Margin requirements for first page
Paper size this page A4

94 pt
1.306 in
33.2 mm

Diesel Engine NOx Emission Modeling Using a New Experiment Design and Reduced Set of Regressors

Gokhan Alcan [*] **Mustafa Unel** [*] **Volkan Aran** [*,**]
Metin Yilmaz [**] **Cetin Gurel** [**] **Kerem Koprubasi** [**]

[*] *Faculty of Engineering and Natural Sciences, Sabanci University Istanbul, Turkey*
[**] *Product Development, Ford OTOSAN, Istanbul, Turkey*

Abstract: In this paper, NOx emissions from a diesel engine are modeled with nonlinear autoregressive with exogenous input (NARX) model. Airpath and fuelpath channels are excited by chirp signals where the frequency profile of each channel is generated by increasing the number of sweeps. Past values of the output are employed only in linear prediction with all input regressors, and the most significant input regressors are selected for the nonlinear prediction by orthogonal least square (OLS) algorithm and error reduction ratio. Experimental results show that NOx emissions can be modeled with high validation performance and models obtained using a reduced set of regressors perform better in terms of stability and robustness.

Keywords: Diesel Engine, NOx Emission, Orthogonal Least Square, Regressor Selection, Sigmoid NARX.

1. INTRODUCTION

40 pt
0.556 in
14.1 mm

Diesel engines are widely preferred in heavy-duty engine market due to their high energy conversion efficiency. Besides the advantage of high efficiency, nitrogen oxides (NOx) and soot emissions are released to environment during standard diesel combustion (Challen and Baranescu, 1999). In order to prevent the hazardous effects of these emissions in environmental pollution, stringent emission regulations limit the maximum acceptable emission values. This leads engine manufacturers to exploit the remaining potential of reducing emissions in both stationary and transient operations. Therefore, depending on the application of control or optimization sufficiently accurate model of diesel engine combustion that provides minimum emission values with maximum power becomes very critical.

NOx emissions are known to have nonlinear input to output relations with fundamental combustion inputs injection timing, air/fuel ratio and exhaust gas recirculation (Heywood,1988). Asprion, Chinellato and Guzzella (2013) proposed a mean-value physics based engine-model to predict NOx emissions. They employed a simplified model using setpoint-relative narrow range inputs. This method is not intended or tested for modeling of whole feasible range for each input dimension. Eventhough their model is very accurate and fast, assumptions and simplications make it impossible to adapt the changes in combustion characteristics. Combustion process in a diesel engine is highly nonlinear, so physics based models don't have enough generalization capabilities due to assumptions and simplifications. Therefore, data driven approaches are widely employed in such cases.

Hirsch, Alberer and Del Re, (2008) proposed a combined grey-box modeling approach in which static maps are identified numerically, but the effect of dominant factors is included on the basis of physical assumptions. Eventhough this structure shows a wide range of validity as well as high accuracy, fit performance of the model in highly dynamic operations is not sufficient enough.

40 pt
0.556 in
14.1 mm

Due to the capabilities of capturing memory effects, Volterra poynomials (Ljung, 1999) were employed to identify nonlinear models. Parametric polynomial Volterra series was utilized by Sakushima et al. (2013) to model the diesel engine emissions which chirp input signals. Increasing number of inputs and the degree of polynomials increases dramatically the number of parameters to be estimated and makes these models difficult to deploy in real systems.

Grahn, Johansson and McKelvey (2014) described the model structures of NOx and soot emissions as local linear regression models where the regression parameters are defined by two dimensional look-up tables. Then they interpreted them as a B-spline function and showed how the globally optimal model parameters can be found by solving a linear least-squares problem. However, this work is evaluated only for steady-state engine operations. Formentin et al. (2014) utilized engine speed and indicated pressure measurement to estimate NOx emissions of a heavy-duty diesel engine. They used a principal component analysis and L2 regularization techniques to derive a simple and reliable estimator. Results of the estimator show sufficient performance in steady-state but improvements are required for transient cycles.

Wiener and Hammerstein modeling of a turbocharged diesel engine is proposed by Perez et al. (2006). However,

68 pt
0.944 in
24 mm

Margin requirements for the other pages
Paper size this page A4

68 pt
0.944 in
24 mm

this method is offered for a small region of operation points and Hammerstein models were problematic due to the saturation of inputs. Boz et al compared the capabilities of linear and nonlinear system identification models for a diesel engine NOx emission and showed the superiority of nonlinear Autoregressive with Exogenous Input (NARX) models for such highly nonlinear process. In that work, only airpath input channels were employed in modeling and up to 80% validation accuracy is achieved.

Roy et al., (2014) employed an artificial neural network structure to estimate CO2, NOx and PM emissons of a Common Rail Diesel Injection (CRDI) type engine. In such studies increasing number of hidden layers and neurons leads to overfit and decrease generalization capabilities.

In this paper, NOx emissions of a diesel engine is modeled with airpath and fuelpath input channels, i.e., fuel quantity, engine speed, rail pressure, start of injection (SOI), Manifold Absolute Pressure (MAP) and Mass Air Flow (MAF). Input channels are excited by chirp signals where the frequency profile of each channel is generated by increasing the number of sweeps. Nonlinear autoregressive with exogenous input (NARX) structure with sigmoid function is utilized to model NOx emission. In order to increase the robustness and decrease the sensitivity of the models to the variation of model parameters, past values of the output are employed only in linear function with all input regressors. The input regressors that are selected due to significance are only used in nonlinear function. To this end, the number of input regressors is reduced by utilizing orthogonal least squares (OLS) algorithm (Billings et al., 1988). Experimental results show that using a reduced set of regressors provide more stable and robust NOx models.

40 pt
0.556 in
14.1 mm

Organization of this paper is as follows: Diesel engine combustion is briefly described in Section 2. Design of experiment for system identification is presented in Section 3. Orthogonal least square algorithm based regressor selection is explained in Section 4. Section 5 is about proposed nonlinear identification of NOx emission. Experimental results are provided in Section 6. Finally, paper is concluded with some remarks and future directions in Section 7.

2. DIESEL ENGINE

Diesel engine combustion consists of two fundamental paths called air and fuel paths. Initially, a compressor sucks the ambient air to the system and mixes with the recirculated exhaust gas. The fresh air flow rate into the engine is referred to Mass Air Flow (MAF). Through the pressurized intake manifold, mixed air is injected into the combustion chamber. The pressure in the intake manifold is defined as Manifold Absolute Pressure (MAP). After the combustion, some of the exhaust gases exit the chamber through exhaust manifold and some of them are recirculated back with the help of controlled valve called Exhaust Gas Recirculation (EGR).

Fuel is pumped into a common pressurized rail, where the pressure is called Rail Pressure. The amount of pumped fuel is measured as Fuel Quantity. The starting angle of the main fuel injection quantity is called Start of Injection (SOI). Impact angle of the exhaust gases to turbine blades is adjusted by a valve called Variable Geometry Turbine (VGT).

Diesel engine used for this study is a 12.7L Ford Otosan heavy duty truck engine with common rail direct injection system, a variable geometry turbocharger and cooled high pressure EGR system.

3. DESIGN OF EXPERIMENT

Fuel Quantity (QNT), engine speed (SPD), Rail Pressure (RailP), SOI, MAP and MAF are chosen as input channels for NOx emission model shown in Figure 1.

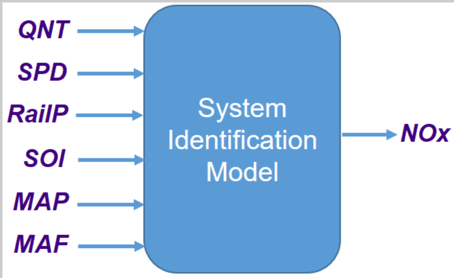

Fig. 1. Inputs and output for system identification model

Periodic signals with changing frequencies have remarkable advantages within the scope of persistent excitation (Ljung, 1999). Therefore input channels were excited by chirp signals, which have sinusoidal waveform with changing frequencies over time given by

$$y(t) = Asin(2\pi(f(t))) \quad (1)$$

40 pt
0.556 in
14.1 mm

where the frequency of the chirp signal can be a linear, quadratic or an exponential function of time. In this work, linear function is employed as

$$f(t) = f_0 + \underbrace{\left(\frac{f_{max} - f_0}{T}\right)}_{k} t \quad (2)$$

where f_{max} is the maximum frequency ($0.5Hz$), f_0 is the initial frequency ($0.05Hz$) and T is the duration time between f_0 and f_{max}. k is defined as chirp ratio.

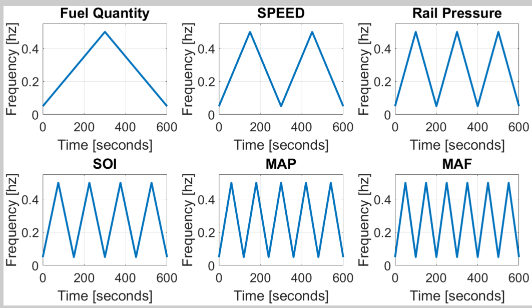

Fig. 2. Frequency profiles of input chirp signals

Boz et al. (2015) proposed reversed sweep frequency profiles for half of the input channels to provide uncorrelated input signals. Differently, in this work we increase the number of sweeps for each input channel (Figure 2). Chirp signals has advantage of having maximum input coverage in minimum testing time while sweeping a significant frequency range. Multi dimensional amplitude coverage is

68 pt
0.944 in
24 mm

Margin requirements for the other pages
Paper size this page A4

68 pt
0.944 in
24 mm

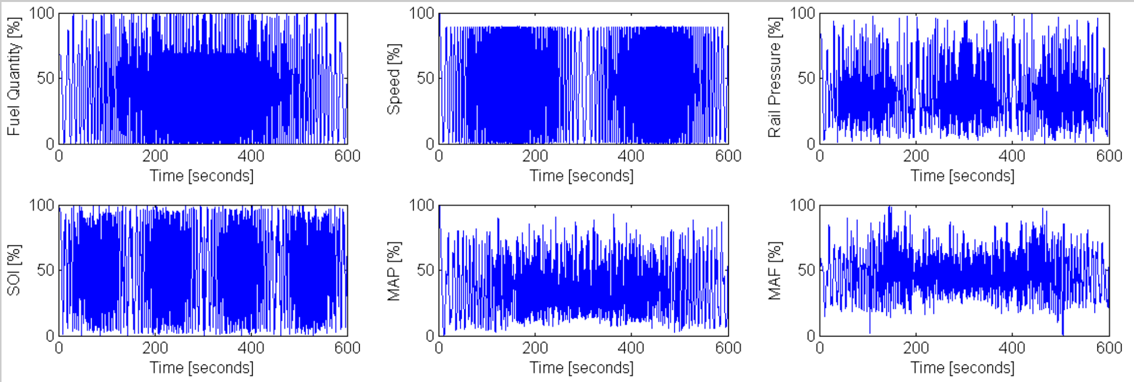

Fig. 3. Normalized input signals

obtained with the DoE design we used. Another reason for choosing chirp input signals over PRBS excitation is that implementation of the DoE on engine dynamometers results with a smoother engine operation. The large steps of PRBS signals causes difficult operation, even though the hardware limitations are used as feasible engine operating limits. Normalized input signals for NOx emissions are shown in Figure 3 for reasons of confidentiality.

In order to check the correlation between input channels, Pearson coefficient for pair-wise channels is calculated as follows and tabulated in Table 1.

$$\gamma = \frac{|<x-\bar{x}, y-\bar{y}>|}{\|x-\bar{x}\|\|y-\bar{y}\|} \tag{3}$$

Table 1. Correlations between input channels

Signals	QNT	SPD	RailP	SOI	MAP	MAF
QNT	1.000	0.036	0.383	0.133	0.509	0.539
SPD	0.036	1.000	0.051	0.069	0.165	0.152
RailP	0.383	0.051	1.000	0.032	0.268	0.240
SOI	0.133	0.069	0.032	1.000	0.005	0.036
MAP	0.509	0.165	0.268	0.005	1.000	0.800
MAF	0.539	0.152	0.240	0.036	0.800	1.000

It should be noted that the frequency profiles shown in Figure 2 are designed as desired references for excitation signals. In Figure 3, it is seen that fuel quantity, speed, rail pressure and SOI input channels are successfully excited in desired manner; but MAP and MAF channels don't show the characteristics of 5 and 6 sweep frequencies. One of the reasons is the limited control capabilities of EGR and VGT valves. Another reason is that MAP and MAF have high interaction during combustion and change their waveforms. Thus the correlation between these signals is higher compared to other input channels (Table 1). However, they are included in the group of input channels for system identification model because both of them have an impact on NOx emissions.

4. REGRESSOR SELECTION WITH OLS ALGORITHM

The generic compact matrix form of linear in the parameters Multi-Input-Single-Output (MISO) systems can be expressed as

$$Y = P\Theta + \Xi \tag{4}$$

where Y is the system response, P is the regression matrix, Θ is the parameters vector to be estimated and Ξ is the modeling error vector.

In fact, some of the regressors in P may be redundant and can be removed to obtain more robust and parsimonious models. In order to determine the terms to be included in a nonlinear model, Billings et al. (1988) introduced forward orthogonal least squares (OLS) algorithm and error reduction ratio.

Assume that the regression matrix P is full rank and can be orthogonally decomposed as

$$P = WA \tag{5}$$

where W is a matrix with orthogonal columns and A is an upper triangular unit matrix. This decomposition can be applied with several orthogonalization procedures such as Gram-Schmidt and Housholder transformations (Chen et al., 1989). One should note that the spaces spanned by the columns of W and P are the same.

$$span\{w_1, w_2, ..., w_m\} = span\{p_1, p_2, ..., p_m\} \tag{6}$$

MISO system in (4) can be rewritten as

$$Y = \underbrace{(PA^{-1})}_{W}\underbrace{(A\Theta)}_{\triangleq G} + \Xi = WG + \Xi \tag{7}$$

where $G = [g_1, g_2, ..., g_m]^T$ can be calculated directly from Y and W as

$$G = (W^T W)^{-1} W^T Y \tag{8}$$

Assume that the past values of the output are not correlated with modeling error Ξ, then the variance of the system response can be calculated as

$$\frac{1}{N} Y^T Y = \frac{1}{N}\sum_{i=1}^{m} g_i^2 w_i^T w_i + \frac{1}{N}\Xi^T \Xi \tag{9}$$

where N is the data length and m is the number of regressors. First part of the output variance is related to regressors because of (6) and the second part is related to unexplained modeling error. Then the regressor selection problem boils down to determining the significant w_i vectors which minimize the variance of unexplained modeling error. To achieve this, error reduction ratio (ERR) is defined as

$$ERR_i = 100 \times \frac{g_i^2 w_i^T w_i}{Y^T Y} \tag{10}$$

40 pt
0.556 in
14.1 mm

40 pt
0.556 in
14.1 mm

68 pt
0.944 in
24 mm

Margin requirements for the other pages
Paper size this page A4
68 pt
0.944 in
24 mm

The details of forward OLS algorithm and error reduction ratio based term and variable selection for nonlinear systems can be found in (Wei et al.,2004).

5. SYSTEM IDENTIFICATION

Diesel combustion NOx formation process is considered to have a nonlinear nature since the linear system identification models can not provide sufficient prediction accuracy to estimate NOx emissions (Boz et al., 2015). Most of the time, validation performances of nonlinear NOx emission models surpass the linear ones. Hence, nonlinear autoregressive with exogenous input (NARX) was selected as modeling structure with sigmoid activation function given by

$$\sigma(x) = \frac{1}{1+e^{-ax}} \tag{11}$$

where a is a positive parameter to be estimated as well.

Modeling structure consists of two parts: nonlinear and linear block (Figure 4). Nonlinear block takes only the input regressors selected by OLS algorithm. Linear block is a single neuron that takes all input and output regressors.

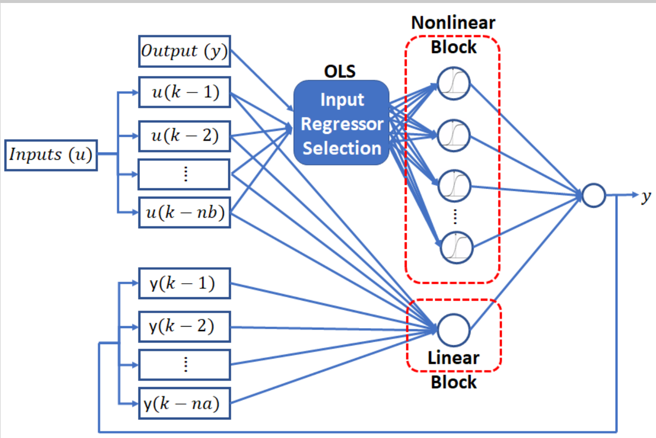

Fig. 4. Sigmoid NARX architecture with OLS based input regressor selection

Removal of output regressors from nonlinear block improves the robustness of the models and enable one to use more than one output regressors in the models. Reducing the number of unnecessary input regressors and obtaining parsimonious models increase the robustness and decrease the sensitivity to parameter changes. Furthermore, deployment of these models to hardware such as Electronic Control Unit (ECU) is much more feasible due to fewer number of estimated parameters. Total number of parameters estimated in proposed architecture is

$$N_P = \Big((N_{NLreg} + 2) \times N_{unit} \Big) + (N_{Lreg} + 1) \tag{12}$$

where N_{NLreg} is the number of regressors used in nonlinear block, N_{Lreg} is the number of regressors used in linear block and N_{unit} is the number of neurons used in nonlinear block.

6. EXPERIMENTAL RESULTS

An exhaustive search for the ranges of parameters given in Table 2 was performed to investigate the performances of the proposed modeling structure. 2970 models were obtained for both proposed structure with regressor selection and the NARX structure without regressor selection. For both cases, output regressors are only employed in linear block.

World Harmonic Transient Cycle (WHTC) tests for two different variant of motor power were also applied for validation purposes.

Table 2. Ranges of parameters

Parameter	Range
na	1-6
nb	2-10
unit	5-15
iteration	20,40,60,80,100

In order to assess the performances of the models after completing the exhaustive search, points are given to the estimated models as follows:

$$Point_i = \frac{100}{3}\left(\frac{est_i}{max(est)} + \frac{val1_i}{max(val1)} + \frac{val2_i}{max(val2)}\right) \tag{13}$$

where est_i is the estimation performance, $val1_i$ is WHTC 1 validation performance and $val2_i$ is WHTC 2 validation performance of the i^{th} model. These performances are calculated by the fit metric given by

$$fit = 100 \times \left(1 - \frac{\|y-\hat{y}\|}{\|y-\bar{y}\|}\right) \tag{14}$$

Note that the number of iterations in running the optimization algorithm is increased by 20 and the model obtained in each 20 iteration is included in 2970 models. In order to get a much more informative statistics, a single model which received the highest point according to (13) is selected out of these 5 models.

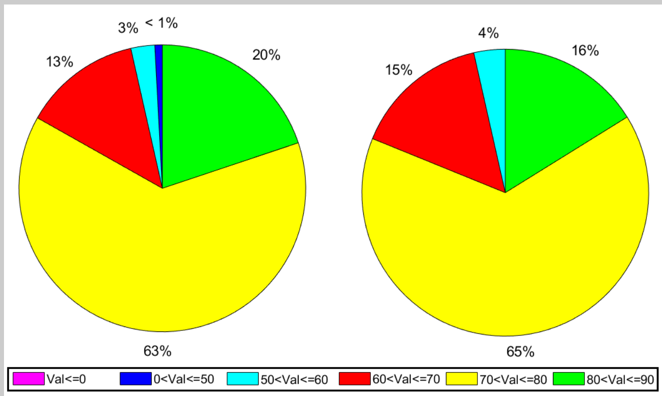

Fig. 5. Estimation performance distribution of the models **Left**: without regressor selection **Right**: proposed

Estimation and validation performance distributions of these models are presented in Figure 5, 6 and 7. Distributions show that the models obtained by proposed modeling structure with less number of parameters show a comparable validation performance as the models obtained by the NARX structure without regressor selection. In these figures, we have used pie charts to visualize that the proposed selection algorithm results with identical or better estimation and validation performances for a batch

40 pt
0.556 in
14.1 mm
40 pt
0.556 in
14.1 mm
68 pt
0.944 in
24 mm

Margin requirements for the other pages
Paper size this page A4
68 pt
0.944 in
24 mm

of modeling trials. In Figure 5, it is shown that the estimation performances of all models obtained by proposed method are above 50%. When all models represented in these distributions are ordered according to points given in (13), the best 6 models are tabulated in Table 3.

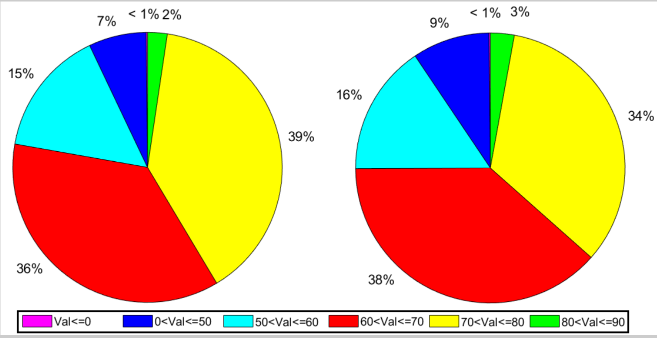

Fig. 6. WHTC 1 validation performance distributions **Left**: without regressor selection **Right**: proposed

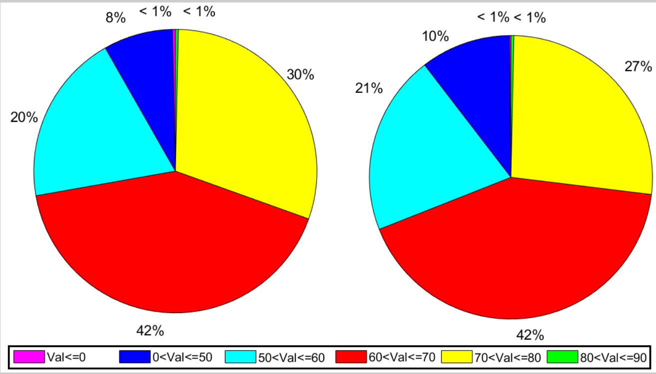

Fig. 7. WHTC 2 validation performance distributions **Left**: without regressor selection **Right**: proposed

Table 3. Best 6 models of proposed structure

na	nb	# of NL Reg	unit	iter	est	whtc1	whtc2
1	7	22	12	100	82,27	84,54	81,74
4	7	22	13	100	82,72	80,40	79,22
4	5	22	15	40	83,44	80,47	77,77
5	5	22	10	100	81,37	79,09	77,95
3	3	15	15	100	79,83	80,71	77,64
3	5	22	15	40	75,83	81,34	80,57

In order to compare these models, the same number of input-output regressors and units are selected with the best iteration and the results are tabulated in Table 4.

Table 4. Corresponding models with the NARX structure without regressor selection

na	nb	unit	iter	est	v1	v3
1	7	12	80	72,53	65,02	61,67
4	7	13	60	82,62	60,95	59,08
4	5	15	100	80,53	77,52	76,36
5	5	10	40	76,21	78,71	78,71
3	3	15	80	80,34	72,71	68,67
3	5	15	80	80,31	77,87	75,39

Table 3 and Table 4 show that the most significant input regressors in nonlinear block prevent the overfit problem and increase the generalization capability of the model. The first model in Table 3 outperforms the corresponding model in Table 4 with 332 parameters whereas there exist 572 parameters in the corresponding model.

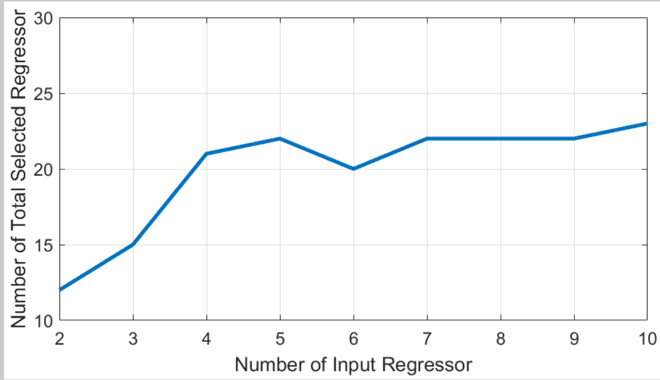

Fig. 8. Total number of selected input regressors

Number of total selected regressors with OLS algorithm is presented in Figure 8. Here, the number of input regressor (nb) is chosen the same for each input channel and regression matrix is constructed accordingly. Figure 8 shows that the total number of selected input regressors converges to 22 regressors.

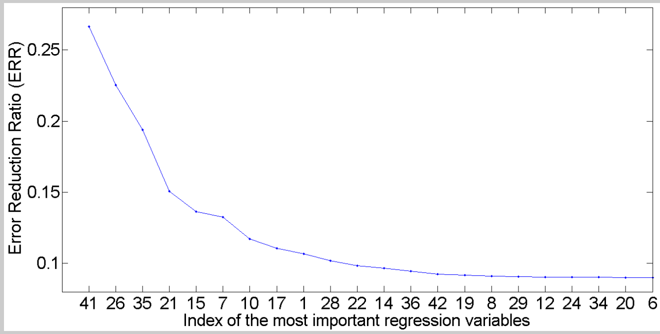

Fig. 9. Determination of significant regressors by error reduction ratio

Figure 9 shows the error reduction ratio corresponding to significant regressors when $nb = 7$ is selected for each input channel. When this ratio converges, indices of most significant regressors are selected and corresponding regressors are extracted to be employed in modeling (Table 5).

Table 5. Extracted regressors for each input channel ($nb = 7$)

MAF	**MAP**	**SPD**	**QNT**	**SOI**	**RailP**
(t-35)	(t-35)	(t-35)	(t-35)	(t-35)	(t-35)
(t-40)	(t-37)	(t-37)	(t-37)	(t-40)	(t-40)
(t-41)	(t-39)	(t-39)	(t-39)	(t-41)	(t-41)
	(t-41)	(t-40)	(t-41)		
		(t-41)			

In Figure 10, 11 and 12, comparison of the model prediction with model training data and validation test data are demonstrated in time plots, respectively. From the validation data comparisons, it is observed that the model accuracy is better on the transient sections of the test cycles compared to the steady state sections. One possible explanation for this phenomennon is that the data used

40 pt
0.556 in
14.1 mm
40 pt
0.556 in
14.1 mm
68 pt
0.944 in
24 mm

Margin requirements for the other pages
Paper size this page A4

68 pt
0.944 in
24 mm

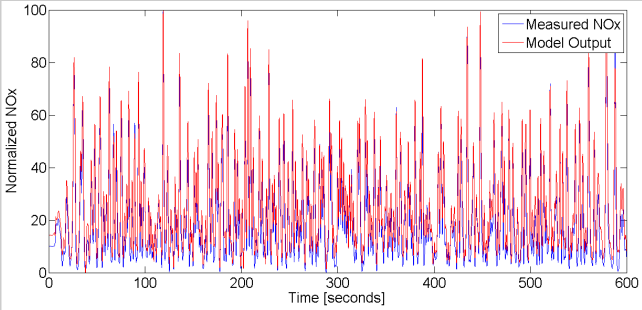

Fig. 10. Estimation performance of the best model in Table 3

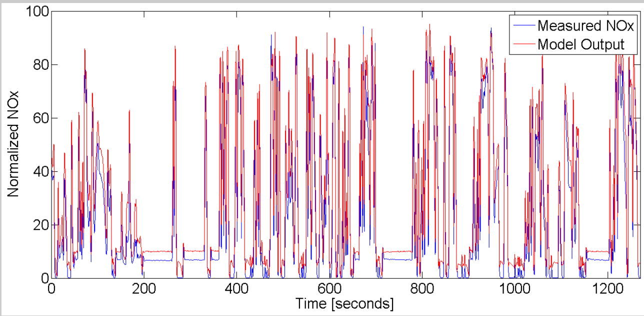

Fig. 11. WHTC-1 validation performance of the best model in Table 3

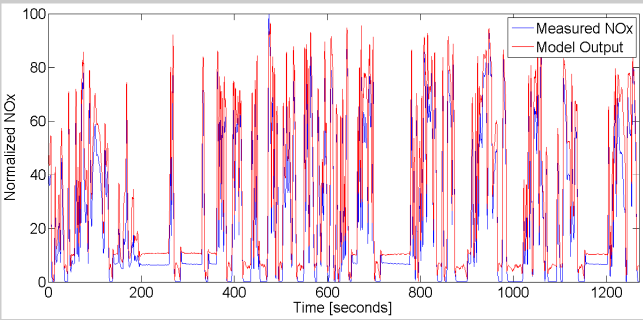

Fig. 12. WHTC-2 validation performance of the best model in Table 3

for model training contains less information about the low frequency response of the system; possibly due to being excited by chirp signals with linearly changing frequency profiles.

7. CONCLUSION

We have now presented a nonlinear ARX structure with reduced nonlinear regressors to model NOx emission of a diesel engine. Employment of output regressors only in linear block and OLS based input regressor selection for nonlinear block provide robustness and decrease sensitivity to model parameter changes. Although performance is similar to other methods, this method decreases training errors and parameter selection problems experienced by calibration development engineers. Using the conventional system identification toolbox, streamlining DoE and regressor selection steps we achieved a convenient engineering solution for diesel emissions modeling with limited testing time. Experimental results are quite promising. However, there are still some problems to tackle such that low frequency components of NOx signals can be learned better with improved design of experiments or preprocessing of measurement data.

40 pt
0.556 in
14.1 mm

ACKNOWLEDGEMENTS

The funding provided by Ford OTOSAN is gratefully acknowledged.

REFERENCES

Asprion, J., Chinellato, O., and Guzzella, L. (2013). A fast and accurate physics-based model for the NOx emissions of Diesel engines. *Applied energy*, 103, 221–233.

Billings, S. A., Korenberg, M. J., and Chen, S. (1988). Identification of non-linear output-affine systems using an orthogonal least-squares algorithm. *International Journal of Systems Science*, 19(8), 1559–1568.

Boz, T., Unel, M., Yilmaz, V. A. M., Gurel, C., Bayburtlu, C., and Koprubasi, K. (2015). Diesel engine NOx emission modeling with airpath input channels. IEEE Industrial Electronics Society, 003382–003387.

Challen, B., and Baranescu, R. (1999) *Diesel engine reference book*. McFarland.

Chen, S., Billings, S. A., and Luo, W. (1989). Orthogonal least squares methods and their application to non-linear system identification. *International Journal of control*, 50(5), 1873–1896.

Formentin, S., Corno M., Waschl, H., Alberer D. and Savaresi, S. M. (2014) NOx Estimation in Diesel Engines via In-Cylinder Pressure Measurement. *IEEE Transactions on Control Systems Technology*, vol. 22,(1), 396–403.

Grahn, M., Johansson, K., and McKelvey, T. (2014). Data-driven emission model structures for diesel engine management system development. *International Journal of Engine Research*, 15(8), 906–917.

Heywood, J. B. (1988) *Internal combustion engine fundamentals*.

Hirsch, M., Alberer, D., and Del Re, L. (2008). Grey-box control oriented emissions models. *IFAC Proceedings Volumes*, 41(2), 8514–8519.

Ljung, L. (1999) *System Identification: Theory for the User*, Prentice Hall.

Perez, E., Blasco, X., Garcia-Nieto, S., and Sanchis, J. (2006). Diesel engine identification and predictive control using Wiener and Hammerstein models. IEEE International Conference on Control Applications, 2417–2423.

Roy, S., Banerjee, R., and Bose, P. K. (2014). Performance and exhaust emissions prediction of a CRDI assisted single cylinder diesel engine coupled with EGR using artificial neural network. *Applied Energy*, 119, 330–340.

Sakushima, N., Wolf, B., Karsten, R., and Mirko, K. (2013). Transient Modeling of Diesel Engine Emissions. *International Journal of Automotive Engineering*, 4(3), 63–68.

Wei, H. L., Billings, S. A., and Liu, J. (2004). Term and variable selection for non-linear system identification. *International Journal of Control*, 77(1), 86–110.

40 pt
0.556 in
14.1 mm

68 pt
0.944 in
24 mm