2) and root mean square error (RMSE) were chosen to evaluate the ML models performance and the learning level of intrinsic law of spark-ignition engines. In order to compare the learning effect of ML models on the intrinsic engine laws, some test data is carefully selected to analyze the applicability of each model from the combustion perspective.

The machine learning approaches used in this study are artificial neural network (ANN), support vector regression (SVR) and random forest (RF) models. The ANN uses backward error propagation algorithm with a 3-9-9-1 neural network structure and the learning rate of 0.01. The SVR model chooses a radial basis function kernel, which maps the samples to a high-dimensional space. The RF model is a combination of bagging idea and random selection feature, 800 decision trees are chosen to vote on the result. And the next section will discuss the predicted results based on three different nonlinear regression models.

3. RESULTS AND DISCUSSION

This section compares the NOx prediction results based on three various machine learning algorithms including artificial neural network (ANN), support vector regression (SVR) and random forest (RF).

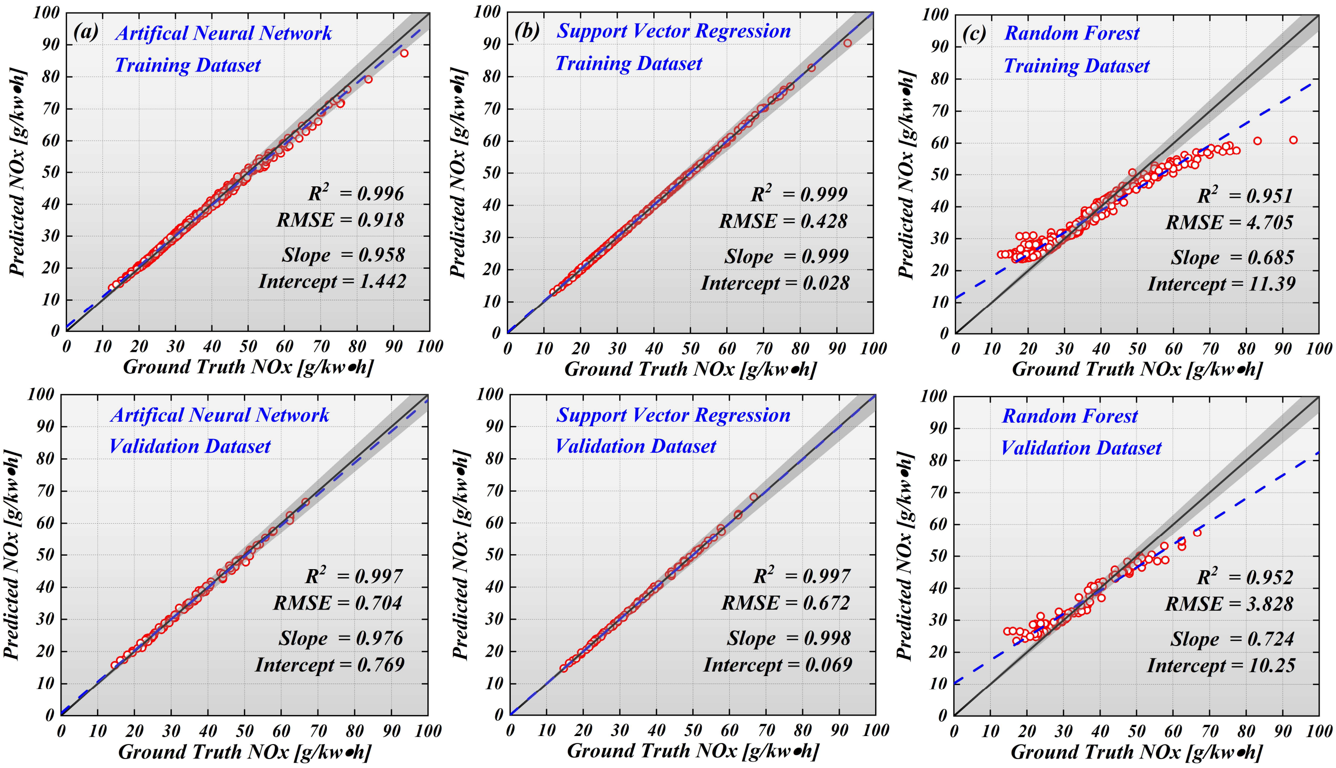

Fig. 2. Comparison of predicted NOx with the ground truth NOx for the training and validation dataset: (a) Artificial neural network; (b) Support vector regression; (c) Random forest

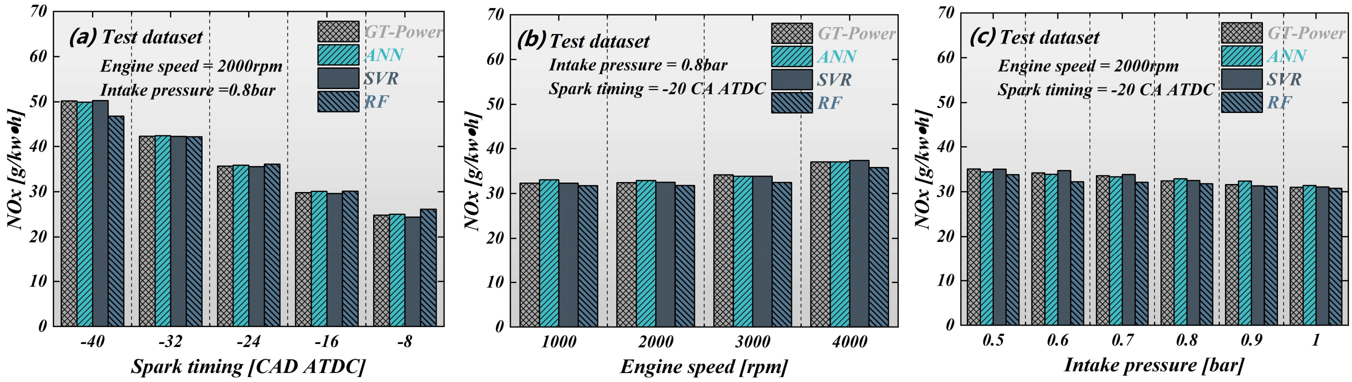

Fig. 3. Comparison of predicted NOx with the simulated values for the test dataset: (a) ST effects; (b) Speed effects; (c) Load effects

Fig. 2 shows the predictions of nitrogen oxides (NOx) based on three different machine learning models (ANN, SVR, and RF) for the training and validation datasets. For the training dataset, the results show that the coefficients of determination (R^2) based on ANN, SVR, and RF for predicting NOx (g/kw-h) are 0.996, 0.999, and 0.951. The higher R^2 and the smaller RMSE indicate better prediction performance. Moreover, if the red hollow circle points are closer to the diagonal, this indicates that the prediction results are more consistent with the ground truth values. And these data points were also fitted using a linear fit, which can be used to assess the degree of agreement between the predicted and actual values. If the slope is closer to 1 and the intercept is closer to 0, it means that the predicted value and the actual value are in better agreement. By comparing the prediction results of different models, it can be found that the prediction points of ANN and SVR are closer to the black line with a slope of 1 than the RF model. Hence, ANN and SVR have better performance than RF model in predicting NOx emissions. By comparing R^2 and RMSE, it is also found that the RF model gives worse results than ANN and SVR. The shading in the figure represents the area where the relative error is within 5%. It can be found that the predicted values of SVR are basically in the middle of the shaded area, as well as most of the values of ANN are within the shaded area. But most of the points of the RF model are located outside the shaded

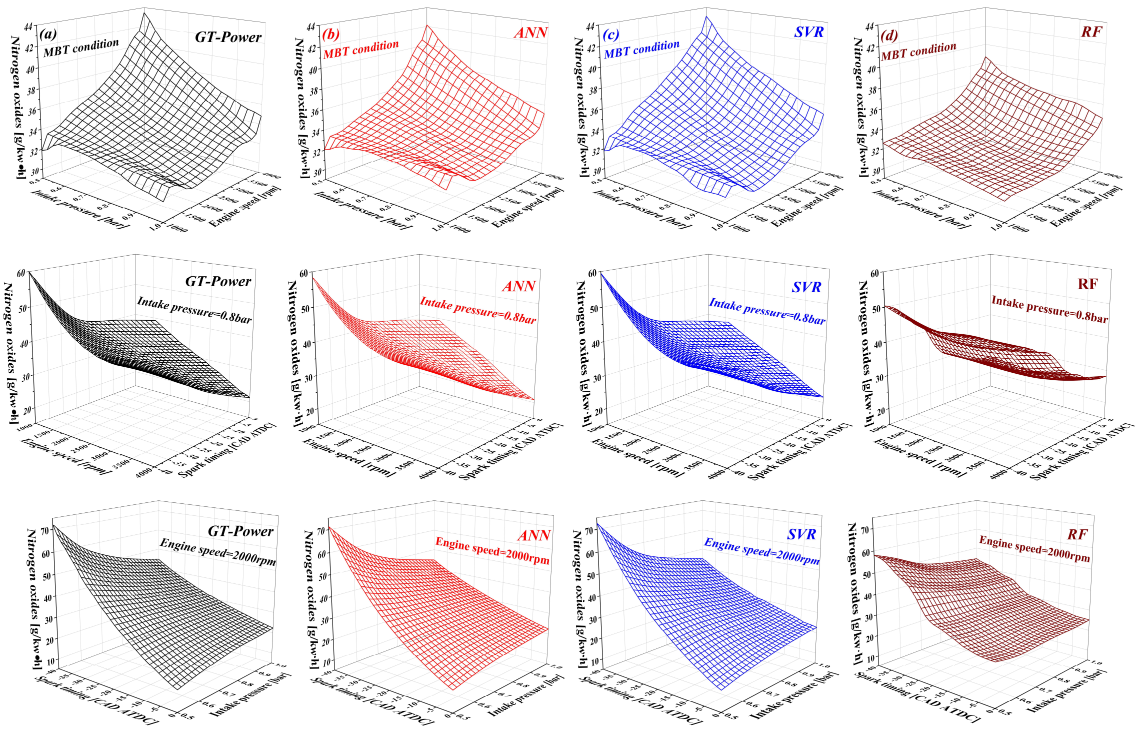

Fig. 4. Comparison of predicted NOx and actual values based on different ML models at MBT condition, at intake pressure and at the speed=2000rpm =0.8bar : (a) GT-Power; (b) ANN; (c) SVR; (d) RF

area, which indicates that the prediction errors of SVR and ANN are smaller than those of the RF model. Next, the prediction results of the ML models on the validation dataset will be analyzed. As expected from the results of training process, the SVR and ANN models had the best agreement of predicted values and results, while the RF model had the worst predicted values, evidenced by lower R^2 and larger RMSE. Moreover, the occupancy of the red hollow points falling in the shaded area, the proximity of slope to 1, and the proximity of intercept to 0 all indicated this point of view. In general, ANN and SVR models can predict NOx emissions well, but RF cannot, probably due to the fact that tree units do not learn the intrinsic laws of SI engine combustion well compared to other ML models.

To further assess the generalization ability of each ML model and whether the small errors in prediction affect the ML models to learn the intrinsic laws of the SI engine. The effects of spark timing, speed and intake pressure on NOx emissions are plotted in Fig. 3.

Fig. 3a indicates that effect of ST on NOx emissions at an engine speed of 2000 rpm and an intake pressure of 0.8 bar. As expected, the NOx decreases with the delay of ST, which is consistent with the combustion law [6]. This is because later spark timing decreased the pressure and temperature of the gasoline engine. Most importantly, all ML models learn the intrinsic engine combustion law well, at least for the operating conditions investigated in this study. Fig. 3b shows the effect of speed on NOx emissions, at the operating condition shown in the figure. As the engine speed increases, NOx emissions increase. This is due to the increase in in-cylinder airflow and flame propagation speed, which leads to increased pressure in the cylinder and more complete combustion, resulting more NOx emissions. Similarly, this law of NOx rises with speed is well learned by these ML models. As for the effect of load on NOx emissions, NOx decreases with larger intake pressure, because the value of g/kw·h was determined by the ratio of NOx concentration to power. Overall, all the models including ANN, SVR and RF can predict the trend well although these models have small errors.

A model with good performance is expected to be able to predict the NOx distribution in the whole operation space. The Fig.4 indicates that both ANN and SVR learn well about the highest and lowest values of the output variable and the patterns of change. However, for the RF algorithm, it can be found that under various operating conditions, the RF predicted map shape is not very consistent with the true one, and the predicted maximum value is smaller and the minimum value is larger compared to the actual value. In other words, the RF model fails for capturing the internal laws of combustion emissions inside the SI engine. Moreover, it can be found that many values predicted by the RF model are within a small interval, which is determined by the nature of decision trees. The RF model decomposes each tree node into smaller subsets according to different characteristics, so that there are many decision trees choosing the same result under one tree node, which may lead to average results. Therefore, RF is not accurate enough for predicting maximum and minimum values, at least for studies on engine-out NOx emissions. However, the RF model also has some advantages, such as it is easy to tune during the training process and is highly resistant to interference. Overall, the prediction accuracy of the RF model is acceptable, just not particularly high. And its hyperparameters are easy tuning, which is faster compared to the other two ML models. And SVR model can predict the trend of NOx with ST, intake pressure, speed as input variables well, which is in line with the relevant studies. Because the noise-free data was obtained by the validated digital twins model operating at various operating conditions. Meanwhile, the anti-noise interference capability of SVR is relatively poor, so ANN may work best if trained with noisy data. However, for the law of simulation data, SVR also performs very well, which is probably due to the absence of noise. Hence, for the choice of most suitable ML model for next generation zero-carbon ammonia/hydrogen SI engines, the ANN and SVR model are recommended to predict engine-out NOx emissions. And in the future, the noisy data will be used as a dataset to further compare the prediction effects of ANN and SVR.

4. SUMMARY AND CONCLUSIONS

The goal of the study was to evaluate the applicability of the ML model in predicting NOx emissions from SI engines fueled with gasoline, which could provide an important reference for the selection of a suitable algorithm in NOx prediction from ammonia/hydrogen engines. This is because the ML methods used to predict NOx emissions are the same for gasoline and ammonia/hydrogen engines. In this paper, three main operating variables (ST, load, and speed) were chosen as input variables and NOx emission was selected as an output parameter. The results indicated that ML algorithms can be used to learn the intrinsic regular connection between the input condition variables and output parameters. And the prediction accuracy of ANN and SVR for NOx output parameters is high, basically within 5% error. However, compared with the other two ML models, the RF model not only has lower prediction accuracy, but also has a larger deviation in the prediction trend graph. This may be due to the structural element of the RF model, which is composed of many tree-like nodes and leaves. By dividing the different features, each tree node will have smaller and smaller subsets, but the voting results of these trees will be consistent for the same class of features. This leads to an overall prediction result biased towards the mean and therefore inaccurate prediction of the peak value. However, the prediction accuracy of the RF model is still acceptable and the hyperparameter adjustment speed of the RF model is high and the interference resistance is strong. Overall, the ANN and SVR models are suitable for predicting NOx emissions from spark-ignition engines based on the dataset in this study. And a comparative analysis of the predictive effects of ANN and SVR using noisy data will be performed in the future.

ACKNOWLEDGEMENT

We are grateful to the Institute of Power Machinery and Vehicle Engineering, College of Energy Engineering, Zhejiang University for providing access to the equipment and software.

CONTACT INFORMATION

Corresponding author: Zhentao Liu
Professor
College of Energy Engineering
Zhejiang University
liuzt@zju.edu.cn

REFERENCE

[1] Berndes G, Abt B, Asikainen A, Cowie A, Dale V, Egnell G, Lindner M, Marelli L, Paré D, Pingoud K, Yeh S. Forest biomass, carbon neutrality and climate change mitigation. From science to policy. 2016;3(7).

[2] Pochet M, Truedsson I, Foucher F, Jeanmart H, Contino F. Ammonia/hydrogen blends in homogeneous-charge compression-ignition engine. SAE technical paper. 2017 ;24:87.

[3] Wan Z, Tao Y, Shao J, Zhang Y, You H. Ammonia as an effective hydrogen carrier and a clean fuel for solid oxide fuel cells. Energy Conversion and Management. 2021 ;228:113729.

[4] Wang R, Chen H, Guan C. Random convolutional neural network structure: An intelligent health monitoring scheme for diesel engines. Measurement. 2021 ;171:108786.

[5] Yang R, Yan Y, Sijia R, Liu Z, Zhang Y, Fu J. A support vector machine model to predict a spark ignition engine performance and emissions. SAE Technical Paper; 2022: 2022-01-0380.

[6] Heywood, J.B. Internal Combustion Engine Fundamentals; McGraw-Hill: New York, NY, USA, 1988.