The Regression Model of NOx Emission in a Real Driving Diesel Vehicle

Hung-Ta Wen, Kuo- Chien Liao, Jau-Huai Lu and Deng-Siang Jhang

May 16, 2020

The regression model of NOx emission in a real driving diesel vehicle

Hung-Ta Wen [1,2] , Kuo- Chien Liao *[1,4], Jau-Huai Lu [1] and Deng-Siang Jhang [3]

[1] Department of Mechanical Engineering, National Chung - Hsing University, Taiwan, R.O.C.
[2] Aerospace Technology Research and Development Center (ATRDC), Taiwan, R.O.C.
[3] Chiayi Motor Vehicles Office, Taiwan, R.O.C.
[4] Department of Aeronautical Engineering, Chaoyang University of Technology, Taiwan, R.O.C.
corresponding author:james19831111@gmail.com

Abstract. The purpose of this study is to use ANN method to train model and predict the NOx emissions of real driving diesel vehicles. In the study, there are two PEMS measurements of driving vehicle's NOx emission data. One of them was divided into three data types: urban, suburbs and highway according to the driving speed. The data such as vehicle speed, vehicle acceleration and EGR are selected as the features values, and the NOx emissions are the target value. Keras and ANN nonlinear autoregressive exogenous model (NARX) are used to train model and predict NOx emissions values. And then make a comparison with the PEMS measurements. The Keras model performances R^2 of first data are 0.9866, 0.9955 and 0.9962 with respect to urban, suburbs and highway types. And The NARX model performance R^2 of second data is 0.99.

Keywords: ANN, PEMS, EGR, NOx, Keras, NARX

1 Introduction

The nitrogen oxides produced by the combustion of diesel engines are one of the air pollutants. And the problem of air pollution on the earth is getting worse. According to the World Health Organization (WHO), in 2018, about 7 million people in the world passed away due to air pollution [1]. This air pollution source is mainly from transportation emission.

In order to reduce effectively and control the emission of nitrogen oxides, both the automobile manufacturer company and the government are betting on considerable research and development expenses and regulations. For example, the EURO 6 emission standard implemented by the European Union in 2014 [2]. It shows that the government and the industry are facing an increasingly serious air pollution problem.

As the regulatory standards are getting higher and higher, taking EURO 6 as an example, the emission standard of nitrogen oxides is 0.08 g/km. The new vehicle needs to pass the laboratory testing standards because of the requirements for market sales, but in fact, after actual driving, the value of kilograms per kilometer measured by the

Portable Emission Measurement System (PEMS) is much higher than the laboratory measurement.

Therefore, the PEMS measurement method is applied to the emission measurement of the vehicle and the pollutant value of the vehicle can be effectively grasped, and the performance of the diesel engine in the real road driving different from the laboratory can be presented. However, the actual road measurement time cost is high, so it is necessary to develop a new method to reduce the measurement cost and effectively control the emission of nitrogen oxides [3].

In order to reduce the cost of measurement and analysis, data analysis techniques are important. Compared with traditional analysis methods, in today's big data era, a large amount of data analysis is no longer difficult. For regression analysis, machine learning technology has been widely used in related topics. Such as Support Vector Machine (SVM), Random Forest (RF) and Artificial Neural Networks (ANN).

Support vector machine (SVM) is used to estimate the fuel consumption of the vehicle and the corresponding factor relationship for multi-variables (such as average speed and driving distance) [4], but the measurement time is up to 1 month, which takes too much time cost [5]. In the machine learning mode, the model is more suitable for estimating the fuel consumption of driving for 10 minutes. If it is necessary to estimate vehicle emissions in real time, the research proposes the use of information and communication technology (ICT) to transmit driving information to the cloud system through the driving of smart phones and calculate pollutant emissions [6].

In addition, it proposes a big data-type of driving fuel consumption estimate, records the emissions of the 1010 trucks at 300 km travel distance, and performs a large amount of data processing using three techniques (SVM, RF, ANN) in the machine learning mode. Successfully established a fuel consumption model [7].

In this study, we propose methods based on Artificial Neural Network. And the results of model performances are good.

2 Method

2.1 Data

The data include two light diesel vehicles measured by PEMS. Vehicle displacement and driving patten are shown in Table 1. The road traffic conditions include general roads and stop traffic lights in the urban, suburbs areas and highways. In car #1, the total distance is divided into three types according to the driving speed. The routes, driving speed and samples of data are listed in Table 1 and from Fig 1 to Fig. 4. For car #2, samples of data are listed in Table 2 and from Fig. 5 to Fig. 8.

Table 1. Vehicle Displacement and Driving Patten

Car	Displacement	Driving time	Driving distance	Samples
#1	*2000*	*92.5 mins*	*80 km*	*5555*
#2	*2200*	*96.5 mins*	*80 km*	*57908*

Table 2. Car #1 Data

Routes	Speed	Data samples
Urban	*0-60 km/hr*	*3215*
Suburbs	*61-90 km/hr*	*1357*
Highway	*91 km/hr above*	*983*

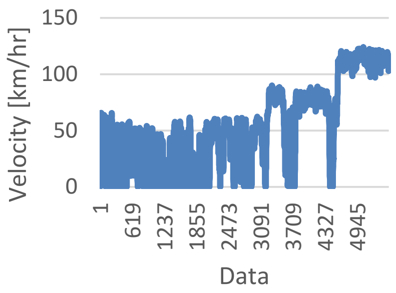

Fig. 1. Car #1 Velocity

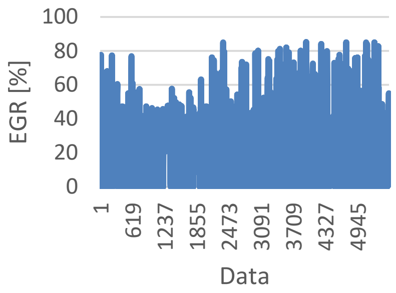

Fig. 2. Car #1 EGR

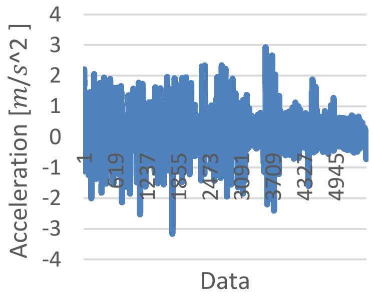

Fig. 3. Car #1 Acceleration

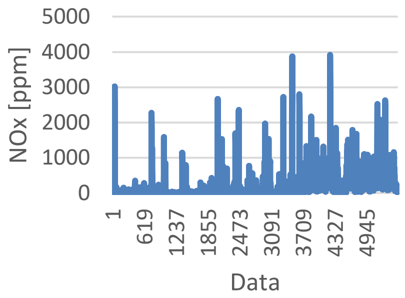

Fig. 4. Car #1 NOx

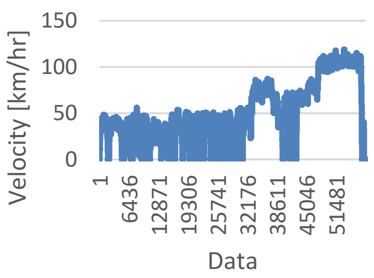

Fig. 5. Car #2 Velocity

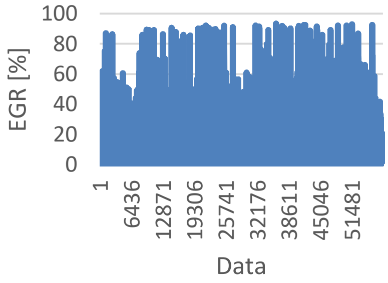

Fig. 6. Car #2 EGR

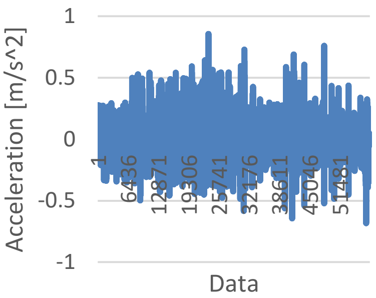

Fig. 7. Car #2 Acceleration

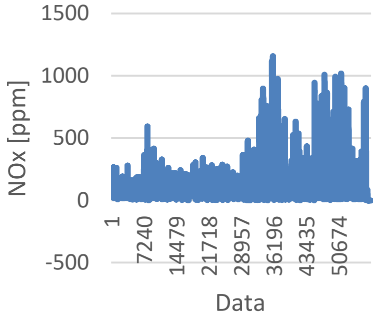

Fig. 8. Car #2 NOx

2.2 Neural network

Deep learning is part of the field of machine learning [8]. The neural network is the architecture, and the unit is the basic structure of the neural network, as shown in Fig 9. The content includes input, weight, deviation value, and output. And the overall structure is also shown in Fig. 10.

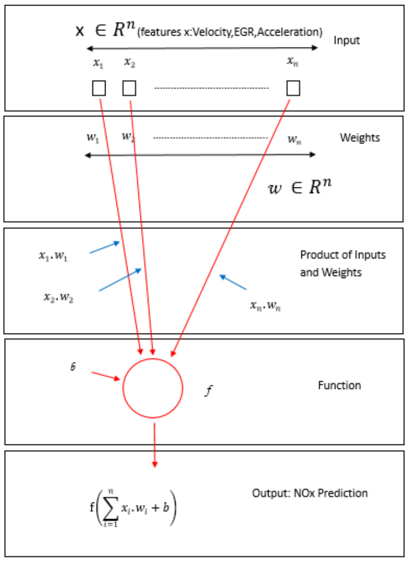

Fig. 9. Neural network unit

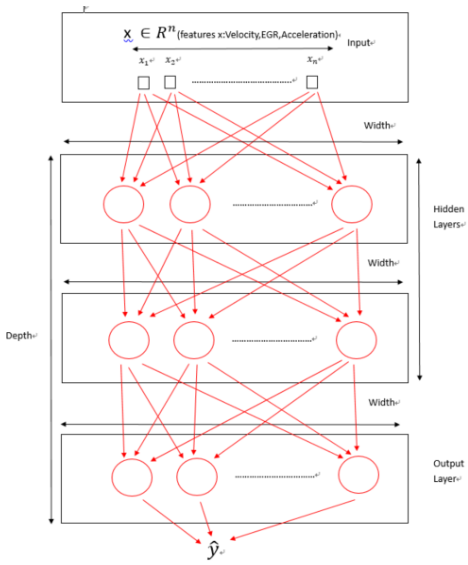

Fig. 10. Structure of neural network

2.3 Neural network model

In car #1 case, Keras open source code written in Python was developed to build the ANN model. And the main Keras model is listed in Table 3.

Table 3. Keras model

Create the model
Model=Sequential( )
Model.add(Dense( ), input_shape=(3,))
Model.add(Activation(activation = 'tanh'))
model.add(Dense( ))
model.compile(optimizer = 'adam', loss = 'mse', metrics = ['accuracy'])

The programming of the code is based on the deep learning model. Loaded in the process of loading data is to pre-organize the data collected by PEMS into Excel *.csv file format, in which the vehicle speed, vehicle acceleration and EGR are input feature values, and NOx is the target value.

The important thing in the compilation model is to choose the appropriate estimator and activation function. The commonly used estimators are AdaGrad, RMSProp, Adam. The selection of the activation function is related to the output of the nonlinear signal. The commonly used activation function is Sigmoid, tanh, ReLu et al. The selection of relevant parameters will affect the optimization degree of the model, and the applicable parameter function needs to be selected as needed.

The evaluation model determines the quality of the model by minimizing the loss function. This can be achieved by adjusting the sample batch-size for each iteration and training the epoch for all samples. The output prediction value is obtained by inputting the characteristic values into the model, and the predicted value can be obtained. The error value can be obtained by comparing the predicted value with the target value. The root mean square error (RMSE), the average absolute error (MAE) and R square (R^2)are used in this study. The formulas are as follows

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(\mathrm{Actual}_i - \mathrm{Predicted}_i)^2}{N}} \tag{2}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|(\mathrm{Actual}_i - \mathrm{Predicted}_i)| \tag{3}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(\mathrm{Predicted}_i - \mathrm{Actual}_i)^2}{\sum_{i=1}^{N}(\overline{\mathrm{Predicted}}_i - \mathrm{Actual}_i)^2} \tag{4}$$

In car #2 case, neural network toolbox under MATLAB environment was developed to build the ANN NARX model. A backpropagation Levenberg-Marquardt (L-M) algorithm using a Multiple Input Single Output (MIMO) layer network was employed to evaluate the performance. The details of the Levenberg-Marquardt (L-M) training of the ANN model could be found in MathWorks website [9].

3 Results and discussion

After the model is built, trained and calculated, the model performance can be evaluated and be listed in Table 4 and Table 5 of car #1 and car #2, respectively. And the error percentage values of the predicted value and the target value can be obtained, results of car #1 and car #2 are listed in Table 6 and Table 7, respectively.

Table 4. Performance of the model

Routes	RMSE	R^2	MAE
Urban	*26.0247*	*0.9866*	*14.9225*
Suburbs	*24.7669*	*0.9955*	*14.9225*
Highway	*21.347*	*0.9962*	*0.1188*

Table 5. Performance of the model

Routes	RMSE	R^2	MAE
All	*12.14*	*0.99*	*6.79*

The car #1 and car #2 predicted values of the different routes driving patterns and the original data are listed in Table 6 and Table 7, respectively. And both error percentage are less than 1%. Fig. 11-13 and 16-17 are curves and scatter plots of results.

In Fig. 11a-13a,16, green line means NOx actual measured data from PEMS and red line means NOx predicted data from model. And in Fig. 11b-13b,17 show the model predictions results. They look like model predicts reasonably well.

And for car #2, Fig. 14, the convergence characteristics of NOx prediction is stable and model offers a better fitting. Fig. 15 shows the errors distribution.

Table 6. Error percentage value of each driving type predicted value and target value

NOx emission g/km	PEMS	Prediction	Error %
Urban	*0.397*	*0.399*	*0.1*
Suburbs	*0.646*	*0.648*	*0.3*
Highway	*1.039*	*1.040*	*0.09*
Average	*0.694*	*0.696*	*0.3*

Table 7. Error percentage value of driving type predicted value and target value

NOx emission g/km	PEMS	Prediction	Error %
Route	*4.72*	*4.69*	*0.6*

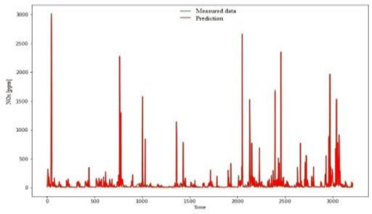

Fig. 11a. Car #1 Urban

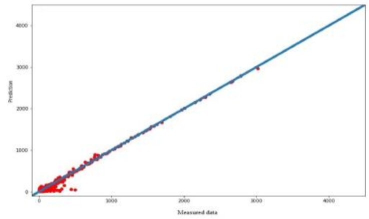

Fig. 11b. Car #1 Urban

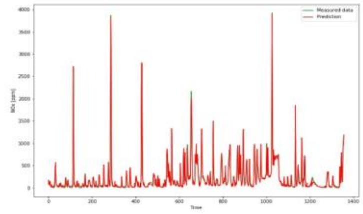

Fig. 12a. Car #1 Suburbs

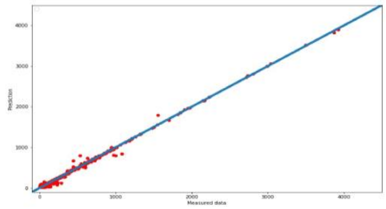

Fig. 12b. Car #1 Suburbs

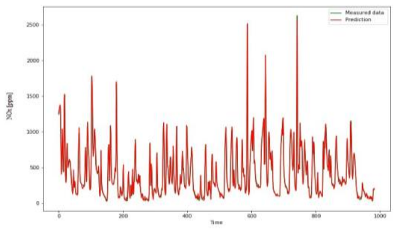

Fig. 13a. Car #1 Highway

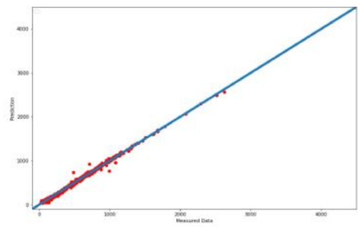

Fig. 13b. Car #1 Highway

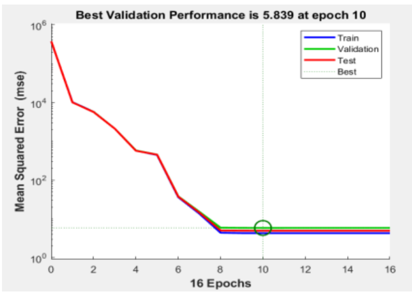

Fig. 14. Car #2 Convergence characteristic

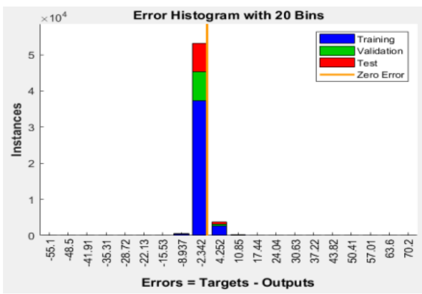

Fig. 15. Car #2 Error histogram

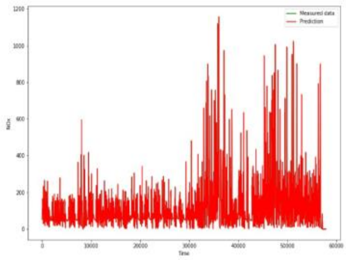

Fig. 16. Car #2

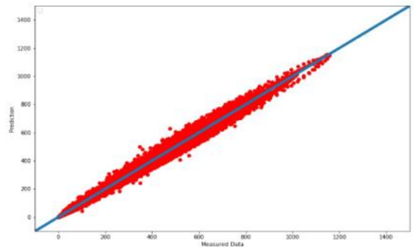

Fig. 17. Car #2

Table 4 indicates the highway driving mode has the smallest values of RMSE, MAE and R^2 , because driving in urban and suburbs areas include stop traffic lights.

4 Conclusion

This study uses ANN method to train model and predict the NOx emissions of real driving diesel vehicles. The following conclusions emerged from the study:

- Two ANN models were successfully built in Python Keras and MATLAB NARX.
- Both models R^2 performances were above 0.98.
- Both NOx prediction emission values were compared with the actual measured value, and the error were within 1%.

References

1. World Health Organization and others, "WHO Global Urban Ambient Air Pollution Database," 2018.
2. https://www.dieselnet.com/standards/eu/ld.php
3. Deng-Siang Jhang, Real road emission measurement and analysis with PEMS for light diesel vehicles: NCHU, 2018.
4. Zeng, W., Miwa, T., & Morikawa, T. (2015). Exploring trip fuel consumption by machine learning from GPS and CAN bus data. Journal of the Eastern Asia Society for Transportation Studies, 11, 906-921.
5. Almér, Henrik. "Machine learning and statistical analysis in fuel consumption prediction for heavy vehicles." (2015).
6. Thibault, Laurent, et al. "A new GPS-based method to estimate real driving emissions." 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2016.
7. Perrotta, Federico, Tony Parry, and Luis C. Neves. "Application of machine learning for fuel consumption modelling of trucks." 2017 IEEE International Conference on Big Data (Big Data). IEEE, 2017.
8. Ketkar, Nikhil. Deep Learning with Python. Apress, 2017.
9. https://www.mathworks.com/matlabcentral/answers/182919-understanding-simple-narx-example 2020/04/05.