IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received April 14, 2021, accepted May 6, 2021, date of publication May 19, 2021, date of current version June 10, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3082030

Modelling Fuel Consumption and NO_x Emission of a Medium Duty Truck Diesel Engine With Comparative Time-Series Methods

MEHMET ILTER OZMEN[1], ABDURRAHIM YILMAZ[2], CEMAL BAYKARA[3], AND OSMAN AZMI OZSOYSAL[1]

[1]Department of Naval Architecture and Marine Engineering, Istanbul Technical University, 34669 Istanbul, Turkey
[2]Department of Mechatronics Engineering, Yildiz Technical University, 34349 Istanbul, Turkey
[3]Faculty of Mechanical Engineering, Istanbul Technical University, 34437 Istanbul, Turkey

Corresponding author: Mehmet Ilter Ozmen (mozmen@itu.edu.tr)

This work was supported by ECEMTAG Control Technologies Turkey as part of data driven modelling of heavy duty diesel engine research and development activities.

ABSTRACT This study focuses on different intelligent time series modelling techniques namely nonlinear autoregressive network with exogenous inputs (NARX), autoregressive integrated moving average with external inputs (ARIMAX), multiple linear regression (MLR), and regression error with autoregressive moving average (RegARMA), applied on a diesel engine to predict NO_x emission and fuel consumption. The experiment data are collected from a six cylinder, four stroke medium duty truck diesel engine, which is integrated on a passenger bus and operated in engine integration tests. NO_x emission and fuel consumption outputs are estimated with the help of input data; exhaust gas recirculation temperature and position, engine coolant temperature, engine speed, exhaust gas pressure, common rail pressure, intake manifold air temperature and pressure, accelerator pedal percentage, engine load, turbocharger variable geometry position and speed, and selective catalytic reduction outlet temperature. NARX artificial time series neural network, MLR, ARIMAX, and RegARMA time series techniques were separately applied for the estimation NO_x emission and fuel consumption outputs. The performance of the models is analyzed and evaluated with Bayesian information criterion (BIC) and root mean square error (RMSE) criteria. When the high cost and time loss of experimental testing are thought, using the intelligent modelling methodology provides far more accurate prediction and fast application abilities to analyze internal combustion engine dynamics for the control and calibration manner. As a result of the comparison of different types of modelling techniques, RegARMA technique comes to the forefront with **6707.6** BIC value with **105.58** RMSE for NO_x emission model and **4026.4** BIC value with **7.93** RMSE for fuel consumption model.

INDEX TERMS Artificial neural network, diesel engine, time-series techniques, NO_x emission modelling, fuel consumption modelling.

I. INTRODUCTION

Diesel engine-originated by Rudolf Diesel in 1982-is an internal combustion engine that converts chemical energy to the mechanical power by way of the compression-ignition principle which allows the highly compressed air has a sudden ignite with the injected diesel fuel [1] that has high ignition capability. This leads to a high compression ratio, hence high power and subsequent relatively low fuel consumption and emissions which rise compression-ignition to prominence.

The associate editor coordinating the review of this manuscript and approving it for publication was Yue Zhang.

Besides the working principle and physical characteristics, diesel engine has additional abilities as pre-post injection to overcome fuel consumption, emissions, and noise [2]. Diesel engine is also equipped with external devices as turbocharger for aspiration, diesel oxidation catalyzer (DOC) to provide extra oxidation on exhaust gases, diesel particulate filter (DPF) to decrease particle matter and soot, selective catalytic reduction (SCR) and exhaust gas recirculation (EGR) to reduce NO_x emissions. Considering the increase of emission restrictions (oncoming Euro 7-2023 for heavy duty vehicles) and fuel economy expectations, investigations are ongoing process for controlling diesel engine especially on the new researching areas. Enhancing the performance and lower the emissions, fuel consumption and NO_x emissions are one of the most important variables to observe and analyze. Fuel consumption represents amount of injected fuel that matches with power efficiency. Thereby the hydrocarbon (HC) and particle matter (PM) emissions are also affected by fuel consumption so the same effect is valid for fuel economy considering the economic standards [3]. NO_x emission is a harmful emission to the environment caused by the oxidation of nitrogen in the air as a diluent at a certain temperature. Unlike gasoline engines, diesel engines operate with air compression, so the most determining pollutants for diesel engines are NO_x emissions [4]. Performance and emission analysis requires testing environment and continuous equipment sources so that the process is not efficient enough in terms of time and cost [5]. Therefore, minimizing the challenging process and creating alternative faster and inexpensive testing environment, wide range of data driven modelling techniques as machine learning based regression and neural network algorithms or system identification structure are used. A large variety of data handling methods are implemented on internal combustion engine modelling since physical modelling is also complicated [6]. Fumo and Biswas detailed linear regression techniques and reveals comprehensive study between linear and nonlinear regression on the subject of renewable and sustainable energy which can be considered as review for the future studies and the regression area [7]. Decomposing the input-output relationship for the nonlinear systems, Perez at al. proposed Hammerstein-Wiener system identification model for the diesel engine [8]. Zambrano *et al.* [9] also investigated on the selective catalytic reduction (SCR) system to model NO_x emissions by means of Hammerstein-Wiener system identification model. The both work represented NO_x behavior well but the model performance is limited with the high nonlinear dynamics of diesel engine. Another type of dynamic nonlinearity estimators is NARMAX (NLARX) models (nonlinear autoregressive moving average model with exogenous inputs). Zito and Landau investigated this black-box nonlinear identification via polynomial NARMAX model to represent the relationship between variable geometry turbine (VGT) and boost pressure [10]. Maass *et al.* predicted NO_x emissions for the heavy duty diesel engine analysed two different experiments, Non-Road-Transient-Cycle (NRTC) and variable calibrations for different operation modes. The model adequately estimated NO_x emissions, however wider training sets and range of features are advised for further developments to overcome the drawbacks caused by nonlinearity [11]. Mahla *et al.* for example, used different type of nonlinear regression named autoregressive integrated moving average (ARIMA) model to estimate performance and variable emissions of biogas fueled diesel engine with 99% of root mean square error (RMSE) convergence [12]. Neural networks also included nonlinear system identification as Alonso *et al.* successfully included Levenberg-Marquardt (LM) training algorithm to generate an artificial neural network (ANN) into genetic algorithm optimization process to reduce fuel consumption and emissions of diesel engine [13]. Similar work is done by Roy *et al.* to estimate NO_x emission, smoke, and brake specific fuel consumption with the change of EGR rates in a single cylinder diesel engine [14]. Comparing the techniques each other, Bilgili and Sahin [15] performed a wind speed prediction according to meteorological variables and compare linear regression, nonlinear regression, and ANN with LM algorithm. Despite linear and nonlinear regression techniques give satisfying outputs, ANN shows better estimation than the other modelling methods.

In this study, different data driven regression techniques are applied to the data which are collected from a six cylinder diesel engine integration tests of a medium duty passenger bus. Different form the dynamometer tests, vehicle tests are completed according to the several engine operation modes as low idle, torque mode, and maximum pedal. The focused outputs are NO_x emission and fuel consumption as the turbocharged diesel engine is equipped with DOC, DPF, SCR, EGR. Different data sets are processed for estimation of NO_x and fuel consumption separately. Boost temperature and pressure, load, coolant temperature, engine speed, common rail pressure, exhaust gas pressure, SCR outlet temperature, EGR temperature, EGR position, turbocharger outlet temperature, turbocharger speed, and VGT position data are evaluated as inputs. MLR, NARX, ARIMAX, RegARMA techniques are performed on the NO_x emission and fuel consumption datasets separately.

II. INPUT SELECTION

The focus of this work is to develop a performance and emission model of a diesel engine with the representative outputs, fuel consumption and NO_x emission. Collected input data are used for the prediction of these outputs separately since there are determinative inputs even if most of the inputs are commonly effect fuel consumption and NO_x emission. The main and common inputs affecting the fuel consumption and NO_x emission are boost pressure, boost temperature, accelerator pedal position, load, coolant temperature, engine speed, rail pressure, turbocharger VGT position, turbocharger speed, EGR position, EGR temperature, and exhaust gas pressure. Turbocharger outlet temperature is specifically used for estimating fuel consumption whereas SCR outlet temperature is used for NO_x estimation.

Selecting proper inputs inside of collecting data, the relationship is evaluated by the reference of literature search. Change of fuel consumption and NO_x emission at different loads according to boost temperature is represented by Yusaf *et al.* [16]. Xin *et al.* analyzed NO_x emission reduction methods and observed the variation of fuel consumption and NO_x emission depend on EGR position and VGT position [17]. Choi *et al.* investigated about a new cooling path control strategy and presented the relationship between coolant temperature, fuel consumption and NO_x emissions [18]. Exhaust gas temperature affects on the performance and emissions is observed by Leahu *et al.* [19]. Damma *et al.* published a study

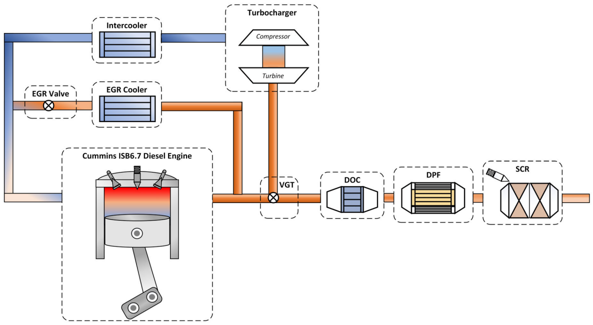

FIGURE 1. Diesel engine schematic with after treatment system.

TABLE 1. Engine specifications.

Engine Model	Medium Duty Passenger Bus Diesel Engine
Engine Type	Diesel, 4-Cycle, Inline 6-Cylinder
Bore and Stroke	4.21 in x 4.88 in
Combustion Ratio	17.3:1
Aspiration	Variable geometry turbocharger
Combustion System	Direct injection
Displacement	6.7 lt
Total Engine Weight	1150 lbs
Maximum Power	242 kW

about low temperature SCR for reduction of NO_x, showing how the SCR usage in different temperature conditions affect NO_x emissions [20]. Advantage of increasing boost pressure include reduced NO_x emissions and combustion efficiency is resulted by Colban *et al.* [21]. Guang *et al.* performed experimental tests proving the direct effect of rail pressure on the fuel consumption and NO_x emissions [22]. Bhure *et al.* pointed out increasing exhaust back pressure affect fuel consumption negatively in contrast decrease NO_x emissions [23]. The experimental tests are performed on the passenger bus as part of engine integration tests based on American Society for Testing and Materials (ASTM) engine test standard. Investigated turbocharged diesel engine that equipped with DOC, EGR, SCR is used for collecting data. Engine control unit of the investigated engine is used to govern the engine and collect the data via sensors. Schematic of diesel engine with exhaust system and engine specifications are shown in Figure 1 and Table 1 respectively.

The data collected from the engine control module (ECM) sensors shows variation as well as different working modes has been processed through the engine operation. The working modes of engine control algorithm are classified based on the way of torque calculation that allow to collect in a wide range and variety of data. Working modes are varied as low speed governor which represents low idle speed governor, automotive governor which represents different accelerator pedal selection, air fuel control derate and maximum throttle

TABLE 2. The range of control variables.

Variable	Min – Max Range
EGR Temperature	63.2 °C — 84.3 °C
EGR Position	0 % — 80 %
Engine Coolant Temperature	65.6 °C — 88.4 °C
Engine Speed	570 rpm — 2427 rpm
Exhaust Gas Pressure	114.05 kpa — 310.95 kpa
Rail Pressure	321 bar — 1741 bar
Boost Temperature	42.7 °C — 57.1 °C
Boost Pressure	0.5 kpa -– 163 kpa
Accelerator Pedal	0 % — 100 %
Load	0 % -– 100 %
Turbocharger VGT Position	4 % — 94 %
Turbocharger Speed	15047 rpm – 115000 rpm
Turbocharger Outlet Temperature	61 °C -– 193 °C
SCR Outlet Temperature	251 °C -– 477 °C
NO_x Emission	0 ppm -– 1306 ppm
Fuel Consumption	0 mg/stroke -– 133.6 mg/stroke

that corresponds torque limiting. The collected data ranges according to different working modes are shown in Table 2.

III. TIME-SERIES TECHNIQUES

Response of the systems based on time change highly depending on previous outputs because of relation between data. This property of the systems causes to model being difficult due to high non-linearity. Different from other non-linear models, past data provide to interrelate next response. In that reason, modelling of the time-dependent systems have additional approaches like recurrent neural networks in comparison with other nonlinear systems. In addition, classical methods like feedforward neural network can be applied to time-dependent systems with reduced metrics. NARX, MLR, ARIMAX, and RegARMA techniques are known as multivariate techniques that allow to predict an output with more than one input. Therefore, these techniques are chosen for the estimation of fuel consumption and NO_x emission outputs of the diesel engine with multiple inputs. The best fitted model is obtained by using the techniques according to metrics, and the coefficients of time series techniques are represented in Table 4.

A. NARX NEURAL NETWORK

NARX neural network is a type of recurrent neural network, with feedback connection, that accepts previous predicted value (parallel architecture) or previous actual value (series-parallel architecture) as inputs when training models. The mathematical background of the NARX model is expressed in 1.

$$y(t) = f(y(t-1), \ldots, y(t-d_y), x(t-1), \ldots, x(t-d_u)) \quad (1)$$

In this equation, $y(t)$ is output time series, d is number of delays and $x(t)$ is input time series [24]. By this way, the shape of NARX neural network is shown in Fig. 2.

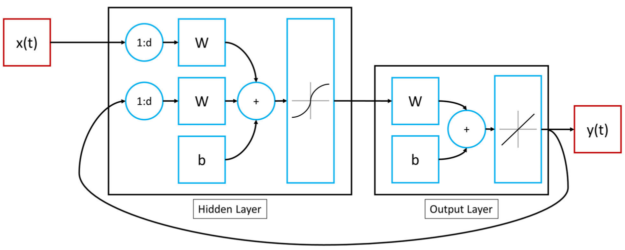

FIGURE 2. Layers and size of NARX neural network model.

TABLE 3. The properties of NARX model.

Property	Value
Train-Test-Validation Ratio	70-15-15
Number of Hidden Neurons	20
Number of Delays	1
Training Algorithm	Levenberg-Marquardt

The properties of the NARX neural network for NO_x and fuel consumption, that is used in this work, are shown in Table 3.

B. MULTIPLE LINEAR REGRESSION

MLR is one of the most well-known time series regression models with multi-variate feature that includes an intercept value and more than one predictor time series with their coefficients (β_i) [25]. i indice defines number of predictors and the equation is represented in 2.

$$y_t = c + X_1\beta_1 + \cdots + X_i\beta_i + \varepsilon_t \tag{2}$$

C. ARIMAX

ARIMA technique based on Box-Jenkins method is a hybrid model that is combination of AR (Autoregressive Process), MA (Moving Average Process), and integral. However, ARIMA technique cannot be used for diesel engine model because of univariate feature. There are more than one features in the model, and ARIMAX is developed to generate multivariate statistical model [26]. ARIMAX formula is shown in 3.

$$\begin{aligned}&\left(1-\phi_1 L - \cdots - \phi_p L^p\right)(1-L)^D y_t \\ &\quad = c + X_1\beta_1 + \cdots + X_r\beta_r + \left(1 + \theta_1 L + \cdots + \theta_q L^q\right)\varepsilon_t\end{aligned} \tag{3}$$

In this formula, p, D, r, and q is represented by autoregressive order, degree of integration, number of predictors and moving average order respectively. Also, c and ε_t hold the constant term and error. If lag polynomial is simplified with lag operator notation, equation takes the form of 4.

$$\phi(L)(1-L)^D y_t = c + \sum_{i=1}^{r}\beta_i X_i + \theta(L)\varepsilon_t \tag{4}$$

$\phi(L)$ and $\theta(L)$ are lag operator notations instead of lag polynomials. Predictors were simplified into the summation.

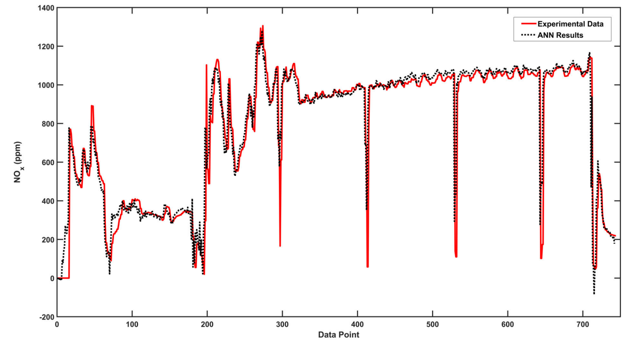

FIGURE 3. Comparison of experimental data and NARX model for NO_X.

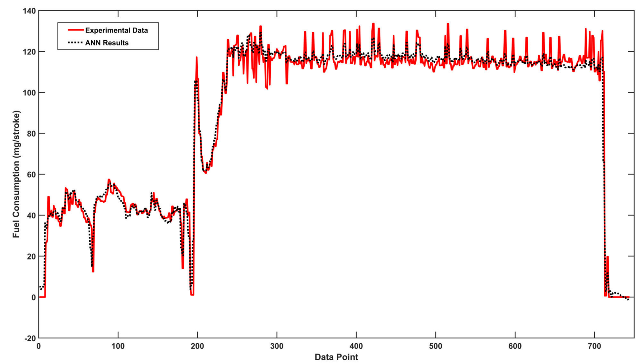

FIGURE 4. Comparison of experimental data and NARX model for fuel consumption.

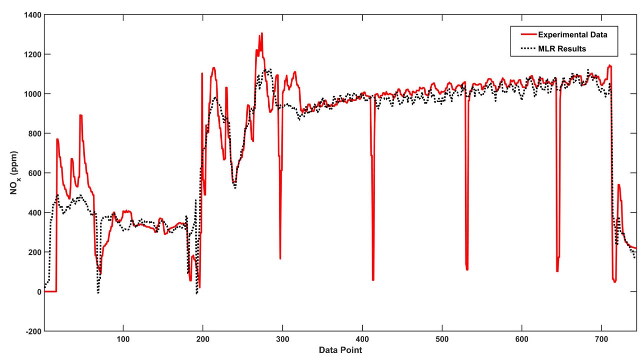

FIGURE 5. Comparison of experimental data and MLR model for NO_X.

D. REGARMA

RegARMA is a model that based on regression with ARMA time series errors. It fits with linear effects of input time series based on multiple linear regression technique. Actually, RegARMA is mixed of multiple linear regression and ARMA. Differently from ARMA, the calculation approach of errors separates from linear model because of assumptions like homoscedasticity [26]. Regression model is given in 5.

$$y_t = c + \sum_{i=1}^{r}\beta_i X_i + \mu_t \tag{5}$$

TABLE 4. Coefficients of MLR, ARIMAX, RegARMA models.

Data / Variable	MLR		ARIMAX		RegARMA	
	NO_x	Fuel Cons.	NO_x	Fuel Cons.	NO_x	Fuel Cons.
Intercept (c)	1.681×10^{3}	3.391×10^{1}	0	0	0	0
EGR Temperature	-3.901×10^{1}	2.085×10^{-1}	-1.533×10^{-1}	1.060×10^{-1}	1.222×10^{1}	3.911×10^{-1}
EGR Position	-6.548	1.180×10^{-1}	-1.966	7.128×10^{-3}	-2.786×10^{-1}	-1.684×10^{-1}
Engine Coolant Temp.	3.131×10^{1}	-9.733×10^{-1}	-1.711×10^{-2}	-5.675×10^{-2}	1.620×10^{1}	-2.268×10^{-1}
Engine Speed	4.722×10^{-1}	-6.933×10^{-2}	1.669×10^{-2}	-3.244×10^{-3}	8.311×10^{-1}	3.152×10^{-2}
Exhaust Gas Pressure	-4.788	2.614×10^{-2}	2.681×10^{-3}	1.443×10^{-2}	1.177×10^{-1}	-1.172×10^{-1}
Fuel Rail Pressure	-1.227×10^{-1}	4.323×10^{-2}	-4.239×10^{-1}	2.693×10^{-3}	-5.561×10^{-2}	7.515×10^{-3}
Boost Temperature	-3.782×10^{1}	-2.209×10^{-1}	-1.481×10^{-1}	9.777×10^{-2}	-3.448×10^{1}	8.119×10^{-2}
Boost Pressure	6.456	8.585×10^{-2}	-1.8×10^{-1}	-3.877×10^{-2}	1.682	5.208×10^{-1}
Exhaust Gas Pressure	-4.788	2.614×10^{-2}	2.681×10^{-3}	1.443×10^{-2}	1.177×10^{-1}	-1.172×10^{-1}
Accelerator Pedal	8.435×10^{-1}	6.386×10^{-1}	2.304×10^{-1}	6.388×10^{-2}	7.950×10^{-1}	-8.229×10^{-2}
Load	9.712×10^{-1}	-1.793×10^{-1}	2.749×10^{-2}	-5.884×10^{-3}	-1.774×10^{-1}	7.682×10^{-1}
Turbocharger VGT Position	5.919	-1.585×10^{-1}	5.983×10^{-4}	-6.682×10^{-3}	-4.011×10^{-1}	1.005×10^{-1}
Turbocharger speed	2.285×10^{-3}	8.038×10^{-4}	6.754×10^{3}	-2.022×10^{-5}	-2.784×10^{-3}	-2.676×10^{-4}
Turbocharger Outlet Temp.	–	1.305	–	-1.568×10^{-1}	–	1.156
SCR Outlet Temperature	1.697	–	1.832×10^{-2}	–	4.016×10^{-1}	–
Moving Average Order 1	–	–	1.162	1.217×10^{-1}	5.401×10^{-1}	1.432×10^{-2}
Moving Average Order 2	–	–	7.082×10^{-1}	–	3.000×10^{-3}	3.303×10^{-2}
Moving Average Order 3	–	–	5.672×10^{-3}	–	–	1.151×10^{-3}
Autoregressive Order 1	–	–	-1.619	-1.392×10^{-1}	1.541	9.999×10^{-1}
Autoregressive Order 2	–	–	4.429×10^{-1}	-2.386×10^{-2}	-5.427×10^{-1}	–
Autoregressive Order 3	–	–	6.899×10^{-2}	6.224×10^{-3}	–	–

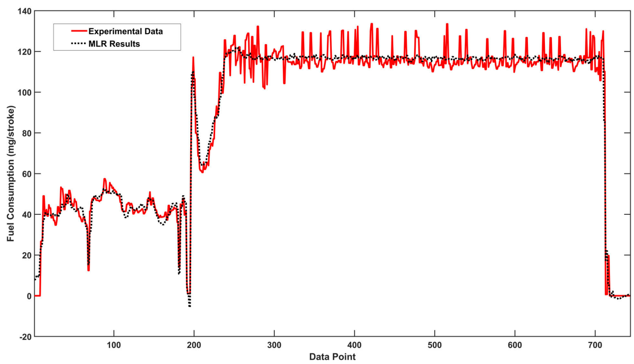

FIGURE 6. Comparison of experimental data and MLR model for fuel consumption.

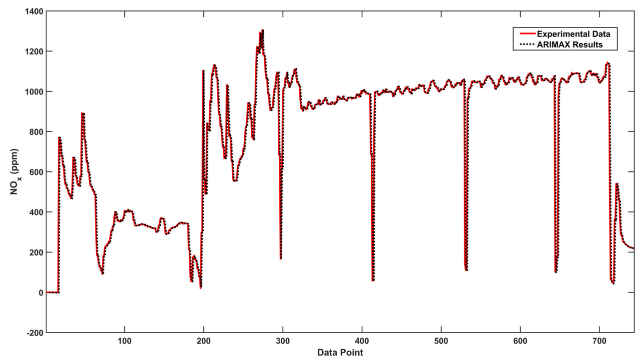

FIGURE 7. Comparison of experimental data and ARIMAX model for NO_x.

c is intercept and μ_t are errors that also named unconditional disturbances. The change of linear model assumptions into error formula is written as 6.

$$\left(1 - \phi_1 L - \cdots - \phi_p L^p\right) \mu_t = \left(1 + \theta_1 L + \cdots + \theta_q L^q\right) \varepsilon_t \tag{6}$$

Simplified form of this formula is stated in 7.

$$\phi(L)\mu_t = \theta(L)\varepsilon_t \tag{7}$$

E. EVALUATION METRICS

Evaluation metrics are used to measure the model performance and led the improvement of the prediction. However, evaluation of the modelling performance with time series techniques requires extra metrics compared with non-time series techniques. In time series data, each of samples have relation between previous and next sample, and information criterion metrics are used to measure goodness of fitting (GoF) as Bayesian information criterion (BIC) [27]. Also, error metrics are used to calculate error value of model as RMSE [28]. In this study, BIC is used for time series metrics and RMSE is used to calculate the error value of the model, and these metrics present that the lower value means good fitting for models. BIC is expressed in 8.

$$BIC = n * log(\frac{SSE}{n}) + p * log(n) \tag{8}$$

TABLE 5. Goodness of fitting of time series techniques.

Metrics \ Data	NARX		MLR		ARIMAX		RegARMA	
	NO_x	Fuel Cons.	NO_x	Fuel Cons.	NO_x	Fuel Cons.	NO_x	Fuel Cons.
BIC	6918.4	4239.9	9616.7	4792.9	6756.6	4240.0	6707.6	4026.4
RMSE	94.03	6.35	147.70	5.75	105.69	7.94	105.58	7.93

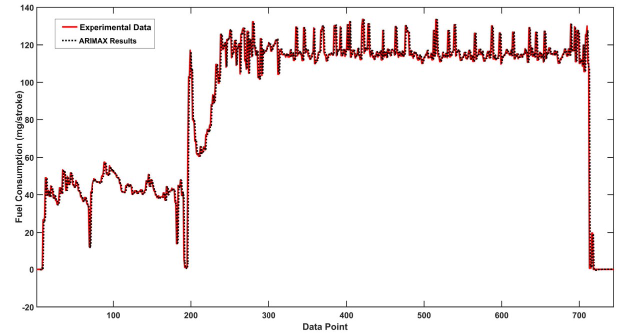

FIGURE 8. Comparison of experimental data and ARIMAX model for fuel consumption.

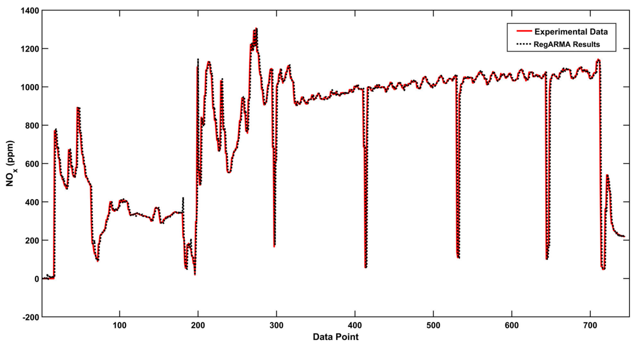

FIGURE 9. Comparison of experimental data and RegARMA model for NO_X.

In the equation, p and n are number of estimated parameters and number of values in the time series data. SSE is sum of squared error and shown in 9.

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{9}$$

Here, y_i is predicted value of model, $\hat{y}_i$ is actual value in dataset and n is size of dataset. Also, another metric, RMSE, is used to calculate error between predicted and actual value shown as 10.

$$RMSE = \sqrt{\frac{SSE}{n}} \tag{10}$$

IV. RESULTS

In this section, the effectiveness of the time series techniques is calculated through metrics and compared in between as it shown in Table 5. BIC is used for the time series models in terms of GoF whereas RMSE is used for error calculation purpose. Representing the evaluation results, NO_x estimation corresponds 6918.4 BIC and 94.03 RMSE value for NARX model, 9616.7 BIC and 147.7 RMSE value for MLR model, 6753.7 BIC and 105.69 RMSE value for ARIMAX and 6707.6 BIC and 105.58 RMSE value for RegARMA. At the end of the comparison, the NARX model has the lowest and MLR has the highest RMSE whereas the other metrics present closer result with NARX. In terms of BIC, again the MLR shows the lowest performance with the highest BIC value because of its simple background caused by absence of lag polynomials. ARIMAX and RegARMA have very close BIC values but RegARMA is better than ARIMAX with a little different. Even if ARIMAX and RegARMA shows enough but not the best performance in terms of RMSE value, because of these techniques are based on autoregressive and moving average method, produced a shifting in models to make the RMSE calculation harder, RegARMA takes the lead for modelling of NO_x with the lowest BIC value and sufficient RMSE value as it represented in 5 as can be observed in Fig. 9. Fuel consumption modelling has a challenging side different from NO_x modelling because of big spikes in time series data. NARX model resulted 4239.9 of BIC and 6.35 of RMSE whereas the of MLR yielded as 4792.9 of BIC and 5.75 of RMSE. In the meantime, ARIMAX resulted 4240.0 of BIC and 7.9386 RMSE and RegARMA resulted 4026.4 of BIC and 7.9310 of RMSE. As a result, MLR has the lowest so the best performance in terms of RMSE but in contrast highest BIC that corresponds lowest time series performance. Investigation on the results shows that the highest estimation performance is belong to RegARMA which is resulted nearly with the ARIMAX in average. Beside this, the reason behind the higher RMSE value resulted for RegARMA and ARIMAX is that the related modelling techniques have shifting caused by the usage of autoregressive and moving average method.

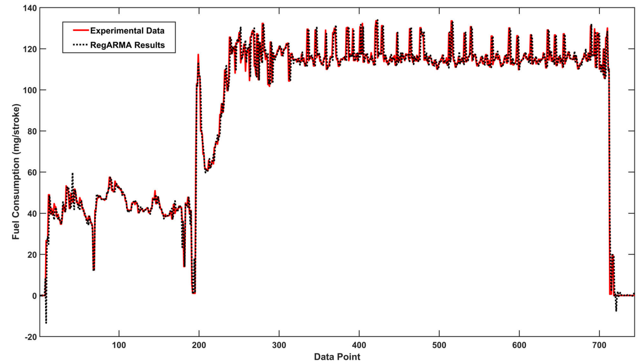

FIGURE 10. Comparison of experimental data and RegARMA model for fuel consumption.

Fittings of models are shown with experimental data, shown as red line, in Fig. 3 and 4 for NARX model, in Fig. 5 and 6 for MLR, in Fig. 7 and 8 for ARIMAX model, in Fig. 9 and 10 for RegARMA model, respectively NO_x and fuel consumption.

V. CONCLUSION

The aim of this study is to represent and compare the performance of different data driven time series modelling techniques that used to predict fuel consumption and NO_x emission which are the most important characteristic variables of diesel engine. The experimental data are collected from medium duty diesel engine at different operation modes which is located in a passenger bus. NARX as an artificial neural network, MLR, ARIMAX, and RegARMA as time series methods are separately applied on a wide variety of data. The effect of input variables is investigated in terms of being meaningful for the output variables to be chosen properly. EGR temperature, EGR position, engine speed, exhaust gas pressure, rail pressure, intake air temperature, intake air pressure, accelerator pedal position, load, turbocharger VGT position, turbocharger speed, turbocharger outlet temperature, SCR outlet temperature are chosen as inputs to be used the estimation of outputs. All of the time series modelling techniques successfully represented the relation between inputs and outputs. To analyze and compare the performance of the models, BIC and RMSE evaluation metrics are progressed. As the result of comparison, RegARMA takes the lead with 4026.4 of BIC and 7.93 RMSE score for fuel consumption and 6707.6 of BIC and 105.58 RMSE score for NO_x emission.

Since the experimental tests are costly and time consuming, time series methods one by one are successfully implemented for the prediction of output variables that allow analyzing performance and emission of diesel engine. At the end, with a wide variety of data, time series methods provide fast and accurate solution in modelling diesel engine characteristics. For further studies, deep learning methods based on recurrent neural network such as long-short term memory (LSTM), gated recurrent unit (GRU) can be used to improve GoF for modelling of diesel engine.

REFERENCES

[1] K. Reif, *Diesel Engine Management Systems and Components* (Bosch Professional Automotive Information). New York, NY, USA: Springer-Verlag, 2014.

[2] F. Z. Aklouche, K. Loubar, A. Bentebbiche, and M. Tazerout, "Effect of pre-injection on the performance of a diesel engine fueled with biogas," *Int. J. Adv. Automot. Technol.*, vol. 2, no. 3, pp. 151–157, 2018.

[3] İ. A. Reşitoğlu, K. Altinişik, and A. Keskin, "The pollutant emissions from diesel-engine vehicles and exhaust aftertreatment systems," *Clean Technol. Environ. Policy*, vol. 17, no. 1, pp. 15–27, Jan. 2015.

[4] L. Cox, *Nitrogen oxides (NO_x) Why How They are Controlled*. Darby, PA, USA: Diane Publishing, 1999.

[5] E. Arcaklioglu and I. Çelikten, "A diesel engine's performance and exhaust emissions," *Appl. Energy*, vol. 80, no. 1, pp. 11–22, 2005.

[6] T. Boz, M. Unel, V. A. M. Yilmaz, C. Gurel, C. Bayburtlu, and K. Koprubasi, "Diesel engine NO_x emission modeling with airpath input channels," in *Proc. 41st Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Nov. 2015, pp. 3382–3387.

[7] N. Fumo and M. A. R. Biswas, "Regression analysis for prediction of residential energy consumption," *Renew. Sustain. Energy Rev.*, vol. 47, pp. 332–343, Jul. 2015.

[8] E. Pérez, X. Blasco, S. García-Nieto, and J. Sanchis, "Diesel engine identification and predictive control using Wiener and Hammerstein models," in *Proc. IEEE Int. Conf. Control Appl.*, Nov. 2006, pp. 2417–2423.

[9] D. Zambrano, S. Tayamon, B. Carlsson, and T. Wigren, "Identification of a discrete-time nonlinear Hammerstein-Wiener model for a selective catalytic reduction system," in *Proc. Amer. Control Conf.*, Jun. 2011, pp. 78–83.

[10] G. Zito and I. D. Landau, "A methodology for identification of narmax models applied to diesel engines," *IFAC Proc. Volumes*, vol. 38, no. 1, pp. 374–379, 2005.

[11] B. Maass, R. Stobart, and J. Deng, "Prediction of NOx emissions of a heavy duty diesel engine with a NLARX model," presented at the SAE Powertrains Fuels Lubricants Meeting, 2009, pp. 1–6, paper 2009-01-2796. [Online]. Available: https://www.sae.org/calendar/techsess/171471.pdf and https://www.sae.org/publications/technical-papers/content/2009-01-2796/

[12] S. K. Mahla, K. S. Parmar, J. Singh, A. Dhir, S. S. Sandhu, and B. S. Chauhan, "Trend and time series analysis by ARIMA model to predict the emissions and performance characteristics of biogas fueled compression ignition engine," *Energy Sour., A, Recovery, Utilization, Environ. Effects*, pp. 1–12, Sep. 2019. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/15567036.2019.1670286

[13] J. M. Alonso, F. Alvarruiz, J. M. Desantes, L. Hernndez, V. Hernndez, and G. Molt, "Combining neural networks and genetic algorithms to predict and reduce diesel engine emissions," *IEEE Trans. Evol. Comput.*, vol. 11, no. 1, pp. 46–55, Feb. 2007.

[14] S. Roy, R. Banerjee, and P. K. Bose, "Performance and exhaust emissions prediction of a CRDI assisted single cylinder diesel engine coupled with EGR using artificial neural network," *Appl. Energy*, vol. 119, pp. 330–340, Apr. 2014.

[15] M. Bilgili and B. Sahin, "Comparative analysis of regression and artificial neural network models for wind speed prediction," *Meteorol. Atmos. Phys.*, vol. 109, nos. 1–2, pp. 61–72, Nov. 2010.

[16] T. F. Yusaf, D. R. Buttsworth, K. H. Saleh, and B. F. Yousif, "CNG-diesel engine performance and exhaust emission analysis with the aid of artificial neural network," *Appl. Energy*, vol. 87, no. 5, pp. 1661–1669, May 2010.

[17] Q. Xin and C. F. Pinzon, *Improving the Environmental Performance of Heavy-Duty Vehicles and Engines: Key Issues and System Design Approaches*. Amsterdam, The Netherlands: Elsevier, 2014. [Online]. Available: https://www.elsevier.com/books/alternative-fuels-and-advanced-vehicle-technologies-for-improved-environmental-performance/folkson/978-0-85709-522-0

[18] K.-W. Choi, K.-B. Kim, and K.-H. Lee, "Investigation of emission characteristics affected by new cooling system in a diesel engine," *J. Mech. Sci. Technol.*, vol. 23, no. 7, pp. 1866–1870, Jul. 2009.

[19] C. I. Leahu, S. Tarulescu, and R. Tarulescu, "The exhaust gas temperature control through an adequate thermal management of the engine," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 444, no. 7, 2018, Art. no. 72016.

[20] D. Damma, P. Ettireddy, B. Reddy, and P. Smirniotis, "A review of low temperature NH_3-SCR for removal of NO_x," *Catalysts*, vol. 9, no. 4, p. 349, Apr. 2019.

[21] W. F. Colban, P. C. Miles, and S. Oh, "Effect of intake pressure on performance and emissions in an automotive diesel engine operating in low temperature combustion regimes," *SAE Trans.*, pp. 957–977, Oct. 2007. [Online]. Available: https://www.sae.org/publications/technical-papers/content/2007-01-4063/

[22] T. Xu-Guang, S. Hai-Lang, Q. Tao, F. Zhi-Qiang, and Y. Wen-Hui, "The impact of common rail System's control parameters on the performance of high-power diesel," *Energy Procedia*, vol. 16, no. C, pp. 2067–2072, 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1876610212003244#:~:text= The%20effect%20of%20common%20rail,while%20NOx%20is%20deteriorating

[23] S. Bhure, "Effect of exhaust back pressure on performance and emission characteristics of diesel engine equipped with diesel oxidation catalyst and exhaust gas recirculation," *Int. J. Vehicle Struct. Syst.*, vol. 10, no. 3, pp. 199–203, Aug. 2018.

[24] T. Lin, B. G. Horne, P. Tino, and C. L. Giles, "Learning long-term dependencies in NARX recurrent neural networks," *IEEE Trans. Neural Netw.*, vol. 7, no. 6, pp. 1329–1338, Nov. 1996.
[25] P. Kennedy, *A Guide to Econometrics*. Hoboken, NJ, USA: Wiley, 2008.
[26] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*. Hoboken, NJ, USA: Wiley, 2015.
[27] G. Schwarz, "Estimating the dimension of a model," *Ann. Statist.*, vol. 6, no. 2, pp. 461–464, Mar. 1978.
[28] J. S. Armstrong and F. Collopy, "Error measures for generalizing about forecasting methods: Empirical comparisons," *Int. J. Forecasting*, vol. 8, no. 1, pp. 69–80, Jun. 1992.

MEHMET ILTER OZMEN was born in Istanbul, Turkey, in May 1990. He received the B.S. degree from the Department of Mechanical Engineering, Istanbul University (IU), in 2010, and the master's degree from the Graduate School of ITU, in 2015. He is currently pursuing the Ph.D. degree with the Faculty of Naval Architecture and Marine Engineering. He studied at the University of Stuttgart, as a case study of Master thesis, from 2012 to 2013.

From January 2016 to July 2017, he worked in Tumosan Engine and Tractor Company, as a Control System Engineer, and AVL Turkey, as a System Engineer, from July 2017 to December 2018. Since December 2018, he has been working with ECEMTAG Control Technologies, as a Senior/Lead Control System Engineer on subjects of engine and transmission modelling and control. He is involved in Ph.D. thesis term with the subjects of diesel engine modelling and control, and machine learning.

ABDURRAHIM YILMAZ was born in Karabuk, Turkey, in 1998. He received the B.S. degree in mechatronics engineering from Yildiz Technical University, Istanbul, in 2020, where he is currently pursuing the M.S. degree in mechatronics engineering.

From 2019 to 2020, he was a Research Engineer with ASELSAN. Since 2016, he has been a Research Assistant with the Advanced Systems Innovation Laboratories, Yildiz Techincal University. He holds one patent. His research interests include machine learning, micro robotics, and digital holography at in-vitro cancer research.

Mr. Yilmaz was a recipient of the Turkey Robotics Conference Best Symposium Paper Award, in 2018.

CEMAL BAYKARA was born in Istanbul, Turkey, in October 1966. He received the B.S. degree from the Department of Mechanical Engineering, Istanbul Technical University (ITU), in 1989, and the Ph.D. degree from the Graduate School of ITU, in 2000. He has been continuing his academic studies at the Faculty of Mechanical Engineering, since 1990.

From 1992 to 1993, he worked with the Ecole Polytechnique de Montreal, as a Predoctorate Researcher. He is currently a full-time Professor and works on the subjects of mechanical construction, powertrain and vehicle dynamics, and control theory.

OSMAN AZMI OZSOYSAL was born in Kayseri, Turkey, in May 1963. He received the B.S. degree from the Department of Naval Architecture and Marine Engineering, Istanbul Technical University (ITU), in 1984, and the Ph.D. degree from the Graduate School of ITU, in 1991. He has been continuing his academic studies at the Faculty of Naval Architecture and Ocean Engineering, since June 1984.

He is currently a full-time professor and works on the subjects of ICE simulations and combustion, finite time thermodynamics and coupled vibration of shaft systems.