

Machine Learning Methods for Emissions Prediction in Combustion Engines with Multiple Cylinders

Hoang Nguyen Khac* Amin Modabberian* Kai Zenger*
Kalle Niskanen* Anton West* Yejun Zhang* Elias Silvola*
Eric Lendormy** Xiaoguo Storm*** Maciej Mikulski*** ****

* *Aalto University, School of Electrical Engineering, Department of Electrical Engineering and Automation, Maarintie 8, 02600 Espoo, Finland. (email: hoang.kh.nguyen@aalto.fi).*

** *Marine Solution Engines, Wärtsilä, Vaasa, Finland (e-mail: eric.lendormy@wartsila.com).*

*** *University of Vaasa, School of Technology and Innovation, Wolffintie 34, 65200 Vaasa, Finland*

**** *University of Vaasa, VEBIC research platform, Wolffintie 34, 65200 Vaasa, Finland*

Abstract: The increasing demand of lowering the emissions of the combustion engines has led to the development of more complex engine systems. This paper presents artificial neural network (ANN) based models for estimating nitrogen oxide (NO_x) and carbon dioxide (CO_2) emissions from in-cylinder pressure of a maritime diesel engine. The architecture of the models is that of Multi-Layer Perceptron (MLP) and Radial Basis Function (RBF) network. The data utilized to train and test the models are obtained from a four-cylinder marine engine. The inputs of the models are chosen as the first principal components of the in-cylinder pressure and engine parameters with highest correlation to aforementioned greenhouse gases. Generalization is performed on the models during the training to avoid overfitting. The estimation result of each model is then compared. Additionally, contribution of each cylinder to the production of emissions is investigated. Results indicate that MLP has a higher accuracy in estimating both NO_x and CO_2 compared to RBF network. The emission levels of each cylinder for both NO_x and CO_2 are mostly even due to the nature of the conventional diesel engine.

Copyright © 2023 The Authors. This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Keywords: machine learning, virtual sensor, green house gases, combustion engine, cylinder pressure

1. INTRODUCTION

The increasing demands for emission reduction in combustion engines has urged research communities to develop more advanced and complex engine systems (Wang et al., 2020; Andriiovych Kuropyatnyk and Victorovych Sagin, 2019; Elkaim and Boyce Jr, 2007). In the marine and off-road sector combustion engines will still be in the key role for a long time because of the high demand of energy density. The need for new engines operating with different fuels with low costs and low emissions calls for new low temperature combustion technology solutions. These are for example Partially Premixed Combustion (PPC), Homogeneous Charge Combustion Ignition HCCI and Reactivity Controlled Combustion Ignition (RCCI) (Bendu and Murugan, 2014). They are promising but difficult technologies, because the physicochemical characteristics in combustion are much more complicated than in traditional diesel engines.

Advanced control methods are needed to control the combustion process, with high constraints set to the emissions. The development of controllers needs modeling and

measured or estimated data from the process. Physical measurement devices are expensive and their performance gradually deteriorates over time because of harsh conditions. Virtual sensors are therefore need to be developed, in order to estimate physical variable values indirectly by safe measurements of other variables. Engine vibration is one variable of interest, because cheap sensors with high accuracy can be used for it. In another paper (Khac et al., 2022) vibration measurements have been used to estimate the cylinder pressure in a combustion engine. The current paper develops the virtual sensor one step further, to estimate nitrogen oxide (NO_x) and carbon dioxide (CO_2) emissions of the engine based on the cylinder pressure data. A combination of these will build a virtual sensor from vibration measurement to emissions.

In this work NO_x and CO_2 emissions of a 4 cylinder marine diesel engine is estimated from in-cylinder pressure utilizing artificial neural networks. The results of different models are then compared in terms of estimation error of aforementioned quantities. Additionally, contribution of each cylinder to the emissions is analyzed.

The structure of the paper is as follows. First of all, the methodology is presented in Section 2. These include the Multi-layer Perceptron (MLP) and the Radial basis function (RBF) networks along with a brief presentation on the principal component analysis (PCA). Section 3 presents the experimental setup, i.e., source of the data and how the data is utilized in this work. Data processing and neural network model development are presented with discussion of the final results in Section 4. Finally, future insights and conclusions are drawn in Section 5 and Section 6 respectively.

2. METHODOLOGY

2.1 Principal Component Analysis

Principal component analysis is a method used for reducing the dimension of large data sets. It is an orthogonal projection of data onto a lower dimensional linear space (Bishop and Nasrabadi, 2006). It aims to preserve as much variation of the data as possible, while reducing the dimension (Shlens, 2014).

In this work, PCA was used to reduce the high dimension of the cylinder pressure traces measured at several operation conditions. The first principal components which explain most of the data were used in the training of the neural network models. Figure 1 shows the first 4 PCA components which explain approximately 98 percent of the variance of the cylinder pressure data.

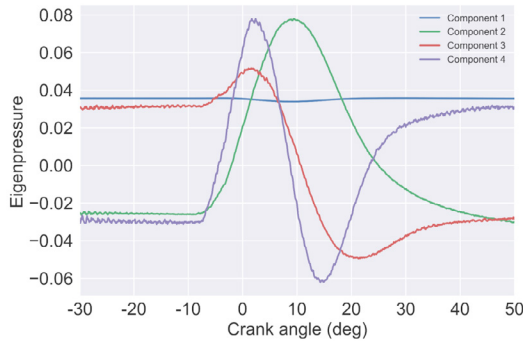


Fig. 1. The first four principal components

2.2 Neural network approaches

Two different types of feed-forward neural networks were used to train the models: the MLP and RBF neural networks.

MLP neural network is a network of interconnected neurons representing the non-linear relationship between input and output vectors, which uses more than one hidden layers without back-propagation (Delashmit et al., 2005). Figure 2 shows an example MLP network with two hidden layers. The activation functions include Rectified Linear unit (ReLU) or softsign functions. ReLU is a piecewise linear function that outputs the input directly, if it is positive, and zero otherwise. On the other hand, softsign is a zero-centered non-linear function, which compresses data into the interval of $[-1,1]$. It is computationally more

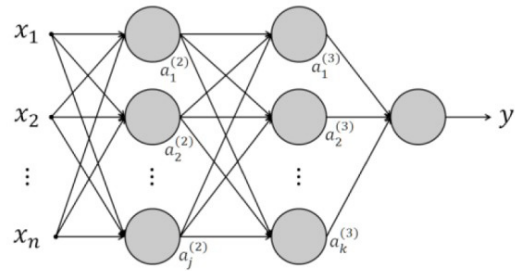


Fig. 2. The MLP network with 2 hidden layers. (Li et al., 2017)

expensive than the ReLU function (Hibat-Allah et al., 2020).

RBF neural network consists of one hidden layer interconnected with an input and an output layer to form a linear combination of the hidden layer signals as shown in Fig 3. There are several activation functions that can be used, but the most popular one is the Gaussian function. It can be expressed as an exponential function of the spreading parameter σ and the Euclidean distance from the input vector to the center of the network as in (1) (Bizon et al., 2011). The output of a RBF neural network is the summation of the weighted radial basis functions of the inputs (Park and Sandberg, 1993).

$$\phi(\|\mathbf{x} - \mathbf{c}_j\|) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{c}_j\|^2}{\sigma^2}\right) \quad (1)$$

in which σ is called the spreading parameter, \mathbf{x} is the input vector, j is the j th neuron in the hidden layer, \mathbf{c}_j is the vector determining the center of the basis function.

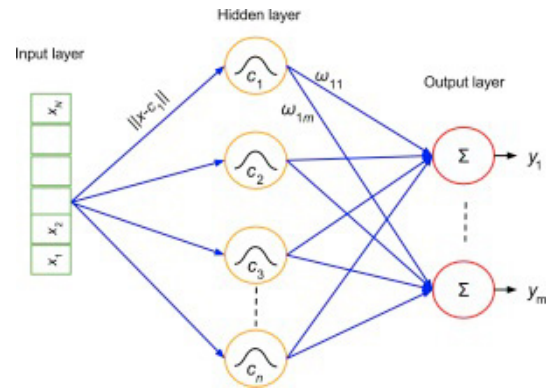


Fig. 3. The architecture of RBF network. (Faris et al., 2017)

3. EXPERIMENTAL SETUP

Cylinder pressure, engine’s parameters and emission data were used to train the virtual sensors. There are two sets of data, one is for training and the other one is for testing. The engine running plan for the training data set was designed by using the Design of Experiment (DoE) method with the Box-Behken design in order to capture the most of the engine’s response at each operating point (Khac and Zenger, 2019). Operating points are defined by engine’s speed and load. By using the Box-Behken design, there are 13 runs with varying parameters (intake pressure, injection timing, rail pressure) at each operating point.

Hence, the training data consists of 182 data samples from 14 operating points which are shown in table 1.

The test run was conducted by running the engine at 4 different settings (varying the engine's parameters) for 8 operating points. Hence, the testing data consists of 32 data samples for each operating points which are shown in table 2.

The data was collected in the University of Vaasa VEBIC engine laboratory from a 4 cylinder common rail diesel engine (WÄRTSILÄ 4L20). The specifications and instrumentation of the engine are given in table 3.

Table 1. Operating points for training data

| Load [kW] | Speed [rpm] | 1000 | 900 | 800 | 700 | 600 |
|-----------|-------------|------|-----|-----|-----|-----|
| 100 | | × | × | × | × | - |
| 150 | | - | - | × | × | - |
| 200 | | × | × | × | × | - |
| 300 | | - | - | × | - | - |
| 400 | | × | × | - | - | - |
| 600 | | × | - | - | - | - |
| 800 | | × | - | - | - | - |

Table 2. Operating points for test data

| Load [kW] | Speed [rpm] | 1000 | 900 | 800 | 700 | 600 |
|-----------|-------------|------|-----|-----|-----|-----|
| 100 | | - | × | - | - | - |
| 150 | | - | - | - | - | × |
| 200 | | × | × | × | × | - |
| 300 | | - | - | - | - | - |
| 400 | | × | - | - | - | - |
| 600 | | × | - | - | - | - |
| 800 | | - | - | - | - | - |

Table 3. WÄRTSILÄ 4L20 engine specifications and instrumentation in VEBIC engine laboratory (Nguyen Khac et al., 2020)

| Technical specs. | Values |
|------------------|-----------------------|
| Cylinder number | 4 |
| Cylinder bore | 200 mm |
| Piston stroke | 280 mm |
| Swept volume | 0.0088 m ³ |
| Rated speed | 1000 rpm |
| Rated power | 800 kW |

(a) Specifications

| Parameters | Devices |
|-------------------|--|
| NO_x | Eco Physics CLD 822 M hr |
| CO, CO_2 | Siemens Ultramat 6 |
| Hydrocarbons | J.U.M.VE7 |
| Cylinder pressure | Kistler 6613CG1 piezoelectric transducer; Kistler KiBox Data Aquisition and post-processing unit |
| Injection timing | Fluke 80i-110s AC/DC current probe |
| Fuel consumption | HBM weight cell, Sartorius X3 |

(b) Instrumentation

4. IMPLEMENTATION AND RESULTS

4.1 Data pre-processing

The raw in-cylinder pressure data cycle was pre-processed on an individual cycle basis. This included filtering, TDC-offset validation and pegging. The pegging corrects the

relative in-cylinder pressure signal to the absolute reference measured at Intake valve port. Such data was further cycle averaged and trimmed from -30 crank angle degree (CAD) to 50 CAD. Meaningful information of emissions formation is mostly contained within the selected crank angle range, and trimming the data also reduces the computational effort in the training of the model (Henningsson et al., 2012). Figure 4 demonstrates the trimmed cylinder pressure data from the test data set.

The emission data is the direct analyzer output (Volume fractions). For all analyzers involved these were measured according to the wet-principle and recorded with 1Hz sampling frequency. For individual steady state operating point the species concentrations were time-averaged over a 60s measurement window.

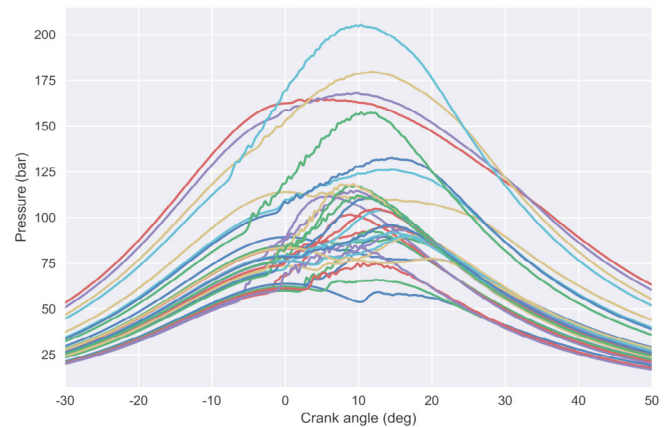


Fig. 4. Cylinder pressure data after pre-processing.

In order to select extra inputs for the neural network model in addition to the cylinder pressure, a Pearson correlation test was conducted to determine the relationships between the recorded engine's parameters and responses. Besides the speed and load, other parameters are: common rail pressure (CR_press), fuel injection timing (MFI), charged air pressure (CA_press), charged air temperature (CA_temp), exhaust gas temperature (Exh_temp), hydrocarbon (HC), and carbon mono-oxide (CO). The test result is described in Fig. 5

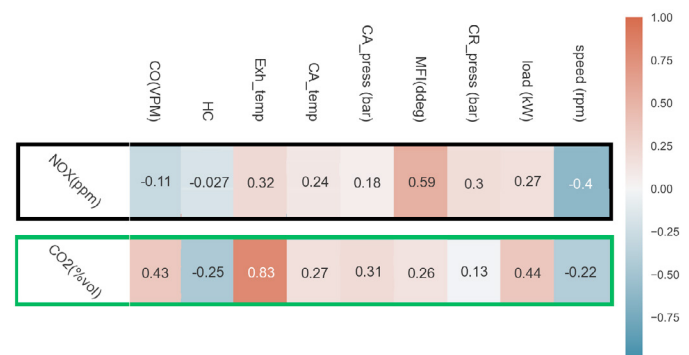


Fig. 5. Correlation test of the measurement data.

According to the test, the NO_x formation is highly correlated with engine speed, load and the fuel injection timing (MFI), highlighted in black rectangle, while the CO_2 was correlated with speed, load, injection timing and exhaust

gas temperature, highlighted in green rectangle. Hence, these parameters were selected as extra inputs for the model, alongside with the principal components of the pressure data.

4.2 Dimension reduction with PCA

In this work, the PCA method was applied to the cylinder pressure data to reduce its dimension. Both of the training and testing data sets were first standardized by subtracting the sample mean and then dividing by the corresponding square root of the sample variance. The meaning of standardization is to make the PCA become scale-invariant. The PCA transform was conducted on both data sets according to (Sánchez, 1982):

$$Y = (X - 1_n \bar{x}^T)G \quad (2)$$

in which X denotes a $n \times p$ data matrix of n independent variables with p observations, \bar{x} denotes the sample mean vector and G denotes the eigenvector matrix of the correlation matrix Σ . Column vectors of G are ordered such that the first vector corresponds to the largest eigenvalue and so forth. The process of taking the eigenvectors from the correlation matrix with the highest eigenvalues is choosing the principal components of the data set. The order of significance of the components is descending with respect to their eigenvalues and those with small eigenvalues can be ignored. With the selected eigenvectors, it is possible to extract the features from the original data set (Shlens, 2014). The features are projections of the data set over the chosen eigenvectors. By doing this, the pressure curve at each operating point can be expressed by as much as the number of the chosen eigenvectors. The same procedure was applied for the test data set. Since each data set contains of pressure data from the four engine cylinders, pressure from each of the cylinders is projected into one set of features. Fig 6 illustrates four features extracted from one cylinder of the training data set. These sets of pressure features are then used as inputs for the neural network model.

4.3 MLP models

MLP neural networks were created using Keras (Chollet et al., 2015), a minimalistic deep learning framework written in Python which runs on top of the machine learning platform TensorFlow (Abadi et al., 2015). Hyper-parameters were optimized using the built-in Hyperband algorithm. Early stopping was used to avoid over-fitting of the model. This is a common regularization technique to determine, when the execution of an iterative algorithm should be terminated. First, the mean square error (MSE) and mean absolute percentage error (MAPE) for training data and test data were monitored during training, when the algorithm was allowed to over fit. When the MSE and MAPE of the test set started to increase, the training was terminated in order to avoid over-fitting. The number of iterations (epochs) at that point was chosen to train the model again. The method is illustrated in Fig. 7

Two separate models were built to estimate the NO_x and the CO_2 emissions with different activation functions. The model developed for NO_x estimation utilizes ReLU

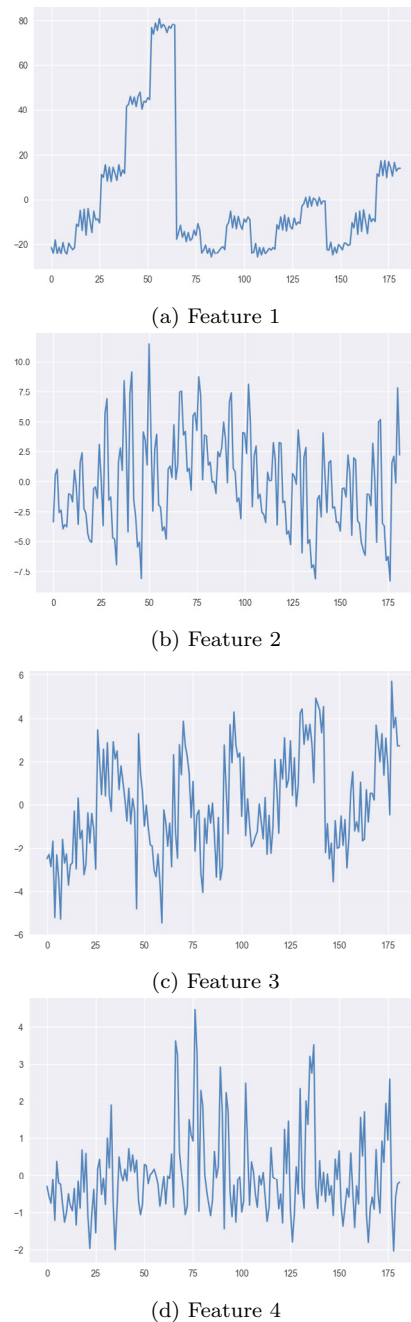


Fig. 6. Four features from one cylinder of the training data set.

function while with the CO_2 model uses soft sign function. Properties and hyper-parameters of these two models are listed in table 4. In the MLP model for NO_x emission estimation, a set of six features selected from each cylinder was chosen as the inputs, together with three extra inputs as engine speed, load and fuel injection timing. That makes it twenty seven inputs in total. The CO_2 's MLP model used the similar input set plus the exhaust gas temperature. That makes it twenty eight inputs in total.

4.4 RBF models

RBF neural networks were created by using the *newrb* function from Deep Learning toolbox in MATLAB. Two important parameters, which require tuning are the max-

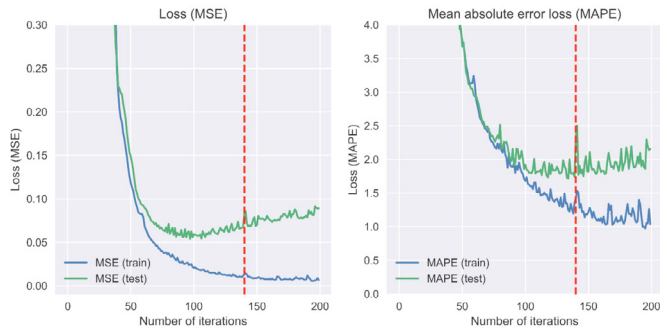


Fig. 7. Early stopping test to prevent over fitting.

imum number of neurons in the hidden layer and the spread of radial basis functions σ . These numbers are tuned by using similar method to the work in (Khac et al., 2022). Two RBF neural network models were developed to estimate NO_x and CO_2 separately. Properties and hyper parameters of these two models are also listed in table 4. The two models were built separately with a same number of four features from each cylinder and the same extra inputs as engine speed, load and fuel injection timing. That makes it nineteen inputs in total for each model.

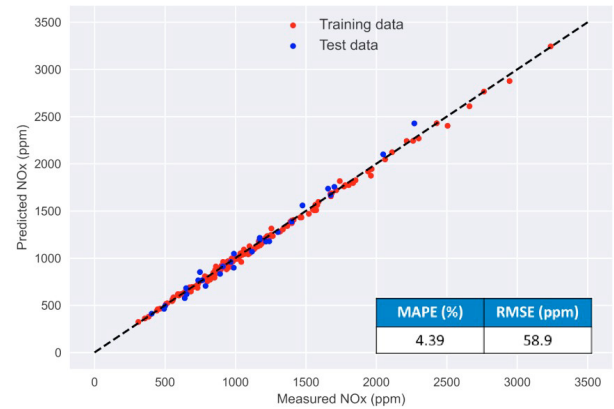
Table 4. Properties and hyper parameters of the neural network models

| Model names | NO_x MLP | CO_2 MLP | NO_x RBF | CO_2 RBF |
|-------------------------------------|------------------|---------------------------------|------------------|------------------|
| Number of principal components used | 6 | 6 | 4 | 4 |
| Extra inputs | Speed, load, MFI | Speed, load, MFI, exhaust temp. | Speed, load, MFI | Speed, load, MFI |
| Total numbers of inputs | 27 | 28 | 19 | 19 |
| Numbers of hidden layers | 3 | 2 | 1 | 1 |
| Number of neurons in hidden layers | 224 | 60 | 150 | 10 |
| Learning rate | 0.0095 | 0.0083 | - | - |
| Activation functions | ReLU | Soft sign | Gaussian | Gaussian |

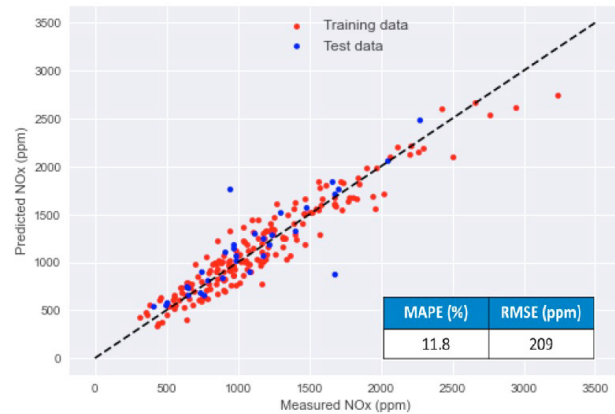
4.5 Results

The prediction and accuracy results of NO_x and CO_2 using MLP and RFB neural network models are demonstrated in Fig 8 and Fig 9 respectively.

Prediction of NO_x from the MLP model has good accuracy with low MAPE, about 4.5 percent. On the other hand, NO_x prediction from the RBF model is not as good as the MLP's result, its MAPE is double (11 percent) and the RMSE is almost quadruple compare to the MLP.



(a) Accuracy of NO_x estimation with MLP.



(b) Accuracy of NO_x estimation with RBF.

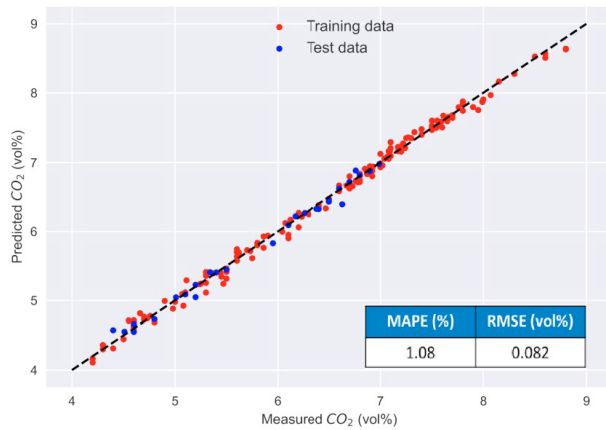
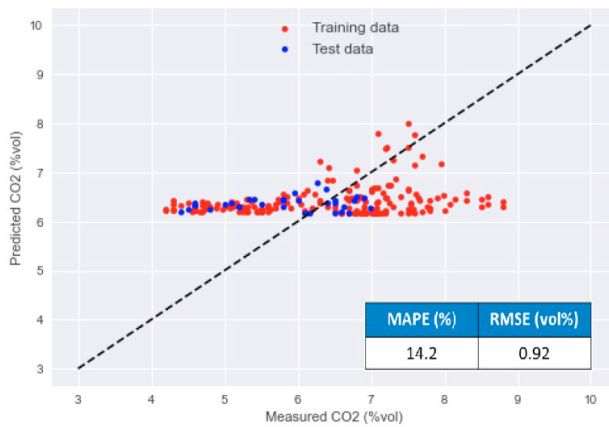
Fig. 8. Accuracy of NO_x estimation.

The result of CO_2 prediction is even more impressive with the MLP model when it yields only one percent of MAPE and very low RMSE (0.082). Nevertheless, the result deteriorated slightly with the RBF model, even though the MAPE and RMSE values are not excessive.

There could be several reasons why the MLP model performed better than the RBF model. One is that the MLP neural network has more hidden layers in its structure than the RBF network and that makes the estimation more accurate. It is shown in table 4 that the number neurons of hidden layers in the MLP network is excessively higher than in the RBF network's hidden layer. Thus it enhances MLP network's capability of calculation of the nonlinear mapping. However, this also increases the computational effort and required resources in training of the MLP models.

Moreover, in the CO_2 models, it can be seen that the RBF network performs considerably worse than the MLP network, which is due to the missing of the exhaust gas temperature input in the RBF model. That input was left out on purpose, and it definitely showed the differences.

Contribution of each cylinder to the production of the emissions was investigated through a validation test by using the two MLP models for NO_x and CO_2 . Instead of using pressure features from all the four cylinders as inputs for the models, it was decided that the pressure features would come from only one cylinder. In other words, it

(a) Accuracy of CO_2 estimation with MLP.(b) Accuracy of CO_2 estimation with RBF.Fig. 9. Accuracy of CO_2 estimation.

was assumed that all the four cylinders are identical and, for example, the engine would have four cylinders of type number 1.

By using this test, it is assumed that the emissions production of the engine comes from only one cylinder type. The test was iterated for all four cylinders, and the result is shown in Fig 10. From the figures, it is understood that the contribution of each cylinder in the emission production of a conventional diesel combustion engine would be approximately equal. In some other engine configurations, where each engine would be operating in a different combustion mode (dual-fuel, gas, etc.), the diversity in each cylinder's contribution to emission production will be more significant.

5. OUTLOOK

The developed neural network model is a cost-effective solution that can be used as a replacement for a physical sensor providing good estimation accuracy and calculation time. The next step is to attach this model to another neural network model developed by Khac et al. (2022). The model in Khac et al. (2022) is capable of estimating in-cylinder pressure from functional vibration measurements. Combination of the two models will then provide the ability to predict in-cylinder pressure and emissions respectively from vibration data. It would then make a complete virtual sensor module for emission measurement.

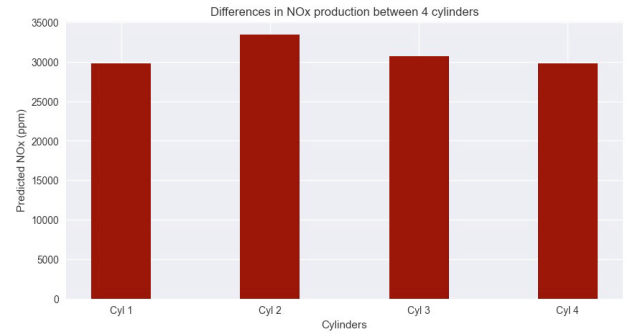
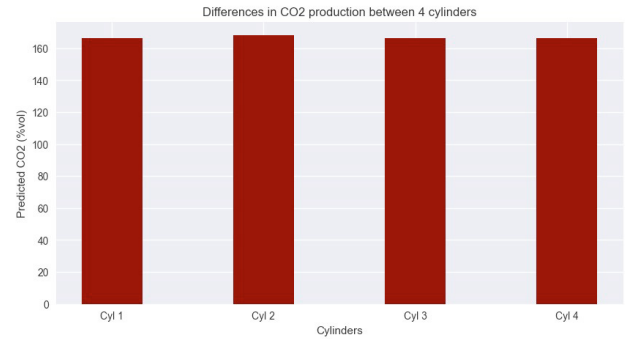
(a) Differences in NO_x production between cylinders.(b) Differences in CO_2 production between cylinders.

Fig. 10. Distribution of each cylinder in emission production.

Additionally, the combined models will be trained with data obtained from reactivity controlled compression ignition (RCCI) combustion process. This black-box model could be later used as a part of a larger gray-box model for model-based RCCI control design.

6. CONCLUSION

In this work, neural network models for emission estimation were successfully developed and tested with measurement data collected from a conventional diesel combustion lab engine. The NO_x and the CO_2 production were predicted by using the in-cylinder pressure data and additional engine parameters such as speed, load, fuel injection timing and exhaust gas temperature.

The PCA method was first implemented to extract the meaningful features from the cylinder pressure data. These features were used as inputs for the estimation model, together with the other selected engine parameters. A few models were developed by using two different approaches: the MLP and the RBF neural networks. The validating test yielded good accuracy with low MAPE and RMSE values. At the same time, it was indicated that MLP neural network provided better estimation results than the RBF network.

Contribution of each cylinder in the production of NO_x and CO_2 was also investigated by using the MLP model. It indicated that each cylinder contributes almost equally in the emission production within this engine configuration, where all four cylinders operate in diesel combustion mode.

ACKNOWLEDGEMENTS

This research is part of the Clean Propulsion Technologies-project (CPT) funded by Business Finland.

The authors also acknowledge the University of Vaasa VEBIC engine laboratory and the following personnels for their support in running the engine tests: Sonja Heikkilä and Olav Nilsson.

REFERENCES

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattemberg, M., Wicke, M., Yu, Y., and Zheng, X. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems. URL <https://www.tensorflow.org/>. Software available from tensorflow.org.
- Andriiovych Kuropyatnyk, O. and Victorovych Sagin, S. (2019). Exhaust gas recirculation as a major technique designed to reduce no emissions from marine diesel engines. *NAŠE MORE: znanstveni časopis za more i pomorstvo*, 66(1), 1–9.
- Bendu, H. and Murugan, S. (2014). Homogeneous charge compression ignition (hcci) combustion: Mixture preparation and control strategies in diesel engines. *Renewable and Sustainable Energy Reviews*, 38, 732–746.
- Bishop, C.M. and Nasrabadi, N.M. (2006). *Pattern recognition and machine learning*, volume 4. Springer.
- Bizon, K., Continillo, G., Mancaruso, E., and Vaglieco, B.M. (2011). Reconstruction of in-cylinder pressure in a diesel engine from vibration signal using a rbf neural network model. Technical report, SAE Technical Paper.
- Chollet, F. et al. (2015). Keras. <https://keras.io>.
- Delashmit, W.H., Manry, M.T., et al. (2005). Recent developments in multilayer perceptron neural networks. In *Proceedings of the seventh Annual Memphis Area Engineering and Science Conference, MAESC*.
- Elkaim, G.H. and Boyce Jr, C.L. (2007). Experimental aerodynamic performance of a self-trimming wing-sail for autonomous surface vehicles. *IFAC Proceedings Volumes*, 40(17), 271–276.
- Faris, H., Aljarah, I., and Mirjalili, S. (2017). Evolving radial basis function networks using moth-flame optimizer. In *Handbook of neural computation*, 537–550. Elsevier.
- Hemmingsson, M., Tunestål, P., and Johansson, R. (2012). A virtual sensor for predicting diesel engine emissions from cylinder pressure data. *IFAC Proceedings Volumes*, 45(30), 424–431.
- Hibat-Allah, M., Ganahl, M., Hayward, L.E., Melko, R.G., and Carrasquilla, J. (2020). Recurrent neural network wave functions. *Physical Review Research*, 2(2), 023358.
- Khac, H.N., Modabberian, A., Zenger, K., Lendormy, É., Mikulski, M., and Hunicz, J. (2022). A neural network approach for reconstructing in-cylinder pressure from engine vibration data. *SAE Technical Papers*.
- Khac, H.N. and Zenger, K. (2019). Designing optimal control maps for diesel engines for high efficiency and emission reduction. In *2019 18th European Control Conference (ECC)*, 1957–1962. IEEE.
- Li, H., Butts, K., Zaseck, K., Liao-McPherson, D., and Kolmanovsky, I. (2017). Emissions modeling of a light-duty diesel engine for model-based control design using multi-layer perceptron neural networks. In *WCX™ 17: SAE World Congress Experience*. SAE International. doi:<https://doi.org/10.4271/2017-01-0601>. URL <https://doi.org/10.4271/2017-01-0601>.
- Nguyen Khac, H., Zenger, K., Storm, X., and Hyvönen, J. (2020). Operational profile based optimization method for maritime diesel engines. *Energies*, 13(10), 2575.
- Park, J. and Sandberg, I.W. (1993). Approximation and radial-basis-function networks. *Neural computation*, 5(2), 305–316.
- Sánchez, J. (1982). Mardia, kv, jt kent, jm bibby: Multivariate analysis. academic press, london-new york-toronto-sydney-san francisco 1979. xv, 518 pp., \$61.00.
- Shlens, J. (2014). A tutorial on principal component analysis. *arXiv preprint arXiv:1404.1100*.
- Wang, Z., Pan, X., Zhang, W., Zhao, Y., Li, H., and Liu, P. (2020). The development trend of internal combustion engine. In *Journal of Physics: Conference Series*, volume 1626, 012139. IOP Publishing.