Data Driven Nonlinear Dynamic Models for Predicting Heavy-Duty Diesel Engine Torque and Combustion Emissions

by

Gökhan Alcan

Submitted to the Graduate School of Sabancı University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

SABANCI UNIVERSITY

July, 2019

Data Driven Nonlinear Dynamic Models for Predicting
Heavy-Duty Diesel Engine Torque and Combustion Emissions

Gökhan Alcan

APPROVED BY

Prof. Dr. Mustafa Ünel
(Thesis Supervisor) .....

Prof. Dr. Mehmet Yıldız .....

Assoc. Prof. Dr. Kemalettin Erbatur .....

Prof. Dr. Şeref Naci Engin .....

Prof. Dr. Metin Gökaşan .....

DATE OF APPROVAL: 12/7/2019

© Gökhan Alcan 2019
All Rights Reserved

Data Driven Nonlinear Dynamic Models for Predicting Heavy-Duty Diesel Engine Torque and Combustion Emissions

Gökhan Alcan

ME, Ph.D. Thesis, 2019

Thesis Supervisor: Prof. Dr. Mustafa Ünel

Keywords: Heavy-Duty Vehicles, Diesel Engine, Combustion Process, Indicated Torque, NOx, Soot, System Identification, Experiment Design, NFIR, NARX, GRU, LSTM

Abstract

Diesel engines' reliable and durable structures, high torque generation capabilities at low speeds, and fuel consumption efficiencies make them irreplaceable for heavy-duty vehicles in the market. However, inefficiencies in the combustion process result in the release of emissions to the environment. In addition to the restrictive international regulations for emissions, the competitive demands for more powerful engines and increasing fuel prices obligate heavy-duty engine and vehicle manufacturers to seek for solutions to reduce the emissions while meeting the performance requirements. In line with these objectives, remarkable progress has been made in modern diesel engine systems such as air handling, fuel injection, combustion, and after-treatment. However, such systems utilize quite sophisticated equipment with a large number of calibratable parameters that increases the experimentation time and effort to find the optimal operating points. Therefore, a dynamic model-based transient calibration is required for an efficient combustion optimization which obeys the emission limits, and meets the desired power and efficiency requirements. This thesis is about developing optimization-oriented high fidelity nonlinear dynamic models for predicting heavy-duty diesel engine torque and combustion emissions.

Contributions of the thesis are: (i) A new design of experiments is proposed where air-path and fuel-path input channels are excited by chirp signals with varying frequency profiles in terms of the number and directions of the sweeps. The proposed approach is a strong alternative to the steady-state experiment based

approaches to reduce the testing time considerably and improve the modeling accuracy in both steady-state and transient conditions. (ii) A nonlinear finite impulse response (NFIR) model is developed to predict indicated torque by including the estimations of friction, pumping and inertia torques in addition to the torque measured from the engine dynamometer. (iii) Two different nonlinear autoregressive with exogenous input (NARX) models are proposed to predict NOx emissions. In the first structure, input regressor set for the nonlinear part of the model is reduced by an orthogonal least square (OLS) algorithm to increase the robustness and decrease the sensitivity to parameter changes, and linear output feedback is employed. In the second structure, only the previous output is used as the output regressor in the model due to the stability considerations. (iv) An analysis of model sensitivities to parameter changes is conducted and an easy-to-interpret map is introduced to select the best modeling parameters with limited testing time in powertrain development. (v) Soot (particulated matter) emission is predicted using LSTM type networks which provide more accurate and smoother predictions than NARX models. Experimental results obtained from the engine dynamometer tests show the effectiveness of the proposed models in terms of prediction accuracies in both NEDC (New European Driving Cycle) and WHTC (World Harmonized Transient Cycle) cycles.

Ağır Vasıta Dizel Motoru Tork ve Yanma Emisyonları Tahmini için
Veriye Dayalı Doğrusal Olmayan Dinamik Modeller

Gökhan Alcan

ME, Doktora Tezi, 2019

Tez Danışmanı: Prof. Dr. Mustafa Ünel

Anahtar kelimeler: Ağır Vasıtalar, Dizel Motor, Yanma Süreci,
Endike Tork, NOx, Soot, Sistem Tanılama, Deney Tasarımı,
NFIR, NARX, GRU, LSTM

Özet

Dizel motorların güvenilir ve dayanıklı yapıları, düşük hızlarda yüksek tork üretme yetenekleri ve yakıt tüketim verimlilikleri, onları piyasadaki ağır hizmet tipi araçlar için vazgeçilmez kılmaktadır. Ancak yanma işlemindeki verimsizlikler emisyonların çevreye salınmasına neden olmaktadır. Emisyonlar için kısıtlayıcı uluslararası düzenlemelere ek olarak, daha güçlü motorlar isteyen rekabetçi talepler ve artan yakıt fiyatları, ağır hizmet tipi motor ve araç üreticilerini, bir yandan performans gereksinimlerini karşılarken bir yandan da emisyonları azaltmak için çözümler aramaya zorlamaktadır. Bu hedefler doğrultusunda, hava işleme, yakıt enjeksiyonu, yanma ve son işlem gibi modern dizel motor sistemlerinde dikkate değer bir ilerleme kaydedilmiştir. Ancak bu tür sistemler, çok sayıda kalibre edilebilir parametreye sahip oldukça karmaşık ekipman kullandıklarından en uygun çalışma noktalarını bulmak için gereken deney süresini ve çabasını artırmaktadır. Bu nedenle, emisyon sınırlarına uyan ve istenen güç ve verimlilik gereksinimlerini karşılayan verimli bir yanma optimizasyonu için dinamik bir model tabanlı geçici kalibrasyon kaçınılmazdır. Bu tez, ağır hizmet dizel motor torkunu ve yanma emisyonlarını tahmin etmek için optimizasyon odaklı yüksek kaliteli doğrusal olmayan dinamik modeller geliştirmekle ilgilidir.

Tezin katkıları şunlardır: (i) Hava yolu ve yakıt yolu giriş kanallarının, taramaların sayısı ve yönleri bakımından değişen frekans profillerine sahip cıvıltı sinyalleri tarafından uyarıldığı yeni bir deney tasarımı önerilmiştir. Önerilen yaklaşım, test süresini önemli ölçüde azaltmak ve hem kararlı durum hem de geçici süreçlerdeki modelleme doğruluğunu geliştirmek adına kararlı durum deneyine dayanan

yaklaşımlara güçlü bir alternatiftir. (ii) NFIR modeli, dinamometreden ölçülen torka ek olarak sürtünme, pompalama ve atalet momentlerinin tahminlerini de dahil eden endike tork tahmini için önerilmiştir. (iii) NOx emisyonlarını modellemek için NARX yapısının iki farklı çeşidi önerilmiştir. İlk yapıda, modelin doğrusal olmayan kısmında kullanılacak girdi regresör seti gürbüzlüğü artırmak ve parametre değişikliklerine duyarlılığı azaltmak için OLS algoritmasıyla azaltılmakta ve çıktı geri beslemesi doğrusal olarak kullanılmıştır. İkinci yapıda, kararlılık hususları nedeniyle sadece bir önceki çıktı, çıktı regresörü olarak modellemede kullanılmıştır. (iv) Parametre değişikliklerine karşı model duyarlılıklarının bir analizi yapılmış ve güç aktarım geliştirmede sınırlı test süresi ile en iyi modelleme parametrelerinin seçilmesi için yorumlanması kolay bir harita sunulmuştur. (v) Kurum (partikül madde) emisyonu NARX modellerinden daha doğru ve pürüzsüz tahminler üretebilen LSTM tipi ağlar kullanarak tahmin edilmektedir. Dinamometre testlerinden elde edilen deneysel sonuçlar, önerilen modellerin hem NEDC (Yeni Avrupa Sürüş Döngüsü) hem de WHTC (Dünya Uyumlulaştırılmış Geçici Döngüsü) döngülerindeki tahmin doğruluğu bakımından etkinliğini göstermektedir.

To my lovely wife...

Acknowledgements

It is a great pleasure to express my sincere gratitude and indebtedness to my thesis advisor Prof. Dr. Mustafa Ünel for his priceless academic and personal guidance, consistent moral and material support, motivation and immense knowledge. It has been an honor and a privilege to complete my Master and Ph.D. studies under his unique supervision. He taught me to enrich a research study with the state-of-the-art methodologies, to be patient and persistent while going deep into a study, and to be planned to set and achieve the milestones on time. Without his exceptional guidance and support, completing this thesis would not have been possible.

I would gratefully thank my thesis committee members: Prof. Dr. Mehmet Yıldız, Assoc. Prof. Dr. Kemalettin Erbatur, Prof. Dr. Şeref Naci Engin and Prof. Dr. Metin Gökaşan for their constructive comments and suggestions on my thesis, spending their valuable time to serve as my jurors.

My sincere thanks go to Volkan Aran, Metin Yılmaz, Çetin Gürel and Dr. Kerem Köprübaşı from Product Development of Ford Otosan for their collaboration in the projects of "System Identification Methods for Diesel Engine Calibration Process". They contributed a lot to my thesis by realizing the design of experiments, making fruitful discussions about the projects, and sharing suggestions and technical information in our publications.

I also want to thank every single member of Control, Vision and Robotics research group Diyar Khalis Bilal, Hammad Zaki, Naida Fetic, Emre Yılmaz, Mehmet Emin Mumcuoğlu and my fellow lab-mates Zeynep Özge Orhan, Uğur Mengilli, Umut Çalışkan, Fatih Emre Tosun and Ayhan Aktaş for their great friendship. I would also like to thank Yusuf Mert Şentürk, Özdemir Can Kara, Ata Otaran, Mert Gülhan and Wisdom Agboh for their weird discussions which provided me with moral support during my Ph.D. studies.

I would like to thank my precious parents Ramazan and Nezahat and my beloved sister Gülşah, for their invaluable love, caring, and never-ending support from the beginning of my life.

Finally, I would like to thank my dearest wife Fatoş Olgun Alcan for her priceless love, encouragement, and constant support, for all the late nights and early mornings, and for keeping me sane throughout my graduate education.

Contents

Abstract iii

Özet v

Acknowledgements viii

Contents ix

List of Figures xii

List of Tables xvii

1 Introduction 1

1.1 Motivation 3

1.2 Contributions of the Thesis 5

1.3 Outline of the Thesis 6

1.4 Publications 7

1.5 Nomenclature 8

2 Literature Review and Background 11

2.1 Diesel Engine Torque Modeling 11

2.2 Diesel Engine NOx Emission Modeling 14

2.3 Diesel Engine Soot Emission Modeling 18

2.4 Modeling of Dynamical Systems 22

2.4.1 Classical System Identification Models 25

2.4.2 RNN Based Modern Network Structures 27

3 Model Based Dynamic Calibration of Diesel Engines 30

3.1 Diesel Combustion Process 32

4 Design of Experiments for Dynamic Calibration of Diesel Engines 34

4.1 Selection of Model Inputs 34

4.1.1 Selection of Regressors 37

4.2 Design of Excitation Signals 39

5 Diesel Engine Torque Modeling 52

5.1 Indicated Torque Estimation 52

5.2 NFIR Modeling 56

6 Diesel Engine NOx Emission Modeling 58

6.1 NARX Modeling 58

6.2 Sensitivity Analysis and Parameter Selection 61

7 Diesel Engine Soot Emission Modeling 66

7.1 GRU Modeling 68

7.2 LSTM Modeling 69

8 Experimental Results **71**

8.1 Experimental Setup 71

8.2 Results and Discussions 74

8.2.1 Torque Modeling 75

8.2.2 NOx Modeling 87

8.2.3 Soot Modeling 100

8.3 Implementation Details 113

9 Concluding Remarks and Future Work **115**

Bibliography **117**

List of Figures

1.1 Diesel engine for heavy-duty vehicles 1

1.2 Model based calibration process 4

2.1 System identification loop for nonlinear dynamical systems 23

2.2 Hammerstein and Wiener block structures 26

2.3 Looped structure of recurrent neural networks 27

2.4 Different types of RNN applications 28

3.1 Evolution of methdologies in model based calibration 31

3.2 Basic structure of a diesel engine 33

4.1 Frequency profiles of linear chirp signals 40

4.2 Normalized input signals (Experiment with 6 inputs) 41

4.3 Frequency profiles of input channels (Experiment 1) 43

4.4 Normalized input signals (Experiment 1) 44

4.5 Normalized input signals (Experiment 2) 44

4.6 Excitation signals used in the experiments for modeling indicated torque and NOx emissions 45

4.7 Coverages of Experiment 1 over validation cycles 46

4.8 Coverages of Experiment 2 over validation cycles 46

4.9 Excitation signals used in the experiments for modeling soot emission 47

4.10 Test 1 inputs and output signals 48

4.11 Test 2 inputs and output signals 48

4.12 Test 3 inputs and output signals 49

4.13 Test 4 inputs and output signals 49

4.14 Determination of intended region of operation 50

4.15 Frequency profiles of chirp input signals (Test 5) 50

4.16 Input signals of an operating point 51

4.17 Test 5 inputs and output signals 51

5.1 Motoring Tests 53

5.2 Representing the friction torque as a quadratic function of speed . . 54

5.3 Determination of inertia torque constant 55

5.4 A typical estimated loss and dyno brake torque 55

5.5 Nonlinear finite impulse response (NFIR) model 56

6.1 A typical NARX structure 59

6.2 NARX structure with OLS based input regressor selection 59

6.3 NARX structure with a single output feedback 60

6.4 Distributions of (a) training, (b) steady-state validation and (c) transient validation performances. 62

6.5 (a) Training, (b) steady-state and (c) transient validation performances by changing the input regressors number 63

6.6 (a) Training, (b) steady-state and (c) transient validation performances by changing the unit number 63

6.7 3D surface plots of (a) training, (b) steady-state and (c) transient validation performances 64

6.8 Obtained map for parameter selection 65

7.1 Logarithmic normalization of soot measurements 67

7.2 Gated Recurrent Unit Network structure 68

7.3 LSTM network structure 69

7.4 Sigmoid and tanh activation functions 70

8.1 Experimental setup 72

8.2 Sampling in the dilution tunnel 73

8.3 Operating principle of the AVL Micro Soot Sensor 74

8.4 Ind. torque model with 4 inputs prediction performances (Ex.1) . . 77

8.5 Ind. torque model with 4 inputs prediction performances (Ex.2) . . 78

8.6 Ind. torque model with 4 inputs prediction performances (Ex.3) . . 79

8.7 Ind. torque model with 4 inputs prediction performances (Ex.4) . . 80

8.8 Ind. torque model with 8 inputs prediction performances (Ex.1) . . 81

8.9 Ind. torque model with 8 inputs prediction performances (Ex.2) . . 82

8.10 Ind. torque model with 8 inputs prediction performances (Ex.3) . . 83

8.11 Ind. torque model with 8 inputs prediction performances (Ex.4) . . 84

8.12 Ind. torque model with 8 inp. prediction performances (E.1+2) . . 85

8.13 Ind. torque model with 8 inp. prediction performances (E.2+3) . . 86

8.14 Est. performance distribution of the models Left:classic Right:proposed 88

8.15 WHTC-1 validation performance distribution of the models Left:classic Right:proposed 88

8.16 WHTC-2 validation performance distribution of the models Left:classic Right:proposed 89

8.17 Total number of selected input regressors 90

8.18 Determination of significant regressors by error reduction ratio . . . 90

8.19 Estimation and validation performances of the best NOx emission model in Table 8.6 91

8.20 NOx emission model prediction performances (Experiment 1) 93

8.21 NOx emission model prediction performances (Experiment 2) 94

8.22 NOx emission model prediction performances (Experiment 3) 95

8.23 NOx emission model prediction performances (Experiment 4) 96

8.24 Training performance of the model with 6 input regressors and 15 units (below: zoomed version) 97

8.25 Validation performances of the model with 6 input regressors and 15 units (a) steady-state cycle (b) transient cycle, (c) zoomed-in version of transient cycle 97

8.26 Training performance of the model with 5 input regressors and 9 units (below: zoomed version) 98

8.27 Validation performances of the model with 5 input regressors and 9 units (a) steady-state cycle (b) transient cycle, (c) zoomed-in version of transient cycle. 98

8.28 Training performance of the model with 8 input regressors and 12 units (below: zoomed version) 99

8.29 Validation performances of the model with 8 input regressors and 12 units (a) steady-state cycle (b) transient cycle, (c) zoomed-in version of transient cycle 99

8.30 Design of experiments in SPD-QNT plane (Test 5) 100

8.31 Prediction performances of the models with 15 units (Case 1) . . . 102

8.32 Prediction performances of the models with 15 units (Case 2) . . . 103

8.33 Training performance of the LSTM soot model 105

8.34 Zoomed version of Figure 8.33 105

8.35 Cumulative training response of the model 105

8.36 NEDC validation performance of the LSTM soot model 106

8.37 Zoomed version of Figure 8.36 106

8.38 Cumulative NEDC validation response of the model 106

8.39 WLTC validation performance of the LSTM soot model 107

8.40 Zoomed version of Figure 8.39 107

8.41 Cumulative WLTC validation response of the model 107

8.42 Training performance of LSTM soot model (T5 - Case 1) 109

8.43 Zoomed version of Figure 8.42 109

8.44 Cumulative training response of the model (T5 - Case 1) 109

8.45 Validation performance of soot model (T5 - Case 1) 110

8.46 Zoomed version of Figure 8.45 110

8.47 Cumulative validation response of the model (T5 - Case 1) 110

8.48 Training performance of LSTM soot model (T5 - Case 2) 111

8.49 Zoomed version of Figure 8.48 111

8.50 Cumulative training response of the model (T5 - Case 2) 111

8.51 Validation performance of soot model (T5 - Case 2) 112

8.52 Zoomed version of Figure 8.51 112

8.53 Cumulative validation response of the model (T5 - Case 2) 112

List of Tables

1.1 European Union emission standards for heavy duty vehicles in steady-state testing 2

1.2 European Union emission standards for heavy duty vehicles in transient testing 3

4.1 Model inputs for indicated torque and combustion emissions 36

4.2 Correlations between input channels 41

4.3 Correlations between designed input signals (Experiment 1) 43

5.1 Coefficients of friction loss torque function ($\tau_{fr}(\omega)$) 54

6.1 Ranges of parameters 61

8.1 Test engine spesifications 73

8.2 Indicated torque models with a single input (QNT) 75

8.3 Indicated torque models with 4 inputs 76

8.4 The best indicated torque models with all experiments 76

8.5 Ranges of parameters 87

8.6 The best 6 models of the proposed structure 89

8.7 Corresponding models with classic structure 89

8.8 Extracted regressors for each input channel ($nb = 7$) 90

8.9 The best NOx emission models for each experiment 92

8.10 Selected model performances using merged experiment (1+2) 92

8.11 Performances of NARX models 101

8.12 Performances of GRU models 101

8.13 Performances of LSTM soot model with merged data 104

8.14 Performances of LSTM soot model with Test 5 108

8.15 Execution times of the indicated torque models 113

8.16 Execution times of the NOx emission models 113

8.17 Execution times of the GRU and LSTM based soot models with Test 5 114

8.18 Execution times of the LSTM based soot model with combined tests 114

Chapter 1

Introduction

Even though the recent advances in emerging electrical engine and battery technologies are creating clean alternatives, diesel engines are still widely preferred for heavy duty vehicles in the market (Figure 1.1). In 2018, the worth of the global diesel engine market was US$ 215 Billion and it is expected to reach a value of US$ 291 Billion by 2024 [1].

FIGURE 1.1: Diesel engine for heavy-duty vehicles

The main reason of the demands for diesel engines is their efficieny in energy conversion. Thanks to its fuel property of being non-volatile, diesel fuel burns gradually, which serves for transmitting better torque and increasing the fuel efficiency. Furthermore, reliability and durability with higher stroke length and higher compression ratios compared to petrol engines help in generating tremendous torque at low engine speeds, which is a crucial need for heavy-duty vehicles.

Due to diesel's fuel property of being non-volatile, the main principle in diesel engines relies on the compression of the air-fuel mixture instead of ignition, which also brings some difficulties to procure optimal engine operating conditions. Inefficiency in the combustion process results in the release of hazardous emissions to the environment such as Nitrogen Oxides (NOx), Particulate Matter (PM) (also known as soot), Hydrocarbons (HC) and Carbon Oxides (COx) [2]. These emissions cause several health problems and contribute to global warming [3]. Azama et al. [3] reported that majority of the diesel engine exhaust outflows are NOx emissions with the rate above more than 50% and the second highest percentage belong to soot emissions.

In order to decrease the level of hazardous emissions produced by new vehicles, the European Union sets some directives ever since 1992. NOx and soot emission standards set by EU from 1992 to 2013 for heavy-duty vehicles in steady-state and transient testing are tabulated in Table 1.1 and 1.2, respectively [4].

Stage	Date	Test	NOx (g/kWh)	PM (g/kWh)
EURO 1	1992, $\leq$85 kW	ECE R-49	8.0	0.612
	1992, >85 kW		8.0	0.36
EURO 2	1996.10		7.0	0.25
	1998.10		7.0	0.15
EURO 3	2000.10	ESC & ELR	5.0	0.10
EURO 4	2005.10		3.5	0.02
EURO 5	2008.10		2.0	0.02
EURO 6	2013.01	WHSC	0.4	0.01

TABLE 1.1: European Union emission standards for heavy duty vehicles in steady-state testing

Stage	Date	Test	NOx (g/kWh)	PM (g/kWh)
EURO 3	2000.10	ETC	5.0	0.16
EURO 4	2005.10		3.5	0.03
EURO 5	2008.10		2.0	0.03
EURO 6	2013.01	WHTC	0.46	0.01

TABLE 1.2: European Union emission standards for heavy duty vehicles in transient testing

Decreasing the limits for emissions necessitated to include more dynamicity in regulation test cycles. As shown in Table 1.2, European Transient Cycle (ETC) was replaced by the World Harmonized Transient Cycle (WHTC) that includes more dynamic characteristics. With Euro 6.2 regulations [5] for light-duty commercial vehicles and passenger cars, the New European Driving Cycle (NEDC) test was also replaced by Worldwide Light-duty Test Cycle (WLTC), which introduces more transient maneuvers and covers a more extensive operating range with the full load region. In addition to these tests, On-Road Emission tests, In-Service Conformity (ISC) and Real Drive Emissions (RDE) tests are also implemented for heavy-duty vehicles, where these tests highlight the importance of transient operations in the optimization of a combustion [6].

1.1 Motivation

Primarily, engine manufacturers consider the efficiency of engine performances such as torque generation and fuel consumption. However, aforementioned restrictive regulations increase the number of constraints in the calibration process. Moreover, the complexity of the equipment utilized in modern diesel engine systems such as air handling, fuel injection, combustion and after treatment increases the effort to find the optimal operating points and obliges to employ high fidelity transient models in the calibration process (Figure 1.2).

Main objectives of a calibration process and the motivation of this thesis in terms of obtaining high fidelity engine models can be summarized under three topics, namely **Emissions**, **Efficiency** and **Expenditure**.

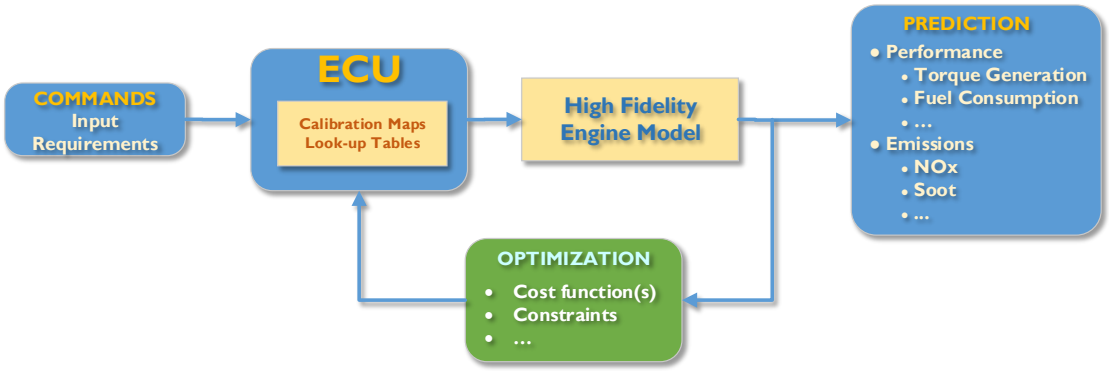


FIGURE 1.2: Model based calibration process

***Emissions*:** European Union emission standards become more and more stringent, and the limits for exhaust emissions are exponentially decaying. As the emission laws getting tougher, engine manufacturers need more hardware and controllers, which increases the number of calibration parameters. In addition to the limits for allowable emissions, types of the test cycles also change and highlight the importance of transient operations. Steady-state calibrations fall behind to meet the requirements of recent legislations, which require dynamic calibration with high fidelity transient exhaust emissions models.

***Efficiency*:** In order to survive in the competition among engine companies, manufacturers must consider the performance of their engines in terms of torque generation and fuel consumption. These considerations reveal themselves as some objective functions in the model-based calibration process.

***Expenditure*:** In steady-state calibrations, the combination of calibration parameters increases dramatically, and performing experiments at every point in the intended operating region sometimes take several weeks. Considering the cost for 1-hour usage of an experimental setup equipped with a dynamometer around €100, the total cost of the experimentation for steady-state calibration become incredibly huge. On the other hand, appropriately designed experiments and sufficiently accurate models that capture the system dynamics in both steady-state and transient cycles can reduce the calibration time dramatically to a few hours.

1.2 Contributions of the Thesis

Contributions of the thesis are highlighted below.

- A new experiment design is proposed where input channels are excited by chirp signals with varying frequency profiles in terms of the number and directions of the sweeps. This is a strong alternative to the steady-state experiment based approaches to reduce testing time considerably and improve modeling accuracy in both steady-state and transient conditions.
- Sufficiently accurate optimization-oriented indicated torque models are obtained using nonlinear finite impulse response (NFIR) structure.
- Two different NARX based NOx emission modeling approaches varying in terms of the input regressor selection and the utilization of output feedback are introduced.
- Sensitivity analysis of the NOx emission models to parameter changes is conducted by generating models for different values of parameters and a parameter selection method using an easy-to-interpret map is proposed. The map can be a convenient means to select the required parameters for diesel engine NOx emissions modeling with limited testing time in powertrain development.
- A new data-driven modeling of diesel engine soot formation using LSTM type networks is developed as an alternative to the classical NARX type recurrent structures which are insufficient to generate accurate and smooth soot predictions.
- The effectiveness of the proposed models in terms of prediction accuracies in both NEDC (New European Driving Cycle) and WHTC (World Harmonized Transient Cycle) cycles are demostrated with real data from the engine dynamometer tests.

1.3 Outline of the Thesis

- **Chapter 2:** reviews the modeling approaches proposed for diesel engine torque and combustion emissions such as NOx and soot, and presents a background about the black-box modeling of dynamical systems as classical system identification models and recurrent neural network based modern network structure.
- **Chapter 3:** highlights the importance of model based dynamic calibration of diesel engines, and explains the diesel combustion process.
- **Chapter 4:** details the design of experiments for dynamic calibration of diesel engines. It includes the model input selections for indicated torque and combustion emissions and selection of regressors for nonlinear autoregressive with exogenous input (NARX) models. A new experiment design that inlcudes chirp based excitation signals for input channels is also presented.
- **Chapter 5:** provides the details of proposed diesel engine torque modeling such as predicting indicated torque using the estimations of friction, pumping and inertia torques in addition to the torque measured from the engine dynamometer, and modeling indicated torque by a particular type of NARX structure called nonlinear finite impulse response (NFIR).
- **Chapter 6:** demonstrates two different NARX based NOx emission modeling approaches varying in terms of the input regressor selection and the utilization of output feedback. It also covers the proposed sensitivity analysis and parameter selection method.
- **Chapter 7:** introduces new data-driven modeling of a diesel engine soot formation using gated recurrent unit (GRU) and long short term memory (LSTM) networks as an alternative to classical NARX type recurrent structures which are insufficient to generate accurate and smooth soot predictions.
- **Chapter 8:** presents the experimental results and discussions regarding the models of diesel engine indicated torque, NOx and soot emissions.

- **Chapter 9:** concludes the thesis with several remarks and indicates the future directions.

1.4 Publications

The following publications are produced from this thesis work.

- Predicting NOx Emissions in Diesel Engines via Sigmoid NARX Models Using A New Experiment Design for Combustion Identification, **G. Alcan**, M. Unel, V. Aran, M. Yilmaz, C. Gurel, K. Koprubasi, *Measurement*, 137, 71-81, 2019.

- Estimating Soot Emission in Diesel Engines Using Gated Recurrent Unit Networks, **G. Alcan**, E. Yilmaz, M. Unel, V. Aran, M. Yilmaz, C. Gurel, K. Koprubasi, *9th IFAC International Symposium on Advances in Automotive Control* (AAC 2019), Orléans, France, June 23-27, 2019.

- Diesel Engine NOx Emission Modeling Using a New Experiment Design and Reduced Set of Regressors, **G. Alcan**, M. Unel, V. Aran, M. Yilmaz, C. Gurel, K. Koprubasi, *IFAC-PapersOnLine*, 51 (15) 168-173, 2018.

- Driving Behavior Classification Using Long Short Term Memory Networks, M. E. Mumcuoglu, **G. Alcan**, M. Unel, O. Cicek, M. Mutluergil, M. Yilmaz, K. Koprubasi, *4th International Conference of Electrical and Electronic Technologies for Automotive* (AEIT AUTOMOTIVE 2019), Torino, Italy, July 2-4, 2019.

- Optimization-Oriented High Fidelity NFIR Models for Estimating Indicated Torque in Diesel Engines, **G. Alcan**, V. Aran, M. Unel, M. Yilmaz, C. Gurel, K. Koprubasi, Journal Paper, (under preparation)

- Modeling Soot Emissions in Heavy-Duty Diesel Engines by LSTM Networks, **G. Alcan**, M. Unel, Journal Paper, (under preparation)

1.5 Nomenclature

Abbreviation	***Description***
ADAM	Adaptive Moment Estimation
AFR	Air-Fuel Ratio
ANN	Artificial Neural Network
ARMAX	Autoregressive Moving Average Exogenous
ARX	Autoregressive Exogenous
BSFC	Break Spesific Fuel Consumption
BTE	Break Thermal Efficiency
CMF	Charge Mass Flow
COx	Carbon Oxides
CRDI	Common Rail Diesel Injection
CVR	Constant Volume Sampling
DI	Direct Injection
DoE	Design of Experiments
ECE	Economic Commission for Europe
ELR	European Load Response
ECU	Engine Control Unit
EGR	Exhaust Gas Recirculation
ERR	Error Reduction Ratio
ESC	European Stationary Cycle
ETC	European Transient Cycle
EU	European Union
FAR	Fuel-Air Ratio
FRF	Frequency Response Function
GMF^{-}	Gas Mass Flow before EGR feed
GRU	Gated Recurrent Unit
HC	Hydrocarbons
HD	Heavy-Duty
IntO2R	Intake Oxygen Ratio
ISC	In-Service Conformity

Abbreviation	***Description***
LOLIMOT	Local Linear Model Tree
LPM	Local Polynomial Models
LPV	Linear Parameter Varying
LSTM	Long Short Term Memory Network
MAF	Mass Air Flow
MAP	Manifold Absolute Pressure
MFB	Mass Fraction Burnt
MISO	Multi Input Single Output
miSOI	Main Start of Injection
miQNT	Main Injection Fuel Quantity
MVM	Mean-Value Model
NARX	Nonlinear Autoregressive with Exogenous Input
NARMAX	Nonlinear Autoregressive Moving Average with Exogenous Input
NEDC	New European Driving Cycle
NFIR	Nonlinear Finite Impulse Response
NN	Neural Network
NOx	Nitrogen Oxides
NRMSE	Normalized Root Mean Squared Error
OLS	Orthogonal Least Square
PCA	Principle Component Analysis
piSOI	Pilot Start of Injection
piQNT	Pilot Injection Fuel Quantity
PM	Particulate Matter
QNT	Total Fuel Quantity
RailP	Rail Pressure
RBF	Radial Basis Function
RDE	Real Drive Emissions
RNN	Recurrent Neural Network
SOI	Start of Injection
SPD	Engine Speed

Abbreviation	***Description***
SVD	Singular Value Decomposition
VGT	Variable Geometry Turbocharger
WHSC	World Harmonized Stationary Cycle
WHTC	World Harmonized Transient Cycle

Chapter 2

Literature Review and Background

The first three sections of this chapter review the modeling approaches proposed for diesel engine torque and combustion emissions such as NOx and soot, respectively. Generally, these studies are categorized based on the level of process information employed in the modeling as white, grey and black box identification. The last section of this chapter examines the black-box modeling of dynamical systems in two categories, namely classical system identification models and recurrent neural network (RNN) based modern network structures.

2.1 Diesel Engine Torque Modeling

An accurate indicated torque model of internal combustion engines is beneficial for the evaluation of engine combustion conditions, engine performance optimization, faults diagnosis, and automotive transmission control. Among the several existing studies, nonlinear observers and neural network-based approaches are much more dominating and commonly used.

Togun and Baysec [7] presented an artificial neural network (ANN) model to predict the brake specific fuel consumption (BSFC) and the torque of a gasoline

engine. They performed several experiments on four-stroke 1.3 gasoline Fiat Tofaş 131 engine for both training and validation. Spark advance, throttle position and engine speed are selected as inputs to model torque and brake specific fuel consumption. Determining the best network architecture is achieved by trying a different number of hidden layers and neurons. They decided to use 2-layer architecture such as 3-13-1 and 3-15-1 NN torque and brake specific fuel consumption, respectively. They also presented genetic programming-based modeling [8] for the same experimental data. Both methodologies show satisfactory validation performances, but they took all the measurements under steady-state conditions, and the performance of the obtained models under transient cycles is not validated.

Rizzoni and Zhang [9] discuss a method for the identification of a nonlinear model of the dynamics relating combustion pressure to crankshaft angular velocity. They claim that the proposed model can be utilized in the implementation of control methods that necessitate estimated values of individual cylinder indicated torque, which is not easily measured directly with low-cost equipment. Their methodology requires an initial known model structure and employs nonlinear programming to extract relevant model parameters. Once the parameters are estimated, they employed the model to design an input observer for the estimation of indicated torque.

A linear parameter varying (LPV) model is proposed by Wei and Del Re [10] as a new approach based on a nonlinear subspace identification method. Engine torque is considered as the output while a nonlinear function of the accelerator pedal position is taken as the input variable. There exist two main approaches to identify LPV systems: the input-output method [11] and the state space method [12]. In [12], estimation of engine system states and system order are realized by the state space method based LPV subspace identification algorithm.

Chen and Moskwa [13] developed a new nonlinear approach to estimate the engine cylinder pressure, the combustion heat release and the engine torque with fast convergence and stability. In their approach, they modify the nonlinear sliding observer gains as a function of crank angle. They applied this technique to

a six-cylinder engine and showed the application possibility of such nonlinear observer to individual pressure estimations. Wang and Chu [14] also applied three nonlinear observers to estimate the indicated torque and compare their results. These nonlinear observers are high gain observer, classical sliding mode observer and second-order sliding mode observer. The validity of the indicated torque observers is examined by a car with four-cylinder in-line fuel-injected 1.8 lt engine and five-speed manual transmission. Based on experimental results, they conclude that the second-order sliding mode observers are more robust and exact, and the classical sliding mode observers are slightly more accurate than the high gain observers. However, the undesired chattering effect is a commonly known disadvantage of these observers.

A non-parametric frequency response functions (FRFs) based method was developed by Ali and Saraswati [15] to determine cylinder pressure, instantaneous indicated torque and load torque using crankshaft speed fluctuations. They initially compute the frequency response functions at various operating points on a grid of engine speed and load. Then, they use a discrete Fourier transform of the crankshaft speed, mean speed and manifold pressure as input set and employ a multilayer neural network to output an FRF between crankshaft speed and cylinder pressure. Later on, they introduced a sliding mode observer [16] to estimate the cylinder pressure using crankshaft speed fluctuations and compared the performances of the new proposed system with FRF based observers. They showed that using sliding mode observer leads to reduce the mean relative error for all the test conditions in comparison with the previous method.

Zhixiong et al. [17] proposed a discrete sliding mode observer to estimate internal combustion engine indicated torque from its crankshaft speed fluctuation. In order to figure out the interaction between the engine torque and instantaneous speed, they established a dynamic model of a six-cylinder internal combustion engine crankshaft. Even though discretization and linearization bring some simplifications and enable one to implement the algorithm in real time, there is still need for a better decomposition method that can solve the nonlinear equations efficiently without simply linearizing the system.

Different from the studies regarding multi-cylinder diesel engines, Zweiri and Seneviratne [18] utilized crankshaft and coupling angular velocity measurements in a nonlinear observer to estimate the indicated torque and load torque of a single-cylinder diesel engine. Instantaneous friction model and the effects of inertia variations of the crankshaft were also employed in the nonlinear dynamic engine model. They showed the stability proof of the observer and satisfactory performances in both simulations and experiments. In order to make indicated torque estimations to be usable in real-time applications, they also employed artificial neural network structure [19] for the engines with a single cylinder.

Brahma et al. [20] investigated a modeling approach for the response of brake torque and NOx emissions of a high-speed common-rail diesel engine. They presented an approach for modeling the combined dynamic influence of fuel quantity and timing on brake torque and NOx in order to come up with multivariate control-oriented models. An existing physical model was extended, and the parameters were identified at different engine operating points. Although torque models' fit accuracy on validation data set is around 87%, resultant signals over time show that the proposed models are still inefficient to capture the steady behavior.

2.2 Diesel Engine NOx Emission Modeling

Engine manufacturers are forced to find new solutions for meeting the requirements of maximum acceptable emission values due to the increasingly tightened emission regulations. In order to reduce the diesel engine NOx emissions, new after-treatment systems [21] are introduced and appropriate combustion control techniques [22] are developed. In addition to the precautions for the dangerous effects of engine emissions, ever increasing prices of fuel and the competition among vehicle manufacturers make it crucial to seek the optimal engine operating conditions. In accordance with these purposes, reducing emissions and increasing the

generated power necessitate the employment of sufficiently accurate engine combustion models in powertrain development, testing and calibration to optimize the engine components [23].

Studies of engine NOx emissions modeling in the literature can be reviewed in three groups as white, grey and black-box modeling according to the level of available information [24]. Some studies focus on the phenomenological nature of the NOx formation and present physics or chemistry-based models, which is called white box modeling. On the other hand, grey-box modeling exploits both empirical data and physical interpretation of the process. Finally, black-box modeling based methods tackle the NOx modeling problem by exploiting the potential of empirical data and try to find a global model or multiple local models which explain the relation between measured input and output signals.

Rezaei et al. [25] presented phenomenological combustion and a tabulated NOx model for a heavy-duty (HD) diesel engine. Local processes such as injection, spray formation, ignition, combustion and NOx formation were described by physical and chemical sub-models. It is seen that the generalization performances of such local physics-based methods are not promising. Perez et al. [26] developed a detailed nonlinear engine simulator based on the first principles, and they proposed identification and control schemes such as Hammerstein Wiener models and model predictive control. However, this method was offered for a small region of operation points and Hammerstein models were problematic due to saturation of inputs. Besides, proposed algorithms were implemented on the engine simulator rather than a real diesel engine. Raptotasios et al. [27] investigated the potential of NOx emission reduction for two-stroke marine engines using EGR. They adapted the multi-zone combustion model developed for high-speed DI diesel engines to a two-stroke marine diesel engine. They showed the significant effect of EGR to reduce NOx emissions, but they validated the proposed model after calibrating the engine only at one load point.

Hirsch et al. [28] proposed a combined grey-box modeling approach for engine

emissions. In their work, the identification of static maps was performed numerically, whereas the cylinder-head temperature and the air path dynamics were included in the model based on physical assumptions. By doing so, they claimed that they avoided the emission models to be too complicated or poorly parametrized to be used online. Different from the most models developed in the literature, their approach is based on pure engine control unit (ECU) data, so their strategy can provide a real-time estimation of NOx emission and particulate matter as a function of ECU outputs. Despite a wide range of validity as well as high accuracy in transient cycles, the fit performance of the proposed model in highly dynamic operations is not sufficient. Quérel et al. [29] presented a Mean-Value Model (MVM) of diesel engine NOx emissions to use in model-based engine control design and online NOx emission estimation. Mainly, a zero-dimensional thermodynamic model was simplified by only selecting the dominant factors in NOx formation. Due to the limited number of experimental data, they only identified the parameters regarding the burned gas temperature model and the mean-value NOx formation model but achieved satisfactory validation performance compared to purely physical models. However, this study is limited in the steady-state cycles, and it is not validated with the transient tests. Asprion et al. [30] also presented a model for the nitrogen-oxide emissions of diesel engines that combines the advantages of the model types of empirical and phenomenological nature, so meets the standard requirements of computationally intensive fields like dynamic optimization and model-based control. They claim that the execution speed is roughly 500 times faster than real-time and relative errors are below 10%. In [31], they enhanced their work and used their method in a mean-value model for the air path, which enabled one to predict the NOx emissions based on the control signals only. Although their model is very accurate and fast, it is not easy to corporate the combustion characteristics changes in such studies.

Due to the assumptions and the simplifications, physical and semi-physical models do not have enough generalization capabilities. Purely data-driven approaches do not consider the knowledge regarding the dynamics of the underlying phenomena, and the models are derived by performing adequate experiments. Benatzky et al.

[32] proposed a static polynomial black-box modeling approach for the design and the evaluation of NOx models for heavy-duty off-road engines. They mainly investigate three approaches, differing in the chosen sets of input regressors. The first method, called ECU approach utilizes only the quantities available on standard production-type ECUs in order to corporate the obtained model on a production engine. Selected inputs in this approach are engine speed, the total amount of fuel injected, air mass flow (MAF), intake manifold absolute pressure (MAP), start of the main injection, rail pressure and exhaust oxygen concentration. Unlike taking input variables directly from air and fuel path, in-cylinder pressure trace is utilized in the second method called SVD approach. In this method, the high dimensionality of the cylinder pressure trace measurement is reduced by a type of principal component analysis (PCA) called singular value decomposition (SVD). Last method (HRL approach) applies only combinations of pressure trace, heat release and engine speed as regressors. Authors compared these three methods and combined them to obtain better models. Proposed approaches perform well in steady state cycles, but it is not ensured to capture the transient behavior due to the usage of static mapping. Moreover, the authors reported the possibility of obtaining large deviations between the measurements and the predictions for higher NOx values based on the sudden engine speed changes.

A linear state-space model for the dynamics of a six-cylinder heavy-duty engine was proposed by Henningsson et al. [33]. The input set of their proposed method consists of fuel injection duration and timing, variable geometry turbo (VGT) and exhaust gas recirculation (EGR). Their model outputs are not only NOx and soot emissions, but also peak pressure derivative, combustion phasing and indicated mean effective pressure. In order to reduce the local linear model number at each operating, Wiener models were employed. The proposed method was targetted to capture the behavior of highly nonlinear dynamical process by utilizing several local linear models and employing a clustering algorithm to select the best subset. However, resultant predictions in high transient cycles were not satisfactory. Formentin et al. [34] proposed a principal component analysis and L2

regularization-based modeling technique to estimate NOx emissions of a heavy-duty diesel engine. Their model includes engine speed and indicated pressure measurements in the input set. Their methodology is suitable for after treatment and closed-loop combustion control since the capability is satisfactory with 0.48% normalized mean error in static cases, but some improvements are required for transient cycles.

As a new input design framework, Boz et al. [35] offered multi-sweep chirp based excitation signals for airpath input channels of a diesel engine. They explored the modeling performances of classical linear and nonlinear system identification techniques. The authors concluded that nonlinear autoregressive with exogenous input (NARX) neural network models are better than linear models for such highly nonlinear process. In their work, NOx emissions were modeled as a function of airpath input channels with a validation accuracy around 80%. Roy et al. [36] presented an artificial neural network based models of NOx, soot and CO2 emissions, Brake Thermal Efficiency (BTE) and Brake Specific Fuel Consumption (BSFC) for a Common Rail Diesel Injection (CRDI) type engine. Load, fuel injection pressure, EGR and fuel injected per cycle were considered as inputs to the network. They claimed that the mean absolute percentage error was in the range of 1.1 – 4.57% with noticeably low root mean square errors under steady-state cycles, but the prediction performance of the proposed model under transient cycles was not validated. Highly dynamic relations between inputs and outputs of a diesel engine may not be captured with the fully connected ANN models due to the lack of recurrent structure.

2.3 Diesel Engine Soot Emission Modeling

High fidelity NOx and soot emission models for transient cycles of diesel engines play a decisive role in the reduction of the experimentation time and cost during the calibration of the engine parameters. Pfeifer et al. [37] claimed that the measurement from both the internal combustion and after-treatment system should be

taken into consideration thoroughly and some significant developments should be done for a safe operation under all conditions in terms of engine and environment in order to meet stringent emission legislation. As in the case of NOx emission modeling, several studies in the literature focus on physics or chemistry-based (white or grey box) and empirical (black box) soot modeling.

Walke et al. [38] utilized fuel consumption, manifold pressure, EGR ratio, exhaust CO2, O2, transient cycle, manifold temperature, volumetric efficiency, SOI and DOI to propose sub-models for cylinder pressure, two zone temperature, NOx and soot emissions. The authors took the Hiroyasu model [39] as the base soot model, and the experimental data was employed in the correction of the model. Tanelli and Maranta [40] compared three semi-empirical soot models namely Moss [41], Lindstest-Leung [42] and Wen [43] for internal combustion engine simulations. In order to predict soot emissions, they also proposed an extension of representative interactive flamelet combustion model. It should be noted that the phenomenological and the CFD models usually utilize crank angle and try to describe the in-cylinder processes in detail, which indeed necessitate high computational cost, so they are not feasible especially for simulating transient cycles of a diesel engine.

Benz et al. [44] utilized a symbolic regression algorithm to derive a nonlinear extended quasi-static model for raw emissions of diesel engines. Genetic programming and artificial neural network (ANN) based input variable selection algorithm were employed to choose the input set. Their raw emission model takes the cylinder mass, the air fraction, rail pressure, 50% and 90% mass fraction burnt (MFB) expressed in crank angle degrees, fuel mean pressure and mean piston speed as inputs and predicts NOx and particulate matter (PM) as outputs. The performances of NOx emission models are entirely satisfactory, but the regression of the soot emission was worse due to the relative high measurement errors. Tschanz et al. [45] presented a control oriented, easily identifiable and portable model for diesel soot emissions. As in the work of [44], the engine-out PM emissions were assumed to be quasi-statically influenced by the conditions inside the cylinder. PM emissions were measured by a photo-acoustic soot sensor (AVL micro soot sensor) and modeled as relative deviations of stationary base maps. In order to

estimate the inputs' effects on the emissions, a polynomial model was used. Easy identifiability, low computational burden, and usage of inputs that are easily obtainable from the ECU are advantageous of the proposed method. However, fewer extrapolation capabilities and a relatively high number of parameters are some certain disadvantages of employing a polynomial approach in such cases. Ericson and Westerberg [46] presented another quasi-stationary approach to model the fuel consumption, CO, HC, NOx and PM emissions of euro 3-class engines. The authors reported the need for transient correction methods to obtain models with better prediction performance. They also suggested to include speed and torque signals in the input set for better generalization capability. Unfortunately, their approaches are primarily useful for predicting the emissions of heavy-duty engines during real driving conditions and are not offered to be used in engine control or optimization. In order to obtain more consistent results and employ the proposed approach in engine control tasks, the delay time estimation must be substantially improved.

Several adaptive local polynomial models constitute a global model structure for soot emissions in work presented by Sequenz et al. [47]. Intake temperature, charge air pressure, the engine operation point, the location of mass fuel burnt (MFB) 50% and the air mass flow rate were considered as model inputs. Mallows' Cp-statistics [48] based regressor selection method was used to select the appropriate parameters and decrease the variance error in the estimation. Admissible set of regressors were employed in the training of local polynomials for each operation point defined by engine speed and injection quantity. In this work, a strong nonlinear connection between soot formation and the air-fuel ratio was highlighted with supporting results. Soot prediction performance of the proposed method is not satisfactory when compared to NOx emission estimations. Mrosek et al. [49] included air path variables at the intake manifold and the location of mass fraction burnt (MFB) 50% in the input set and proposed a stationary batch process for modeling of the simplified combustion process and the emission formation. Diesel engine emissions were modeled under steady-state cycles at the points defined by engine speed and the desired fuel injection quantity. In their work, raw emission outputs

such as NOx and soot were modeled as a function of the air path signals and the combustion characteristic (MFB50) by using local polynomial models (LPM) [47]. Experimental results regarding NOx emissions were quite satisfactory, but the prediction performance of the soot models still need to be improved.

Hafner et al. [50] utilized a special local linear model tree (LOLIMOT) with radial basis function (RBF) [51] in the stationary soot formation modeling and they compared two different input set. In the first set the engine control settings, injected fuel quantity, injection angle and engine speed were included in the input set, whereas the selected inputs in the second model were the characteristics of the measured cylinder pressure signal. The authors reported that the static soot models with only cylinder pressure characteristics had shown a comparable performance with the first model. Grahn et al. [52] proposed another local linear regression models for diesel NOx and soot emissions of a 5-cylinder Volvo passenger car in their work. Two-dimensional look-up tables defined the regression parameters of local linear models. The authors then interpreted them as a B-spline function and present to find globally optimal model parameters by solving a linear least-squares problem. The authors reported that the global equivalence ratio is a dominant input channel in soot emission formation and the other inputs proposed in soot modeling are engine speed, injected fuel amount, injection timing and partial pressure of oxygen in the intake. In steady-state operation, their soot models were verified with an average relative error of 29%. Atkinson et al. utilized neural networks in transient engine emissions modeling [53, 54]. The authors selected engine speed, engine temperature, MAP, manifold air temperature and the engine control inputs to model the emissions such as NOx, soot, HC, CO and CO2. However, the operational range of this study is very narrow and different from the other emissions PM prediction performances still need some improvements.

2.4 Modeling of Dynamical Systems

System identification process is described as **white**, **grey** or **black** box based on the level of available information [24]. When the system's dynamics are fully known or can be derived by employing physical laws, the identification process is called ***white-box***. In that case, only the parameter values of the system are identified. In ***grey-box*** identification, combination of partial knowledge of the dynamics and experimental measurements constitute a model. Finally, in ***black-box*** modeling, no knowledge regarding the dynamics of the system is available, and models are derived by performing experiments. Although it is useful to incorporate the available knowledge into the identification process, generally black-box identification is preferred to obtain more generic and widely applicable models.

Knowledge about the characteristics of the process plays a decisive role in the realization of system identification procedures. Once a sinusoidal signal at any frequency excite a system if the output is also a sinusoidal signal with the same frequency than the system is called **linear**, otherwise, **nonlinear** [55]. If the current output of a process depends on both the current and the earlier values of the stimuli, the process is called **dynamic** [56]. This thesis focuses on the identification procedures for nonlinear dynamical systems such as heavy-duty diesel engine torque and combustion emissions.

Identification of nonlinear dynamical systems is a looped process starts with the selection of model inputs. Then it is followed by the design of excitation signals for these input channels. After appropriate experiments performed on the system to be identified, a typical problem of the system identification paradigm is that given observed data $D = \{(\mathbf{x}^i, y^i), i = 1, ..., N\}$ where $\mathbf{x}^i$ is the input vector and y^i is the output of the i^{th} measurement taken from a system, find a description (model) of the system that generated the data. Architecture proposals for the models can be investigated in two groups, such as parametric and nonparametric approaches, which will be detailed later on. Once the properties of the selected approach are assigned, the model is trained with the observed data using various optimization

techniques. Finally, the obtained model is validated with different tests. If it does not pass the validation, i.e., perform insufficient prediction accuracy depending on the intended application, the procedure is backtracked step-by-step. Figure 2.1 depicts the black-box system identification loop for nonlinear dynamical systems.

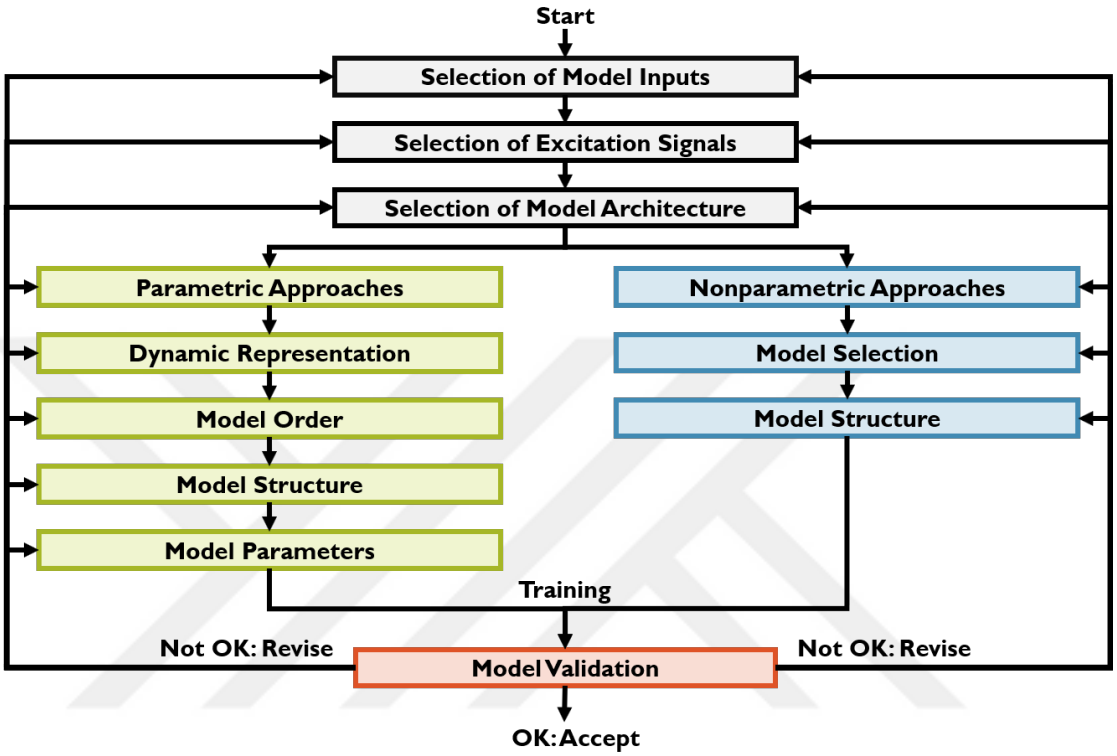


FIGURE 2.1: System identification loop for nonlinear dynamical systems

Selection of Model Inputs: In many mechanical processes, available insights regarding the physics of the process determine the relevant possible set of model inputs. However, increasing number of potential input variables leads to increase in the complexity of the model and decrease the cost efficiency. In such cases, many redundant input variables could be included in the process. One alternative could be trying all input combinations and selecting the best performed set, but obviously this may not always be feasible. In order to exclude non-relevant inputs, supervised, e.g., evolutionary algorithms [57], or unsupervised, e.g., principal component analysis [58], input selection strategies can be employed.

Selection of Excitation Signals: Since black-box modeling solely depends on observed signals, the measurements are the most important source. Process behavior must be represented within the measure data set. To be able to achieve

this, system must be persistently excited by appropriately designed input signals.

Selection of Model Architecture: According to George Box [59] "*Essentially all models are wrong, but some are useful.*". There is no global true model for nonlinear dynamic processes, but the main aim is to find the most useful model depending on the intended application. Model complexity and desired target accuracies are determined by intended use such as simulation, optimization, control or fault detection. Model architectures can be divided into two groups: parametric and nonparametric approaches.

***Parametric Approaches*:** If a finite number of parameters are used in the identification process, it is called parametric modeling. In such cases, dynamical representation is chosen as an initial step. NARX or NARMAX representations are commonly employed in one-step prediction applications. Then the model order and the structure complexity are determined by considering the intended use again. In low order systems, significant behaviors of dynamical system might not be modeled. One should be very careful when increasing the complexity of the model because of bias/variance dilemma.

***Nonparametric Approaches*:** The model of an unknown system is identified without structural information. Basically, nonparametric system identification can be divided into two groups [60]. First group assumes that the system can be modeled by a linear or nonlinear combination of some basis functions with some unknown coefficients. By doing so, system identification problem is transformed to the problem of finding a suitable basis function and coefficient estimation. Obviously, this approach is sensitive to the choice of that basis function. Second group estimates the nonlinear system point by point locally.

The first part of the following section is about classical system identification methodologies which are utilized for the modeling of diesel engine torque and NOx emission in the literature. Due to the incapabilities of classical system identification techniques in obtaining high fidelity soot emission modeling, recurrent neural network (RNN) based modern network structures are studied for soot modeling

within the scope of this thesis. Therefore, the second part of this section gives some background about RNN structures.

2.4.1 Classical System Identification Models

Classical system identification techniques can be grouped into three categories based on their structure as polynomial, block structure, and neural networks. Kolmogorov-Gabor and Volterra series [61] are the earliest classical approaches in nonlinear system identification. Kolmogorov-Gabor models include the regressors from both input and output of the system, whereas in Volterra series only the input regressors are employed. Volterra series/polynomials dominated the early works in nonlinear system identification due to their memory capabilities. Volterra series in discrete time can be written as

$$\begin{aligned} y(k) = h_0 + \sum_{m_1=1}^{M} h_1(m_1)u(k-m_1) \sum_{m_1=1}^{M}\sum_{m_2=1}^{M} h_2(m_1,m_2)u(k-m_1)u(k-m_2) \\ + \sum_{m_1=1}^{M}\sum_{m_2=1}^{M}\sum_{m_3=1}^{M} h_3(m_1,m_2,m_3)u(k-m_1)u(k-m_2)u(k-m_3) + ... \end{aligned} \tag{2.1}$$

where $u(k)$ and $y(k)$ are the measured input and output, respectively. h_l is the l^{th} order Volterra kernel, i.e., l^{th} order nonlinear impulse response.

In polynomial based modeling approaches, dimensionality is a fundamental problem. In order to represent the underlying nonlinear behaviors, mappings in high-dimensions could be obtained by increasing the polynomial order. Higher order polynomials are so sensitive in term of oscillations, especially in dynamical systems. Lastly, polynomial models have unlimited gain for extrapolation, which necessitates being extremely careful in such cases.

Problems of using Volterra polynomials in modeling dynamical system make them unattractive and motivated researchers to look for different model structures. As a result of this situation, block-structured nonlinear models were introduced [62]. A

Hammerstein model consists of a static single-valued nonlinear element and then a linear dynamic (Figure 2.2). On the other hand, a Wiener model consists of a linear dynamic first and then a static nonlinearity (Figure 2.2).

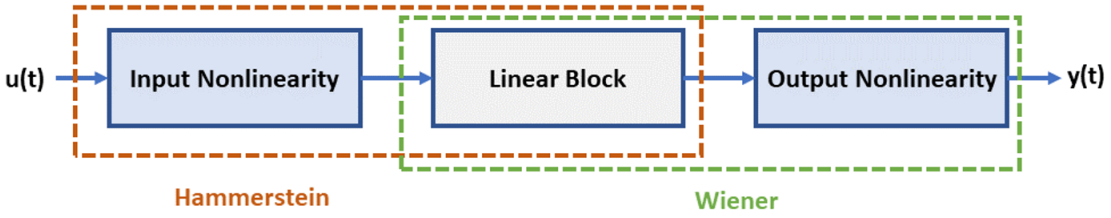


FIGURE 2.2: Hammerstein and Wiener block structures

There exist several static nonlinearities to be used in Hammerstein-Wiener structures such as piecewise linear, one-dimensional polynomial, saturation, dead zone and network structures such as sigmoid and wavelet. In a generic form, many existing linear and nonlinear model types such as ARX, ARMAX, Volterra, and Hammerstein-Wiener can be viewed as special cases of NARMAX [63] defined as

$$\begin{aligned} y(k) = f\Big(&y(k-1), y(k-2), ..., y(k-n_y), \\ &u(k-d), u(k-d-1)..., u(k-d-n_u), \\ &e(k-1), e(k-2)..., e(k-n_e)\Big) + e(k) \end{aligned} \tag{2.2}$$

where $u(k)$,$y(k)$ and $e(k)$ are the system input, output and noise signals, respectively. n_y, n_u and n_e are the maximum lags for these signals. $f(.)$ is the nonlinear function, and d is the pure time delay. Since the output of the system can not be generated as an immediate response of the input, time delay should be at least 1. Nonlinear function $f(.)$ can be realized as power-form polynomial models, rational models, neural networks, fuzzy logic-based models, wavelet expansions, and radial basis function networks.

Over the past few decades, artificial neural networks (ANN) [64] have become very popular and have been applied in various applications. This supervised learning method mainly relies on fully connected neurons and during the training phase, weights are learned iteratively. Optimization process enables one to find optimal parameters which emulate the given data set. Depending on the number of selected hidden layers, they are called either a single-layer network or multi-layer network.

In a system identification perspective, neural networks can be employed to predict the output of the system at the next sampling point. In order to capture the dynamical relations between inputs and output, past values of these variables called "regressors" are also included in the input set of the network. Such networks are called **N**onlinear **A**uto**R**egressive with e**X**ogenous input (NARX) networks [65]. In this thesis, different variations of NARX models are developed for diesel engine torque and NOx emissions, which will be detailed in Chapter 5 and 6,respectively.

2.4.2 RNN Based Modern Network Structures

Simple recurrent neural networks (RNN) were proposed by Elman [66, 67] for the first time in 1990. He proposed to employ the time in the modeling implicitly with its effects rather than explicitly using it. Different from feedforward neural networks which do not have any notion of order in time, RNNs consider the sequences along the time to generate the output. As shown in Figure 2.3, RNN structures have networks with a loop inside to extract the hidden information caused by the time.

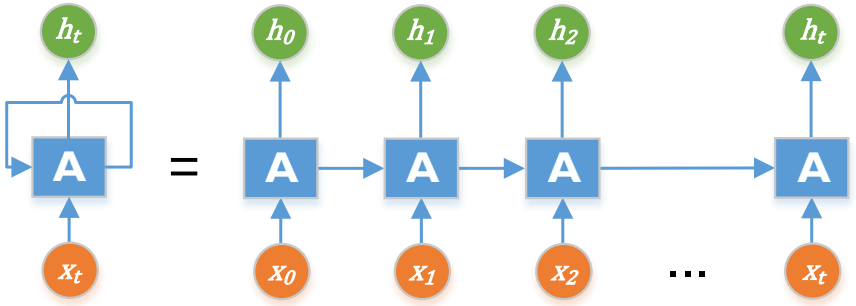


FIGURE 2.3: Looped structure of recurrent neural networks

A simple RNN structure consists of three layers, namely input, hidden, and output layer. Input layer includes a sequence of vectors. The memory capability of the network is defined by the hidden layer, where the states of the system are represented as a set of weights and biases. It should be noted that one of the important lack of NARX structures presented in Section 2.4.1 is being limited to output feedback only. However, this lack is compensated with a looped hidden

layer in RNN. Finally, the output layer of an RNN can be a single value or a sequence of vectors again depending on the scenario.

One of the most important challenges of RNN structures is the problem of vanishing and exploding gradient. In order to model long-term dependencies in a sequence without such problems, more complicated RNN structure called Long-Short Term Memory (LSTM) network was proposed by Hochreiter and Schmidhuber in 1997 [68]. Later on, Gerz et al.[69] extended the LSTM structure with the addition of forget gate.

Cho et al. [70] proposed a new type of LSTM called Gated Recurrent Unit. In this structure, cell and hidden states are merged, forget, and input gates are also combined as an update gate. Since GRU structure has fewer parameters compared to LSTM, it is computationally more efficient [71, 72].

Corporating the effects of time in modeling process makes it possible to use RNN in most of the sequence based applications that can be categorized as one to one, one to many, many to one and many to many as illustrated in Figure 2.4.

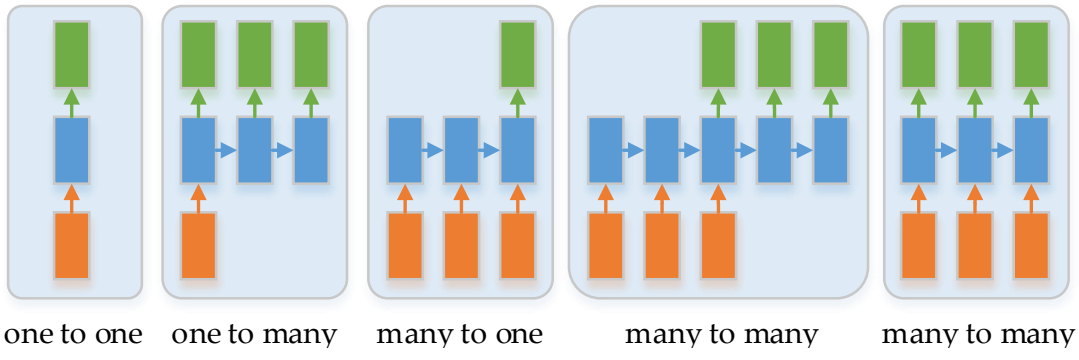


FIGURE 2.4: Different types of RNN applications

Image classification using RNN [73] is an example for one to one application. In this case, the pixel values of an image are considered as a sequence, and a single output (classification result) is obtained by RNN calculations. Image captioning [74] deals with the problem of understanding a single image and generating sentences about it, which is an excellent example for one to many applications of RNN. Sentiment analysis [75] is an example of many to one application of RNN, where the input is a set of words such as the comments of a customer and the

output is a metric that represents the sentiment of the customer as positive or negative. Based on the application, RNN based sentiment analysis can be considered as a classification or regression problem. Language translation [76] is a popular and hot topic in many to many RNN applications. In such applications, the size of the outputs and input need not be always the same. However, there exist some of many to many applications that the lengths of the output and input are synchronized, such as dense video labeling [77].

Details of LSTM and GRU structures, and employing them in the modeling of a dynamical process such as soot emission in diesel engines will be explained in Chapter 7.

Chapter 3

Model Based Dynamic Calibration of Diesel Engines

Calibration of an engine involves the steps to search for optimal engine operating points based on some performance goals and constraints. Strategies for such optimization objectives can be improving the fuel consumption efficiency and increasing the torque generation capability while minimizing the exhaust emissions to meet the requirements over a legislative cycle.

Let's consider a calibration process of a diesel engine over only four parameters such as Injection Timing, Injection Pressure, EGR and VGT, and engine operating region is a discrete speed and load domain represented by a 10×10 matrix for simplicity. Once we assume that each calibratable parameters vary independently and all of them have ten different levels, totally around 10^6 different experimental set points would be required [53]. If we add three more calibration parameters in this process, the total number of potential experiments will increase by a factor of 1000. It should be noted that every single test takes a few minutes until the system converges to steady-state values, which necessitate several weeks to complete the calibration in the whole intended engine operating region.

A considerable amount of studies has been conducted on the steady-state calibration of heavy-duty diesel engines ever since the declaration of first European

Union emission standards in 1992. Stuhler et al. [78] developed some toolboxes that involves an adaptive online design of experiment approach to decrease the calibration costs while handling the complexity of the system. Proposed algorithm selects the experimental set values adaptively, but there exists still a limited number of steady-state points. Nozaki et al. [79] presented a rule-based algorithm to increase the number of DoEs by narrowing down the design range. The authors report that the algorithm does not guarantee the optimal solution and high level of engine calibration skills are required to benefit from the algorithm. Simulation-based calibration algorithms were also proposed to utilize an engine software to decrease the steady-state calibration efforts [80].

Even though the studies on steady-state calibration try to develop approaches to decrease the time by finding the optimal experimental set values, allowable exhaust emission limits drop dramatically, and new test cycles which highlight the transient behaviors are implemented for new vehicles with upcoming regulations. Therefore, these approaches become inadequate in the calibration process and need to be improved. Evolution of methodologies in model-based calibration by changing legislation [81] is depicted in Figure 3.1.

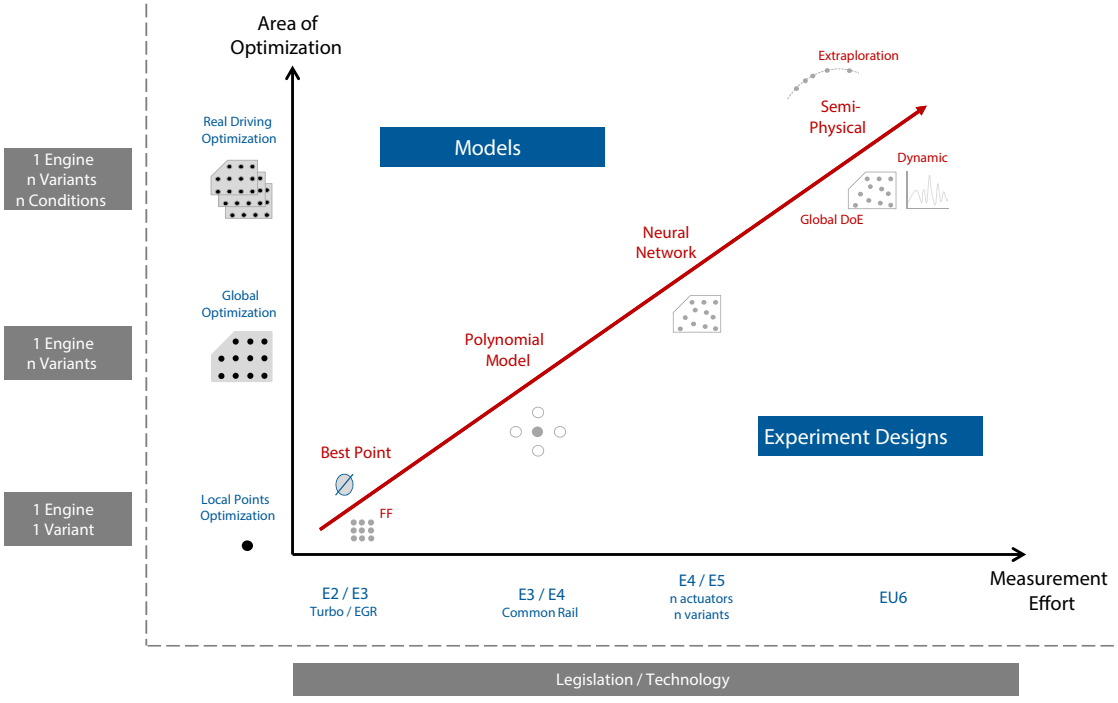


FIGURE 3.1: Evolution of methdologies in model based calibration

As the emission standards set by European Union gets tougher, area of optimization also widens, which necessitates increasing model complexity and enhancing the DoEs from local steady-states to global dynamic experiments (Figure 3.1). Brahma et al. [82] claimed that a dynamic model based transient calibration is inevitable with the recent Europen Union directives. According to them, a dynamic calibration process should answer these three questions:

1. Which empirical model structures are the most suitable for the torque and combustion emissions of a diesel engine?
2. How can the selected model structure capture transient behavior rather than the only static or quasi-static operations?
3. Is the size of the model structure appropriate to manage in the optimization process?

Offering an efficient global transient based design of experiment and selecting a suitable model structure with the answers of the questions mentioned above rely on the understanding of the underlying phenomena. Therefore, a diesel combustion process will be briefly explained in the following section.

3.1 Diesel Combustion Process

Diesel engine combustion involves two fundamental paths, namely air and fuel paths. The basic structure of a diesel engine is illustrated in Figure 3.2, where the ambient air is represented as blue, the red path is the dirty outflow after combustion, and the green path is the mixture of ambient air and the recirculated exhaust gas [35].

At the beginning of the air path, the ambient air is sucked to the system with a compressor, and Mass Air Flow (MAF) measures the amount of sucked air. The ambient air and recirculated exhaust gas are then mixed and injected into

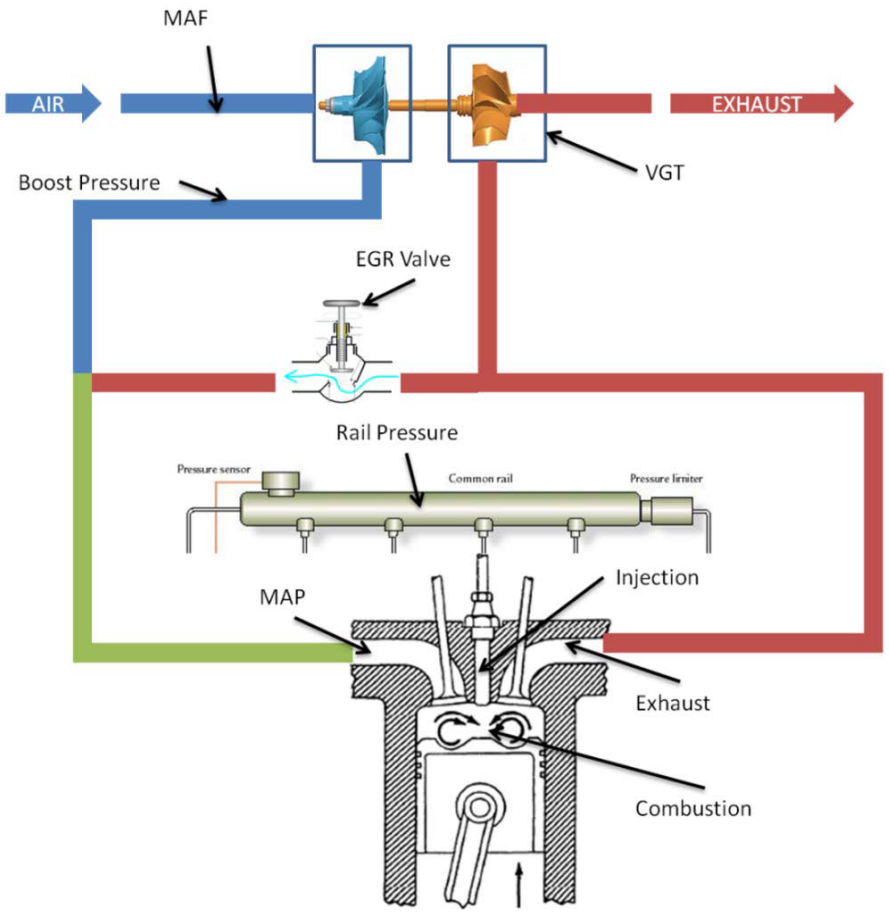


FIGURE 3.2: Basic structure of a diesel engine

the combustion chamber through the intake manifold. The intake manifold is pressurized, and it is defined as Manifold Absolute Pressure (MAP). After the combustion period, it is known that there exists still some oxygen concentration in the outflow. Therefore some portion of this gas is recirculated back by a valve called Exhaust Gas Recirculation (EGR). The rest of the exhaust gases which are not recirculated back exit through the exhaust pipeline.

Fuel path of the engine combustion involves the controller regarding the injection of diesel fuel into the chamber. Initially, the fuel is pumped into a pressurized common rail. The crank angle called Start of Injection (SOI) controls the total amount of fuel injected into the system. In addition to the main injection system, pilot injections are also employed to increase controllability and efficiency of the combustion process.

Chapter 4

Design of Experiments for Dynamic Calibration of Diesel Engines

Main steps of the system identification loop are the selection of related model inputs, the choice of excitation signals, proposing an appropriate modeling architecture and finally validating the obtained model with a suitable test set as illustrated in Figure 2.1. In this chapter, the first two steps of the identification loop for diesel engine torque and combustion emissions will be detailed, namely the selection of model inputs and designing the excitation signals.

4.1 Selection of Model Inputs

Components in a diesel engine combustion can be investigated in two groups such as air and fuel path as described in Section 3.1. Since the air and fuel path of a diesel engine have an essential role in torque generation and NOx formation, variables in these paths were included in the input set such as Manifold Absolute Pressure (MAP), Mass Air Flow (MAF), rail pressure (RailP), engine speed (SPD), main and pilot injection fuel quantities (miQNT and piQNT), main and pilot start

of injections (miSOI and piSOI). These variables are also the parameters desired to be calibrated in the optimization process. Utilization of pilot injections as an addition to the main injection increases the controllability of the torque generation and NOx formation processes and reduce the noise in combustion. However, it also makes the optimization problem more complex by increasing the number of parameters to be calibrated. Therefore, in some scenarios, pilot injections are kept constant for simplicity.

Since the working principle of the diesel engines depends on the compression, the ratio between air and fuel has a dominating role in the efficiency of combustion. Thus, air-to-fuel (AFR) ratio was intended to include in the input set of soot models. However, it is not guaranteed that the fuel will always be injected, so the total amount of fuel quantity can be zero in some cases, and the ratio of air-to-fuel can be infinite. On the other hand, ambient air is always sucked to the chamber when the engine is working, so MAF is always positive and fuel-to-air ratio (FAR) will always be greater than or equal to zero. FAR is calculated as

$$FAR(t) = \frac{1}{AFR(t)} = \frac{QNT(t)}{MAF(t)} \tag{4.1}$$

where QNT is the total of main and pilot injection fuel quantities. Different from indicated torque and NOx emission models, intake oxygen ratio (IntO2R), which is the amount of unburned oxygen during combustion was also included in the input set in addition to FAR. In the calculations of IntO2R, no delay is introduced by the assumption of the slow variations in exhaust oxygen concentration, and it is also assumed that the oxygen-enrichment of the intake air is %21 [83]. IntO2R is calculated by the following equations

$$Exh_{02R}(t) = 0.21\frac{MAF(t) - AFR_s QNT_{inj}(t)}{MAF(t) + QNT_{tot}(t)} \tag{4.2}$$

$$EGR_{flow}(t) = CMF(t) - GMF^{-}(t) \tag{4.3}$$

$$EGR_{O2}(t) = EGR_{flow}(t)Exh_{02R}(t) \tag{4.4}$$

$$IntO2R(t) = \frac{EGR_{O2}(t) + 0.21GMF^{-}(t)}{CMF(t)} \tag{4.5}$$

where AFR_s is the stoichiometric air-fuel ratio, which is 14.5 for diesel fuels, QNT_{inj} is the amount of fuel quantity which is assumed to be converted to torque, QNT_{tot} is the total amount of injected fuel, CMF is the charge mass flow, and GMF^- is the gas mass flow before EGR feed.

Finally, rail pressure, engine speed, main injection quantity and the main start of injection were also added to the input set of the soot models. Input selections for indicated torque and combustion emissions are tabulated in Table 4.1.

Indicated Torque	**NOx Emission**	**Soot Emission**
MAP	MAP	FAR
MAF	MAF	IntO2R
RailP	RailP	RailP
SPD	SPD	SPD
miQNT	miQNT	miQNT
miSOI	miSOI	miSOI
piQNT	piQNT	
piSOI	piSOI	

TABLE 4.1: Model inputs for indicated torque and combustion emissions

Here the EGR value is indirectly taken into consideration using MAF and MAP setpoints. EGR valve control is enabled for the entire duration of the tests to meet the MAF setpoint, and MAP is controlled using the turbocharger (VGT) actuator. Since MAP directly determines the amount of total gas flow into the engine considering volumetric efficiency; the difference between total gas flow and MAF is equal to the actual EGR flow into the engine. The rationale for using MAF and MAP as reference signals instead of valve positions is to be able to cover a more realistic engine operating range in terms of air-fuel ratio (AFR) and EGR flow. EGR ratio coverage is also monitored after the tests to ensure that we achieve informative experiments in terms of magnitude coverage.

Since the torque generation and the emission formation in a diesel engine are dynamical systems, past values of the inputs and possibly output, which are called regressors are also utilized in the modeling process. In this thesis, variations of NARX models were employed in the modeling of indicated torque and NOx emission, whereas variations of LSTM structure were selected as modeling architectures

due to the insufficiencies of NARX models in soot modeling. LSTM structure does not need any a priori knowledge about the number of regressors, whereas it is required to define regressor numbers for each input channel in a NARX model. In the following section, regressor selection methods followed in this thesis will be detailed.

4.1.1 Selection of Regressors

It is possible to select different lags and different amount of regressors for each input of a NARX model depending on the a priori knowledge about the process. However, it is not easy to designate such details of regressors, especially in a fully black-box identification process. Therefore, an equal number of regressors (n_b) were selected for each input channels as a simple way.

Selecting equal regressor number for each channel may result in redundancy in some cases and increase the number of parameters to be estimated. Thus, an orthogonal least square (OLS) algorithm with error reduction ratio was implemented to obtain more light-weight NOx emission models by eliminating unnecessary regressors.

Assuming the NOx emission formation process as a linear Multi-Input-Single-Output (MISO) system, the generic compact matrix form of it can be written as

$$Y = P\Theta + \Xi \tag{4.6}$$

where P is the regression matrix, Θ is the vector of parameters, Y is the NOx prediction and Ξ is modeling error vector. Redundant regressors in P can be removed by determining the most effective regressors by forward orthogonal least squares (OLS) algorithm and error reduction ratio [84].

Assuming P is a full rank matrix and the orthogonally decomposed as

$$P = WA \tag{4.7}$$

where W is a matrix with orthogonal coloumns and A is an upper triangular unit matrix. Gram-Schmidt and Housholder transformations [85] can be applied to perform this decomposition. It should be noted that the spaces spanned by the columns of W and P are the same.

$$span\{w_1, w_2, ..., w_m\} = span\{p_1, p_2, ..., p_m\} \tag{4.8}$$

MISO system in (4.6) can be rewritten by using the matrix A as

$$Y = \underbrace{(PA^{-1})}_{W}\underbrace{(A\Theta)}_{\triangleq G} + \Xi = WG + \Xi \tag{4.9}$$

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

$$G = (W^T W)^{-1} W^T Y \tag{4.10}$$

Assuming the past values of the output are not correlated with the modeling error Ξ, then the variance of the NOx emission predictions can be calculated as

$$\frac{1}{N} Y^T Y = \frac{1}{N} \sum_{i=1}^{m} g_i^2 w_i^T w_i + \frac{1}{N} \Xi^T \Xi \tag{4.11}$$

where N is the length of the samples and m is the number of regressors, which is initially selected equal for each input channel. Second part of the summation in (4.11) is the variance of the modeling error and the first part is related to regressors because of the relation shown in (4.8).

Significant w_i vectors which minimize the variance of modeling error determines the most effective regressors. These regressors are selected from the w_i vectors resulted in higher error reduction ratio (ERR), which is defined as

$$ERR_i = 100 \times \frac{g_i^2 w_i^T w_i}{Y^T Y} \tag{4.12}$$

More details of the OLS and ERR based term and variable selection for nonlinear systems can be found in [86].

4.2 Design of Excitation Signals

Once the related inputs are selected for the modeling process, the next step is the design of excitation signals for these inputs, which has an essential role in the accuracy of parameter estimation. Depending on the available information about the process, model-based and model-free approaches are utilized in the design of excitation signals [87]. A priori model of the underlying process is utilized to adjust the excitation signals for the system in model-based approaches [88]. On the other hand, these requirements should be met for a model-free approach:

- The whole intended input space should be covered as much as possible by using space filling methods [89].
- If the system is dynamic, the frequency range of the underlying process should also be encapsulated [90].

To meet these necessities, chirp signals are commonly preferred in the identification of nonlinear dynamic systems thanks to their persistent excitation capabilities. Furthermore, they have a lower crest factor (4.13), which indicates that they inject much more power into the system than the signals having the same peak value and a higher crest factor [91].

$$C_r^2 = \frac{max_t u^2(k)}{\lim\limits_{N\to\infty} \frac{1}{N} \sum\limits_{k=1}^{n} u^2(k)} \tag{4.13}$$

Indicated torque and combustion emissions of a diesel engine are considered to have nonlinear complex dynamics that requires to have a model-free design of excitation signals. Because of aforementioned advantages of periodic signals, input channels were excited by chirp signals, which have sinusoidal waveform with changing frequencies over time given by

$$y(t) = Asin(2\pi f(t)) \tag{4.14}$$

where the frequency of the chirp signal can be a linear (4.15), quadratic (4.16) or a logarithmic (4.17) function of time.

$$\text{Linear} \quad : f(t) = f_{min} + \underbrace{\left(\frac{f_{max} - f_{min}}{T}\right)}_{k} t \quad = f_{min} + kt \tag{4.15}$$

$$\text{Quadratic} \quad : f(t) = f_{min} + \underbrace{\left(\frac{f_{max} - f_{min}}{t_{max}^2}\right)}_{k} t^2 \quad = f_{min} + kt^2 \tag{4.16}$$

$$\text{Logarithmic} \; : f(t) = f_{min} \underbrace{\left(\left(\frac{f_{max}}{f_{min}}\right)^{\frac{1}{t_{max}}}\right)}_{k}^{t} \quad = f_{min} k^t \tag{4.17}$$

where f_{max} is the maximum frequency, f_{min} is the minimum frequency, t_{max} is the time that the system operates at maximum frequency, and T is the duration time between f_{min} and f_{max}.

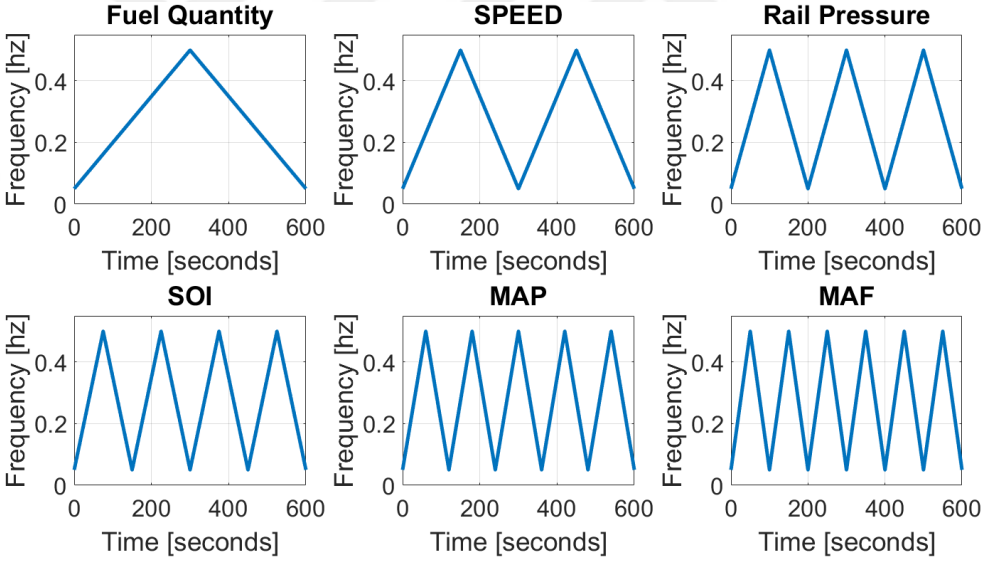


FIGURE 4.1: Frequency profiles of linear chirp signals

In the initial experiments for NOx emission modeling, the pilot injection was kept constant, and other six input channels such as MAF, MAP, QNT, SPD, RailP and SOI were excited by linear chirp signals. In the identification of multi-input systems, it is very critical to have a low correlation between input signals to obtain maximum informative experiments. In order to decouple the input channels, the number of the sweeps in frequency profiles of the chirp signals is selected as an increasing manner (Figure 4.1).

In order to check the correlation between input channels, the Pearson coefficient for pair-wise channels is calculated as follows and presented in Table 4.2.

$$\gamma = \frac{| < x - \bar{x}, y - \bar{y} > |}{\|x - \bar{x}\| \|y - \bar{y}\|} \tag{4.18}$$

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

TABLE 4.2: Correlations between input channels

It should be noted that the frequency profiles shown in Figure 4.1 are designed for the desired reference excitation signals. Normalized input signals acquired from the experimental setup are depicted in Figure 4.2. It is seen that fuel quantity,

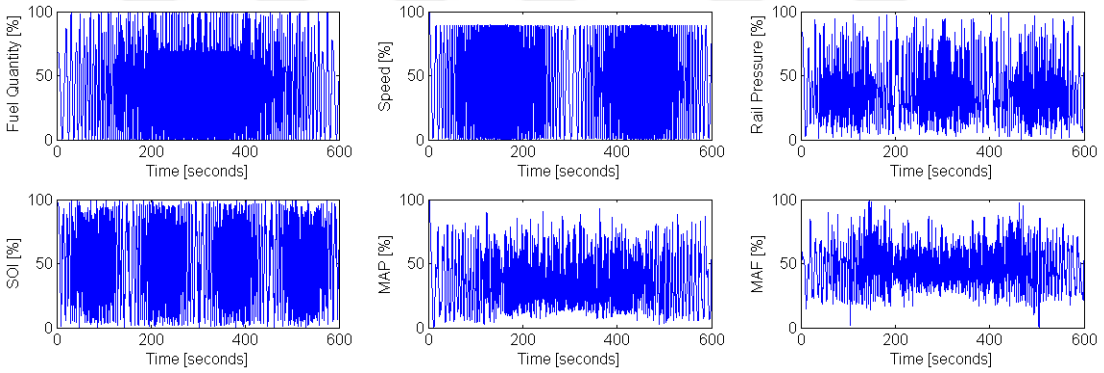


FIGURE 4.2: Normalized input signals (Experiment with 6 inputs)

speed, rail pressure and SOI input channels are successfully excited in the desired manner but MAP and MAF channels do not show the characteristics of five and six 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 4.2). However, they are included in the group of input channels for the system identification model because both of them have an impact on NOx emissions.

This six-input experiment was utilized in a NARX based NOx emission modeling, where the most crucial regressors were selected by an OLS and ERR algorithm. Experimental results regarding this work will be presented in Section 8.2.2. It has been shown that the transient behavior for NOx emissions were captured adaquetly with a chirp based experiment design. However, there were some offsets in the steady-state regions, which causes inaccuracies in steady-state validation cycles.

Later on, two separate experiments with different input designs based on chirp and ramp-hold signals are conducted and then merged to cover both steady-state and transient cycles of the combustion process. In these experiments, the pilot injection was also activated, so the input channels of main fuel injection quantity and the main start of injection were also excited. With the proposed experiment design, it is targeted to reduce testing time dramatically and improve modeling accuracy in both steady-state and transient operating conditions as an alternative to the steady-state DoE based modeling approaches.

In the first experiment called Experiment 1, quadratic chirp signals were designed for all the input channels in order to excite the system adequately in both high and low frequencies. The number of forward sweeps is increased in the frequency profiles of SPD, MAF, miSOI and piQNT, respectively. In the same way, reversed sweeps are applied for miQNT, MAP, railP and piSOI, respectively. The frequency range of the input signals is determined by investigating the World Harmonic Transient Cycle (WHTC) [92] and the New European Driving Cycle (NEDC) [93]. Short-time Fourier transform of the designed input signals depicts the frequency profiles of the inputs in Figure 4.3.

Pearson coefficients for pair-wise channels tabulated in Table 4.3 and the designed signals are depicted in Figure 4.4. Pearson coefficient is 0 when two signals are completely different and 1 when the signals have maximum similarity. Table 4.3 shows that the desired input signals are sufficiently decoupled, since all the coefficients for pair-wise different channels are less than 0.1.

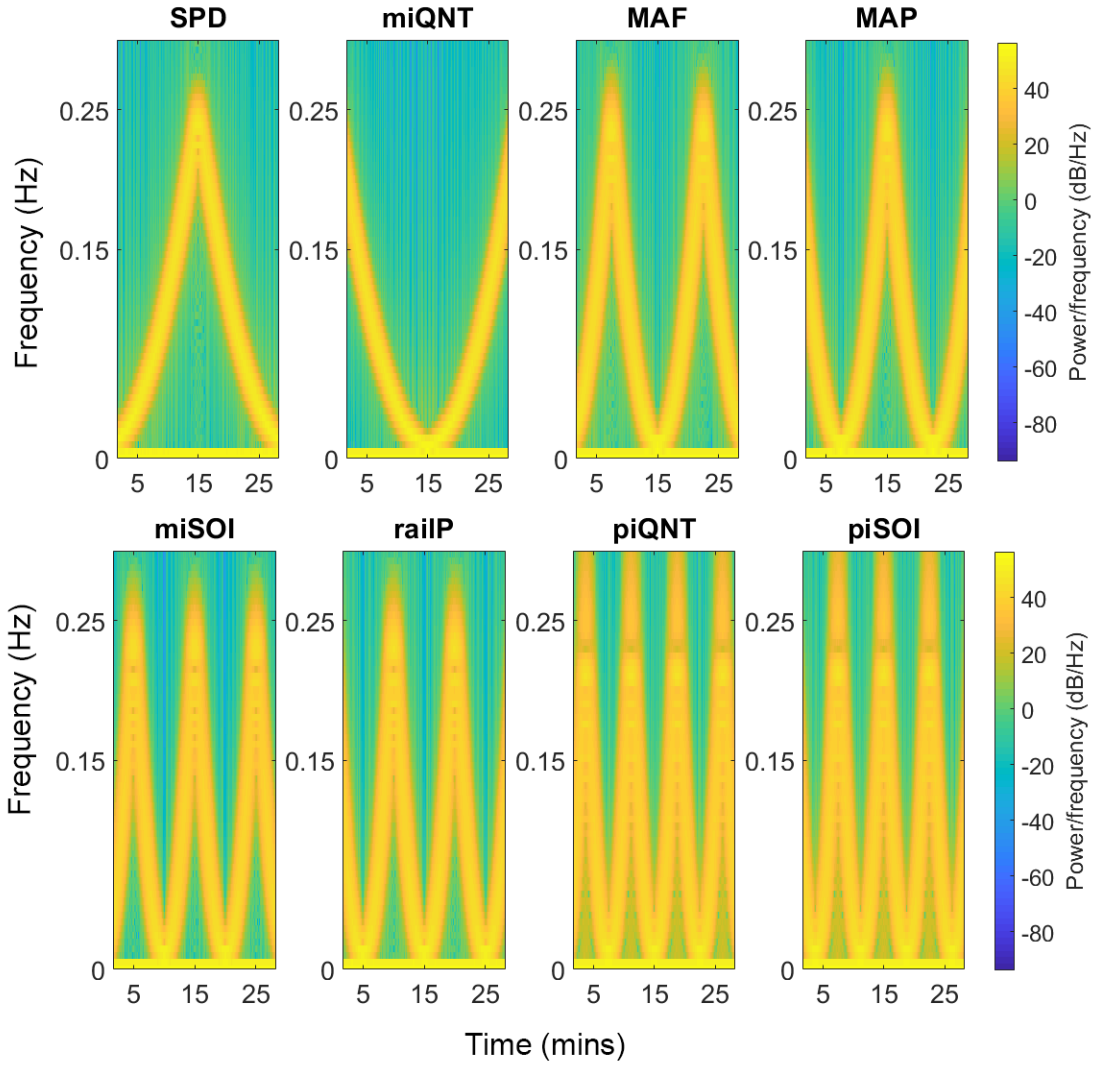


FIGURE 4.3: Frequency profiles of input channels (Experiment 1)

	SPD	MAP	MAF	railP	miQNT	piQNT	miSOI	piSOI
SPD	1	0.050	0.092	0.017	0.033	0.004	0.085	0.050
MAP	0.050	1	0.046	0.003	0.050	0.057	0.003	0.022
MAF	0.092	0.046	1	0.108	0.092	0.057	0.108	0.022
railP	0.017	0.003	0.108	1	0.085	0.065	0.057	0.025
miQNT	0.033	0.050	0.092	0.085	1	0.004	0.017	0.050
piQNT	0.004	0.057	0.057	0.065	0.004	1	0.065	0.082
miSOI	0.085	0.003	0.108	0.057	0.017	0.065	1	0.025
piSOI	0.050	0.022	0.022	0.025	0.050	0.082	0.025	1

TABLE 4.3: Correlations between designed input signals (Experiment 1)

In order to increase the prediction accuracy in the steady-state cycle of a diesel engine combustion, an additional experiment with different excitation signals for engine speed and main injection fuel quantity were designed as Experiment 2.

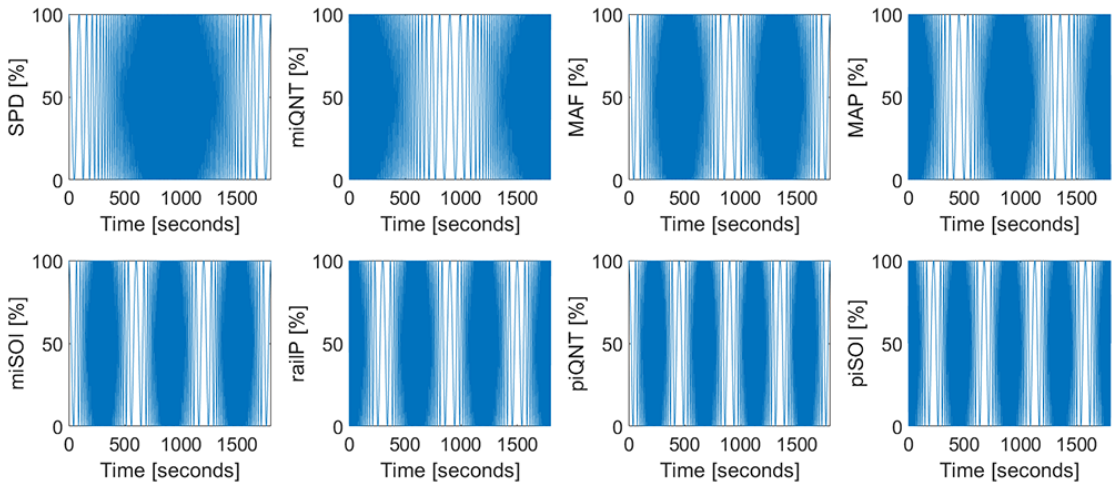


FIGURE 4.4: Normalized input signals (Experiment 1)

Different from the first experiment, engine speed was excited by a ramp-hold signal at different levels and the frequency profile of the main injection fuel quantity channel was changed to a single forward sweep. By doing so, it was targeted that the speed - fuel quantity operation region in Experiment 2 contains much lower frequencies than Experiment 1. Duration of each experiment was 30 minutes and the signals were collected in 10 Hz. Both experiments were merged and employed in the training. Designed input signals of Experiment 2 are shown in Figure 4.5.

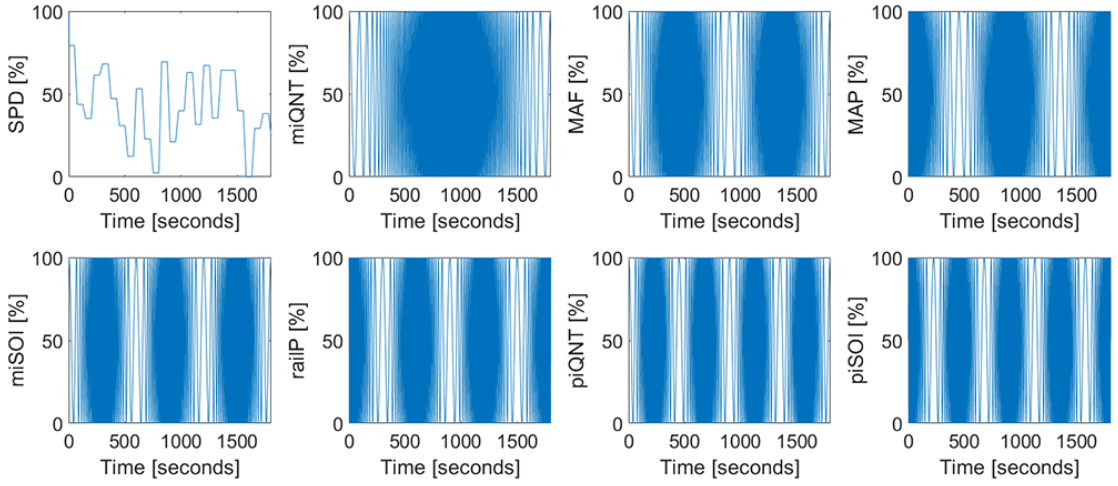


FIGURE 4.5: Normalized input signals (Experiment 2)

It should be noted that Experiment 3 and 4 were also performed with the same type excitation signals of Experiment 1 and 2, instead their duration time was 120 minutes (Figure 4.6).

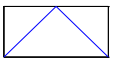
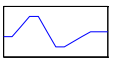
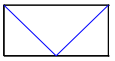
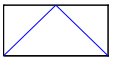
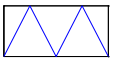
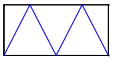
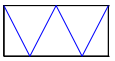
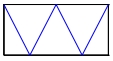
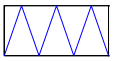
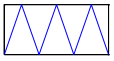
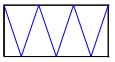
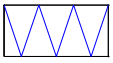
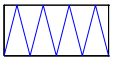
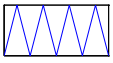
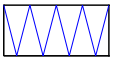
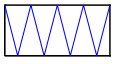
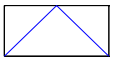
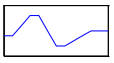
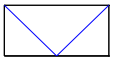
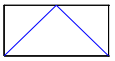
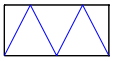
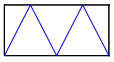
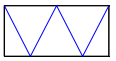
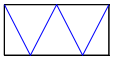
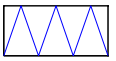
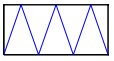
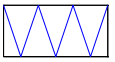
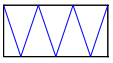
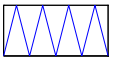
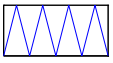
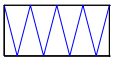
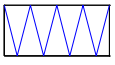
	Experiment 1 (*30 min.*)		Experiment 2 (*30 min.*)	
SPD	Quadratic Chirp 0.01 - 0.25 Hz		Ramp-Hold 23 Points	
miQNT	Quadratic Chirp 0.01 - 0.25 Hz		Quadratic Chirp 0.01 - 0.25 Hz	
MAF	Quadratic Chirp 0.01 - 0.25 Hz		Quadratic Chirp 0.01 - 0.25 Hz	
MAP	Quadratic Chirp 0.01 - 0.25 Hz		Quadratic Chirp 0.01 - 0.25 Hz	
miSOI	Quadratic Chirp 0.01 - 0.25 Hz		Quadratic Chirp 0.01 - 0.25 Hz	
RailP	Quadratic Chirp 0.01 - 0.25 Hz		Quadratic Chirp 0.01 - 0.25 Hz	
piQNT	Quadratic Chirp 0.01 - 0.25 Hz		Quadratic Chirp 0.01 - 0.25 Hz	
piSOI	Quadratic Chirp 0.01 - 0.25 Hz		Quadratic Chirp 0.01 - 0.25 Hz	
	Experiment 3 (*120 min.*)		Experiment 4 (*120 min.*)	
SPD	Quadratic Chirp 0.01 - 0.25 Hz		Ramp-Hold 23 Points	
miQNT	Quadratic Chirp 0.01 - 0.25 Hz		Quadratic Chirp 0.01 - 0.25 Hz	
MAF	Quadratic Chirp 0.01 - 0.25 Hz		Quadratic Chirp 0.01 - 0.25 Hz	
MAP	Quadratic Chirp 0.01 - 0.25 Hz		Quadratic Chirp 0.01 - 0.25 Hz	
miSOI	Quadratic Chirp 0.01 - 0.25 Hz		Quadratic Chirp 0.01 - 0.25 Hz	
RailP	Quadratic Chirp 0.01 - 0.25 Hz		Quadratic Chirp 0.01 - 0.25 Hz	
piQNT	Quadratic Chirp 0.01 - 0.25 Hz		Quadratic Chirp 0.01 - 0.25 Hz	
piSOI	Quadratic Chirp 0.01 - 0.25 Hz		Quadratic Chirp 0.01 - 0.25 Hz	

FIGURE 4.6: Excitation signals used in the experiments for modeling indicated torque and NOx emissions

Coverages of Experiment 1 and 2 over WHSC (steady-state) and WHTC (transient) validation cycles are depicted in Figure 4.7 and 4.8. Figure 4.7 shows that a wide range of engine speed is covered in Experiment 1 thanks to chirp signals, whereas the controlled area over validation cycles are covered using a ramp-hold based engine speed in Experiment 2 as depicted in Figure 4.8.

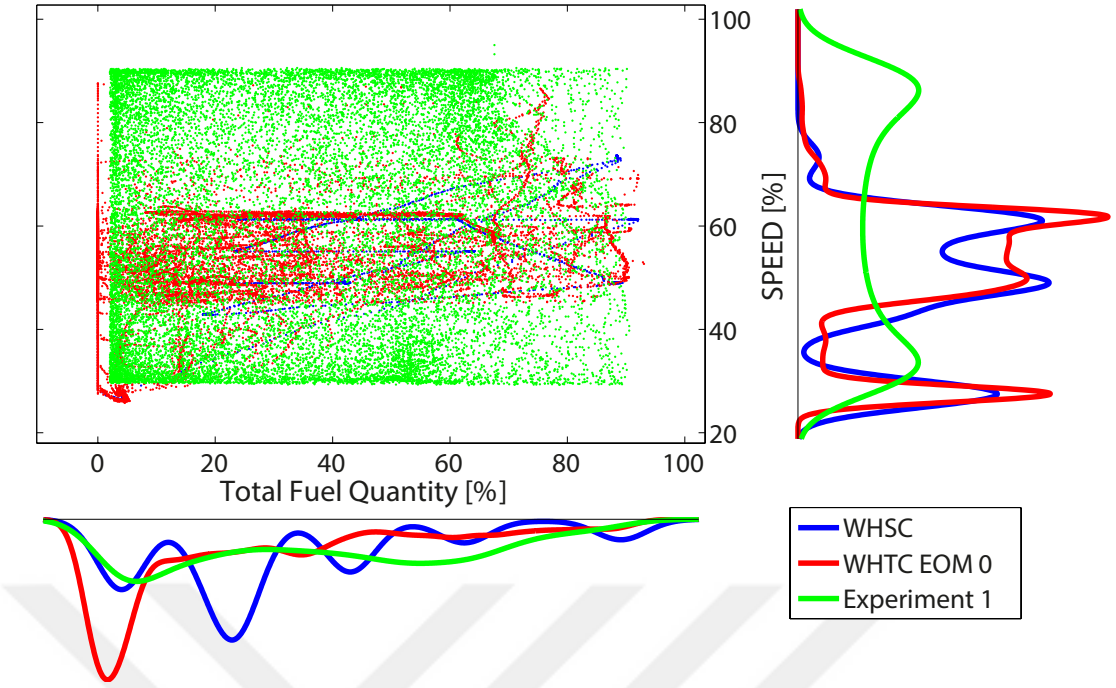


FIGURE 4.7: Coverages of Experiment 1 over validation cycles

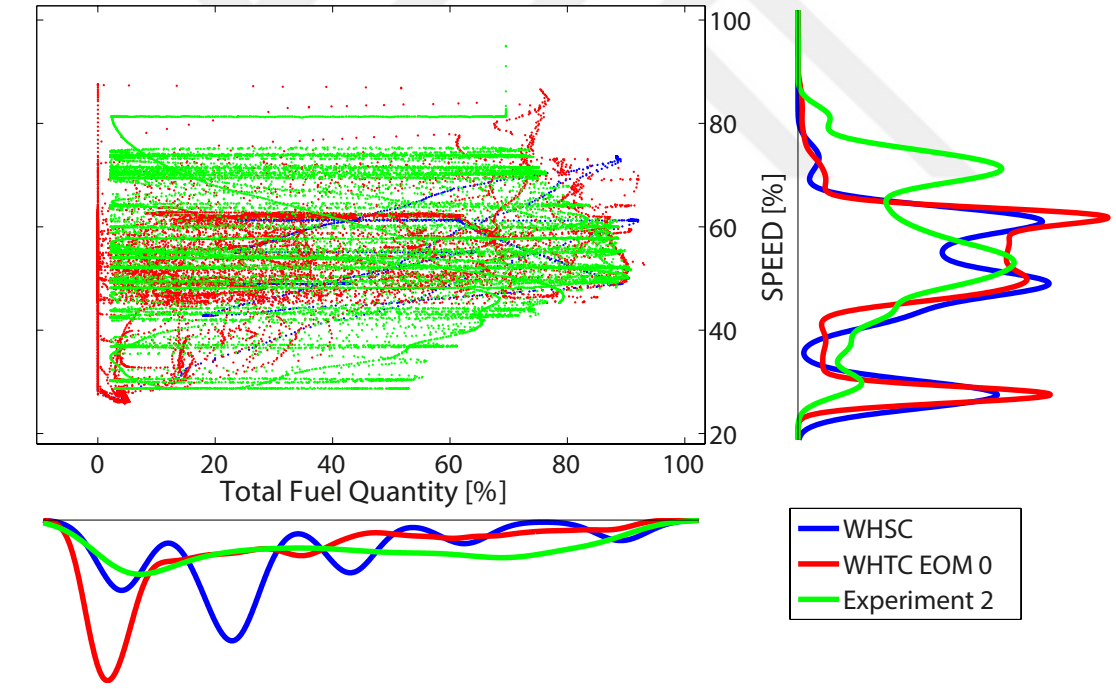


FIGURE 4.8: Coverages of Experiment 2 over validation cycles

In order to obtain accurate soot models to be validated with WLTC and NEDC cycles, 4 different tests were designed. Types of the signals that excited the input channels such as SPD, miQNT, MAF, MAP, miSOI and railP were presented in Figure 4.9 and the time plots of the inputs and output signals for all tests are

depicted in Figure 4.10 - 4.13.

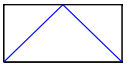
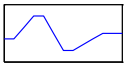
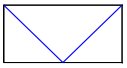
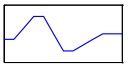
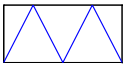
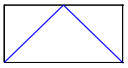
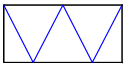
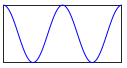
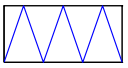
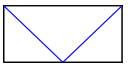
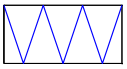
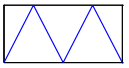
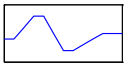
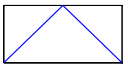
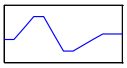
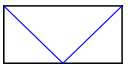
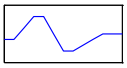
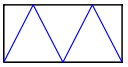
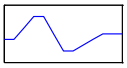
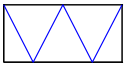
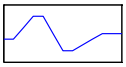
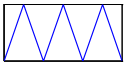
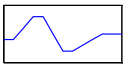
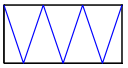
	Test 1		Test 2	
SPD	Quadratic Chirp 0.01 - 0.25 Hz		Ramp-Hold 21 Points	
miQNT	Quadratic Chirp 0.01 - 0.25 Hz		Ramp-Hold 21 Points	
MAF	Quadratic Chirp 0.01 - 0.25 Hz		Quadratic Chirp 0.01 - 0.25 Hz	
MAP	Quadratic Chirp 0.01 - 0.25 Hz		Sine 0.02 Hz	
miSOI	Quadratic Chirp 0.01 - 0.25 Hz		Quadratic Chirp 0.01 - 0.25 Hz	
RailP	Quadratic Chirp 0.01 - 0.25 Hz		Quadratic Chirp 0.01 - 0.25 Hz	
	Test 3		Test 4	
SPD	Ramp-Hold 277 Points		Quadratic Chirp 0.01 - 0.5 Hz	
miQNT	Ramp-Hold 277 Points		Quadratic Chirp 0.01 - 0.5 Hz	
MAF	Ramp-Hold 277 Points		Quadratic Chirp 0.01 - 0.5 Hz	
MAP	Ramp-Hold 277 Points		Quadratic Chirp 0.01 - 0.5 Hz	
miSOI	Ramp-Hold 277 Points		Quadratic Chirp 0.01 - 0.5 Hz	
RailP	Ramp-Hold 277 Points		Quadratic Chirp 0.01 - 0.5 Hz	

FIGURE 4.9: Excitation signals used in the experiments for modeling soot emission

All the input channels in Test 1 and Test 4 were excited by second order chirp signals. Frequency range for Test 1 was 0.01 – 0.25 Hz whereas in Test 4 it was 0.01 – 0.5 Hz. In order to decrease the correlations between input channels, the directions and the number of the sweeps in frequency profiles were chosen differently.

There are 21 different operating points for SPD and miQNT in Test 2 and the duration of the experiment in every operating point is 340 seconds. In these points, MAF, miSOI and railP input channels were excited by chirp signals which have different frequency profiles due to the directions and the number of the sweeps

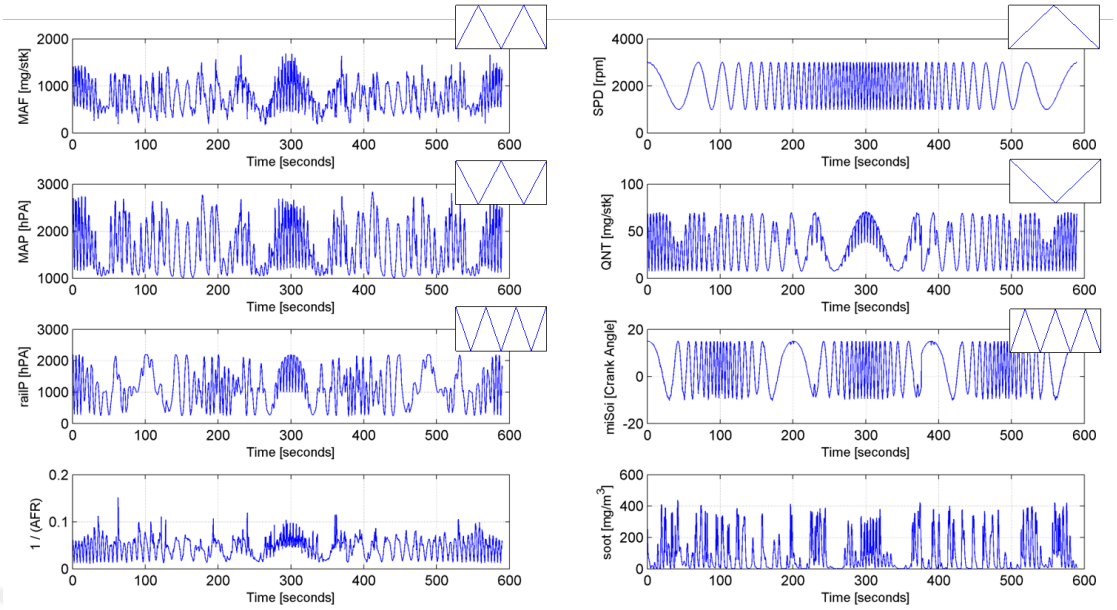


FIGURE 4.10: Test 1 inputs and output signals

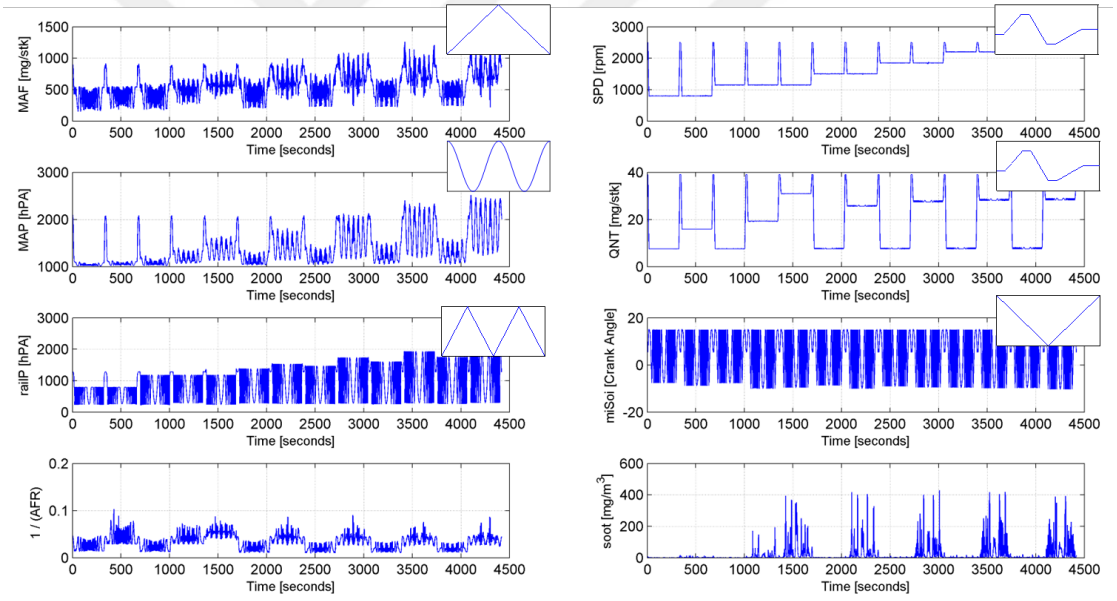


FIGURE 4.11: Test 2 inputs and output signals

between 0.01 and 0.25 Hz. In previous studies, it is observed that MAF and MAP input channels were highly correlated, and they have always similar patterns even if they are excited by different types of chirp signals. Therefore, MAP channel was excited by a sine wave with the frequency of 0.02 in Test 2. Lastly in Test 3, all input channels were excited by 15 seconds hold and 3 seconds ramp input signals in 277 different points. It can be claimed that the types of the excitations signals in Test 3 is suitable for steady-state calibration purposes. In this set of

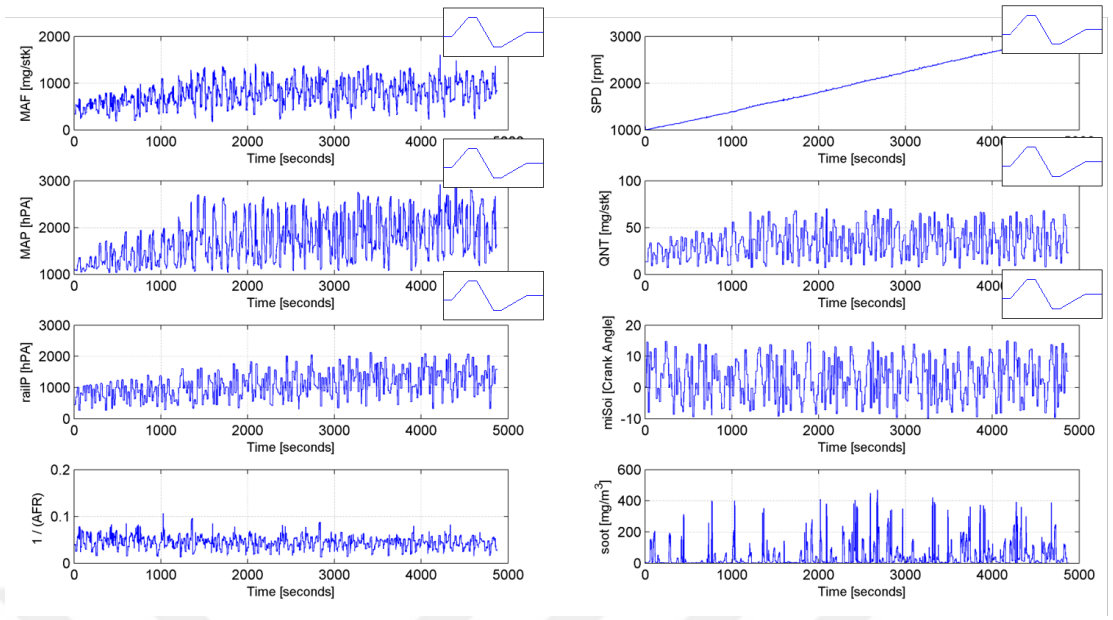


FIGURE 4.12: Test 3 inputs and output signals

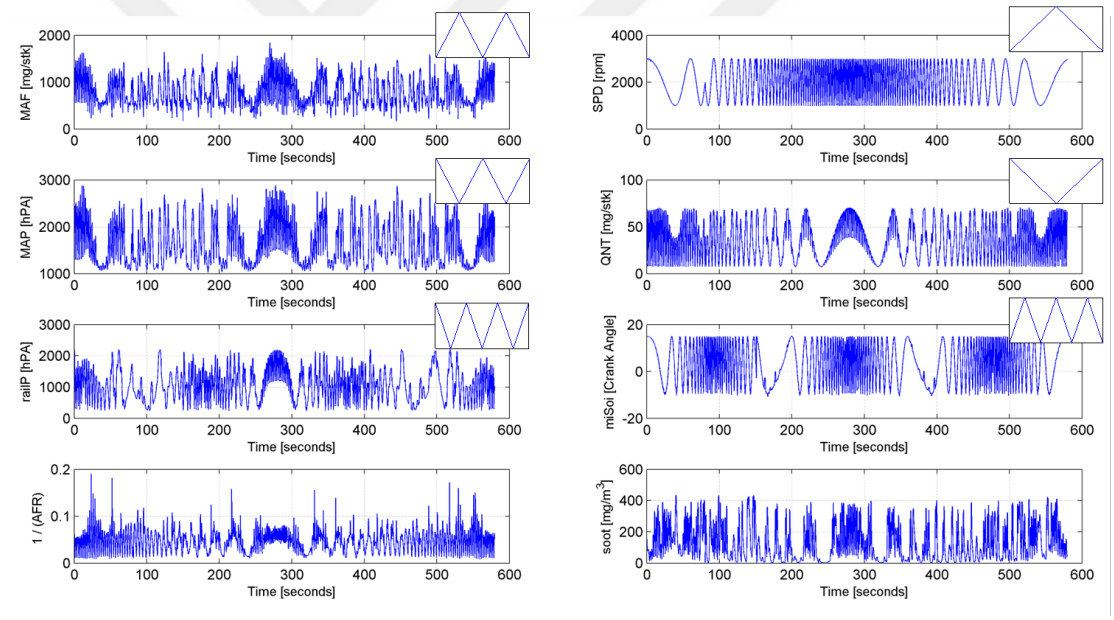


FIGURE 4.13: Test 4 inputs and output signals

experiments, Test 1 and 4 are 10 minutes, Test 2 is around 120 minutes and Test 3 is around 85 minutes. Due to some troubles in data acquisition and the limits of the engine, only 13 of 21 points in Test 2 were usable which duration is around 73 minutes. In addition to these tests which will be used in training of the models, WLTC and NEDC cycles were also performed with the same engine for validation purposes.

For Test 5, a diesel engine's intended region of operation was determined by 30 different points in engine speed - injected fuel quantity plane, where the engine speed and the injected fuel quantity are selected in the ranges of 800-2200 rpm and 4-30 mg/stroke, respectively (Figure 4.14).

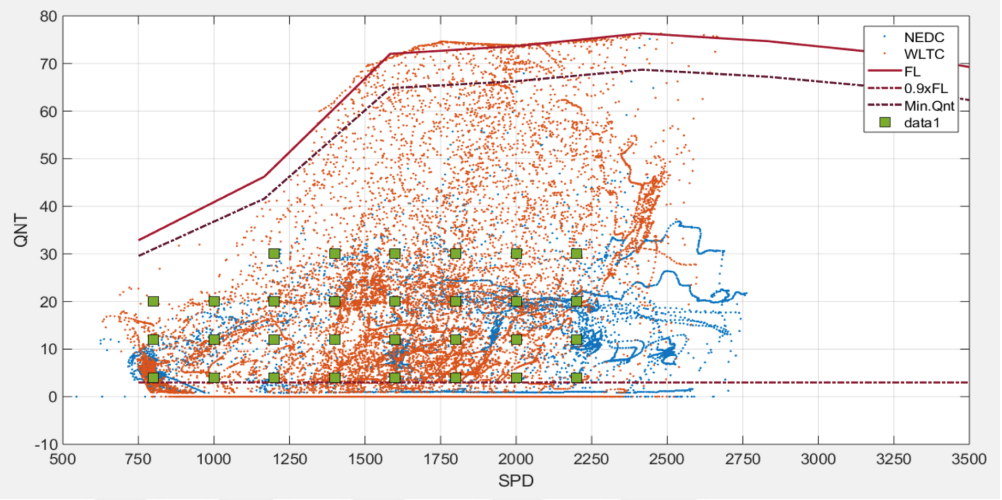


FIGURE 4.14: Determination of intended region of operation

The input channels other than SPD and QNT were excited by quadratic chirp signals as shown in (4.16) with $f_{max} = 0.25$ Hz and $f_{min} = 0.05$ Hz. The frequency profiles of the chirp input signals were presented in Figure 4.15.

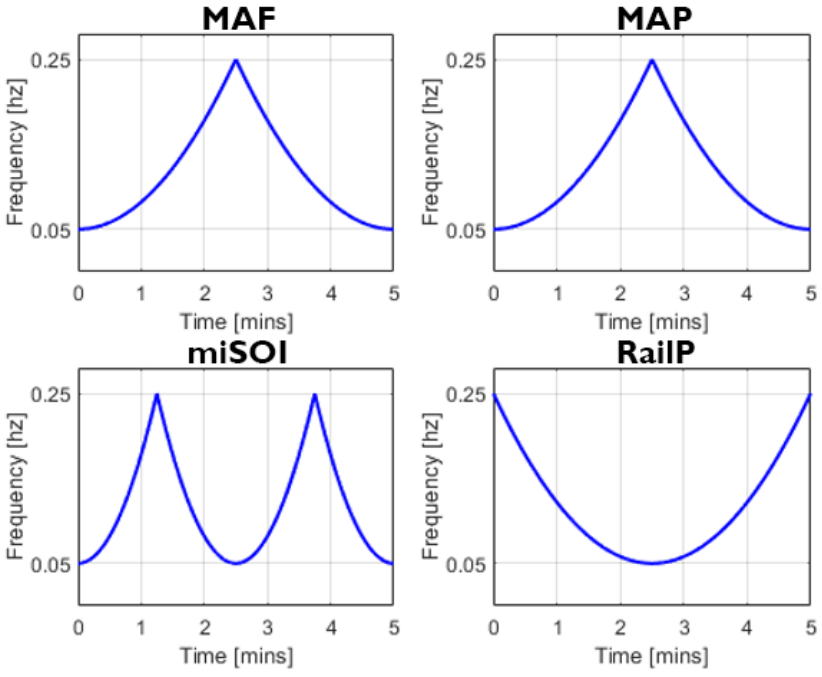


FIGURE 4.15: Frequency profiles of chirp input signals (Test 5)

Duration of each experiment is 5 minutes and the signal acquisition frequency is 10 Hz. Figure 4.16 presents a sample time plot of an operating point out of 30.

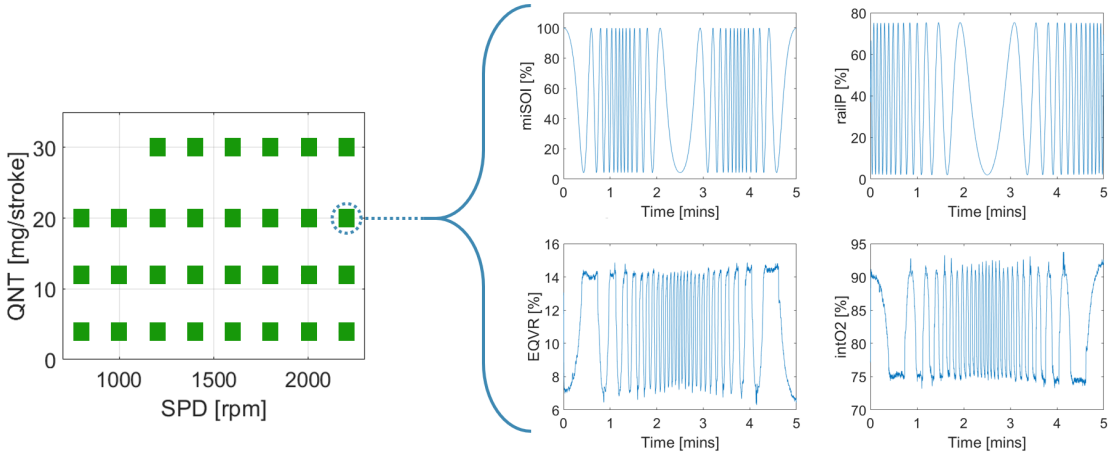


FIGURE 4.16: Input signals of an operating point

Figure 4.17 illustrates all the signals acquired in 30 operating points.

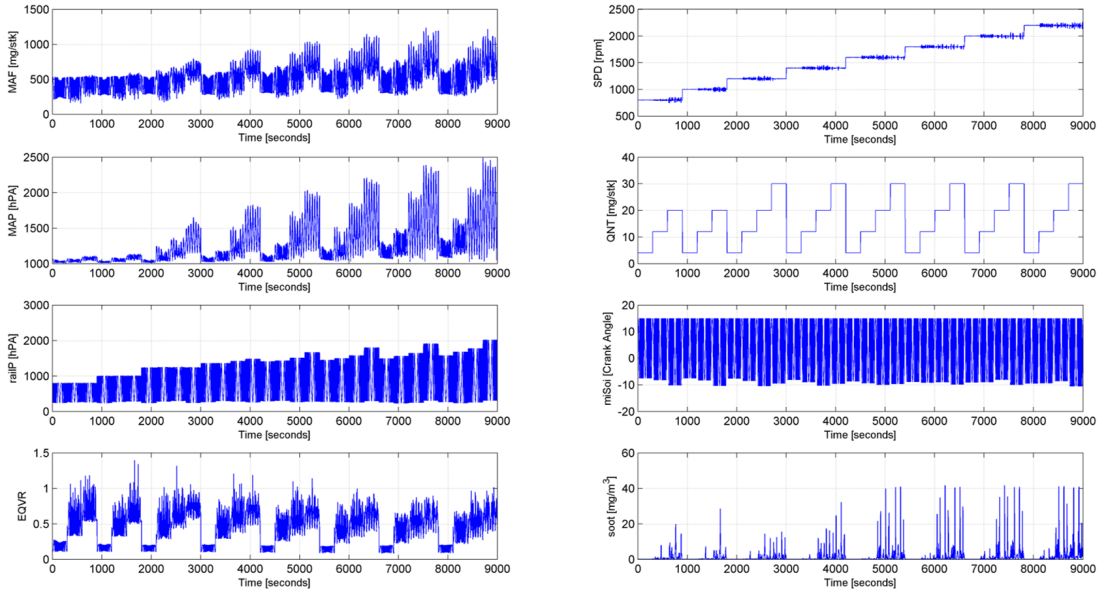


FIGURE 4.17: Test 5 inputs and output signals

Chapter 5

Diesel Engine Torque Modeling

In order to balance the conflict between the stringent emission legislation and high demand on the performances in terms of torque generation with fuel economy, an efficiently accurate torque model of internal combustion engines is required. In this chapter, indicated torque estimation by considering the losses during combustion and nonlinear finite impulse response (NFIR) based modeling approach will be detailed.

5.1 Indicated Torque Estimation

Torque generation in a diesel engine is a dynamic and nonlinear process, where the combustion-induced pressure causes the reciprocating motion of the piston [94]. Then a transmission mechanism is utilized to convert this motion to a rotational one. Some portions of the generated torques are lost in this transmission mechanism as friction, pumping and inertia torques. Indicated torque (τ_{ind}) is the ideal torque generated by combustion without any losses and can be calculated as

$$\tau_{ind} = \tau_{dyno} + \tau_{pump} + \tau_{fric} + \tau_{iner} \tag{5.1}$$

where τ_{dyno} is the breaking torque of the engine dynamometer, τ_{pump} is the pumping loss torque, τ_{fric} is the friction loss torque, and τ_{iner} is the inertia torque. Among

these torque signals, only the breaking torque of the engine is directly measured from the dynamometer. In the prediction of the pumping torque, the pumping mean effective pressure which is measured by in-cylinder sensors is utilized as follows

$$\hat{\tau}_{pump} = \frac{p_{me} V_d}{n_c 2\pi} \tag{5.2}$$

where p_{me} is the mean effective pressure in pascal, V_d is the displacement volume in cubic metre (12.7×10^{-3}), and n_c is the number of revolutions per power stroke, where in this case $n_c = 2$ for such a 4-stroke engine.

In order to decrease the number of unknown torque variables from two to one, motoring tests that include different speeding and slowing profiles between 600 and 2100 rpm have been performed Figure 5.1. Duration of the ramp speed signals was adjusted to 600, 300, 60, 30, 20, 10, 5, 3, and 2 seconds. During motoring tests, the gas pedal was kept constant at zero position, so it is assumed that no fuel is injected to the system and indicated torque is not generated.

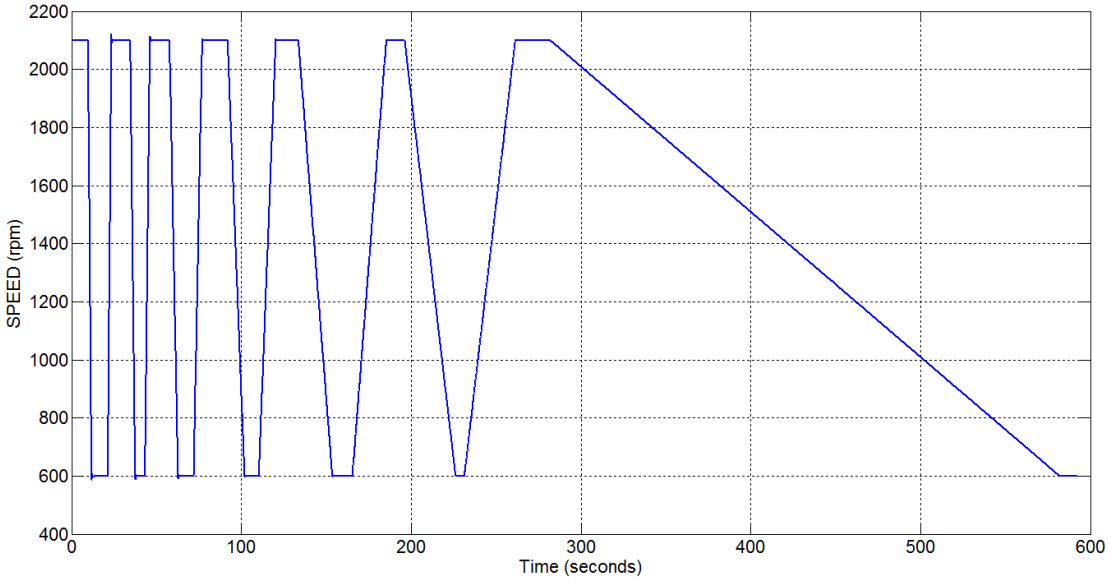


FIGURE 5.1: Motoring Tests

Since the speed is not changing dramatically in the period of 600 seconds, the inertia torque is assumed to be zero, and the measurements from this part of the test were utilized in the estimation of friction torque as follows

$$\hat{\tau}_{fric} = \tau_{dyno} + \hat{\tau}_{pump} \tag{5.3}$$

A second order polynomial is fitted to the friction loss torque estimations calculated from (5.5) with the accuracy of $R^2 = 0.9886$ (Figure 5.2). By this way, the friction torque is interpreted as a quadratic function of speed as follows

$$\tau_{fr}(\omega) = \alpha_1\omega^2 + \alpha_2\omega + \alpha_3 \tag{5.4}$$

where the coefficients (α_i) in (5.4) are tabulated in Table 5.1.

$\alpha_1 = -10^{-5}$	$\alpha_2 = -0.0106$	$\alpha_2 = -91.766$

TABLE 5.1: Coefficients of friction loss torque function $(\tau_{fr}(\omega))$

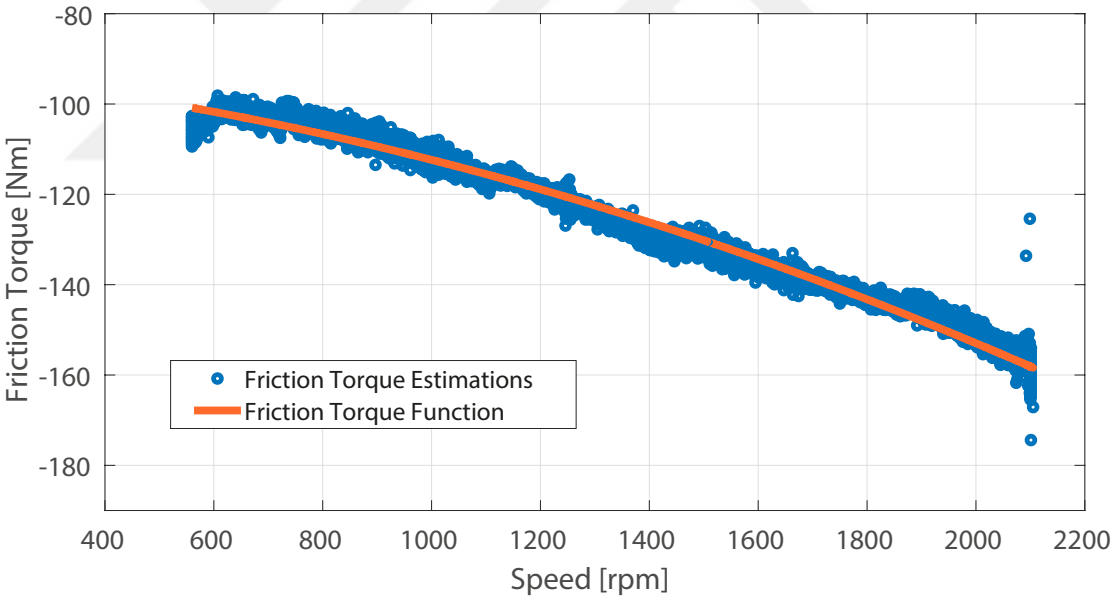


FIGURE 5.2: Representing the friction torque as a quadratic function of speed

Finally, inertia torque can be calculated using the estimations of friction and pumping loss torques as follows

$$\hat{\tau}_{iner} = \tau_{dyno} + \hat{\tau}_{pump} + \tau_{fr}(\omega) \tag{5.5}$$

The inertia torque can be calculated as a function of engine speed as

$$\tau_{in}(\omega) = J\frac{d\omega}{dt} \tag{5.6}$$

where J is the torque constant and it is determined as a value (6.12) which minimizes the error between $\hat{\tau}_{iner}$ and $\tau_{in}(\omega)$ as shown in Figure 5.3.

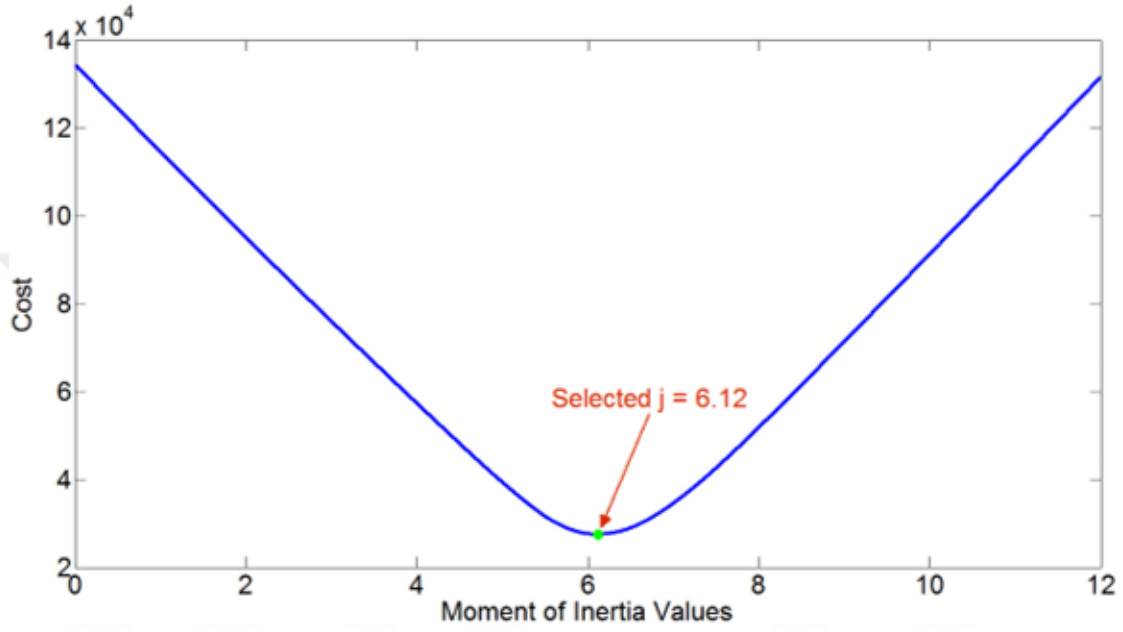


FIGURE 5.3: Determination of inertia torque constant

Figure 5.4 depicts a typical estimated loss and dyno brake torque signals.

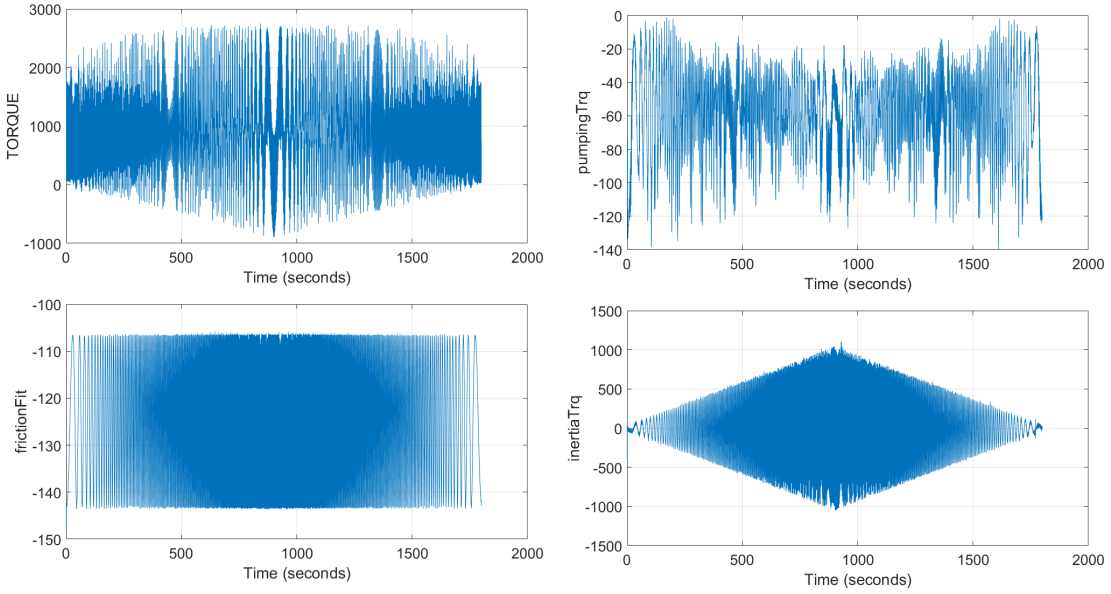


FIGURE 5.4: A typical estimated loss and dyno brake torque

5.2 NFIR Modeling

Since high prediction accuracies of torque models in both training and validation are very crucial to employ them in the optimization process, it is targetted to model indicated torque by including the estimations of friction, pumping and inertia torques in addition to the torque measured from the engine dynamometer. In this thesis, a particular type of NARX structure called nonlinear finite impulse response (NFIR) model (Figure 5.5) is proposed to model the indicated torque of a diesel engine.

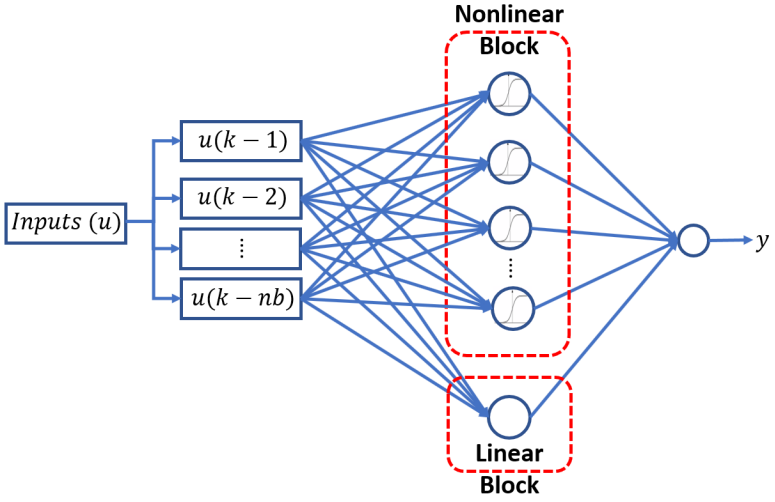


FIGURE 5.5: Nonlinear finite impulse response (NFIR) model

Different from a NARX structure, feedback from the output is not employed in the modeling of NFIR structures. The mathematical expression of a NFIR model can be expressed as

$$y(k) = w_0 x(k) + b_0 + \sum_{i=1}^{m} \phi_i\Big(w_i x(k) + b_i\Big) + e(k) \tag{5.7}$$

where $y(k)$ is the output, $x(k)$ is the regression vector consists only of previous inputs, w_i's are the weights and b_i's are the bias parameters in nonlinear block, w_0 and b_0 are weights and bias parameters in linear block, $e(k)$ is the modelling error, m is the number of units used in the nonlinear block, and $\phi_i(.)$ is the activation function. Nonlinear activation function (ϕ) is selected as a sigmoid function that

can be written as

$$\sigma(x) = \frac{1}{1 + e^{-ax}} \tag{5.8}$$

where a is a positive parameter to be estimated as well. Since the system can not respond instantaneously when the inputs are applied, there is at least 1 sample pure delay between inputs and output, in the NARX type structures. In the NFIR based indicated torque models, pure time delay was decided as 1 by considering the fast fast interactions of inputs in the torque generation mechanism. For instance, there is a direct relation between the fuel quantity and the torque as follows

$$\dot{m}_f = \frac{\omega_e.\tau_e}{n_f.Q_{lhv}} \tag{5.9}$$

where $\dot{m}_f$ is the fuel mass flow rate, ω_e is the engine speed, τ_e is the effective engine brake torque, n_f is the fuel conversion efficiency and Q_{lhv} is the fuel lower heating value.

The parameters in the NFIR model are estimated iteratively by using a combination of the line search algorithms such as subspace Gauss-Newton least squares search, adaptive subspace Gauss-Newton search, Levenberg-Marquardt least squares search and steepest descent least squares search [95].

Chapter 6

Diesel Engine NOx Emission Modeling

Increasingly stringent emission regulations are driving the evolution of internal combustion engine technologies. Especially NOx emissions are a focal point in the advancement of a diesel engine and after-treatment technologies, influencing overall vehicle development and hardware costs. Accurate prediction of NOx emissions is critical for fuel economy optimization in diesel engines.

In this chapter, two different NARX based NOx emission modeling approaches will be detailed. Furthermore, sensitivity analysis and parameter selection methods will be presented for the NARX models with linear output feedback.

6.1 NARX Modeling

Due to the highly nonlinear combustion process, it is showed that linear system identification models are not sufficient enough to model the NOx emissions [35]. Most of the time, validation performances of nonlinear NOx emission models surpass the linear ones. Therefore, a single layer recurrent nonlinear autoregressive with exogenous input (NARX) model was selected as modeling structure for NOx emissions. A typical NARX structure is depicted in Figure 6.1.

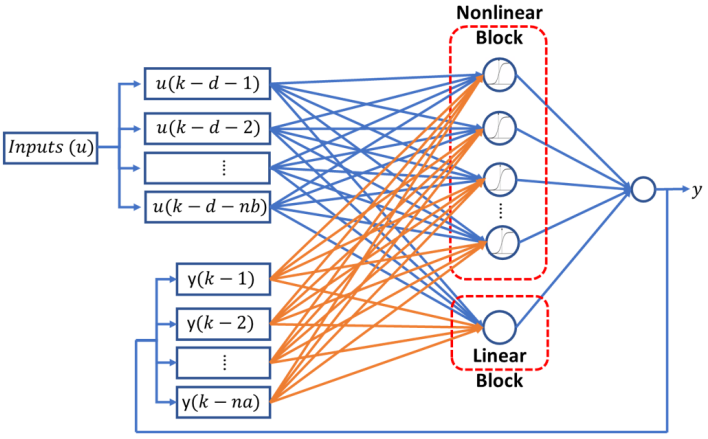


FIGURE 6.1: A typical NARX structure

d is the time delay between inputs and output, nb and na are the number of input and output past values used in the network, respectively. The mathematical expression of a single layer recurrent NARX network can be expressed as

$$y(k) = w_0 x(k) + b_0 + \sum_{i=1}^{m} \phi_i \Big(w_i x(k) + b_i \Big) + e(k) \tag{6.1}$$

where $y(k)$ is the output, $x(k)$ is the regressor set, w_i's are the weights and b_i's are the bias parameters in nonlinear block, w_0 and b_0 are weights and bias parameters in linear block, $e(k)$ is the modelling error, m is the number of units used in the nonlinear block, and $\phi_i(.)$ is the activation function.

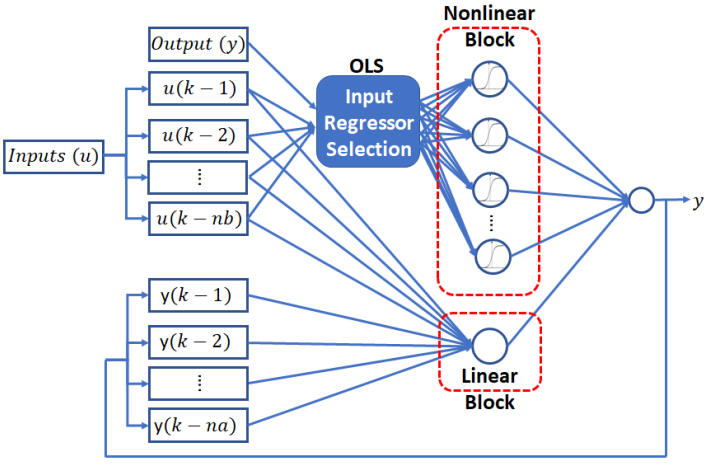


FIGURE 6.2: NARX structure with OLS based input regressor selection

Figure 6.2 illustrates the proposed NARX modeling using reduced set of regressors.

Modeling structure consists of two part such as nonlinear and linear block. Nonlinear block takes only the input regressors selected by OLS algorithm explained in Section 4.1.1. Linear block is a single neuron that takes all input and output regressors. Removal of output regressors from nonlinear block improves the stability of the models and enable one to use more than one output regressors in 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 less number of estimated parameters. Total number of parameters estimated in the proposed architecture is

$$N_P = \Big((N_{NLreg} + 2) \times N_{unit}\Big) + (N_{Lreg} + 1) \tag{6.2}$$

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.

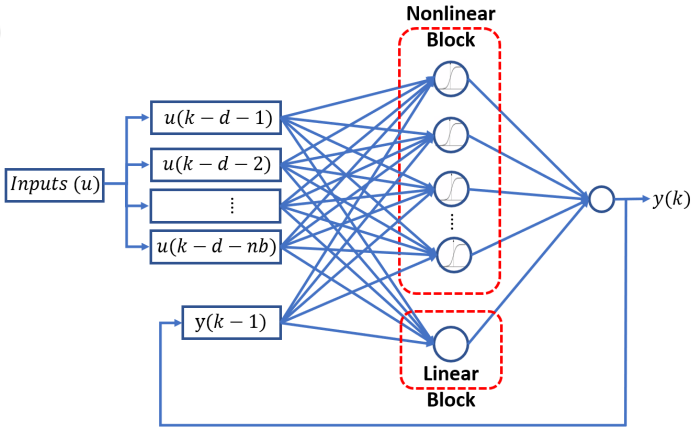


FIGURE 6.3: NARX structure with a single output feedback

Nelles [51] showed that general stability of the models with output feedback is not guaranteed. It is also experimentally verified that using the increasing number of output regressors results in the divergence of the optimization algorithm very frequently during the parameter estimation process. Therefore, as an alternative to regressor reduction, a single past value of the output feedback is proposed to include in the regressor set (Figure 6.3).

6.2 Sensitivity Analysis and Parameter Selection

To investigate the training and the validation performances, models for the ranges of parameters given in Table 6.1 were generated with the NARX model presented in Figure 6.3. The input regressor number and the unit number are the parameters of the model structure, whereas the maximum number of bisections used for line search along the search direction and the iteration number are the parameters of the optimization algorithm.

Parameter	**Range**
Input Regressor Number	3-10
Unit Number	5-15
Maximum Bisections	4-7
Iteration	20,40,60,80,100

TABLE 6.1: Ranges of parameters

It should be noted that an equal number of regressors are taken from each input channel. During the parameter estimation process, optimization algorithm searches for the best set of parameters that minimizes the error between the predicted and the measured output. Optimization runs for 20 iteration steps, and the obtained model is evaluated with training and validation (both steady-state and transient) tests. The last estimated model is utilized as an initial model for the next estimation up to 100 iterations. By doing so, 1760 models were obtained. Once the parameters regarding the model structure (input regressor and unit numbers) are fixed, there will be 20 models due to the variation of optimization parameters. In order to assess the performances of all models and determine the best optimization parameters, scores in terms of points (P) are calculated as the weighted average of training and validation performances as follows

$$P = \frac{1}{3}\Big(\omega_1 fit_{train} + \omega_2 fit_{NEDC} + \omega_3 fit_{WLTC}\Big) \tag{6.3}$$

where fit_{train} is the training fit performance, fit_{NEDC} and fit_{WLTC} are the validation performances for steady-state (NEDC) and transient (WLTC) tests, respectively. These performances are calculated by the fit metric given by

$$fit = 100 \times \left(1 - \frac{\|y - \hat{y}\|}{\|y - \bar{y}\|}\right) \tag{6.4}$$

The weights in (6.3) (ω_1, ω_2 and ω_3) are calculated as

$$\omega_1 = \frac{100}{max(fit_{train})}, \quad \omega_2 = \frac{100}{max(fit_{NEDC})}, \quad \omega_2 = \frac{100}{max(fit_{WLTC})} \tag{6.5}$$

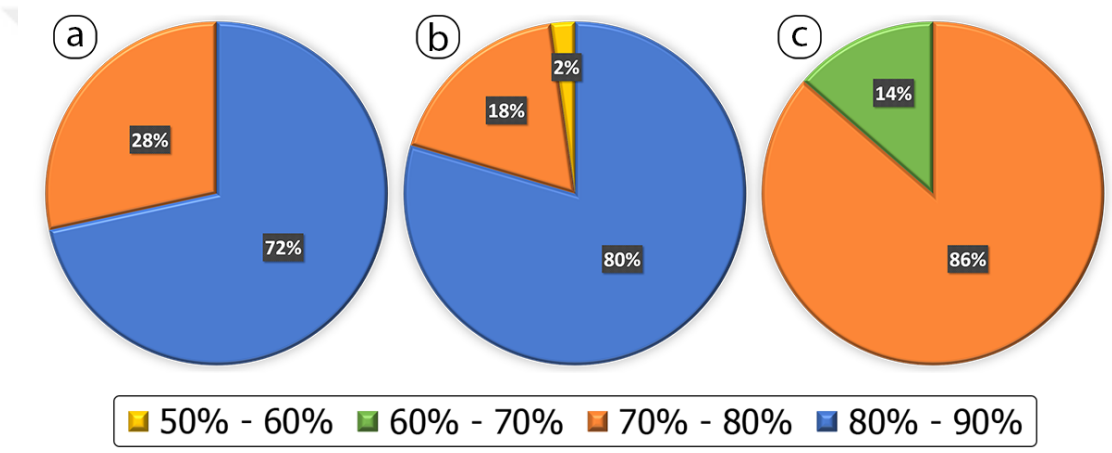


FIGURE 6.4: Distributions of (a) training, (b) steady-state validation and (c) transient validation performances.

After the determination of the best optimization parameters for each pair of unit and regressor numbers, 88 different models are obtained. Distributions of training, steady-state and transient validation performances of those 88 models are depicted in Figure 6.4. Each slice of the pie chart represents the percentage of the models which have the fit performance in the same interval. Figure 6.4(a) shows that the training fit accuracies of the all models are higher than 70% and 72% of the models have the training fit accuracy between 80% and 90%. Distributions of the validation performances show that 80% of the models have the fit accuracy between 80% and 90% under steady-state cycles and 86% of the models have the fit accuracy between 70% and 80% under transient cycles (Figure 6.4(b)-(c)). It should be noted that the WLTC includes higher loads and velocity distributions

than the NEDC [96], which makes it more challenging to obtain fit accuracies higher than 80% under transient cycles.

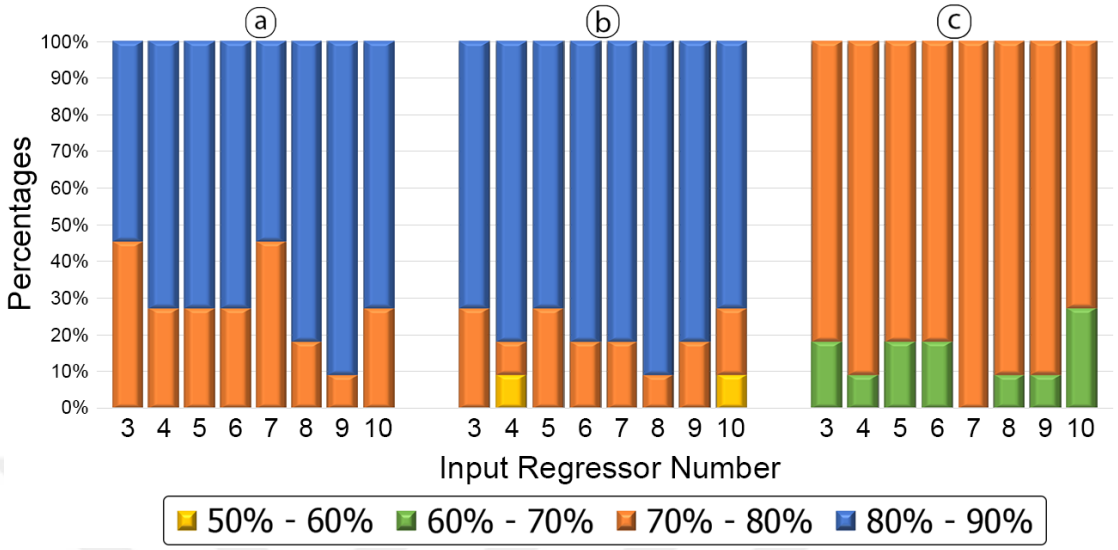


FIGURE 6.5: (a) Training, (b) steady-state and (c) transient validation performances by changing the input regressors number

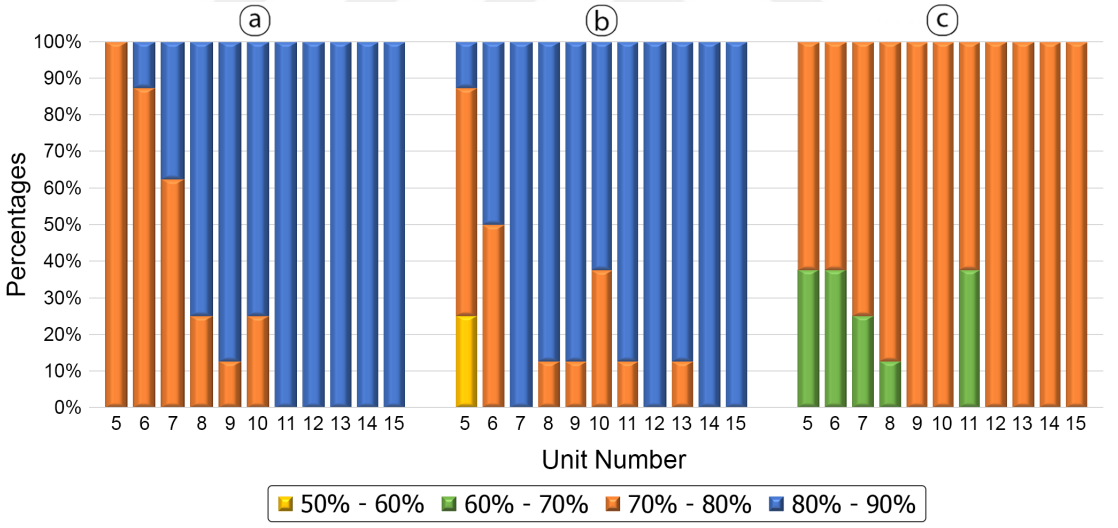


FIGURE 6.6: (a) Training, (b) steady-state and (c) transient validation performances by changing the unit number

Distributions of the training and the validation performances by changing the input regressor and the unit numbers are presented in Figure 6.5 and 6.6, respectively. These distributions reveal the influences of model structure parameters on the training and the validation performances. Figure 6.5 shows that 8 or 9 input regressors would be a reasonable choice to obtain high training and validaton accuracies. The models with at least 11 units in the hidden layer achieved the

training accuracy higher than 80% (Figure 6.6). Furthermore, all the models with 14 or 15 units have the highest training and validaton accuracies.

In order to investigate the pairwise effects of model structure parameters, 3D model performances are depicted in Figure 6.7, where the third dimension is the best fit performance of the model and the first two dimensions are the corresponding unit and regressor numbers.

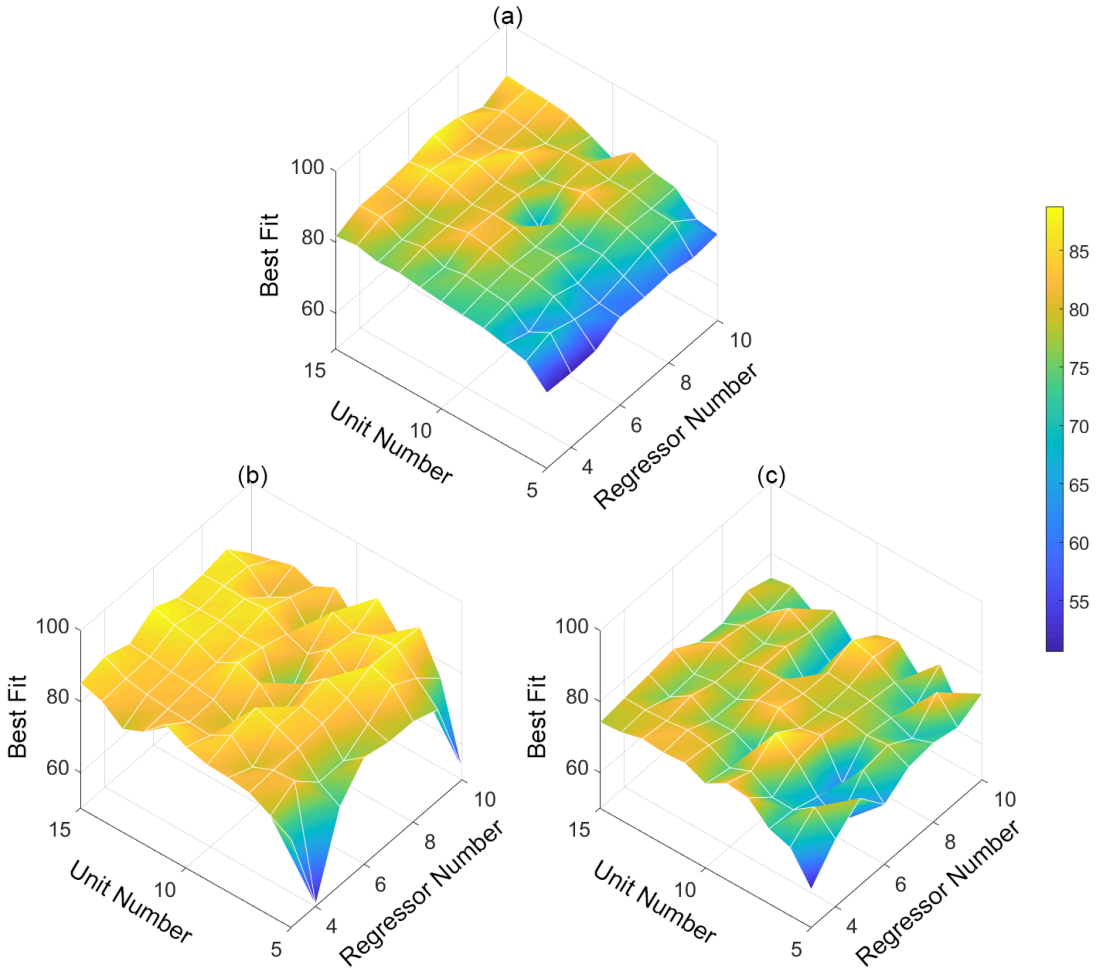


FIGURE 6.7: 3D surface plots of (a) training, (b) steady-state and (c) transient validation performances

The yellow regions represent the high fit accuracies, whereas the regions of low fit accuracies are shown in blue (Figure 6.7). Thus, the problem of finding a good model with high training and validation performances boils down to the searching for the peak points in the presented graphs.

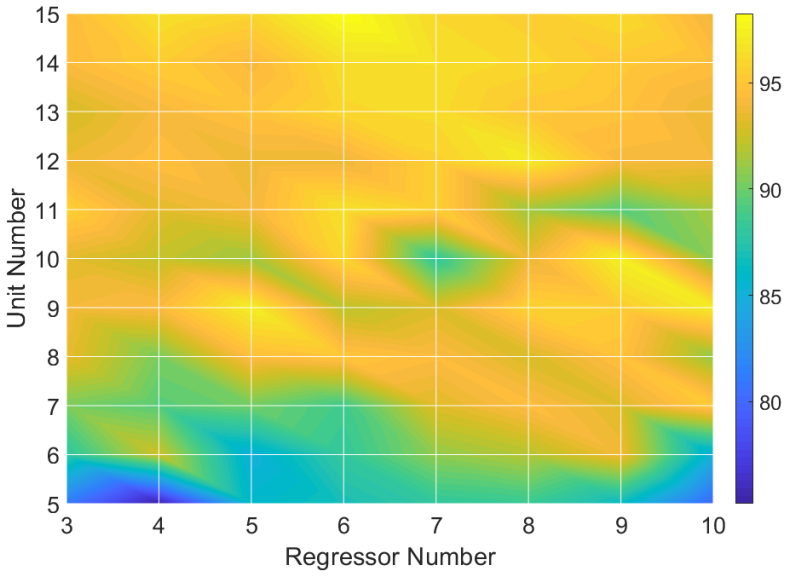


FIGURE 6.8: Obtained map for parameter selection

Since it may not always be easy to find the best model that is valid for each graph, the points given in (6.3) are calculated for these models by considering their performances and a single easy-to-interpret map for parameter selection is obtained (Figure 6.8). Obtained map can be a convenient engineering solution to select the required parameters for diesel engine emissions modeling with limited testing time in powertrain development. It should be noted that the regions with similar yellow colors indicate the places of robust models with high performances. The map in Figure 6.8 also demonstrates the sensitivity of the models for parameter changes. For instance, it would be recommended to select the center point of a large yellow area to obtain a robust model with high validation accuracies.

Chapter 7

Diesel Engine Soot Emission Modeling

Azama et al. [3] reported that the second highest percentage of the diesel engine exhaust outflows belong to soot emissions. As in the case of NOx emissions, soot is also a result of the inefficient combustion process. It is also known that there is a strong relationship between the soot formation and the ratio between air and fuel quantity [52]. However, establishing the complete relations between combustion inputs and soot emissions is indeed an extremely challenging task, since soot formation highly depends on many interactive physical processes.

Within the scope of this thesis, gated recurrent unit (GRU) and long short term memory (LSTM) type recurrent neural networks (RNN) are proposed to model soot emission of a heavy-duty diesel engine. RNNs are artificial neural networks where the connections between neurons form a directed loop. This structure allows the modeling of dynamic temporal behaviors. Unlike feedforward neural networks, RNNs also use their own hidden memory in the modeling process. In a conventional neural network, all inputs (and outputs) are assumed to be independent of each other, but, in a RNN model, each element of an array performs the same task by repeating it. The main purpose of recurrent neural networks is to use sequential information. This makes it possible to use RNN structures to predict input signals

such as time series. Furthermore, unlike NARX networks, RNN structures does not require predetermination of parameters such as pure delay between input and output signals and the number of regressors thanks to their memory capabilities.

In order to capture the nonlinear behavior of soot emissions, logarithmic normalization was applied to soot measurements as follows

$$\hat{x} = \begin{cases} log(x) & x > 1 \\ 0 & |x| \leq 1 \\ -log(-x) & x < -1 \end{cases} \tag{7.1}$$

Figure 7.1 shows the raw measurements and the logarithmically normalized signal of soot emissions. It has been seen that the logarithmic normalization cancels the sudden spikes and results in a smoother signal.

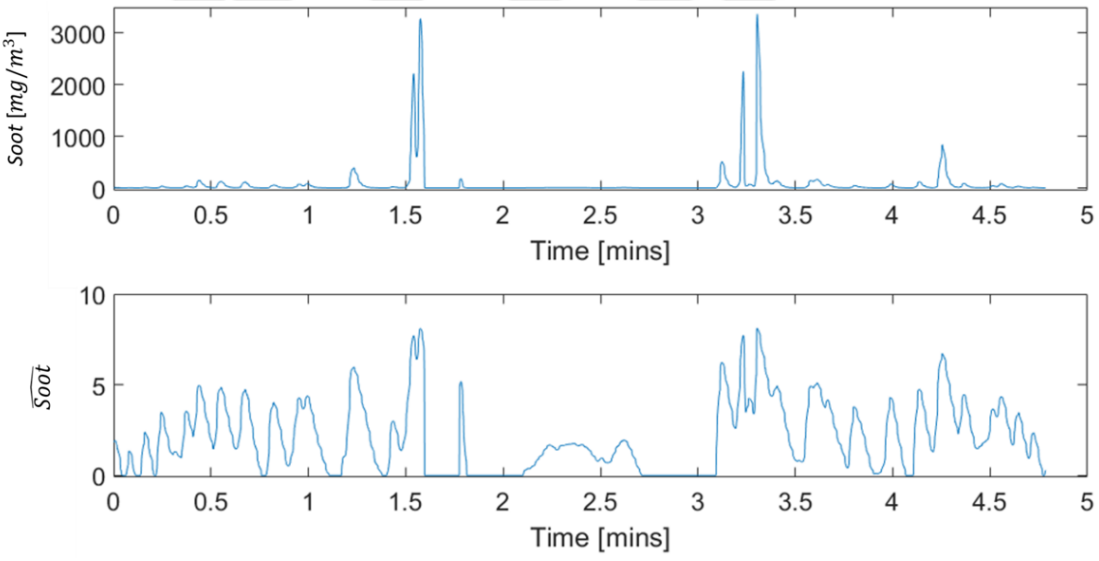


FIGURE 7.1: Logarithmic normalization of soot measurements

It should be noted that there might not be an exponential relationship between soot measurements and model inputs. Here, logarithmic normalization is applied to filter the acquired suddenly changing measurements and extract meaningful features. Since logarithm is a bijective function, it is always possible to apply the inverse operation.

In the following sections, GRU and LSTM modeling structures will be described.

7.1 GRU Modeling

A typical GRU neural network includes an input layer, a recurrent hidden layer, and an output layer. Different from classical neural networks, GRU networks have memory capabilities with two gates called: reset gate and update gate (Fig. 7.2).

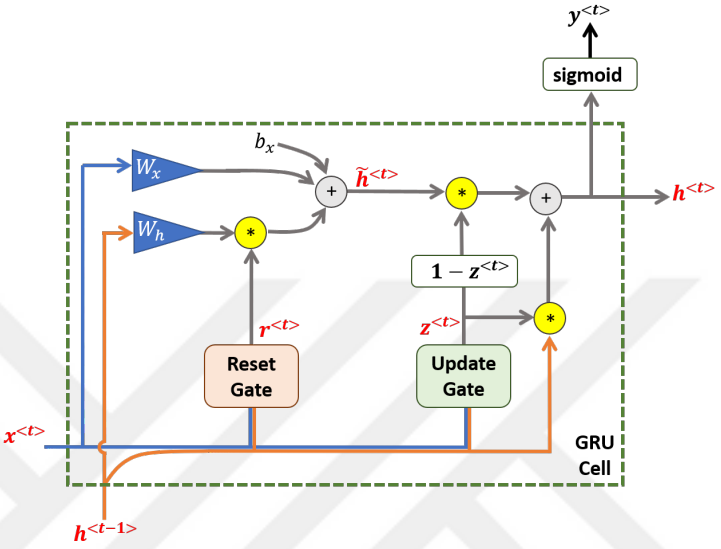


FIGURE 7.2: Gated Recurrent Unit Network structure

These gates are independent neural networks with the same dimension and sigmoid activation functions. The input to these gates is the concatenation of system measurements and the memory content of previous GRU cell ($h^{<t-1>}$). Reset gate adjusts the information to discard from the cell and update gate determines the values from the input to update the current memory content ($h^{<t>}$). A typical GRU network can be implemented by the following equations

$$z^{<t>} = \sigma\Big(W_u[h^{<t-1>}, x^{<t>}] + b_u\Big) \tag{7.2}$$

$$r^{<t>} = \sigma\Big(W_r[h^{<t-1>}, x^{<t>}] + b_r\Big) \tag{7.3}$$

$$\tilde{h}^{<t>} = tanh\Big(W_x x^{<t>} + \big(r^{<t>} * W_h h^{<t-1>}\big) + b_x\Big) \tag{7.4}$$

$$h^{<t>} = \big(z^{<t>} * h^{<t-1>}\big) + \Big(\big(1 - z^{<t>}\big) * \tilde{h}^{<t>}\Big) \tag{7.5}$$

$$y^{<t>} = \sigma\Big(W_y h^{<t>} + b_y\Big) \tag{7.6}$$

where W_u, W_r, W_x, W_h, W_y are the weight matrices and b_u, b_r, b_x, b_y are the bias vectors of corresponding operations. $\sigma(.)$ denotes the sigmoid activation function (5.8). $(*)$ is the element-wise multiplication. $x^{<t>}$ and $y^{<t>}$ denote the input vector and output prediction at time t, respectively. In this application, the GRU network is employed in a regression problem, therefore the GRU cell is followed by a fully connected regression layer in the output layer (7.6). Generally, the activation function of regression layers in neural network structures is a linear function. However, in this study, a sigmoid activation function is utilized in the output layer which results in smoother signals.

7.2 LSTM Modeling

Long short term memory (LSTM) networks are the most complex version of RNNs. An LSTM cell is the smallest unit of the structure and it is depicted in Figure 7.3.

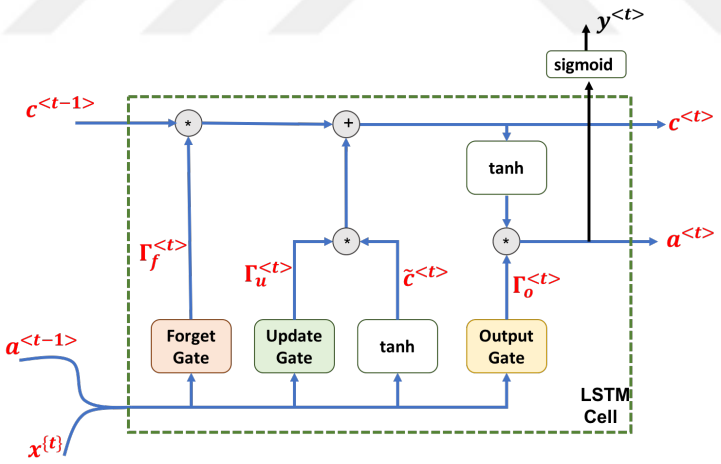


FIGURE 7.3: LSTM network structure

Inputs of the single LSTM cell are the previous state vector ($a^{<t-1>}$), the previous memory vector ($c^{<t-1>}$) and the current measurement vector ($x^{<t>}$). LSTM cell outputs the current state vector ($a^{<t>}$) and current memory vector ($c^{<t>}$). Thanks to the memory capability of the LSTM cell, some parts of the information extracted from measurements are stored and some of them are forgotten. The amount of

the information to be forgotten or to be stored for upcoming cell is decided by three different small neural networks called "Forget Gate", "Update Gate" and "Output Gate". A typical LSTM network can be implemented by the following equations

$$\Gamma_f = \sigma\Big(W_f[a^{<t-1>}, x^{<t>}] + b_f\Big) \tag{7.7}$$

$$\Gamma_u = \sigma\Big(W_u[a^{<t-1>}, x^{<t>}] + b_u\Big) \tag{7.8}$$

$$\tilde{c}^{<t>} = tanh\Big(W_c[a^{<t-1>}, x^{<t>}] + b_c\Big) \tag{7.9}$$

$$\Gamma_o = \sigma\Big(W_o[a^{<t-1>}, x^{<t>}] + b_o\Big) \tag{7.10}$$

$$c^{<t>} = \Gamma_u * \tilde{c}^{<t>} + \Gamma_f * c^{<t-1>} \tag{7.11}$$

$$a^{<t>} = \Gamma_o * tanh(c^{<t>}) \tag{7.12}$$

$$y^{<t>} = \phi\Big(W_y a^{<t>} + b_y\Big) \tag{7.13}$$

where W_f, W_u, W_c, W_o, W_y are the weight matrices and b_f, b_u, b_c, b_o, b_y are the bias vectors of corresponding operations. The activation function in the output layer $(\phi(.))$ is selected as a linear function or a sigmoid function depending on the modeling requirements and the amplitude of soot measurements.

Time plots of the activation functions employed in the equations of LSTM structure are presented in Figure 7.4. The parameter a in (5.8) designates the slope of sigmoid function. Once a is adjusted as a learnable parameter, a remarkable increase in soot modeling accuracies is observed. All the learnable parameters of the GRU and LSTM models are estimated by a stochastic optimizer called adaptive moment estimation (Adam)[97].

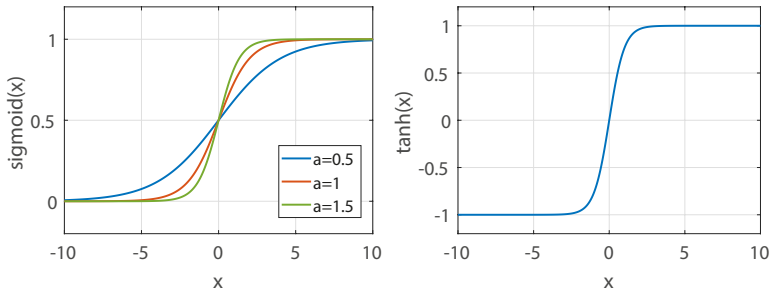


FIGURE 7.4: Sigmoid and tanh activation functions

Chapter 8

Experimental Results

In this chapter, the setup utilized in the experimentations during this PhD study is introduced, and experimental results and discussions regarding the models of diesel engine indicated torque, NOx and soot emissions are presented.

8.1 Experimental Setup

The experimental measurements are performed on Ford Otosan's Ecotorq 13L Euro 6 diesel engine. The diesel engine combustion composes of two fundamental paths named air and fuel paths. On its air-path, the engine has high-pressure Exhaust Gas Recirculation (EGR) routing and swing vane Variable Geometry Turbine (VGT) turbocharger as shown in Figure 8.1. The components in the experimental setup are: (1) Air Filter, (2) MAF sensor, (3) Compressor, (4) Turbine, (5) VGT, (6) Exhaust pipe, (7) Charge air cooler, (8) EGR cooler, (9) EGR valve, (10) Exhaust manifold, (11) Exhaust gas measurement system, (12) Pressurized rail, (13) Injector, (14) TMAP sensor, (15) Fuel tank, (16) High pressure fuel pump, and (17) AVL Dynamometer.

The common pressurized fuel rail, solenoid injectors and high-pressure fuel pump are components of the fuel system. The ambient fresh air which flows through air filter is measured before entering to the compressor. This air flow measurement is

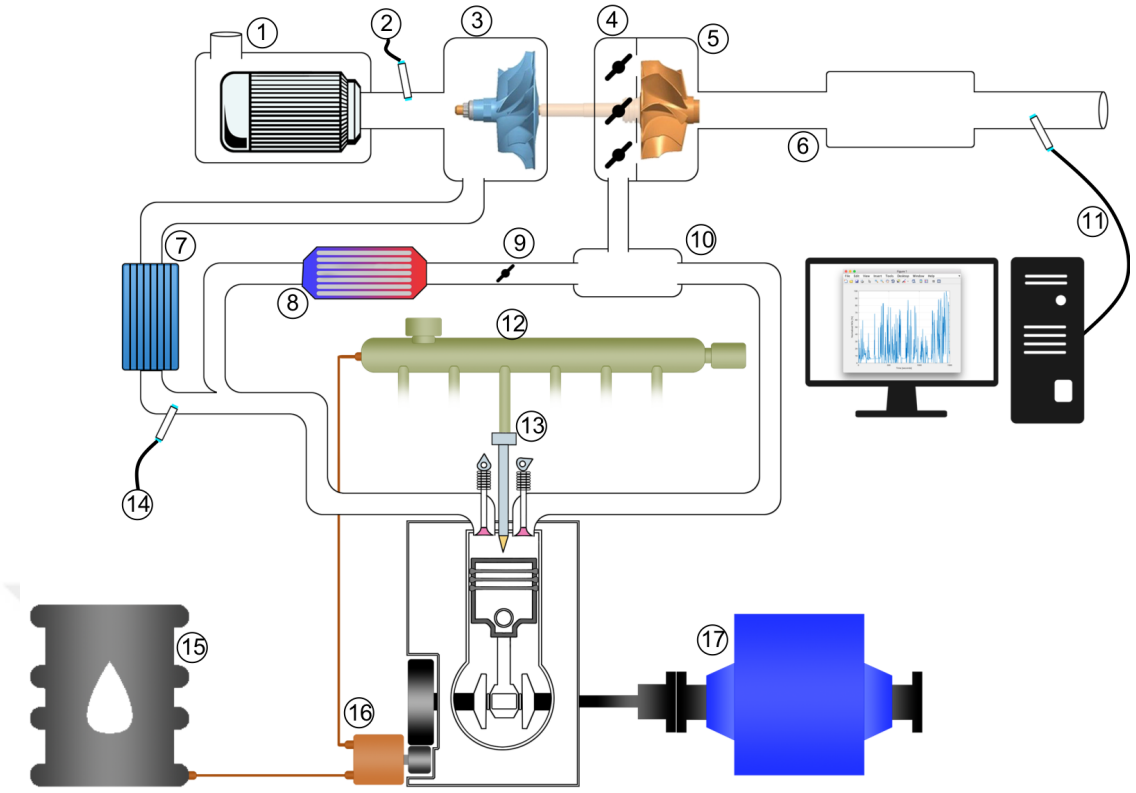


FIGURE 8.1: Experimental setup

stated as Mass Air Flow (MAF). After the compressor, its pressure is increased to a regulated level, and in order to boost the volumetric efficiency of the engine, the compressed air flow is forwarded to the charge air cooler which reduces the flow temperature to a regulated level. Before the EGR mixing point, the temperature and the pressure of the cooled charger air are measured. The cooled charger air flow which is mixed with EGR flow is injected into the combustion chamber through the pressurized intake manifold. The pressure of the mixed air flow in the intake manifold is named as Manifold Absolute Pressure (MAP). After the combustion, through the exhaust manifold the exhaust gases exit the chamber and certain portion of them are recirculated back via EGR valve.

Fuel is injected into a pressurized rail. The fuel amount and the pressure in the rail are measured as Fuel Quantity (QNT) and Rail Pressure (RailP), respectively. The starting angle of the main fuel injection quantity is defined as Start of Injection (SOI). Impact angle of the exhaust gases to turbine blades is regulated by a valve called Variable Geometry Turbine (VGT) turbochargers. The exhaust

gas expanded in the turbine is sent to the exhaust line and its soot content is measured by the AVL 483 Micro Soot Sensor. Micro soot sensor is a photoacoustic system designed for the measurement of black carbon aerosols. It is a suitable test bench with high sensitivity down to $\mu g/m^3$ level, and good linearity and reproducibility[98]. Specifications of the test engine are listed in Table 8.1.

Type	: Ecotorq 12.7 lt.
Power	: 480 PS (353 kW) /1800 d/d
Torque	: 2500 Nm/1000-1200 d/d
Compression Ratio	: 17 ± 0.5 : 1
Cylinders	: 6 in line
Cylinder Volume	: 12740 cc
Cylinder Bore	: 130 mm
Cylinder Stroke	: 160 mm
Turbocharger	: VGT
EGR	: High Pressure, Cooled
Fuel System	: Common Rail
Rail Pressure	: Max. 2500 bar

TABLE 8.1: Test engine spesifications

Soot measurements can be made by taking samples from the engine outlet, exhaust pipe or dilution tunnel. When soot samples are taken from the motor output, the particle morphology is simple but the measurement application is very complex. When samples are taken from the dilution tunnel, particle morphology is complex but the measurement application is very simple. Therefore, during the experimentations, soot samples were taken from the dilution tunnel with AVL Micro Soot sensor [99] by Constant Volume Sampling (CVS) method as shown in Figure 8.2.

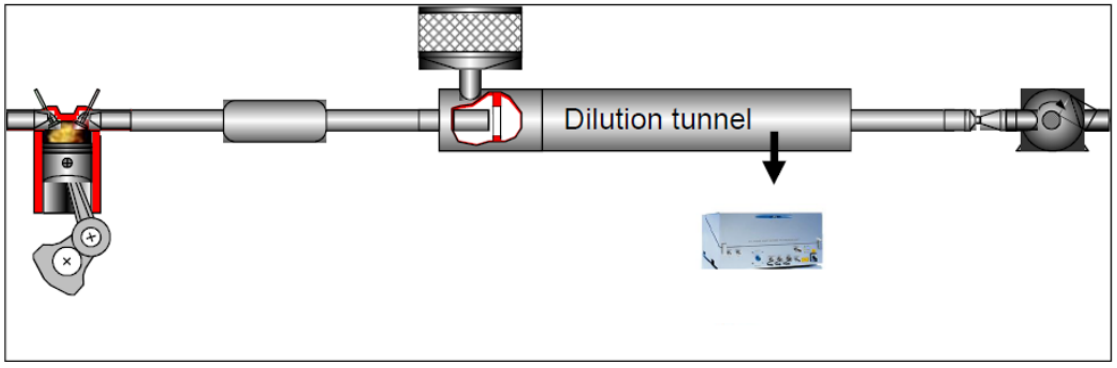


FIGURE 8.2: Sampling in the dilution tunnel

The AVL Micro Soot Sensor is based on the principle of photoacoustic measurement [99]. In this method, highly absorbing soot particles called "blacking" are irradiated to the modulus. Periodic heating and cooling, and thus the expansion and contraction of the carrier gas, can be evaluated as sound waves and detected by microphones (Figure 8.3).

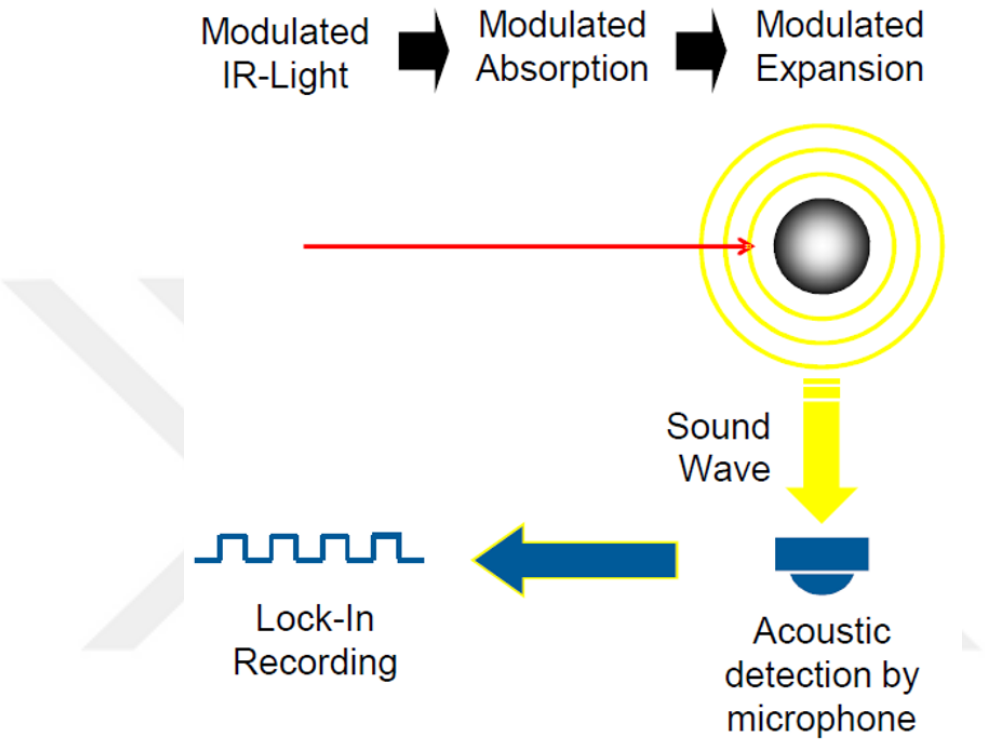


FIGURE 8.3: Operating principle of the AVL Micro Soot Sensor

Clean air does not produce a signal, so the presence of soot-laden air or exhaust gas increases the proportion of the measured signal proportionally to the volume of soot. This is an important advantage over opasymmetric methods. Thus, a significant increase in measurement accuracy can be achieved.

8.2 Results and Discussions

In this section, training and validation performances of the torque and combustion emissions are presented, and the experimental results are discussed with the time plots of the model responses.

8.2.1 Torque Modeling

It was a priori known that the fuel quantity and indicated torque are highly correlated (5.9) in diesel combustion, so initially single input (QNT) indicated torque models with NFIR structure (Figure 5.5) were obtained. The pure delay between inputs and output was determined as 0.1 seconds, i.e., 1 sample since the acquisition frequency is 10 Hz. Utilized data in modeling belong to 4 different experiments (Figure 4.6). Models were obtained with the same parameters, and their performances are calculated by the fit metric calculated as

$$fit = 100 \times \left(1 - \frac{\|y - \hat{y}\|}{\|y - \bar{y}\|}\right) \tag{8.1}$$

Training, validation and cross-validation performances of a single input indicated torque models are tabulated in Table 8.2.

	Inputs	Output	nb	unit	Iteration	Estimation %	WHSC %	WHTC EOM0 %	WHTC EOM4 %	CV T2 %	CV T3 %	CV T4 %
T1	QNT	indicatedTrq	5	10	80	81,22	84,93	82,53	82,92	84,35	78,30	79,27
	QNT	indicatedTrq	5	15	60	81,75	84,78	83,90	84,19	84,68	78,92	82,55
	QNT	indicatedTrq	8	15	80	83,50	84,17	84,69	84,78	85,40	80,88	84,40
	QNT	indicatedTrq	3	10	60	81,28	83,62	84,07	84,16	84,35	78,81	82,08

	Inputs	Output	nb	unit	Iteration	Estimation %	WHSC %	WHTC EOM0 %	WHTC EOM4 %	CV T1 %	CV T3 %	CV T4 %
T2	QNT	indicatedTrq	5	10	80	86,50	89,85	84,63	84,94	80,06	78,27	84,79
	QNT	indicatedTrq	5	15	60	86,55	89,85	84,12	84,51	79,71	77,88	84,97
	QNT	indicatedTrq	8	15	80	86,35	88,73	83,17	83,74	78,21	76,62	82,66
	QNT	indicatedTrq	3	10	60	85,77	88,84	85,26	85,33	80,07	77,87	83,87

	Inputs	Output	nb	unit	Iteration	Estimation %	WHSC %	WHTC EOM0 %	WHTC EOM4 %	CV T1 %	CV T2 %	CV T4 %
T3	QNT	indicatedTrq	5	10	80	79,30	86,31	84,67	84,76	81,27	85,16	83,27
	QNT	indicatedTrq	5	15	60	79,60	85,87	83,51	83,63	81,24	85,17	83,74
	QNT	indicatedTrq	8	15	80	81,31	85,32	84,21	84,37	82,63	85,53	84,84
	QNT	indicatedTrq	3	10	60	78,95	82,95	83,81	84,23	80,85	83,92	82,84

	Inputs	Output	nb	unit	Iteration	Estimation %	WHSC %	WHTC EOM0 %	WHTC EOM4 %	CV T1 %	CV T2 %	CV T3 %
T4	QNT	indicatedTrq	5	10	80	85,01	87,85	83,24	83,49	78,75	85,60	77,30
	QNT	indicatedTrq	5	15	60	85,30	88,16	84,51	84,47	78,82	85,83	77,24
	QNT	indicatedTrq	8	15	80	85,61	89,09	83,86	84,00	76,32	86,15	75,10
	QNT	indicatedTrq	3	10	60	84,60	86,84	84,47	84,85	79,19	85,19	77,36

TABLE 8.2: Indicated torque models with a single input (QNT)

In order to see the effects of calibratable inputs in indicated torque generation, the input channels MAP, MAF and miSOI are included in the input set, one-by-one. The training, validation and cross-validation performances of initially obtained four inputs indicated torque models are tabulated in Table 8.3. Based on these modeling experiences, four inputs (QNT, MAP, MAF, miSOI) indicated torque models are shown to provide better training and validation accuracies when compared to single input models.

	Inputs	Output	nb	unit	Iteration	Estimation %	WHSC %	WHTC EOM0 %	WHTC EOM4 %	CV T2 %	CV T3 %	CV T4 %
T1	QNT MAP MAF miSOI	indicatedTrq	4	8	80	84,84	91,74	86,36	86,05	88,28	81,96	86,25
	QNT MAP MAF miSOI	indicatedTrq	3	5	80	83,86	91,98	86,37	85,87	87,08	80,81	85,09
	QNT MAP MAF miSOI	indicatedTrq	5	6	40	84,89	93,02	87,37	86,82	88,15	82,01	86,11
	QNT MAP MAF miSOI	indicatedTrq	6	15	60	85,34	89,52	86,23	86,13	86,74	81,68	85,18

	Inputs	Output	nb	unit	Iteration	Estimation %	WHSC %	WHTC EOM0 %	WHTC EOM4 %	CV T1 %	CV T3 %	CV T4 %
T2	QNT MAP MAF miSOI	indicatedTrq	4	8	80	89,87	94,34	86,75	85,87	80,20	78,81	87,47
	QNT MAP MAF miSOI	indicatedTrq	3	5	80	88,12	93,17	87,45	86,90	80,44	78,27	86,05
	QNT MAP MAF miSOI	indicatedTrq	5	6	40	88,85	94,39	87,30	85,87	80,02	78,57	86,84
	QNT MAP MAF miSOI	indicatedTrq	6	15	60	90,42	94,92	86,94	85,93	73,32	73,92	88,04

	Inputs	Output	nb	unit	Iteration	Estimation %	WHSC %	WHTC EOM0 %	WHTC EOM4 %	CV T1 %	CV T2 %	CV T4 %
T3	QNT MAP MAF miSOI	indicatedTrq	4	8	80	82,53	90,83	85,78	85,03	84,00	87,38	86,28
	QNT MAP MAF miSOI	indicatedTrq	3	5	80	82,00	89,92	85,55	85,01	83,99	87,35	85,81
	QNT MAP MAF miSOI	indicatedTrq	5	6	40	82,86	92,18	86,18	85,88	84,16	88,06	86,72
	QNT MAP MAF miSOI	indicatedTrq	6	15	60	83,58	89,65	85,38	84,78	84,89	88,03	86,60

	Inputs	Output	nb	unit	Iteration	Estimation %	WHSC %	WHTC EOM0 %	WHTC EOM4 %	CV T1 %	CV T2 %	CV T3 %
T4	QNT MAP MAF miSOI	indicatedTrq	4	8	80	87,28	90,37	85,97	85,75	74,17	87,61	73,99
	QNT MAP MAF miSOI	indicatedTrq	3	5	80	86,85	91,10	86,58	85,86	81,29	87,34	79,76
	QNT MAP MAF miSOI	indicatedTrq	5	6	40	87,60	92,21	86,70	85,75	79,36	87,82	77,98
	QNT MAP MAF miSOI	indicatedTrq	6	15	60	88,52	91,34	86,04	85,82	77,66	88,34	77,38

TABLE 8.3: Indicated torque models with 4 inputs

To obtain high fidelity indicated torque models, which consist of full calibratable parameters, engine speed, rail pressure and the variables of pilot injections were also included in the input set. Then, indicated torque models with a specific range of parameters were obtained and the best models which have the highest training and validation accuracies were tabulated in Table 8.4 and their responses in time plots are presented in Figures 8.4 - 8.13.

Input Num.	Experiment	*nb*	Unit	Iter.	Train.	WHSC	WHTC	Figure
4	1	6	12	40	86,27	94,12	88,2	8.4
4	2	3	13	80	89,12	94,93	88,46	8.5
4	3	5	4	40	82,62	94,49	87,73	8.6
4	4	2	12	60	85,84	94,28	87,75	8.7
8	1	8	9	80	88.91	93.71	86.51	8.8
8	2	3	8	40	91.35	93.03	84.18	8.9
8	3	5	9	80	86.78	92.78	86.51	8.10
8	4	3	11	80	89.22	92.22	81.51	8.11
8	1+2	5	8	80	89.90	94.30	87.18	8.12
8	2+3	3	10	60	86.09	93.70	87.33	8.13

TABLE 8.4: The best indicated torque models with all experiments

Experimental results show that by merging the experiments, the best eight input indicated torque models are trained with the accuracy above 86%, and the steady-state and transient validation performances are above 93% and 87%, respectively.

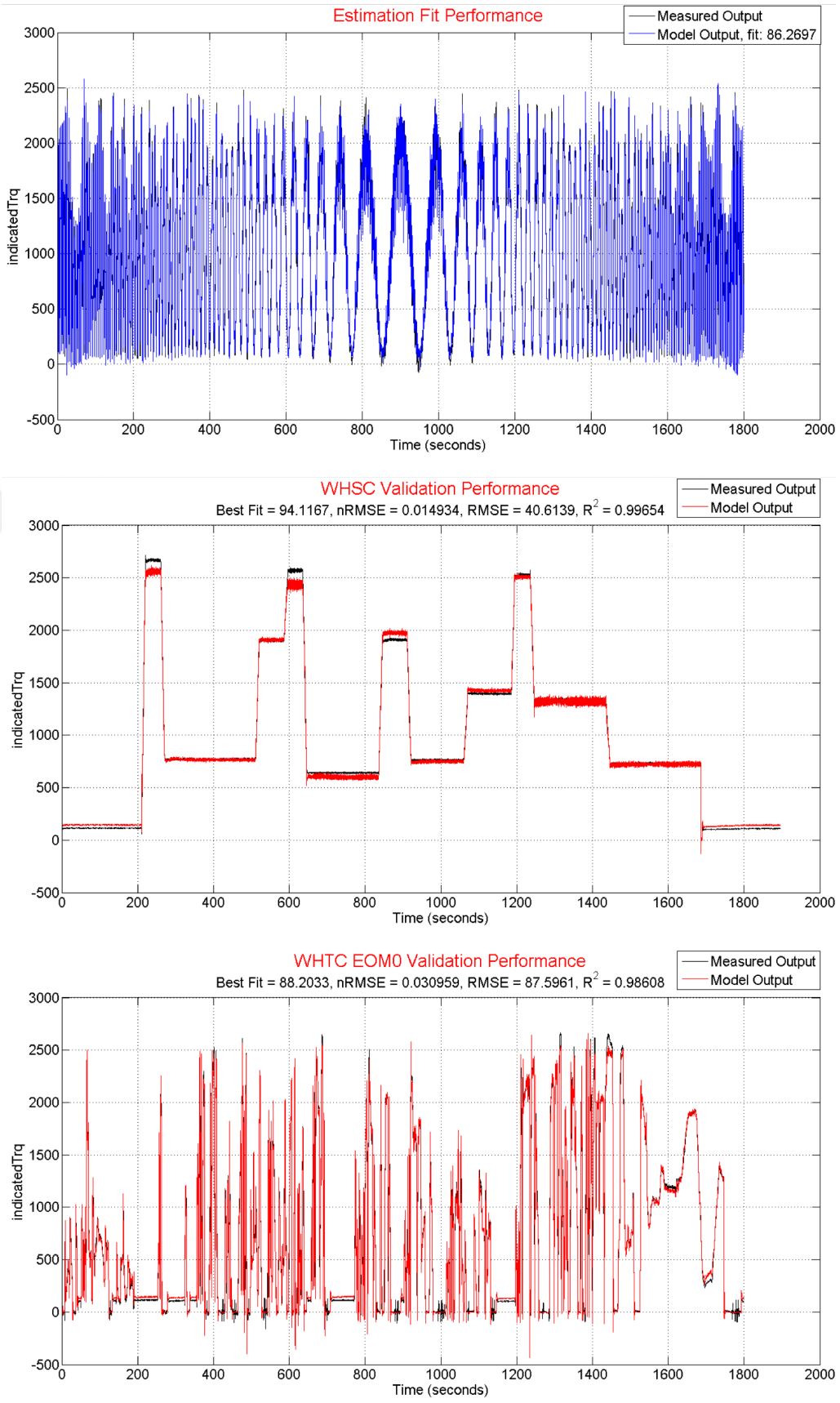


FIGURE 8.4: Ind. torque model with 4 inputs prediction performances (Ex.1)

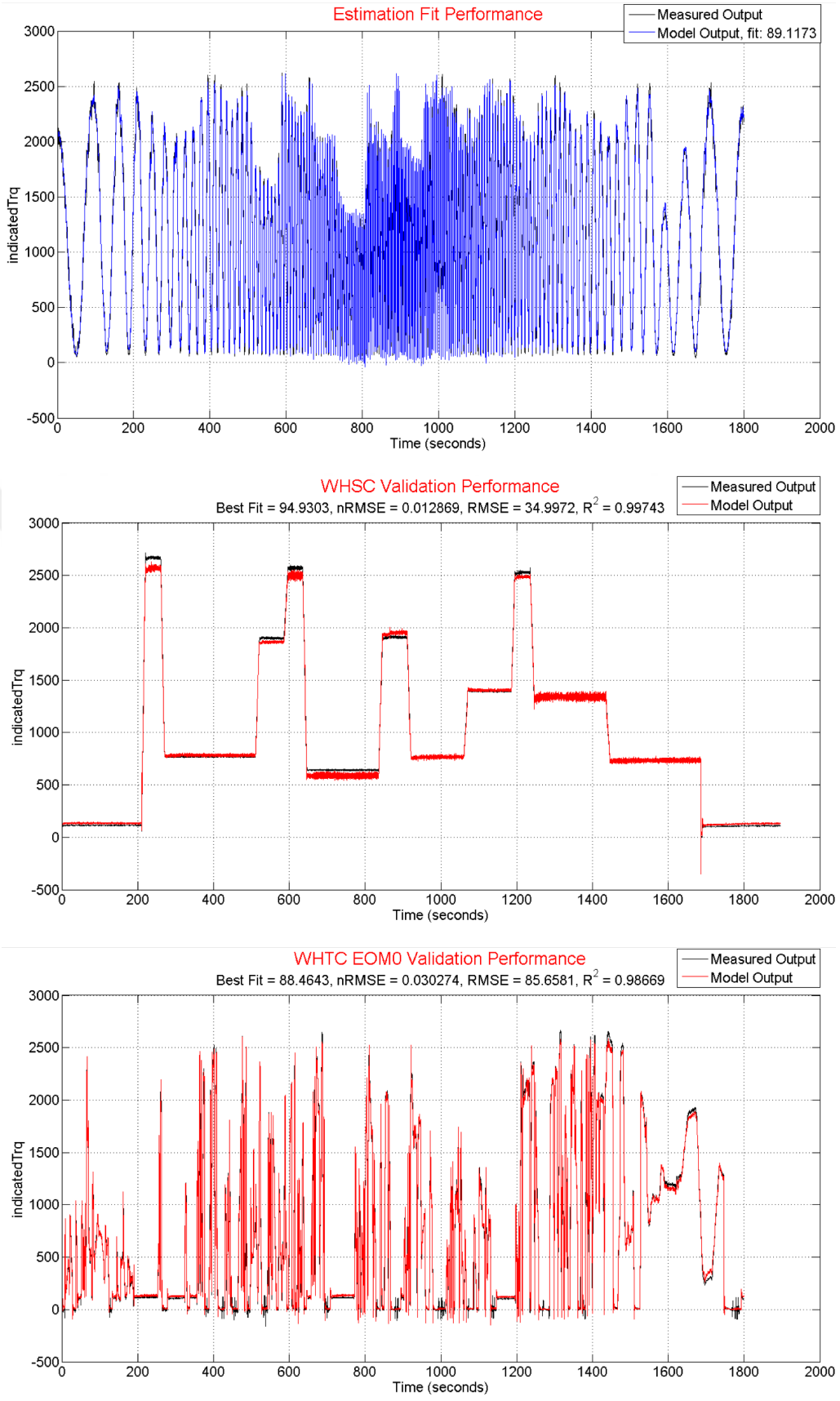


FIGURE 8.5: Ind. torque model with 4 inputs prediction performances (Ex.2)

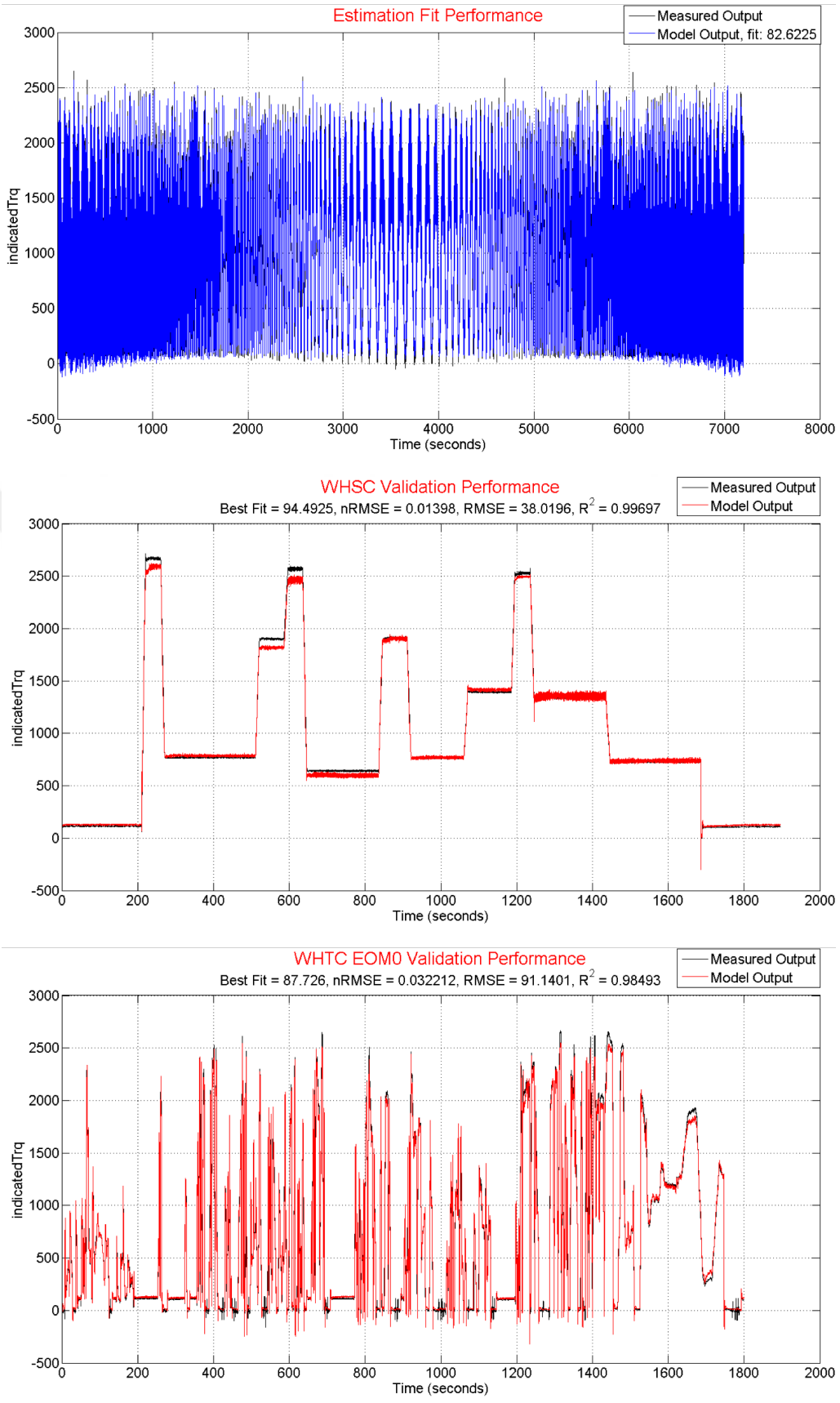


FIGURE 8.6: Ind. torque model with 4 inputs prediction performances (Ex.3)

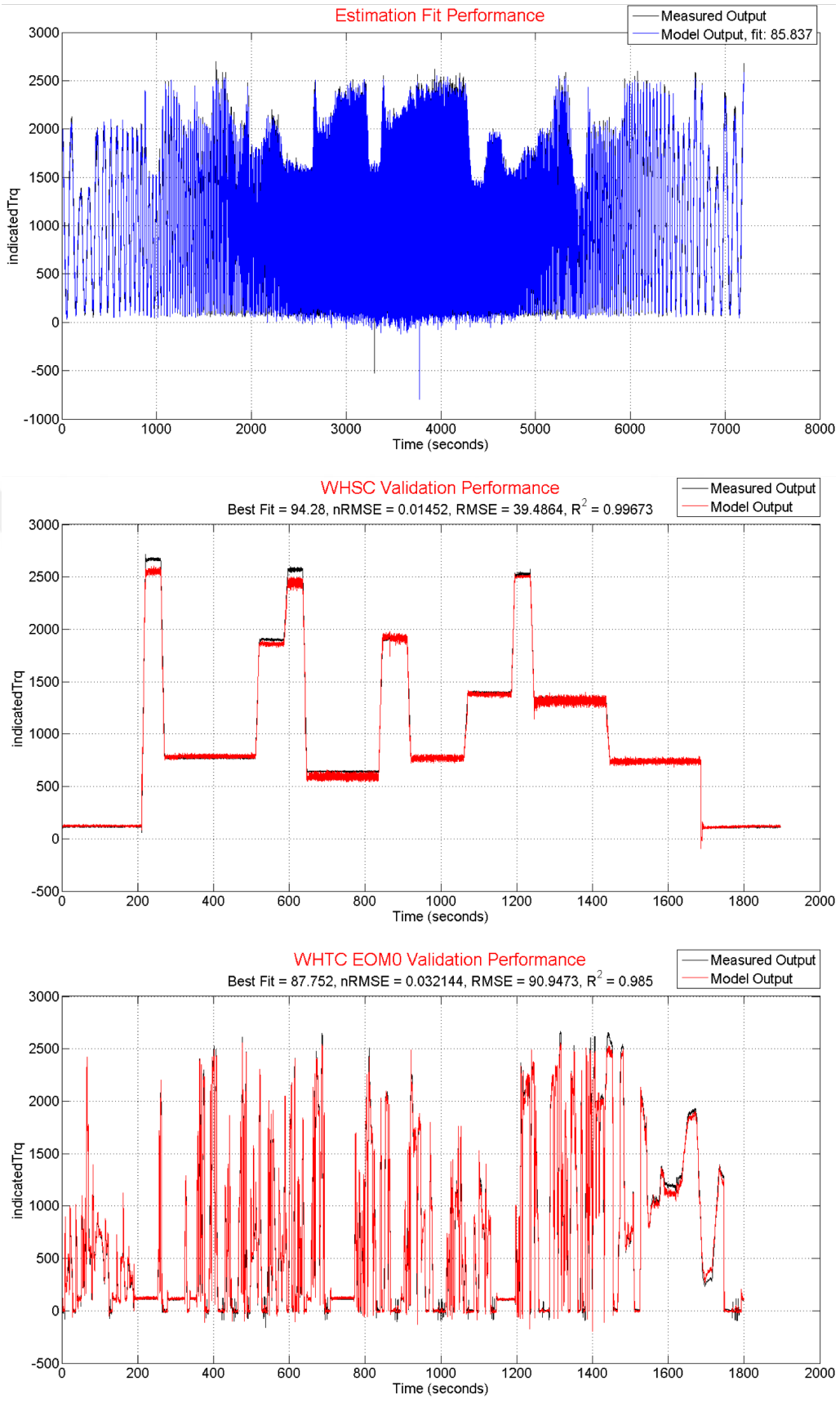


FIGURE 8.7: Ind. torque model with 4 inputs prediction performances (Ex.4)

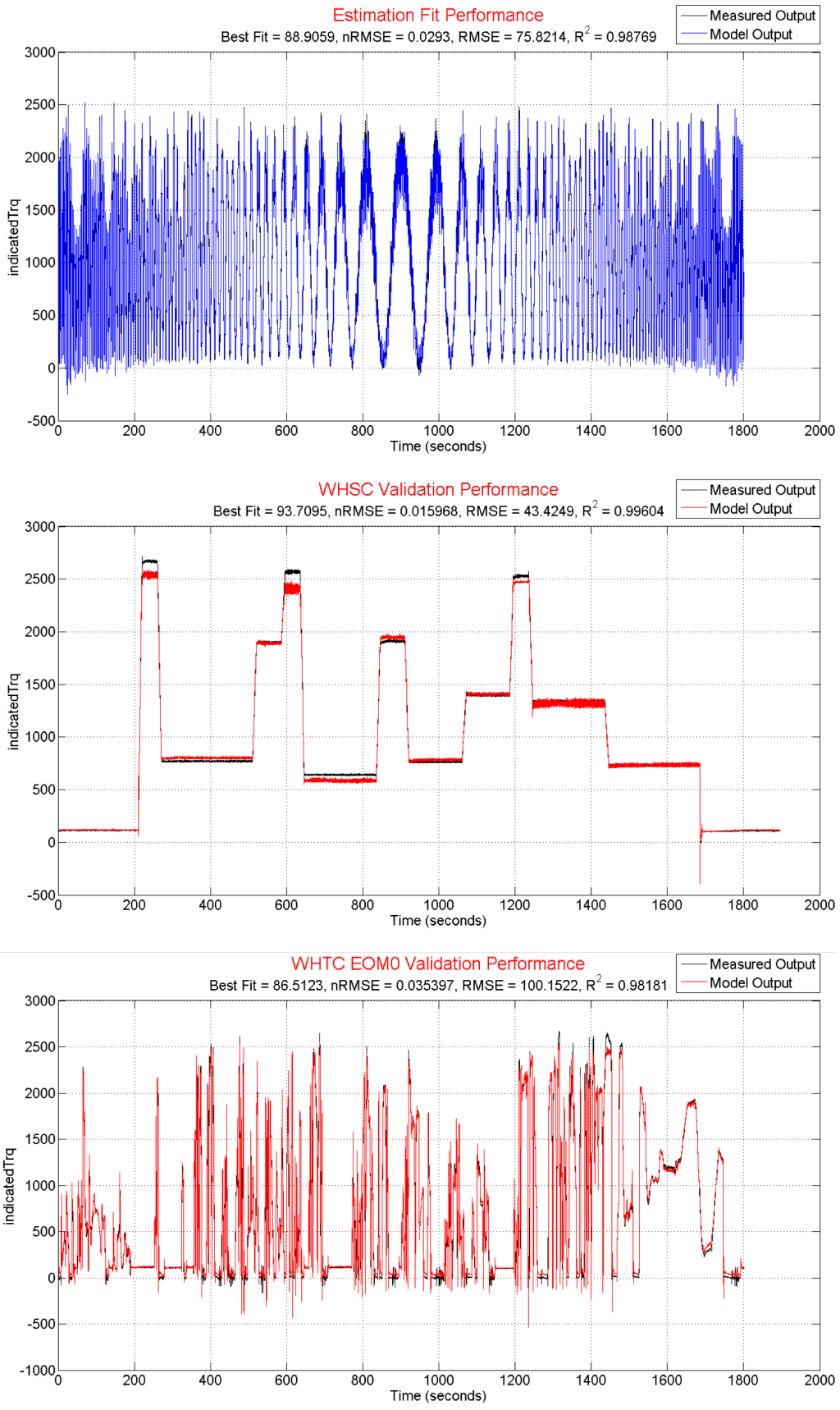


FIGURE 8.8: Ind. torque model with 8 inputs prediction performances (Ex.1)

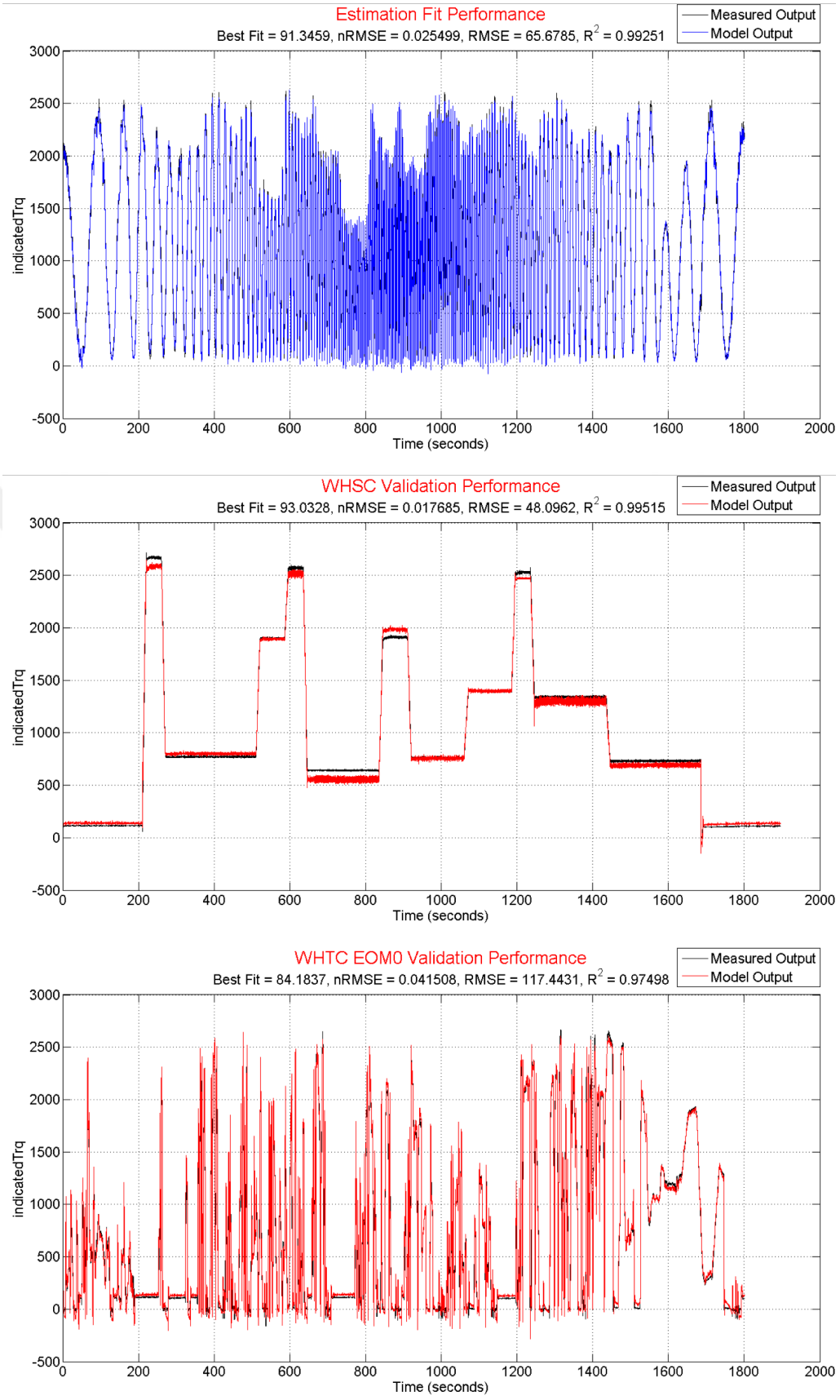


FIGURE 8.9: Ind. torque model with 8 inputs prediction performances (Ex.2)

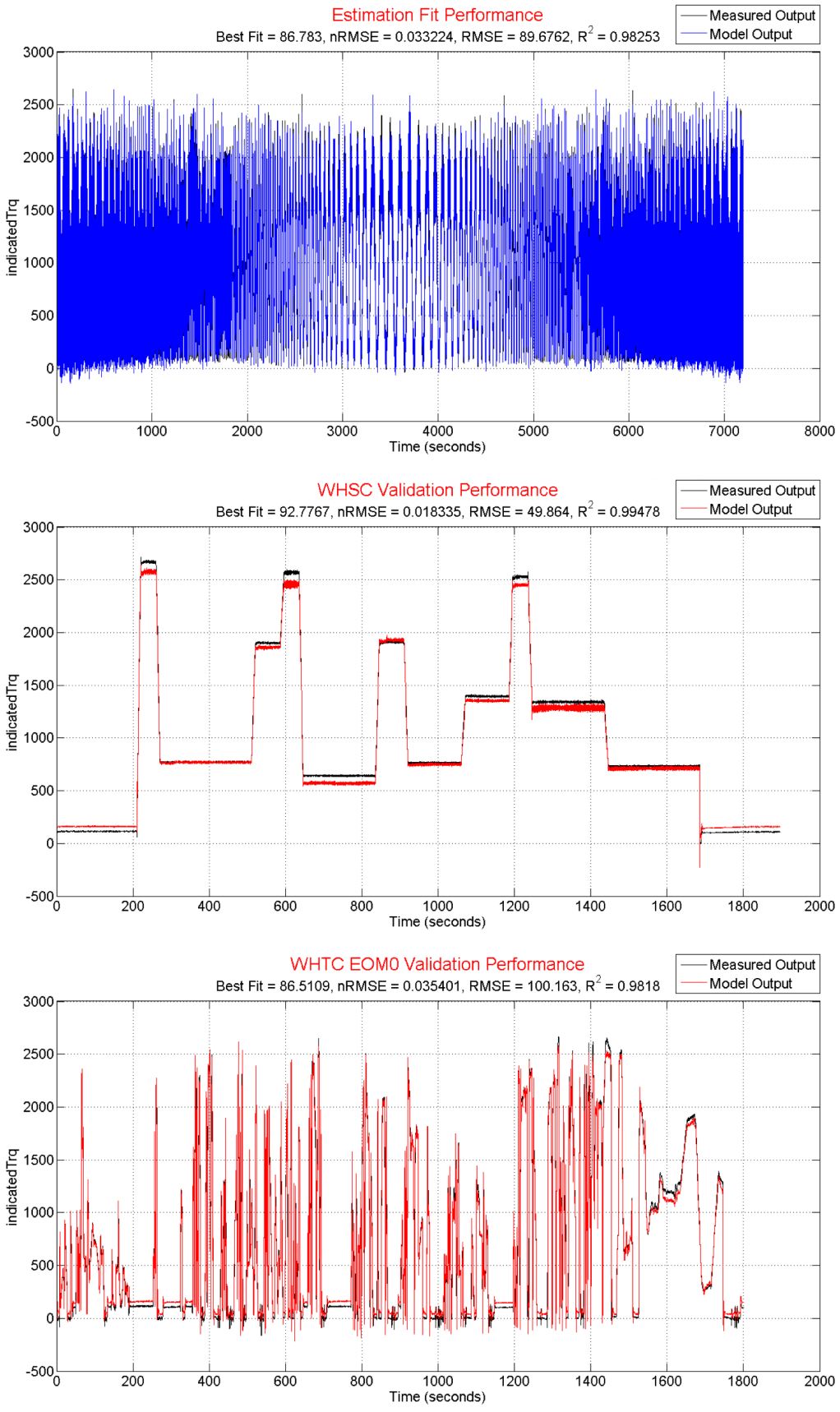


FIGURE 8.10: Ind. torque model with 8 inputs prediction performances (Ex.3)

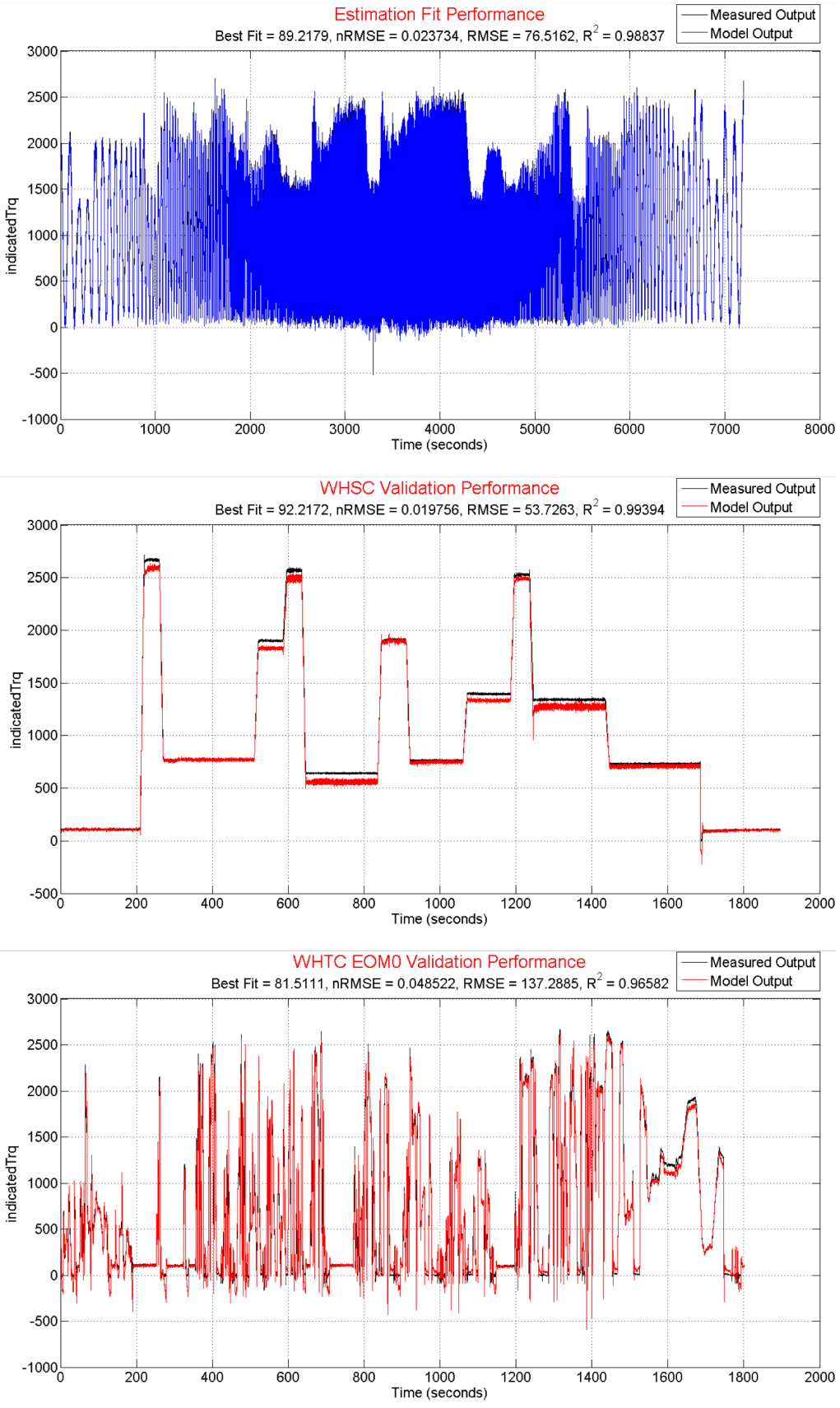


FIGURE 8.11: Ind. torque model with 8 inputs prediction performances (Ex.4)

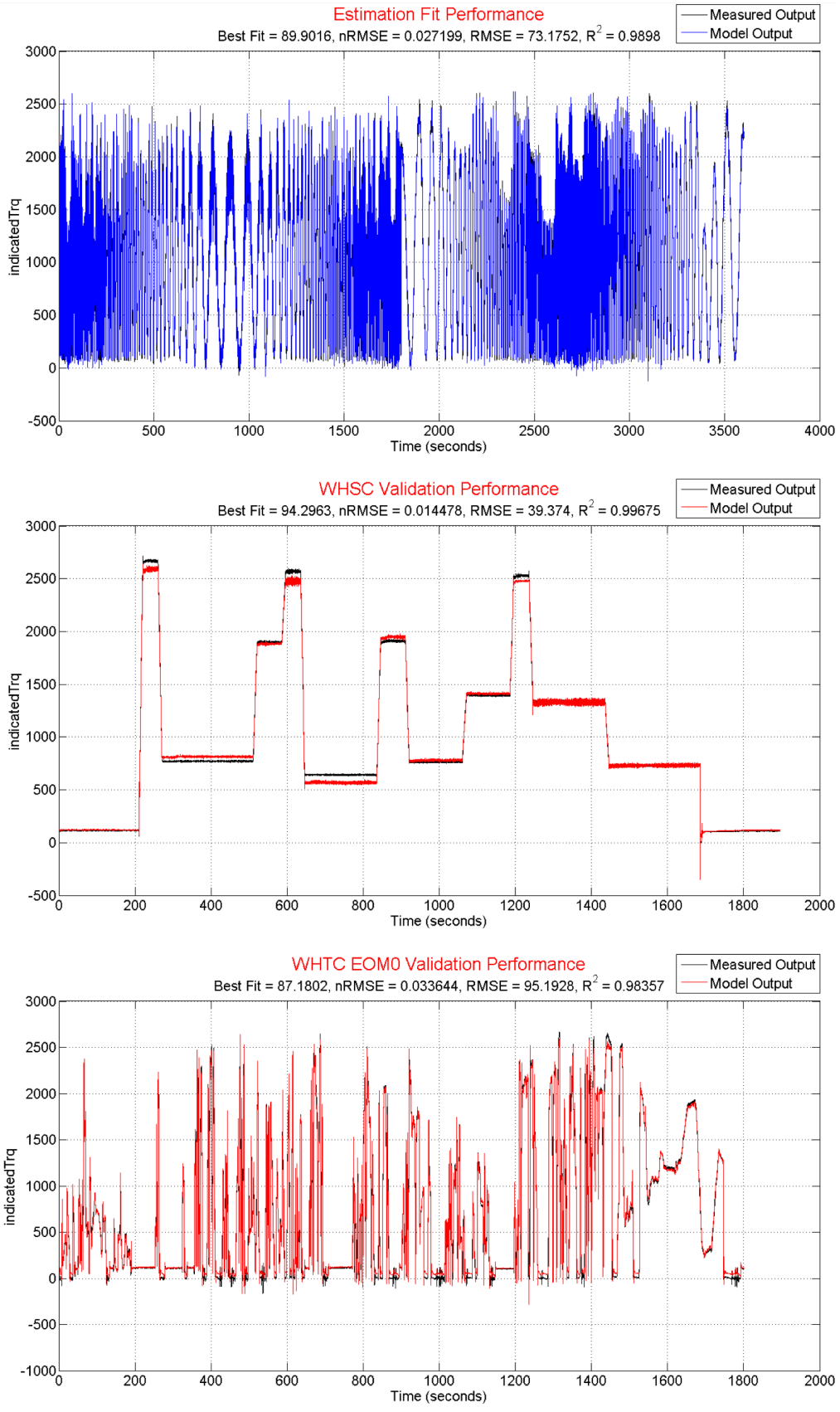


FIGURE 8.12: Ind. torque model with 8 inp. prediction performances (E.1+2)

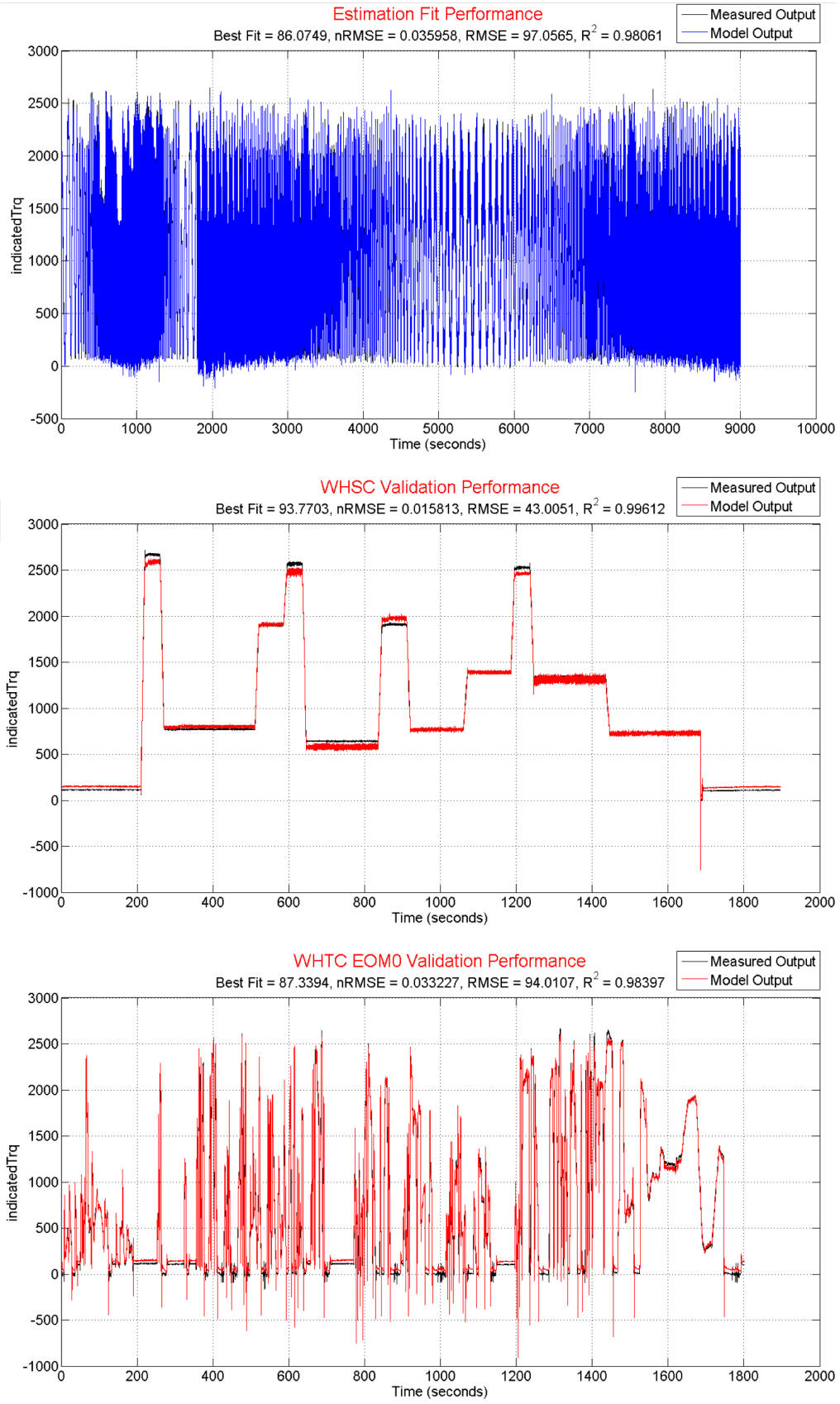


FIGURE 8.13: Ind. torque model with 8 inp. prediction performances (E.2+3)

8.2.2 NOx Modeling

An exhaustive search for the ranges of parameters given in Table 8.5 was performed to investigate the performances of the proposed modeling structure. 2970 models were obtained for both proposed structure with OLS based regressor selection and classical structure without regressor selection. For both cases, output regressors are only employed in the linear block. 6 input experimental data (Figure 4.2) are utilized to train the models. World Harmonic Transient Cycle (WHTC) tests for two different variants of motor power were also applied for validation purposes.

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

TABLE 8.5: Ranges of parameters

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

$$Point_i = \frac{100}{3}\Big(\frac{est_i}{max(est)} + \frac{val1_i}{max(val1)} + \frac{val2_i}{max(val2)}\Big) \tag{8.2}$$

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 (8.1). Note that the iteration of each model is increased by 20 and the model obtained in each 20 iteration is counted in 2970 models. In order to get a much more informative statistics, a single model which took the highest point according to (8.2) is selected out of these 5 models.

Estimation and validation performance distributions of these models are presented in Figure 8.14, 8.15 and 8.16. Distributions show that the models obtained by the proposed modeling structure with less number of parameters show the same validation performances as classical ones. In Figure 8.14, it is shown that the estimation performances of all models obtained by the 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 8.6.

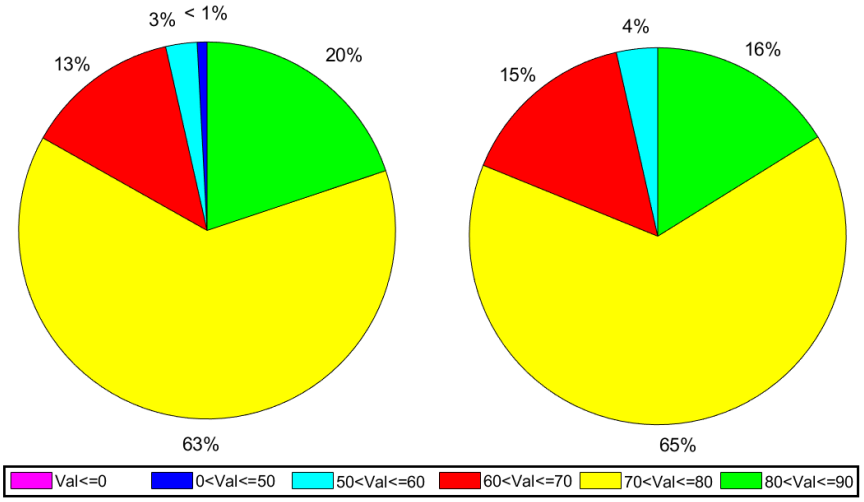


FIGURE 8.14: Est. performance distribution of the models Left:classic Right:proposed

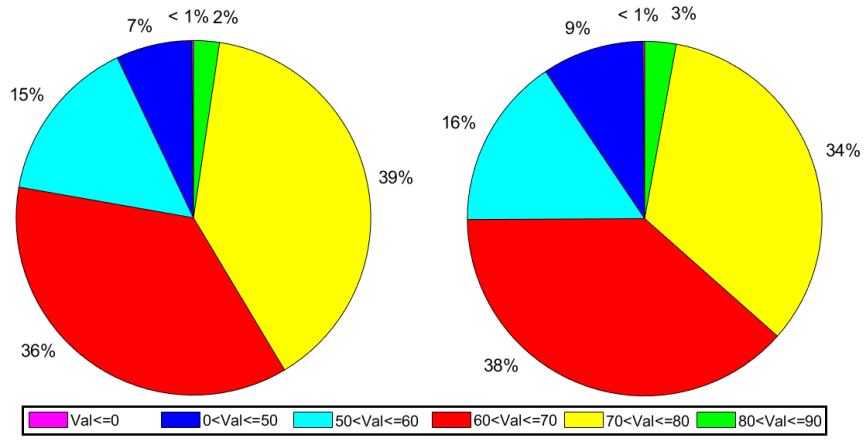


FIGURE 8.15: WHTC-1 validation performance distribution of the models Left:classic Right:proposed

In order to compare these models, the same number of input-output regressors and units are selected with the best iteration number and tabulated in Table 8.7.

Tables 8.6 and 8.7 shows that required significant input regressors in nonlinear block prevents the overfit problem and increase the generalization capability of the model. First model in Table 3 outperforms the corresponding model in Table 4 with 332 parameters whereas there exist 572 parameters in corresponding model.

The number of total selected regressors with OLS algorithm is presented in Figure 8.17. Here, the number of input regressor (nb) is chosen the same for each input

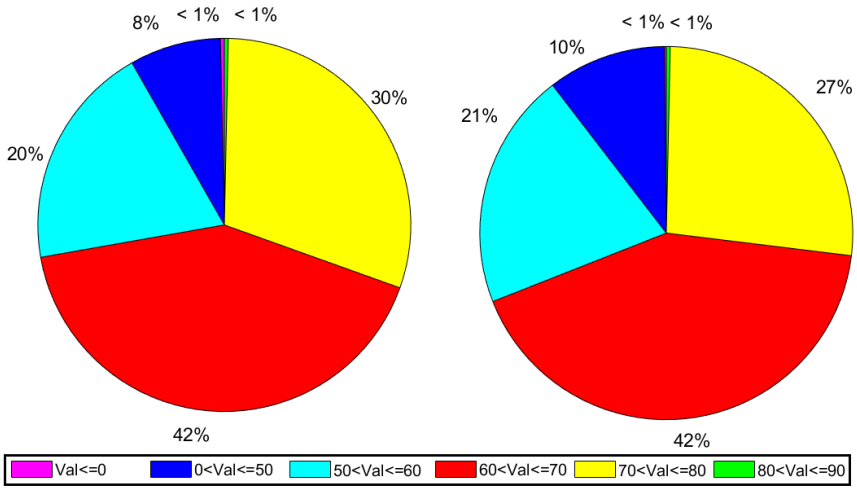


FIGURE 8.16: WHTC-2 validation performance distribution of the models Left:classic Right:proposed

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

TABLE 8.6: The best 6 models of the proposed structure

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 8.7: Corresponding models with classic structure

channel and regression matrix is constructed accordingly. Figure 8.17 shows that selected total number of input regressors converges to 22 regressors. Figure 8.18 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 employ them in modeling (Table 8.8).

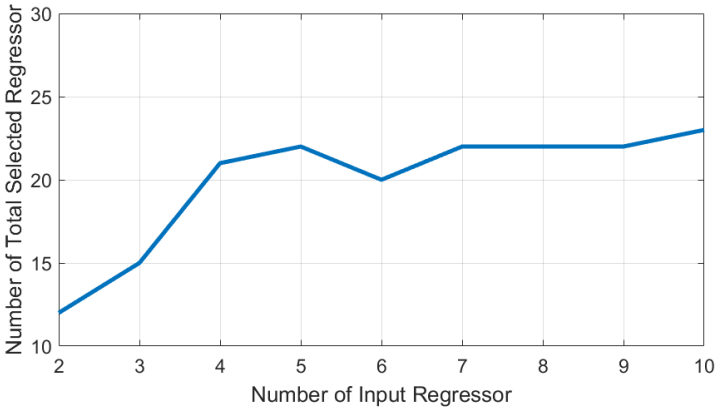


FIGURE 8.17: Total number of selected input regressors

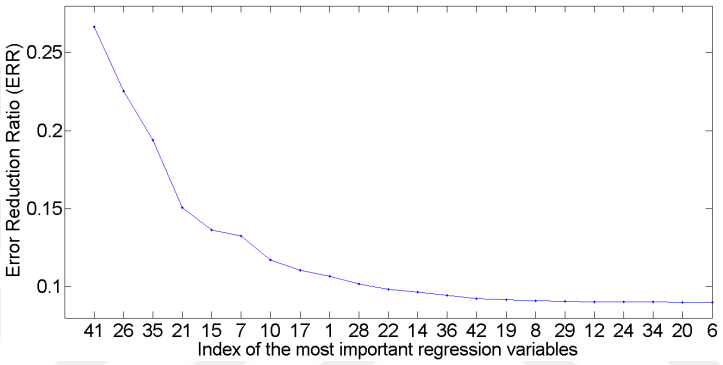
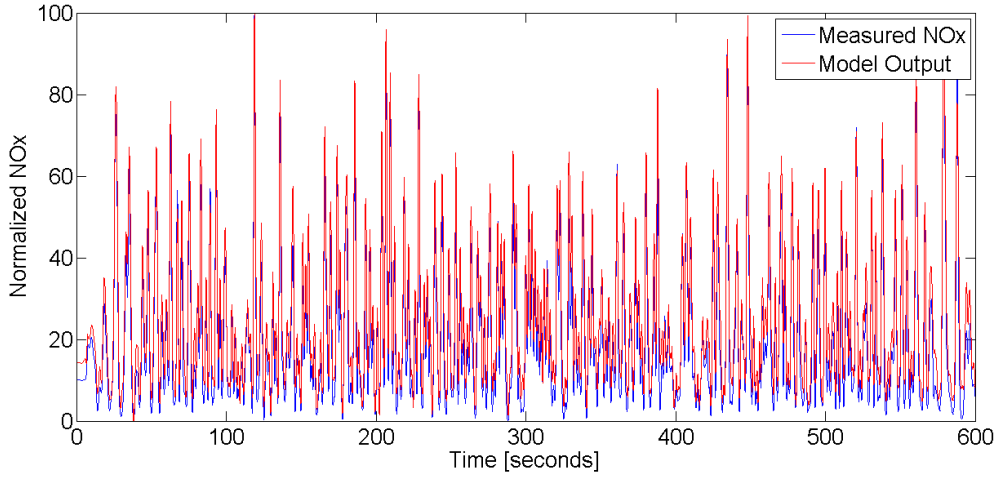


FIGURE 8.18: Determination of significant regressors by error reduction ratio

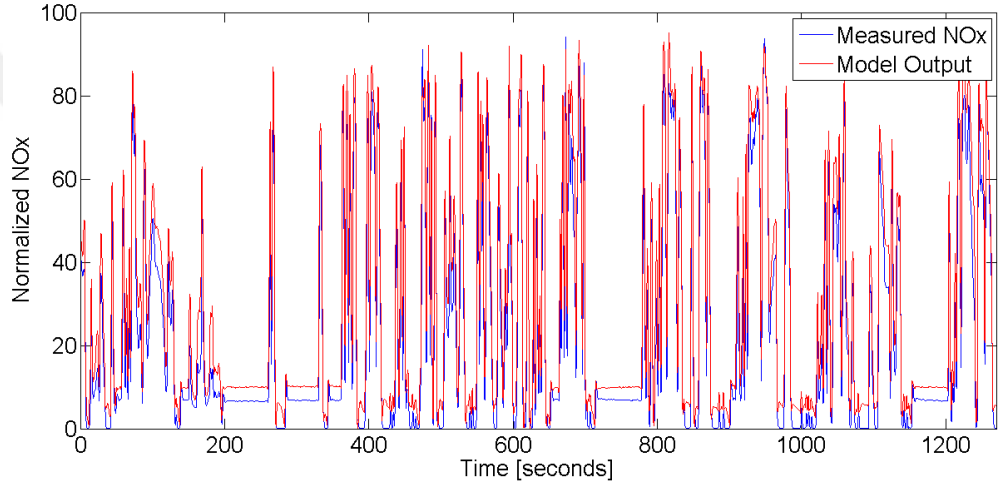
MAF	**MAP**	**SPD**	**QNT**	**SOI**	**RailP**
(t-35)	(t-35)	(t-35)	(t-35)	(t-35)	(t-35)
(t-40)	(t-37)	(t-37)	(t-37)	(t-40)	(t-40)
(t-41)	(t-39)	(t-39)	(t-39)	(t-41)	(t-41)
	(t-41)	(t-40)	(t-41)		
		(t-41)			

TABLE 8.8: Extracted regressors for each input channel ($nb = 7$)

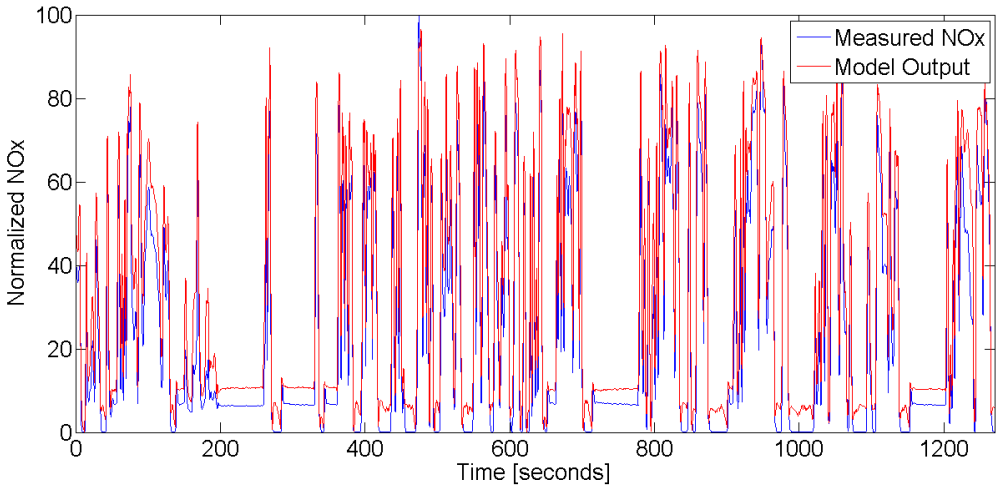
In Figure 8.19, estimation and validation performances are depicted in time plots. It is seen that the obtained model performs better in transient operations. In Figures 8.19(b) and 8.19(c), small errors are encountered in steady parts of the WHTC tests. This may be caused by the characteristics of training signals because all input channels are excited by chirp signal with linearly changing frequency profiles and it may lead to provide less information about low frequencies.



(a) Estimation performance



(b) WHTC-1 validation performance



(c) WHTC-2 validation performance

FIGURE 8.19: Estimation and validation performances of the best NOx emission model in Table 8.6

Another NARX structure presented in Figure 6.3 is employed in NOx emission modeling using the experimental data shown in Figure 4.6. Performances of the best NOx emission models for each experiment are assessed in Table 8.9.

Experiment	nb	unit	iteration	estimation	WHSC	WHTC	Figure
1	5	15	60	84,72	81,45	71,38	8.20
2	6	6	60	75,82	76,37	68,05	8.21
3	6	9	60	78,35	78,23	77,17	8.22
4	5	8	40	80,25	78,86	68,50	8.23

TABLE 8.9: The best NOx emission models for each experiment

Performances of some selected models by utilizing the map in Figure 6.8 are tabulated in Table 8.10. The best model according to P points (6.3) employs 6 input regressor number and 15 units. The model was trained by 84,92% fit accuracy, and the steady-state and the transient validation accuracies are 88,75% and 76,96%, respectively.

Regressor Number	Unit Num.	Maximum Bisections	Iterations	Training (%)	NEDC (%)	WLTC (%)
6	15	5	40	84,92	88,75	76,96
5	9	4	40	81,61	86,59	79,53
8	12	5	60	85,48	86,40	75,74
7	14	7	60	84,21	85,54	76,24
9	9	4	20	82,18	83,96	77,01
5	6	7	100	76,74	75,74	64,56
4	5	6	100	70,93	50,63	68,95

TABLE 8.10: Selected model performances using merged experiment (1+2)

In Figure 8.24, 8.26 and 8.28 training performances, in Figure 8.25, 8.27 and 8.29, steady-state and transient validation performances of the best three models in Table 8.10 are demonstrated in time plots, respectively. Results in Figure 8.25, 8.27 and 8.29 show that the model predictions are very accurate in the transient regions of the validation tests compared to steady-state regions. It should be noted that all of the input channels (except engine speed in Experiment 2) are excited by chirp signals during the training phase, which naturally enabled the model to learn the transient behavior properly. However, majority of the models still have the steady-state validation accuracy between 80%-90%. Indeed, it is still a challenging task to perform an experiment to model both steady-state and transient behavior of the diesel engine NOx emissions.

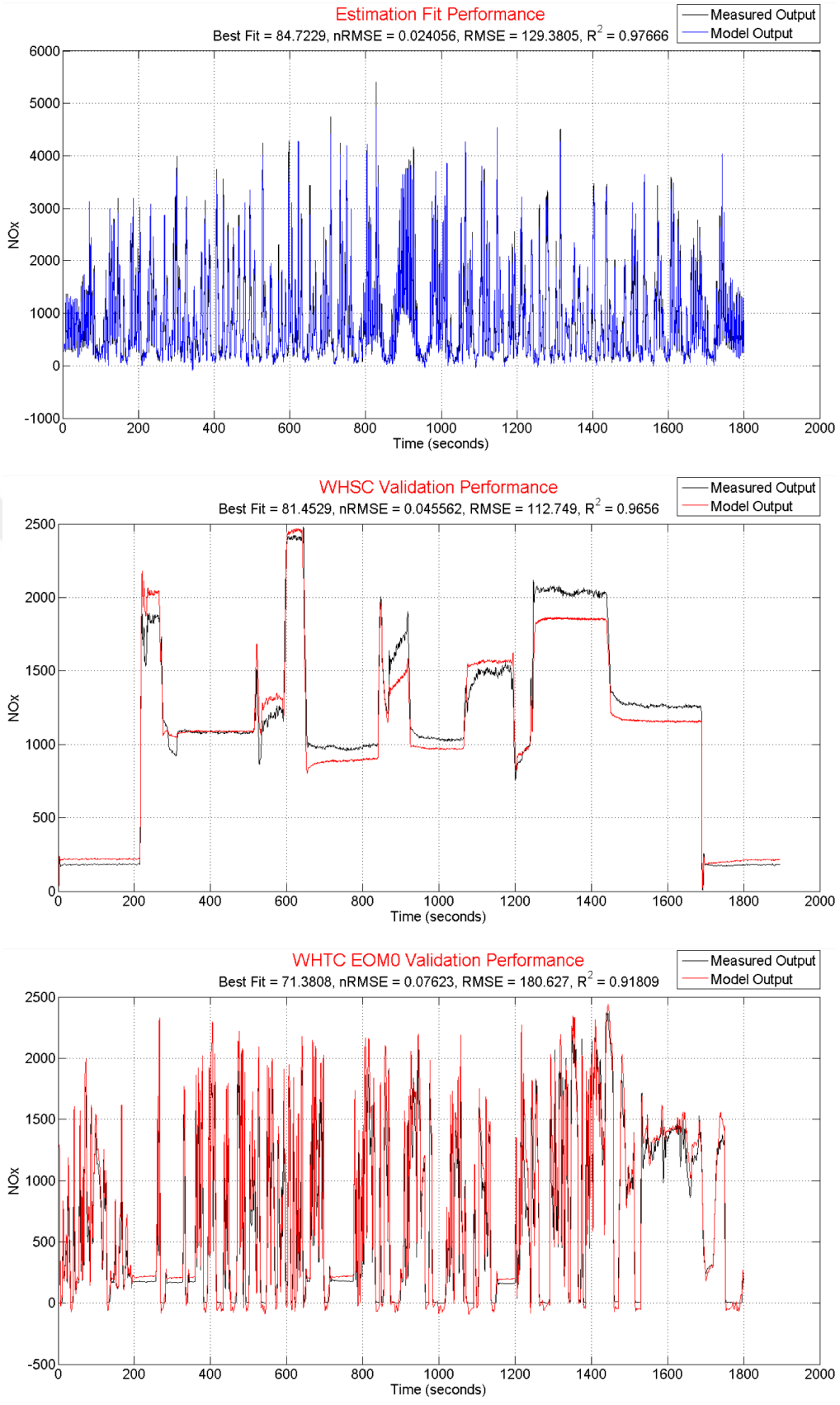


FIGURE 8.20: NOx emission model prediction performances (Experiment 1)

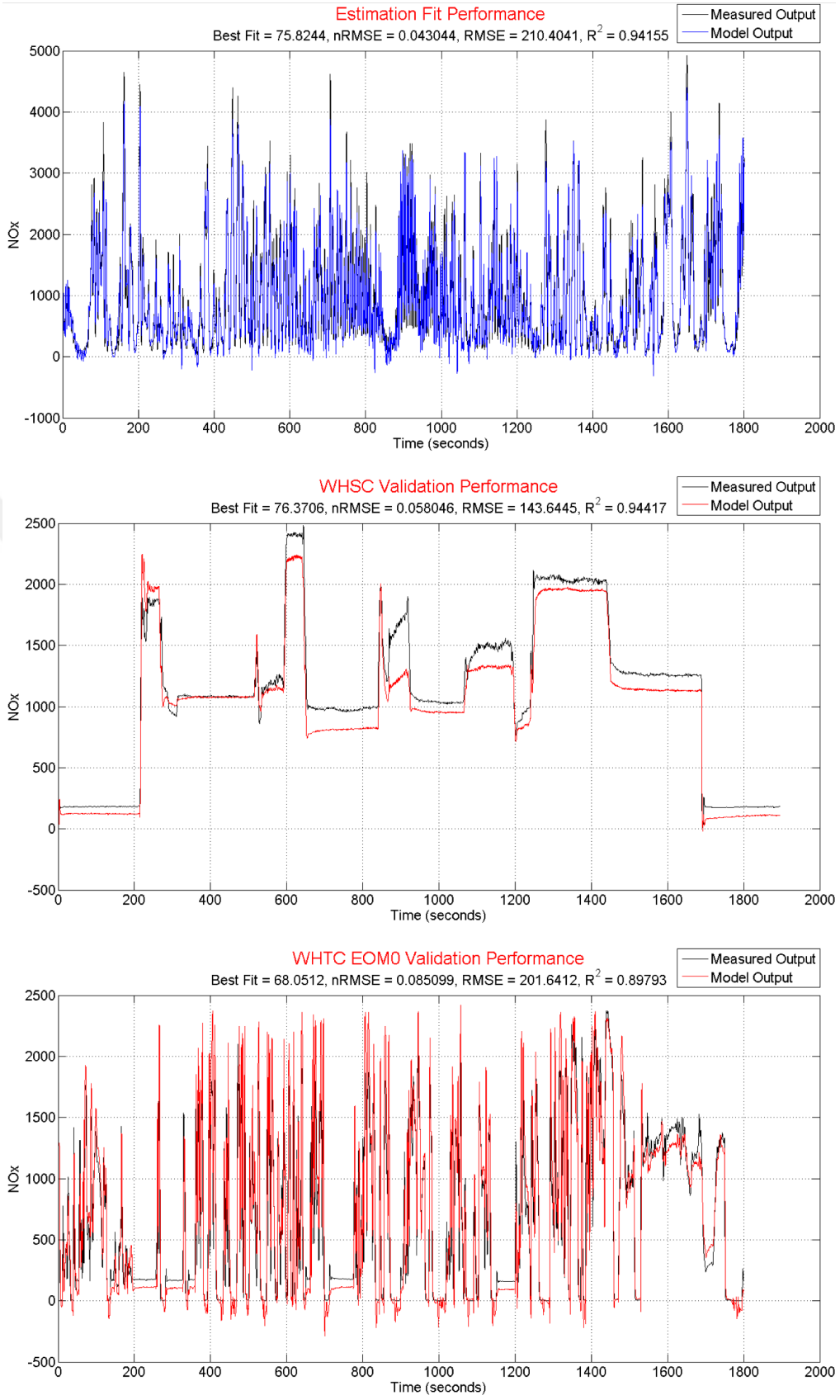


FIGURE 8.21: NOx emission model prediction performances (Experiment 2)

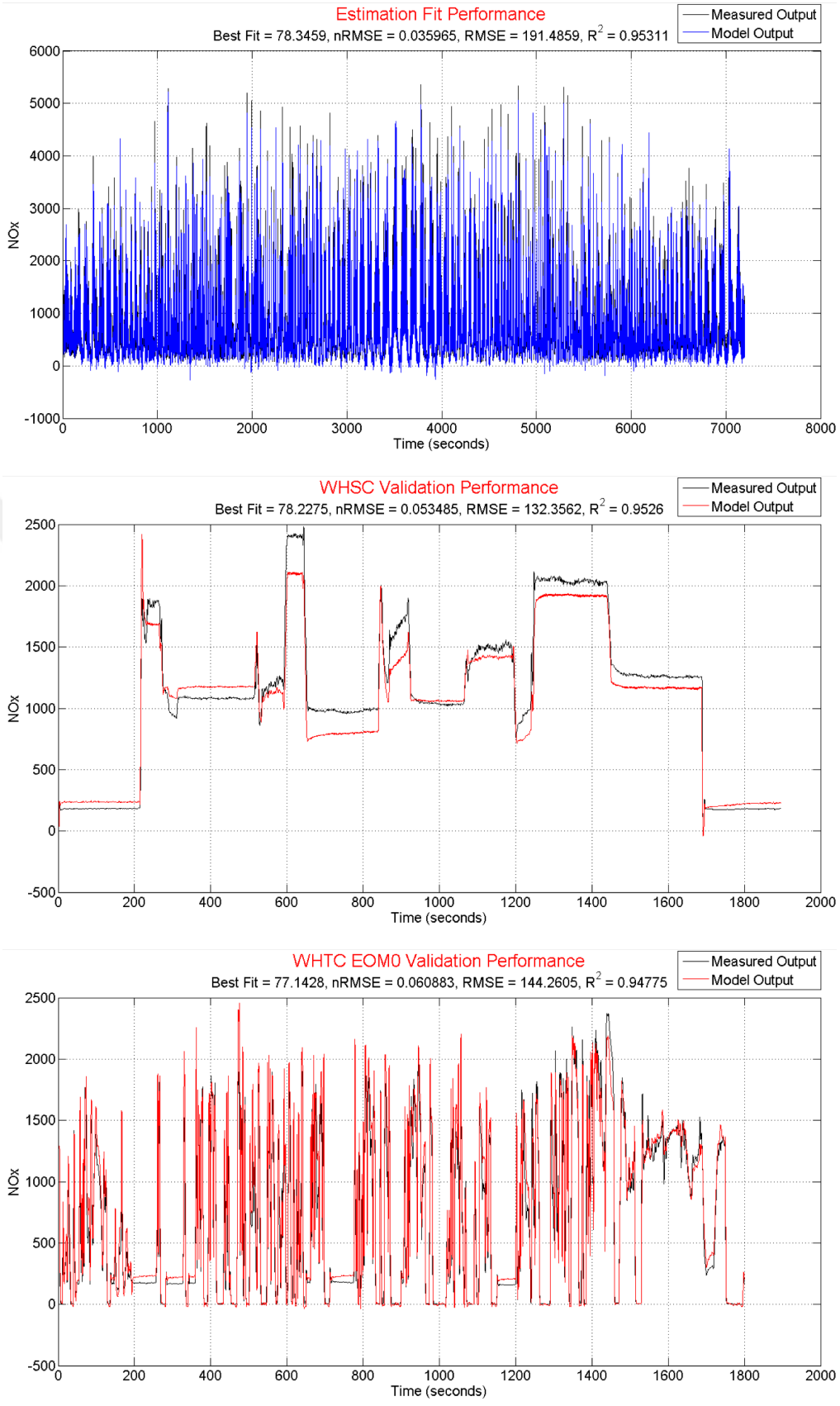


FIGURE 8.22: NOx emission model prediction performances (Experiment 3)

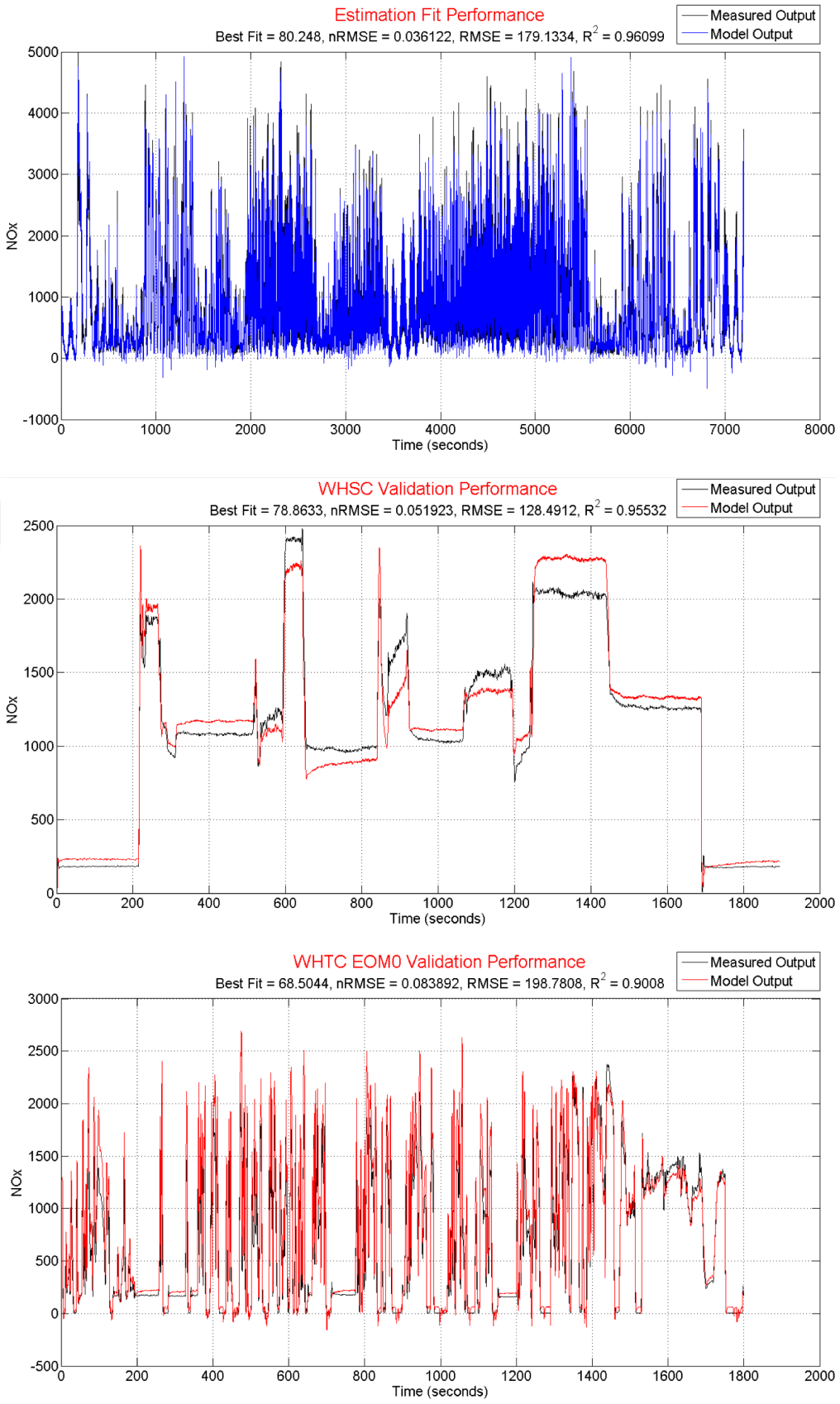


FIGURE 8.23: NOx emission model prediction performances (Experiment 4)

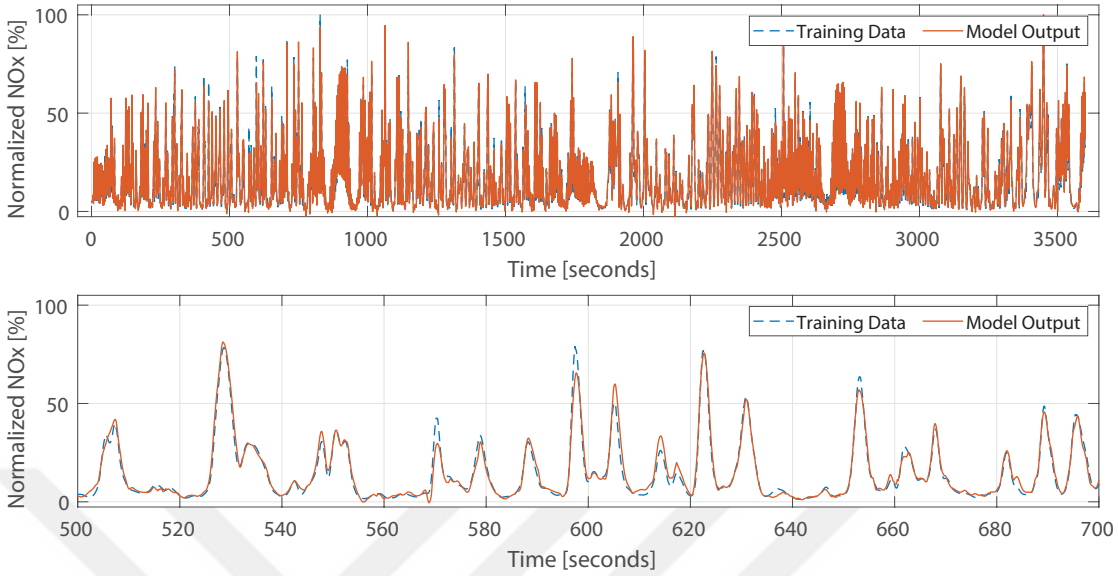


FIGURE 8.24: Training performance of the model with 6 input regressors and 15 units (below: zoomed version)

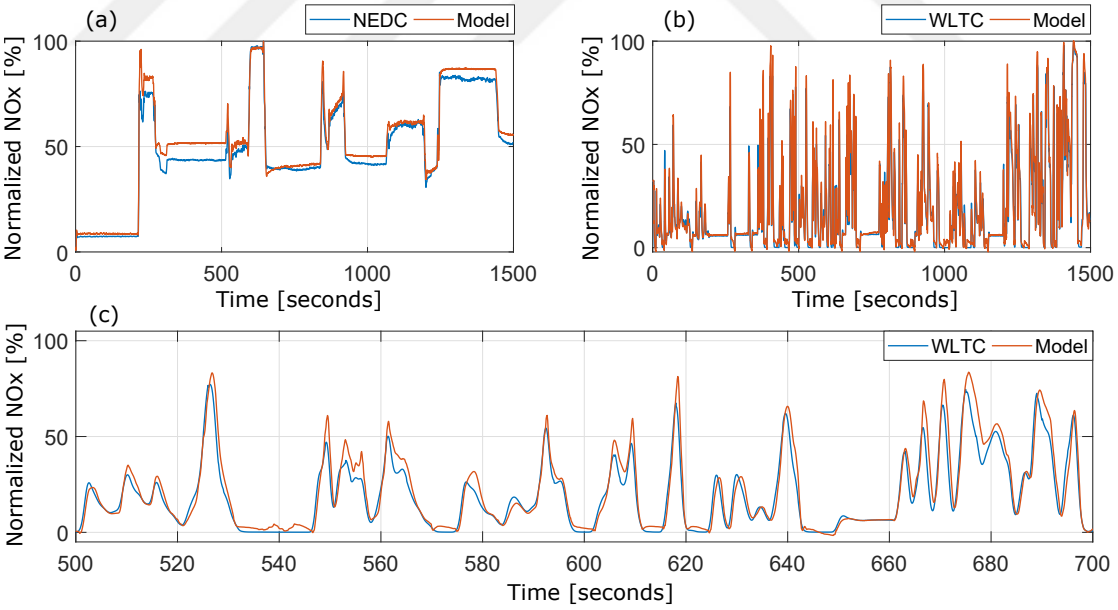


FIGURE 8.25: Validation performances of the model with 6 input regressors and 15 units (a) steady-state cycle (b) transient cycle, (c) zoomed-in version of transient cycle

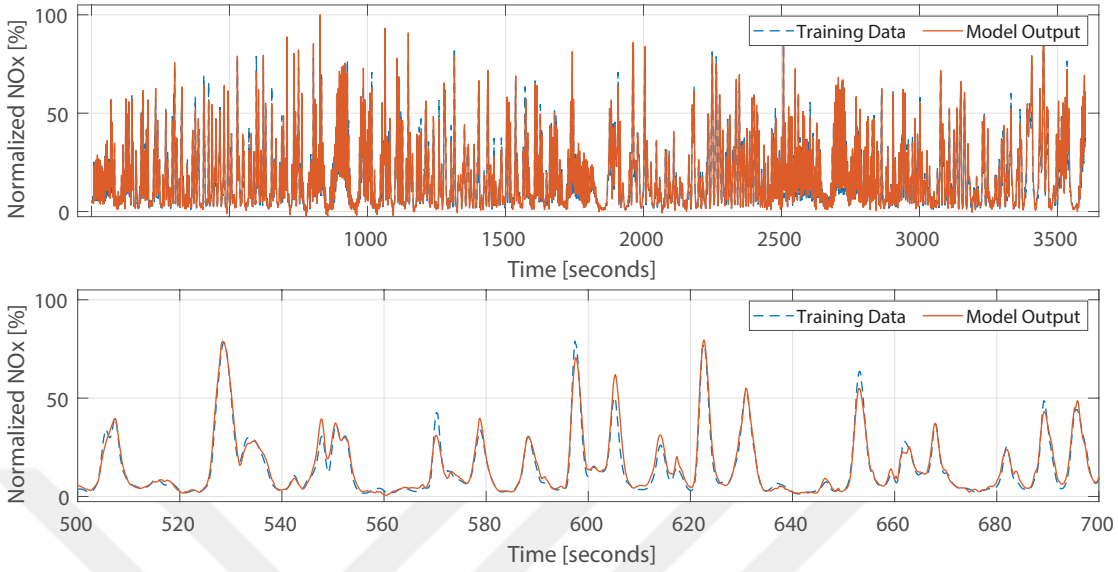


FIGURE 8.26: Training performance of the model with 5 input regressors and 9 units (below: zoomed version)

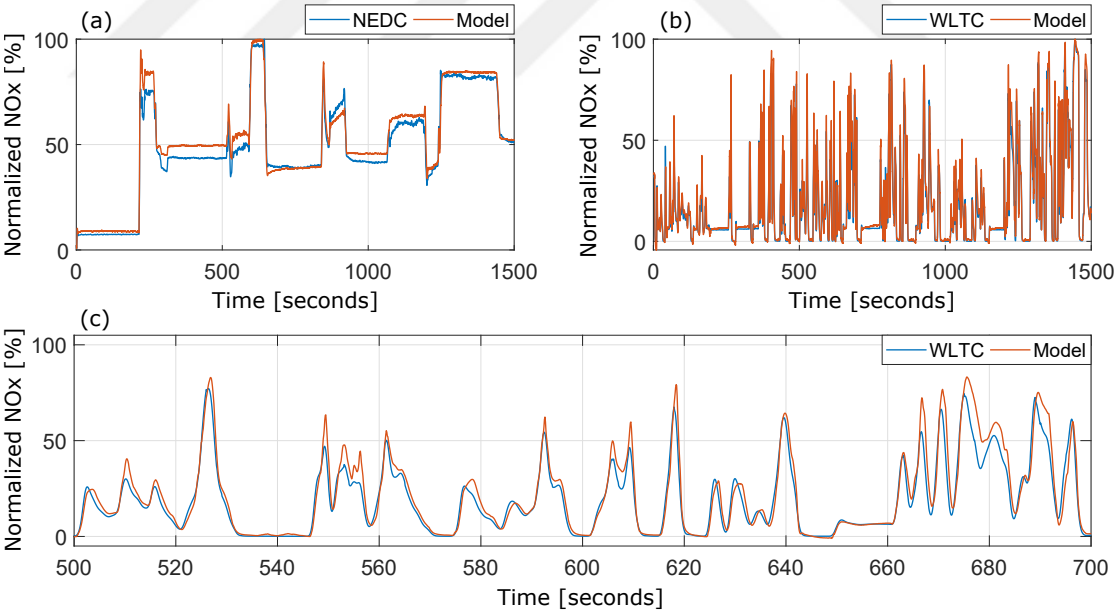


FIGURE 8.27: Validation performances of the model with 5 input regressors and 9 units (a) steady-state cycle (b) transient cycle, (c) zoomed-in version of transient cycle.

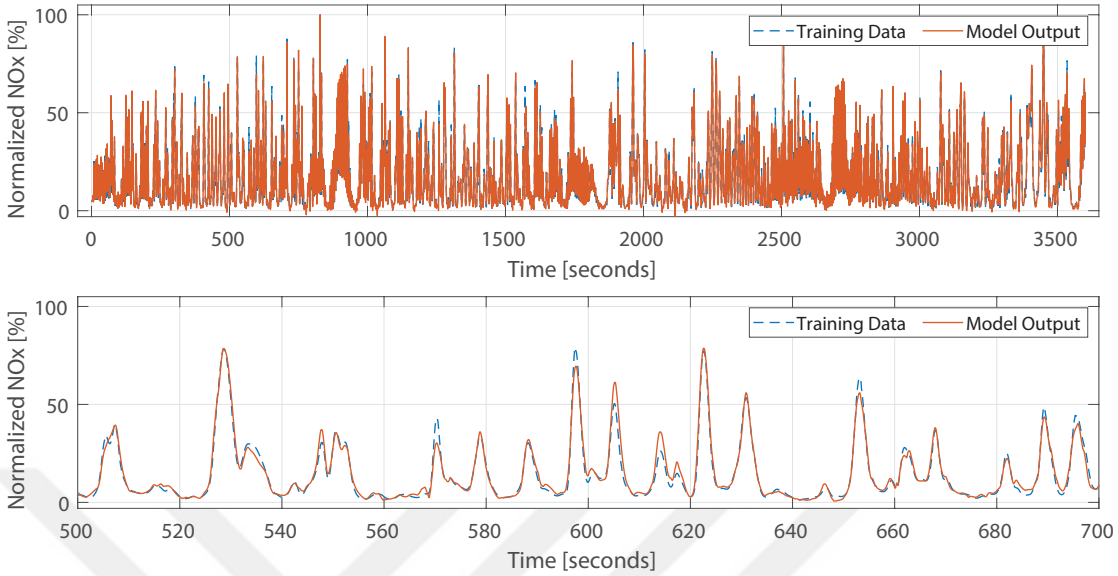


FIGURE 8.28: Training performance of the model with 8 input regressors and 12 units (below: zoomed version)

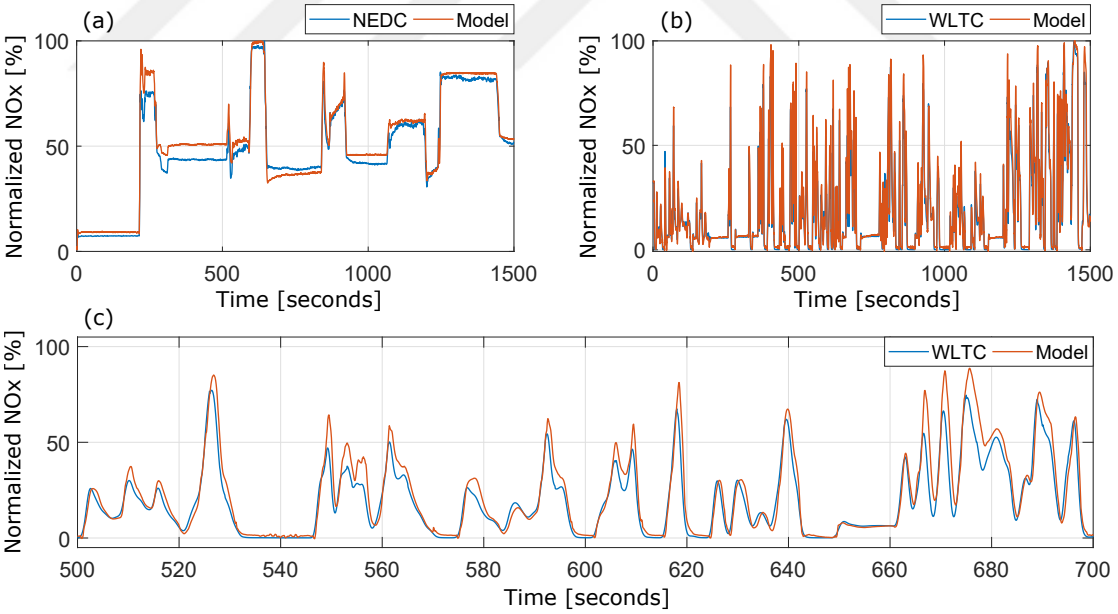


FIGURE 8.29: Validation performances of the model with 8 input regressors and 12 units (a) steady-state cycle (b) transient cycle, (c) zoomed-in version of transient cycle

8.2.3 Soot Modeling

An heavy-duty diesel engine's intended region of operation was determined by 30 different points in engine speed - injected fuel quantity plane, where the engine speed and the injected fuel quantity are selected in the ranges of 800-2200 rpm and 4-30 mg/stroke, respectively (Figure 4.14). 10 of 30 operation points were chosen for validation purposes in two different ways (Figure 8.30). In Case 1, the validation points were randomly selected. In Case 2, the validation points were selected from the boundary of the intended region of operations.

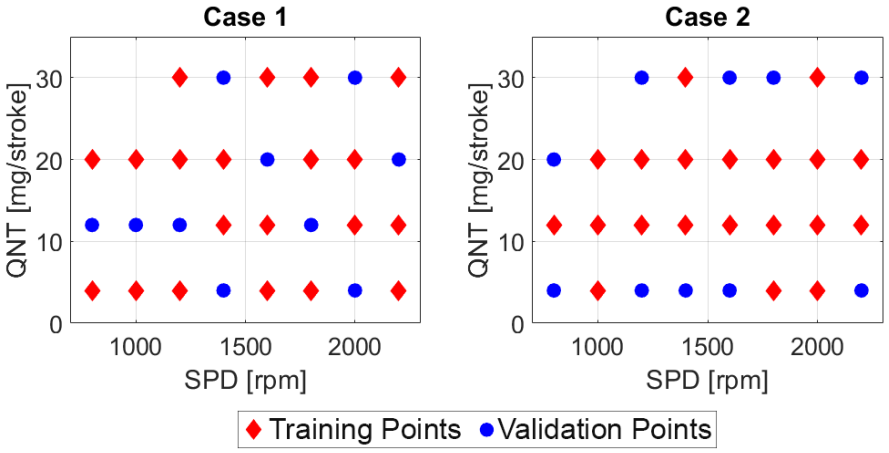


FIGURE 8.30: Design of experiments in SPD-QNT plane (Test 5)

In order to compare the prediction performances of the proposed GRU soot modeling approach, NARX network based soot modeling is also performed. The NARX structure has several hyperparameters including the number of output regressors, the number of input regressors (nb) and the number of delays between inputs and output (Section 6.1). Since the proposed GRU modeling structure does not include feedback from the estimated output and captures the relations between time delayed input signals automatically, the number of output regressors in NARX models is set to zero and the number of input regressors is selected in the range of 2 to 5. Moreover, the amount of delay between inputs and output is decided as 3.5 seconds by considering the physical conditions of the experimental setup. Levenberg-Marquardt and Adaptive Moment Estimation (ADAM) optimizers are employed in the parameter estimation process of NARX and GRU modeling structures, respectively.

The performances of obtained models are assessed by the fit metric (8.1) and normalized root mean squared error (NRMSE) given by

$$NRMSE = \frac{1}{\big(max(y) - min(y)\big)} \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}} \tag{8.3}$$

Training and validation performances of some representative models obtained by NARX and GRU structures are tabulated in Table 8.11 and Table 8.12.

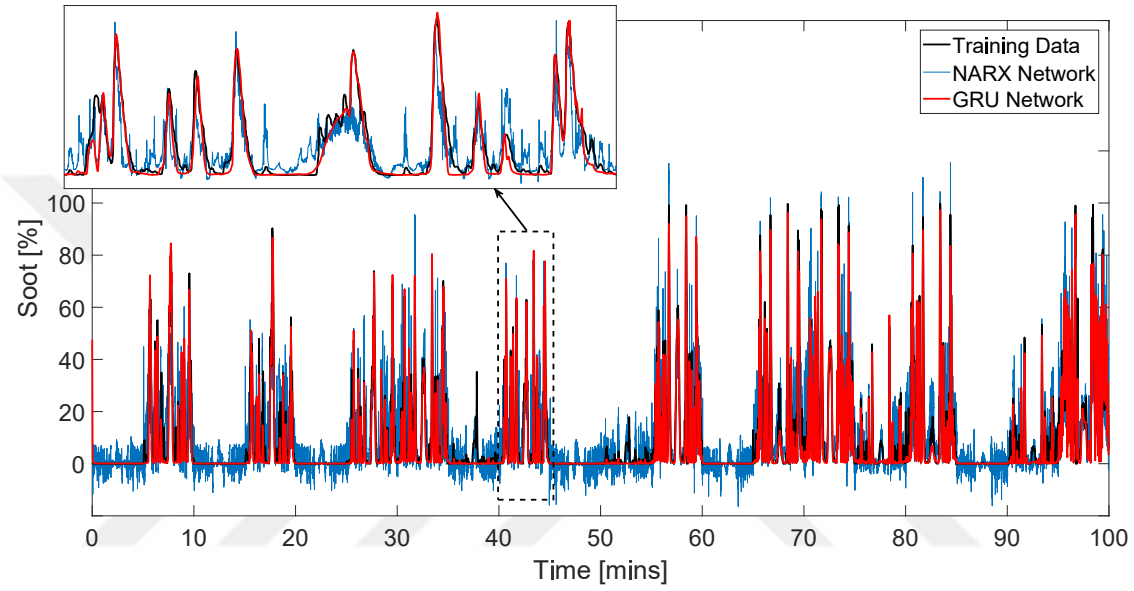
Case	**nb**	**Unit**	**Training**		**Validation**	
			Fit (%)	NRMSE	Fit (%)	NRMSE
1	5	10	46.89	0.0888	38.95	0.1017
1	3	15	49.70	0.0841	47.59	0.0873
1	2	20	48.87	0.0855	45.30	0.0911
1	5	25	55.45	0.0745	44.83	0.0919
1	5	30	54.48	0.0761	40.40	0.0992
1	2	50	52.79	0.0789	41.00	0.0982
2	5	10	46.60	0.0862	38.69	0.1088
2	3	15	45.46	0.0881	43.94	0.0995
2	5	20	51.99	0.0775	40.46	0.1057
2	5	25	54.75	0.0731	40.83	0.1050
2	5	30	55.07	0.0726	43.03	0.1011
2	3	50	59.48	0.0654	40.93	0.1049

TABLE 8.11: Performances of NARX models

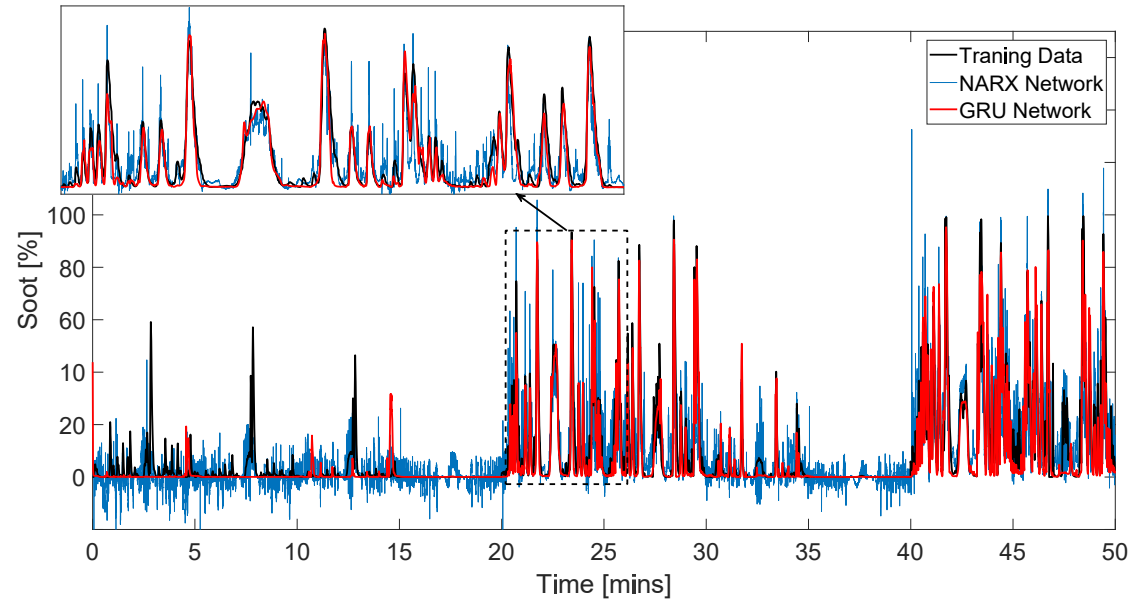
Case	**Unit**	**Training**		**Validation**	
		Fit (%)	NRMSE	Fit (%)	NRMSE
1	10	76.19	0.0391	53.28	0.0802
1	**15**	**77.17**	**0.0373**	**57.37**	**0.0686**
1	20	68.53	0.0517	54.86	0.0739
1	25	63.95	0.0586	50.37	0.0773
1	30	74.74	0.0419	53.72	0.0718
1	50	73.03	0.0434	53.59	0.0725
2	10	59.72	0.0611	47.75	0.0877
2	**15**	**76.99**	**0.0368**	**52.53**	**0.0807**
2	20	49.23	0.0885	42.97	0.1070
2	25	75.67	0.0379	49.06	0.0854
2	30	54.17	0.0707	47.08	0.0929
2	50	57.42	0.0726	45.58	0.1086

TABLE 8.12: Performances of GRU models

Although the NARX soot models are trained at most 59% and the models are validated with at most 47% fit accuracy, training performances of GRU models are able to reach up to 77%, corresponding validation performances exceeds 55%. Furthermore, training and validation NRMSE values are around 0.037 and 0.069 in the selected GRU model, however, NRMSE values are always above 0.65 for training and validation of NARX models. The best results of GRU structure for both cases are achieved when the number of units is selected as 15.

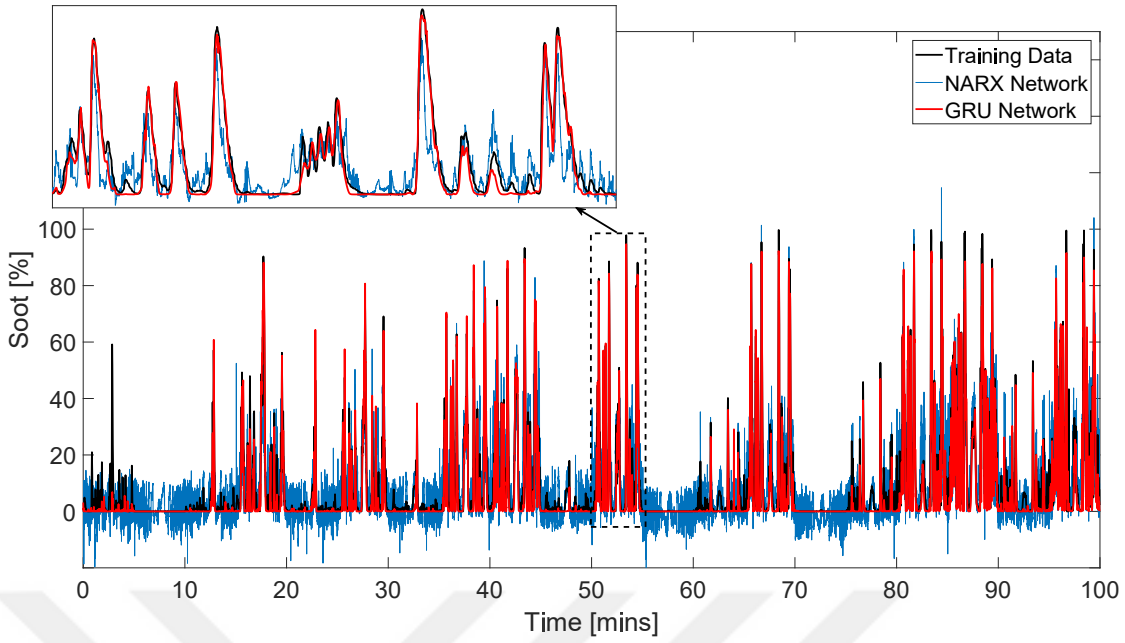


(a) Training performances

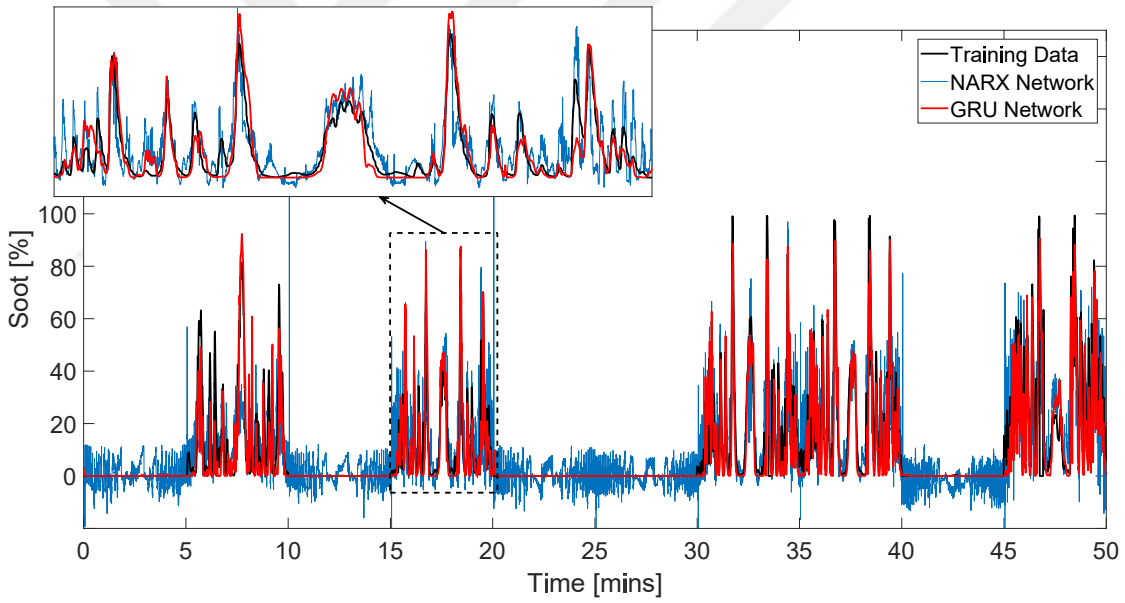


(b) Validation performances

FIGURE 8.31: Prediction performances of the models with 15 units (Case 1)



(a) Training performances



(b) Validation performances

FIGURE 8.32: Prediction performances of the models with 15 units (Case 2)

Time plots of NARX and GRU models with 15 units are depicted in Figures 8.31 and 8.32. Since the number of input regressors and the pure time delay between inputs and output are pre-defined and can not be adjusted during parameter estimation in a NARX structure, the predictions of NARX models are highly oscillatory and their accuracies are not good at steady-state operating points compared to transient sections. Furthermore, it is not easy to determine the model parameters

that result in a smooth response by fixing these values, particularly in soot emissions modeling. However, the predictions of GRU models are smoother thanks to the gated structure which enables the model to adjust the effects of former input signals adaptively with update and reset gates. Moreover, the prediction accuracies of GRU models are better on both steady-state and transient sections of test cycles. In Figure 8.31(b), it is seen that the prediction performance of GRU model in the first 10 minutes is inadequate because the engine works in its idle state with speed less than or equal to 1000 rpm, and soot emission dynamics differ in this speed range. In order to increase the prediction performances, separate models for different ranges of engine speed can be trained.

LSTM structure is much more complex than GRU with additional gates (Section 7.2). This kind of structures are suitable to model "big data". Therefore, performed all tests (1 - 4) are merged and used in training of LSTM soot model (Figure 4.9). Number of the unit used in LSTM structure was also set to 15 as in the case of GRU models. During the optimization process fixed learning rate (0.01) was utilized and models were converged after around 5000 – 10000 iterations. After several trials with changing different input sets, unit number and learning strategies, parameters were tuned, and the best model obtained is presented in Table 8.13. Signals of SPD, QNT, RailP, miSOI, EQVR and IntO2R parameters were included in the input set for this model.

Test	Unit	Training	NEDC	WLTC
1+2+3+4	15	71,08	50,23	53,99

TABLE 8.13: Performances of LSTM soot model with merged data

Training, NEDC and WLTC validation accuracies of the model are 71%, 50% and 54%, respectively. Model responses in time plots are presented in Figure 8.33 – 8.41. The experimental results are quite satisfactory and unique in the as a global soot emission modeling which achieves above 50% validation accuracies in both steady-state and transient cycles. Figures 8.33 – 8.41 illustrates that obtained soot responses are quite smooth and the cumulative errors are always less than or equal to 10%, which makes it possible to utilize such structure in optimization-oriented tasks.

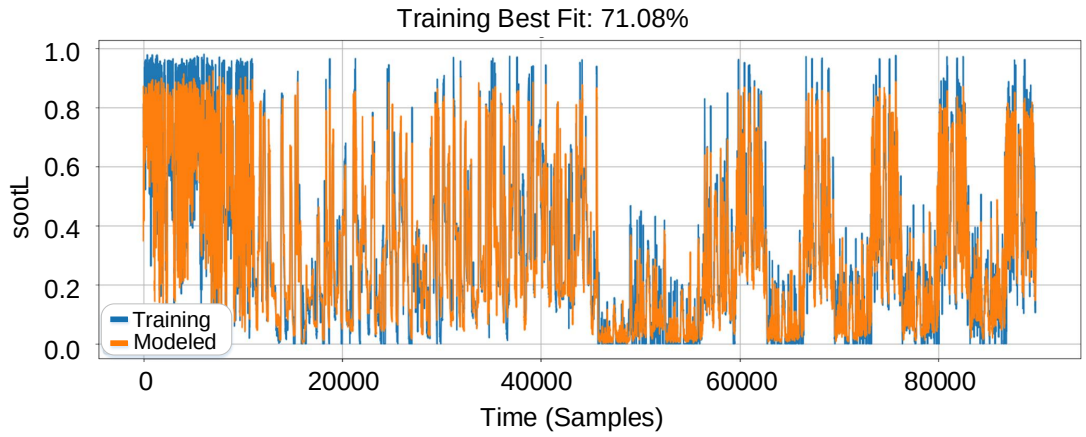


FIGURE 8.33: Training performance of the LSTM soot model

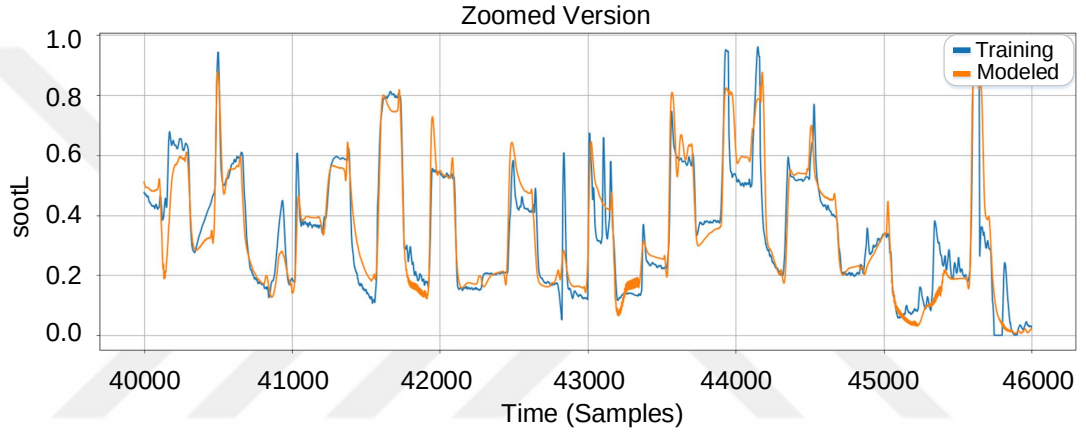


FIGURE 8.34: Zoomed version of Figure 8.33

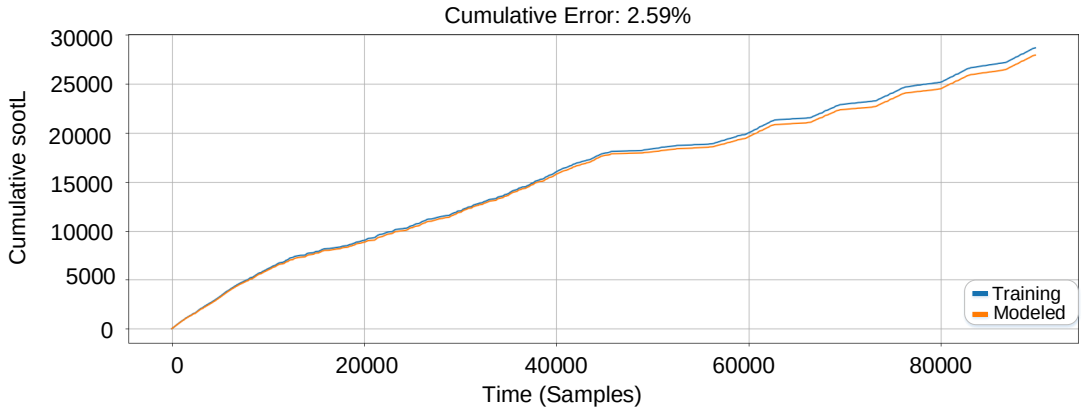


FIGURE 8.35: Cumulative training response of the model

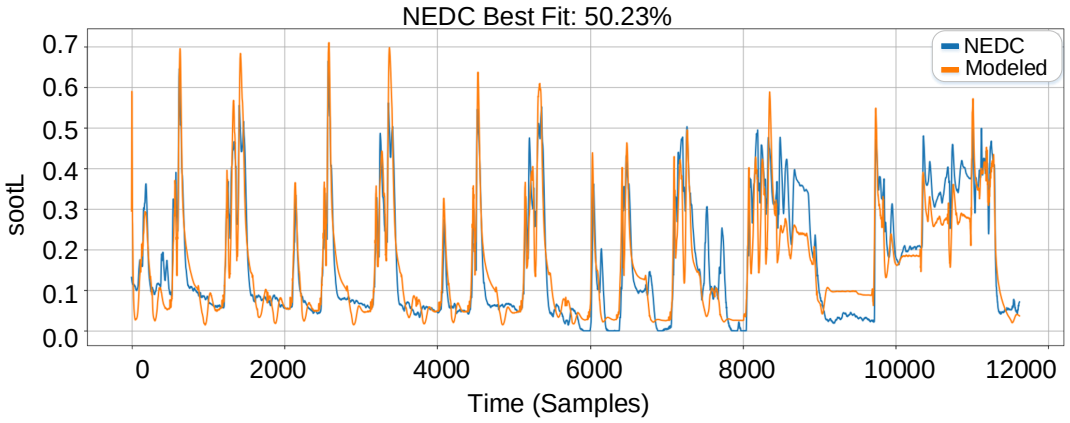


FIGURE 8.36: NEDC validation performance of the LSTM soot model

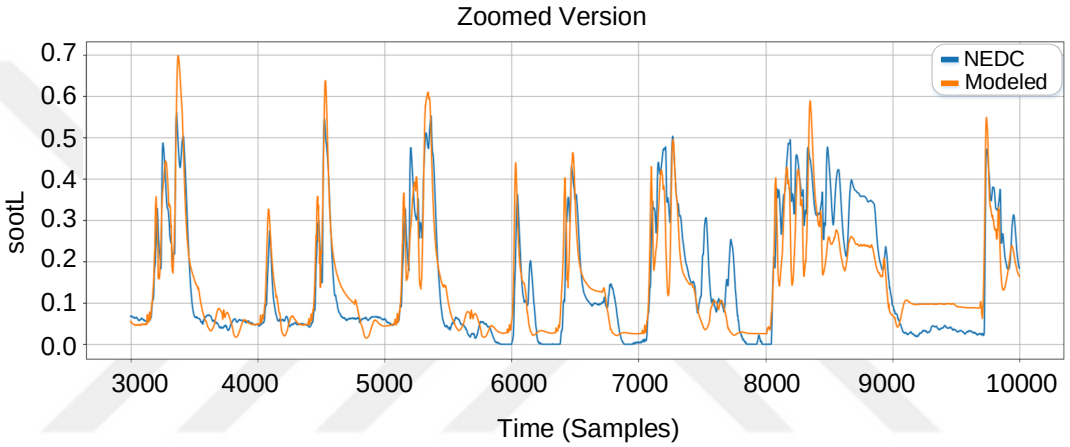


FIGURE 8.37: Zoomed version of Figure 8.36

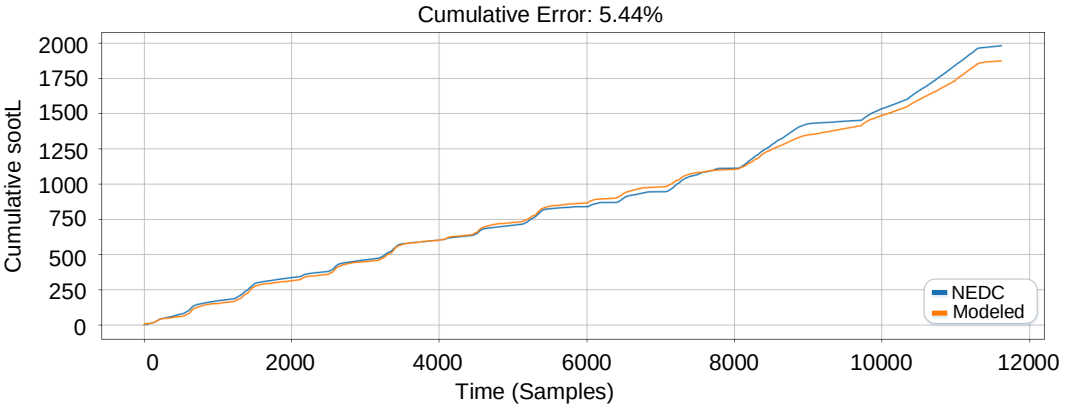


FIGURE 8.38: Cumulative NEDC validation response of the model

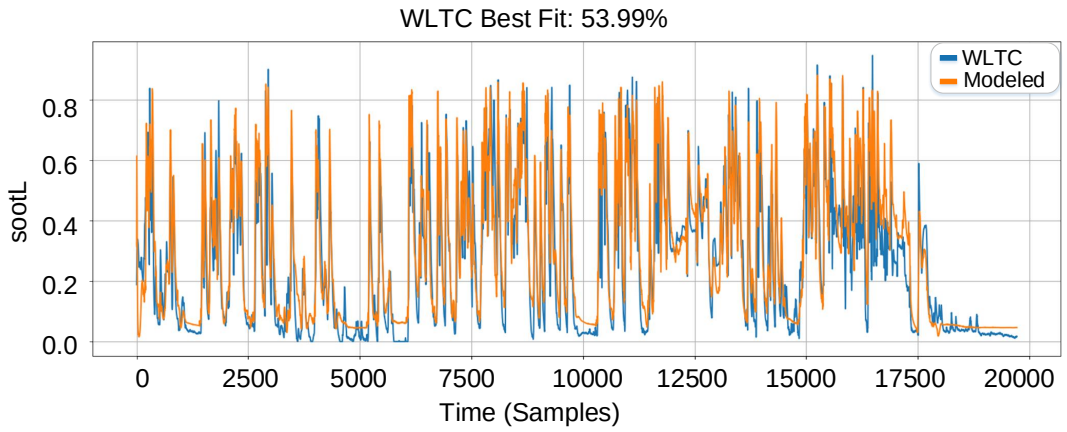


FIGURE 8.39: WLTC validation performance of the LSTM soot model

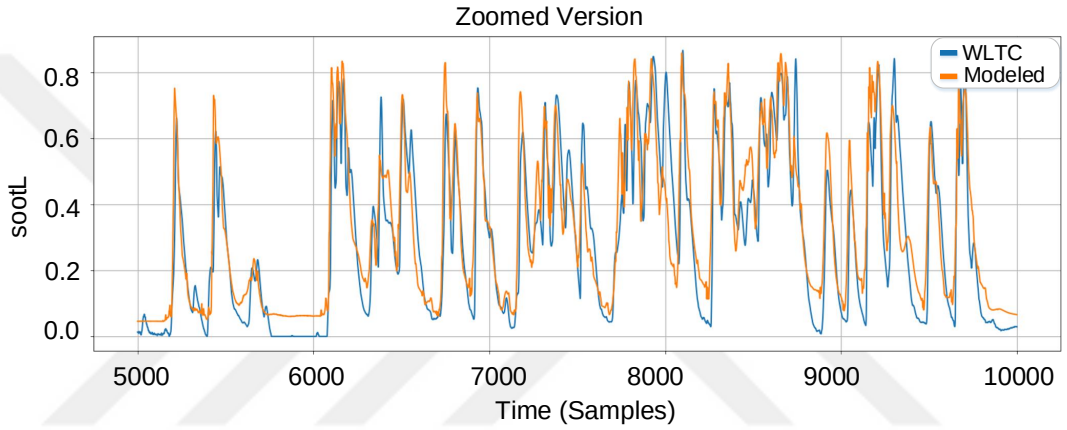


FIGURE 8.40: Zoomed version of Figure 8.39

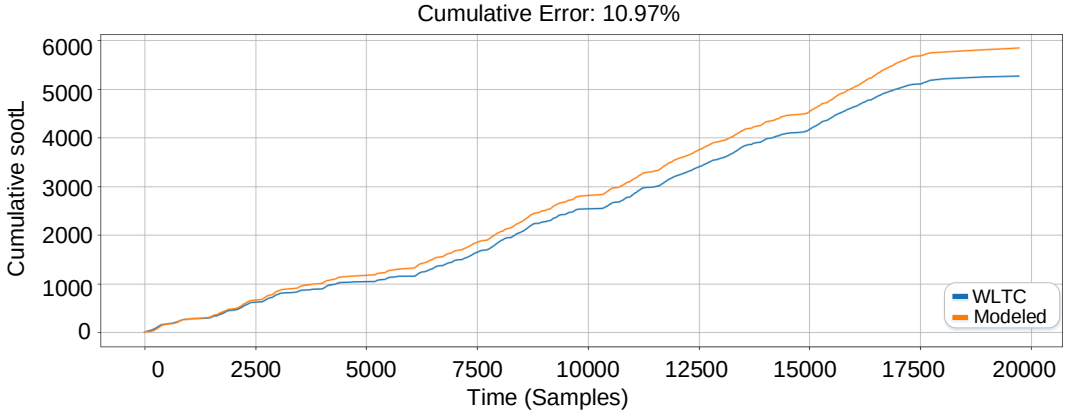


FIGURE 8.41: Cumulative WLTC validation response of the model

Test 5 (Figure 4.17) was also utilized in the training of LSTM soot models. Since the working conditions of that experiment were quite different than Test 1 - 4, WLTC and NEDC validation cycles were not utilized in the validation of the models obtained by that experiment with 30 points. Therefore, 10 of 30 operation points were chosen for validation purposes in two different ways such as selecting randomly or selecting from the boundary of the intended region of operations (Figure 8.30). It has been noticed that the validation set in Case 2 is much more challenging that Case 1 due to data from the boundaries.

Different combinations of input channels such as SPD, QNT, MAP, MAF, miSOI, RailP, EQVR, IntO2R and PMax were investigated in the modeling of logarithmically normalized soot emissions. Finally, the input set including SPD, QNT, RailP, miSOI EQVR and IntO2R was decided. After the trials of several modeling parameters such as the number of unit and iterations, the best of the obtained models for both cases are tabulated in Table 8.14.

Test	Unit	Training	Validation
5 - Case 1	15	78,93	64,54
5 - Case 2	15	72,01	60,83

Table 8.14: Performances of LSTM soot model with Test 5

Training and validation accuracies of the models are above 70% and 60%, respectively. Cumulative error in the performance graphs are consistently below 10%. Figures 8.42 - 8.53 presents the training and the validation performances of these models.

It should be noted that this work explores the possibility of employing a GRU or LSTM network in soot emissions modeling for an extensive range of engine speeds. We do not claim any optimality regarding structure of the network used in this work. Therefore obtained results are not optimal in any sense, however they are satisfactory when compared to the existing results in the literature.

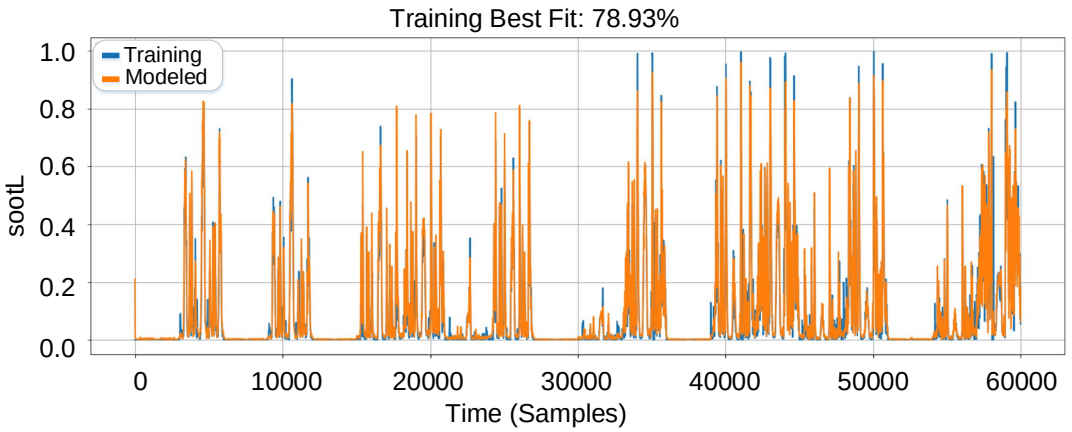


FIGURE 8.42: Training performance of LSTM soot model (T5 - Case 1)

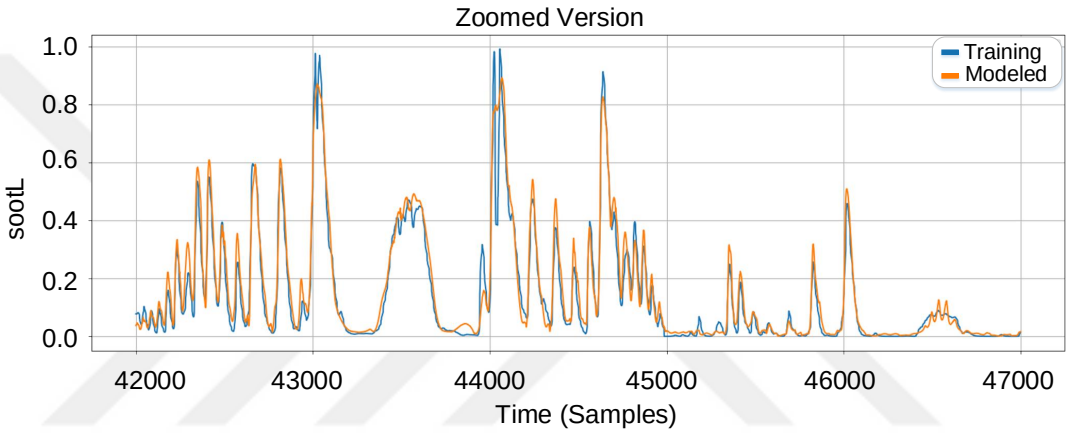


FIGURE 8.43: Zoomed version of Figure 8.42

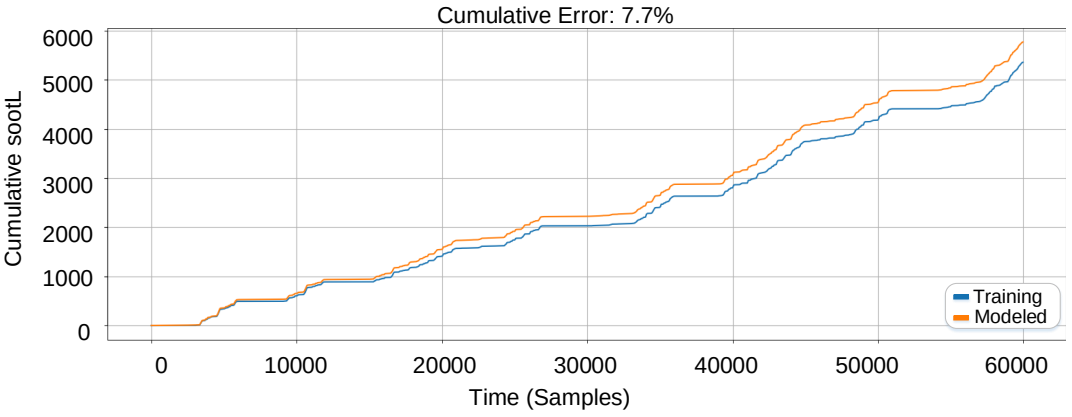


FIGURE 8.44: Cumulative training response of the model (T5 - Case 1)

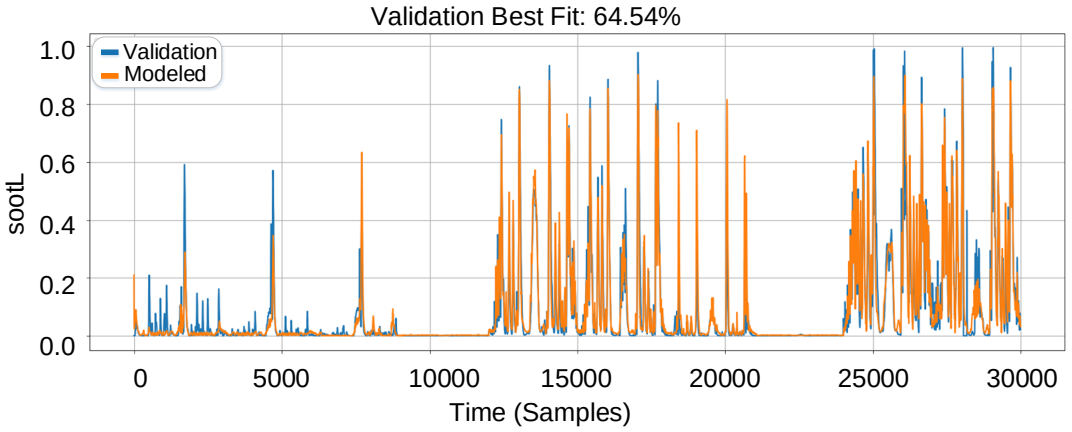


FIGURE 8.45: Validation performance of soot model (T5 - Case 1)

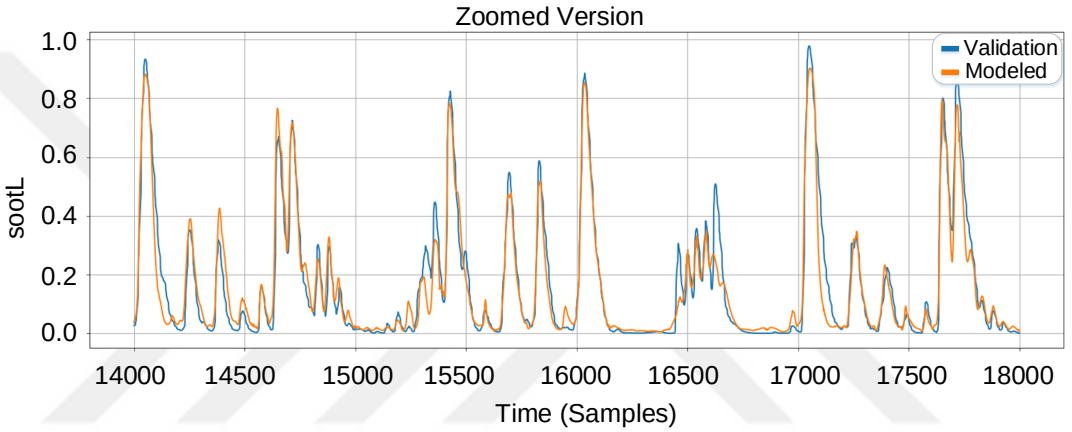


FIGURE 8.46: Zoomed version of Figure 8.45

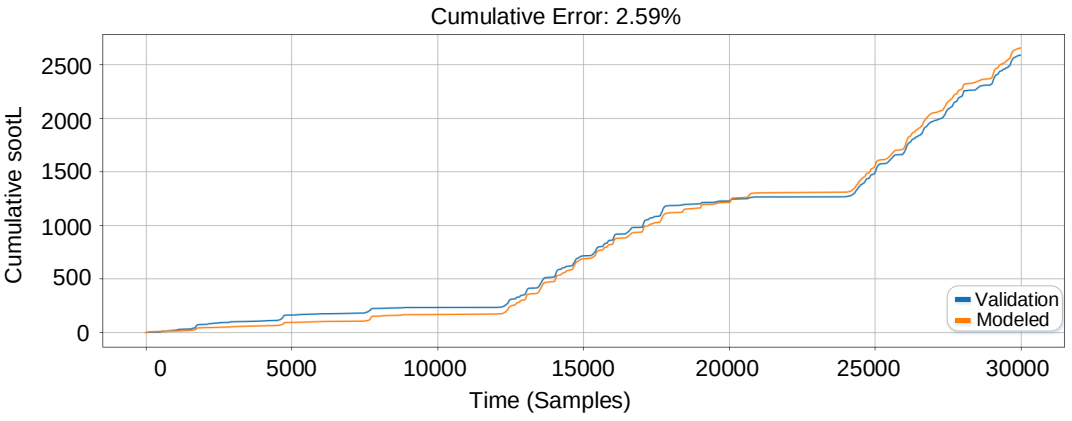


FIGURE 8.47: Cumulative validation response of the model (T5 - Case 1)

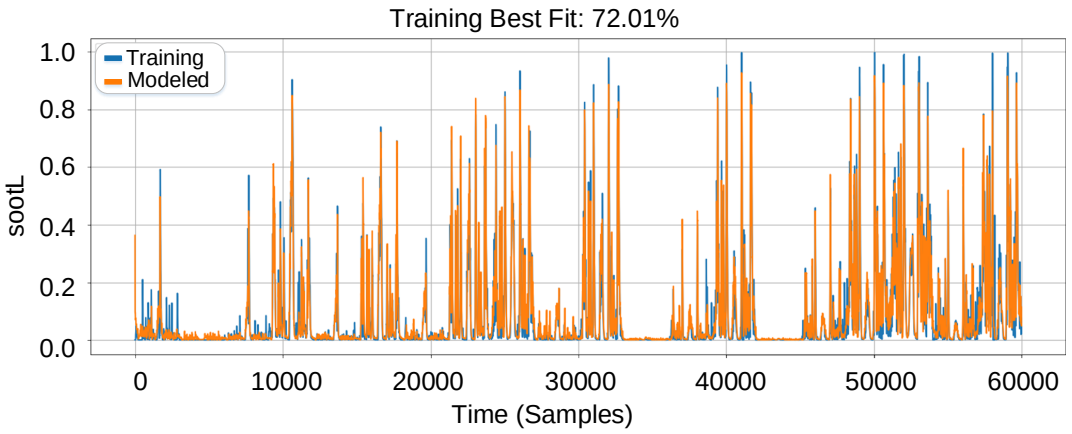


FIGURE 8.48: Training performance of LSTM soot model (T5 - Case 2)

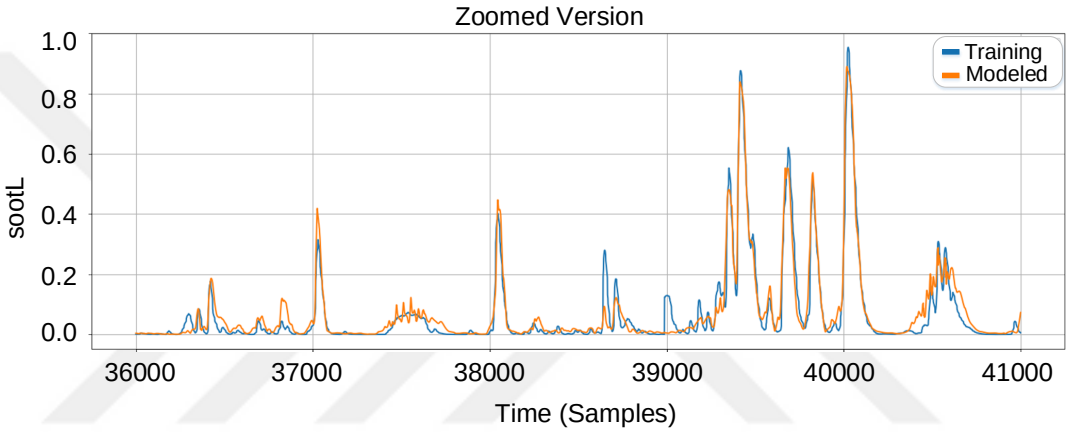


FIGURE 8.49: Zoomed version of Figure 8.48

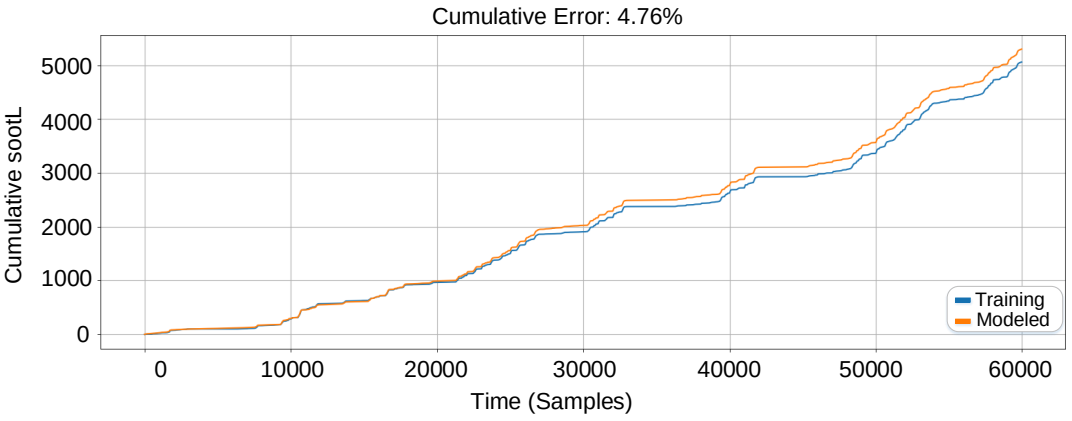


FIGURE 8.50: Cumulative training response of the model (T5 - Case 2)

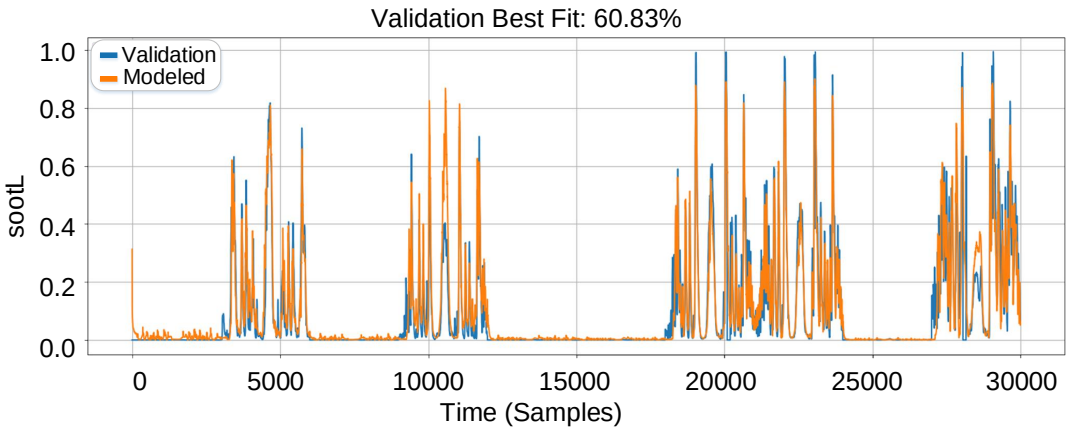


FIGURE 8.51: Validation performance of soot model (T5 - Case 2)

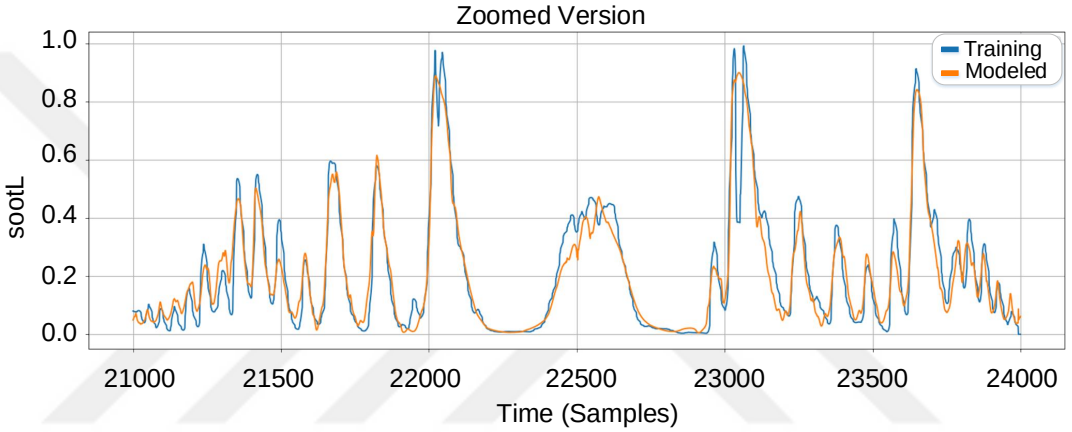


FIGURE 8.52: Zoomed version of Figure 8.51

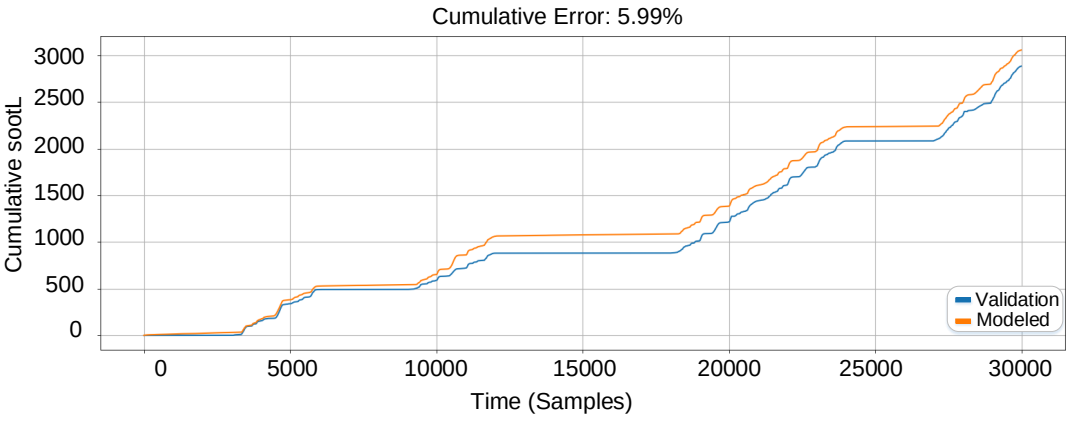


FIGURE 8.53: Cumulative validation response of the model (T5 - Case 2)

8.3 Implementation Details

NFIR and the variants of NARX models developed in this thesis have been implemented by using MATLAB's System Identification Toolbox. Optimization runs for 20 iteration steps, and the obtained model is evaluated with training and validation tests. The last estimated model is utilized as an initial model for the next estimation up to 100-120 iterations. Total execution times of the presented indicated torque and NOx emission models are tabulated in Table 8.15 and Table 8.16, respectively.

Experiment	*nb*	Unit	Iteration	Execution Time	Figure
1	8	9	80	∼5.6 min.	8.8
2	3	8	40	<1 min.	8.9
3	5	9	80	∼3.4 min.	8.10
4	3	11	80	∼4 min.	8.11
1+2	5	8	80	∼2.6 min.	8.12
2+3	3	10	60	∼3.1 min.	8.13

Table 8.15: Execution times of the indicated torque models

Experiment	*nb*	Unit	Iteration	Execution Time	Figures
1	5	15	60	∼4.3 min.	8.20
2	6	6	60	∼2.7 min.	8.21
3	6	9	60	∼10.4 min.	8.22
4	5	8	40	∼4.8 min.	8.23
1+2	6	15	40	∼7.4 min.	8.24,8.25
1+2	5	9	40	∼2.7 min.	8.26,8.27
1+2	8	12	60	∼13.4 min.	8.28,8.29

Table 8.16: Execution times of the NOx emission models

GRU and LSTM based soot models have been implemented in Python environment by using several libraries such as Tensorflow, Numpy, Scipy, Matplotlib and Pandas. When compared to NFIR and NARX models, there are more parameters to be estimated, and the back-propagation of the models are much more complex

in GRU and LSTM models. Furthermore, the number of the samples in the training set is higher than the NARX models due to the need for larger datasets. All these factors increase the execution time dramatically to several hours.

Training data was separated into several mini-batches. Each of the mini-batch was utilized once in regular turn for every iteration. When every mini-batch was employed in training for one epoch, the sequence of the mini-batches starts again from the beginning. Models were evaluated with training and validation tests, and the parameters were saved for every 2-5 epochs. Total execution times of the presented GRU and LSTM based soot emission models are tabulated in Table 8.17 and Table 8.18, respectively.

Model	Case	Epoch	Execution Time	Figures
GRU	1	1460	∼7.1 hours	8.31
GRU	2	1490	∼7.2 hours	8.32
LSTM	1	1575	∼16.2 hours	8.42 - 8.47
LSTM	2	1490	∼11.4 hours	8.48 - 8.53

Table 8.17: Execution times of the GRU and LSTM based soot models with Test 5

Model	Test	Epoch	Execution Time	Figures
LSTM	1+2+3+4	176	∼11.3 hours	8.33 - 8.41

Table 8.18: Execution times of the LSTM based soot model with combined tests

Developed codes in this thesis have been executed in a workstation with Intel(R) Xeon(R) CPU E5-1650 processor which operates at 3.50 GHz speed, and has a 32 GB RAM.

Chapter 9

Concluding Remarks and Future Work

In this thesis, optimization-oriented high fidelity nonlinear dynamic models are developed for heavy-duty diesel engine torque and combustion emissions. In order to reduce the testing time and achieve high validation performances under both steady-state and transient cycles, a new design of experiments is proposed where input channels are excited by chirp signals with varying frequency profiles in terms of the number and directions of the sweeps.

To attain accurate indicated torque predictions, estimations of loss torques such as friction, pumping and inertia, and the torque measured from the engine dynamometer are utilized in a nonlinear finite impulse response (NFIR) model. Steady-state and transient validation performances of the obtained indicated torque models are entirely satisfactory to use them in a dynamic calibration process with around 94% and 87% accuracies, respectively.

Nonlinear modeling of diesel engine NOx emissions based on sigmoid NARX models is proposed with different variations in terms of the input regressor selection and the utilization of output feedback. Furthermore, sensitivity analysis of the models to parameter changes is conducted by generating models for different values of parameters and a parameter selection method using an easy-to-interpret

map is proposed. The map can be a convenient means to select the required parameters for diesel engine NOx emissions modeling with limited testing time in powertrain development. Experimental results show that the steady-state and the transient validation accuracies for the majority of the obtained models are higher than 80% and 70%, respectively. The accuracies of the selected best models are sufficiently high for calibration development purposes. Using these models, engineers can employ optimization routines to calibrate Fuel-Air control setpoint functions automatically. Since the purpose of the developed models is calibration optimization, having a model that offers good sensitivity to variation in combustion inputs is very important. Currently, this model development process is not considered as a potential replacement for physical NOx sensors.

Finally, a new data-driven modeling of diesel engine soot formation using the gated recurrent unit (GRU) and long short term memory (LSTM) networks are introduced as an alternative to classical NARX type recurrent structures which are insufficient to generate accurate and smooth soot predictions. Utilization of a logarithmic normalization on soot measurements in the parameter estimation process reduces model sensitivity to sudden variations in soot measurements and increases the modeling accuracies. Prediction performances of GRU based soot models are quite satisfactory. GRU models are trained and validated with the accuracies above 70% and 50%, respectively. GRU soot models surpass the traditional NARX based models in both steady-state and transient test cycles. Moreover, a soot model with 71% training, 50% NEDC (steady-state) and 54% WLTC (transient) validation accuracies was obtained for the first time using LSTM structure.

As a future work, a global LSTM based modeling approach which involves the diesel engine torque and all combustion emissions is targetted. It is believed that some further modifications in an LSTM structure such as changing the architectures of the gates adaptively or using multiple LSTMs in series/parallel configurations according to input types will capture more dynamical relations and increase the modeling accuracies of challenging processes like soot emissions.

Bibliography

[1] "Diesel engine market: Global industry trends, share, size, growth, opportunity and forecast 2019-2024," 2019. IMARC Group.

[2] R. Baranescu and B. J. Challen, "Diesel engine reference book," *Warrendale, PA: Society of Automotive Engineers, 1999. 700*, 1999.

[3] A. Azama, S. Alia, and A. Iqbala, "134. emissions from diesel engine and exhaust after treatment technologies," in *4th International Conference on Energy, Environment and Sustainable Development*, 2016.

[4] "Eu: Heavy-duty truck and bus engines." `https://www.dieselnet.com/standards/eu/hd.php`. Revision: 2019.03, Accessed: 2019-06-25.

[5] "Ece – united nations," *Regulation No. 83–07*, 2015.

[6] C. Gurel, E. Ozmen, M. Yilmaz, D. Aydin, and K. Koprubasi, "Multi-objective optimization of transient air-fuel ratio limitation of a diesel engine using doe based pareto-optimal approach," *SAE International Journal of Commercial Vehicles*, vol. 10, no. 2017-01-0587, pp. 299–307, 2017.

[7] N. K. Togun and S. Baysec, "Prediction of torque and specific fuel consumption of a gasoline engine by using artificial neural networks," *Applied Energy*, vol. 87, no. 1, pp. 349–355, 2010.

[8] N. Togun and S. Baysec, "Genetic programming approach to predict torque and brake specific fuel consumption of a gasoline engine," *Applied Energy*, vol. 87, no. 11, pp. 3401–3408, 2010.

[9] G. Rizzoni and Y. Zhang, "Identification of a non-linear internal combustion engine model for on-line indicated torque estimation," *Mechanical Systems and Signal Processing*, vol. 8, no. 3, pp. 275–287, 1994.

[10] X. Wei and L. Del Re, "Lpv identification of a diesel engine torque model," *IFAC Proceedings Volumes*, vol. 36, no. 16, pp. 1405–1410, 2003.

[11] B. Bamieh and L. Giarré, "Identification for a general class of lpv models," *IFAC Proceedings Volumes*, vol. 33, no. 15, pp. 277–282, 2000.

[12] V. Verdult and M. Verhaegen, "Identification of multivariable linear parameter-varying systems based on subspace techniques," in *Proceedings of the 39th IEEE Conference on Decision and Control (Cat. No. 00CH37187)*, vol. 2, pp. 1567–1572, IEEE, 2000.

[13] S. Chen and J. Moskwa, "Application of nonlinear sliding-mode observers for cylinder pressure reconstruction," *Control Engineering Practice*, vol. 5, no. 8, pp. 1115–1121, 1997.

[14] Y. Wang and F. Chu, "Application of non-linear observers to on-line estimation of indicated torque in automotive engines," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 219, no. 1, pp. 65–75, 2005.

[15] S. A. Ali and S. Saraswati, "Cycle-by-cycle estimation of cylinder pressure and indicated torque waveform using crankshaft speed fluctuations," *Transactions of the Institute of Measurement and Control*, vol. 37, no. 6, pp. 813–825, 2015.

[16] S. A. Ali and S. Saraswati, "A comparison of sliding mode and frf-based observers for cylinder pressure estimation of spark ignition engine," *Transactions of the Institute of Measurement and Control*, vol. 39, no. 2, pp. 163–172, 2017.

[17] L. Zhixiong, G. Zhiwei, H. Chongqing, and L. Aihua, "On-line indicated torque estimation for internal combustion engines using discrete observer," *Computers & Electrical Engineering*, vol. 60, pp. 100–115, 2017.

[18] Y. Zweiri and L. Seneviratne, "Diesel engine indicated and load torque estimation using a non-linear observer," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 220, no. 6, pp. 775–785, 2006.

[19] Y. H. Zweiri, D. Seneviratne, *et al.*, "Diesel engine indicated torque estimation based on artificial neural networks," in *2007 IEEE/ACS International Conference on Computer Systems and Applications*, pp. 791–798, IEEE, 2007.

[20] A. Brahma, D. Upadhyay, A. Serrani, and G. Rizzoni, "Modeling, identification and state estimation of diesel engine torque and nox dynamics in response to fuel quantity and timing excitations," in *Proceedings of the 2004 American Control Conference*, vol. 3, pp. 2166–2171, IEEE, 2004.

[21] M.-F. Hsieh and J. Wang, "No and no2 concentration modeling and observer-based estimation across a diesel engine aftertreatment system," *Journal of Dynamic Systems, Measurement, and Control*, vol. 133, no. 4, p. 041005, 2011.

[22] R. Sindhu, G. A. P. Rao, and K. M. Murthy, "Effective reduction of nox emissions from diesel engine using split injections," *Alexandria engineering journal*, vol. 57, no. 3, pp. 1379–1392, 2018.

[23] C. Bertram, R. Rezaei, B. Tilch, and P. van Horrick, "Development of an euro vi engine using model-based calibration," *MTZ worldwide*, vol. 75, no. 10, pp. 4–9, 2014.

[24] J. Stoev and J. Schoukens, "Nonlinear system identification—application for industrial hydro-static drive-line," *Control Engineering Practice*, vol. 54, pp. 154–165, 2016.

[25] R. Rezaei, F. Dinkelacker, B. Tilch, T. Delebinski, and M. Brauer, "Phenomenological modeling of combustion and nox emissions using detailed tabulated chemistry methods in diesel engines," *International Journal of Engine Research*, vol. 17, no. 8, pp. 846–856, 2016.

[26] E. Perez, X. Blasco, S. Garcia-Nieto, and J. Sanchis, "Diesel engine identification and predictive control using wiener and hammerstein models," in *2006 IEEE Conference on Computer Aided Control System Design, 2006 IEEE International Conference on Control Applications, 2006 IEEE International Symposium on Intelligent Control*, pp. 2417–2423, IEEE, 2006.

[27] S. I. Raptotasios, N. F. Sakellaridis, R. G. Papagiannakis, and D. T. Hountalas, "Application of a multi-zone combustion model to investigate the nox reduction potential of two-stroke marine diesel engines using egr," *Applied Energy*, vol. 157, pp. 814–823, 2015.

[28] M. Hirsch, D. Alberer, and L. Del Re, "Grey-box control oriented emissions models," *IFAC Proceedings Volumes*, vol. 41, no. 2, pp. 8514–8519, 2008.

[29] C. Quérel, O. Grondin, and C. Letellier, "Semi-physical mean-value nox model for diesel engine control," *Control Engineering Practice*, vol. 40, pp. 27–44, 2015.

[30] J. Asprion, O. Chinellato, and L. Guzzella, "A fast and accurate physics-based model for the nox emissions of diesel engines," *Applied energy*, vol. 103, pp. 221–233, 2013.

[31] J. Asprion, O. Chinellato, and L. Guzzella, "Optimisation-oriented modelling of the nox emissions of a diesel engine," *Energy conversion and management*, vol. 75, pp. 61–73, 2013.

[32] C. Benatzky, S. Stadlbauer, S. Formentin, A. Schilling, and D. Alberer, "Indicated pressure-based data-driven diesel engine nox modeling," *international journal of engine research*, vol. 15, no. 8, pp. 934–943, 2014.

[33] M. Henningsson, K. Ekholm, P. Strandh, P. Tunestål, and R. Johansson, "Dynamic mapping of diesel engine through system identification," in *Identification for automotive systems*, pp. 223–239, Springer, 2012.

[34] S. Formentin, M. Corno, H. Waschl, D. Alberer, and S. M. Savaresi, "Nox estimation in diesel engines via in-cylinder pressure measurement," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 1, pp. 396–403, 2013.

[35] T. Boz, M. Unel, V. Aran, M. Yilmaz, C. Gurel, C. Bayburtlu, and K. Koprubasi, "Diesel engine nox emission modeling with airpath input channels," in *IECON 2015-41st Annual Conference of the IEEE Industrial Electronics Society*, pp. 003382–003387, IEEE, 2015.

[36] 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," *Applied Energy*, vol. 119, pp. 330–340, 2014.

[37] A. Pfeifer, M. Krueger, U. Gruetering, and D. Tomazic, "Us 2007-which way to go? possible technical solutions," tech. rep., SAE Technical Paper, 2003.

[38] N. Walke, M. Nandgaonkar, and N. Marathe, "Nox, soot, and fuel consumption predictions under transient operating cycle for common rail high power density diesel engines," *Journal of Combustion*, vol. 2016, 2016.

[39] H. Hiroyasu and T. Kadota, "Models for combustion and formation of nitric oxide and soot in direct injection diesel engines," *SAE transactions*, pp. 513–526, 1976.

[40] C. Tanelli and S. Maranta, "Modeling soot emissions in diesel engines by means of semi-empirical approaches and a combustion model based on detailed chemistry," 2015.

[41] J. Moss, C. Stewart, and K. Young, "Modeling soot formation and burnout in a high temperature laminar diffusion flame burning under oxygen-enriched conditions," *Combustion and flame*, vol. 101, no. 4, pp. 491–500, 1995.

[42] K. M. Leung, R. P. Lindstedt, and W. Jones, "A simplified reaction mechanism for soot formation in nonpremixed flames," *Combustion and flame*, vol. 87, no. 3-4, pp. 289–305, 1991.

[43] Z. Wen, S. Yun, M. Thomson, and M. Lightstone, "Modeling soot formation in turbulent kerosene/air jet diffusion flames," *Combustion and Flame*, vol. 135, no. 3, pp. 323–340, 2003.

[44] M. Benz, C. H. Onder, and L. Guzzella, "Engine emission modeling using a mixed physics and regression approach," *Journal of engineering for gas turbines and power*, vol. 132, no. 4, p. 042803, 2010.

[45] F. Tschanz, A. Amstutz, C. H. Onder, and L. Guzzella, "A real-time soot model for emission control of a diesel engine," *IFAC Proceedings Volumes*, vol. 43, no. 7, pp. 222–227, 2010.

[46] C. Ericson, B. Westerberg, and R. Egnell, "Transient emission predictions with quasi stationary models," tech. rep., SAE Technical Paper, 2005.

[47] H. Sequenz, M. Mrosek, and R. Isermann, "Stationary global-local emission models of a cr-diesel engine with adaptive regressor selection for measurements of airpath and combustion," *IFAC Proceedings Volumes*, vol. 43, no. 7, pp. 246–251, 2010.

[48] C. L. Mallows, "Some comments on c p," *Technometrics*, vol. 15, no. 4, pp. 661–675, 1973.

[49] M. Mrosek, H. Sequenz, and R. Isermann, "Control oriented nox and soot models for diesel engines," *IFAC Proceedings Volumes*, vol. 43, no. 7, pp. 234–239, 2010.

[50] M. Hafner, M. Schüler, O. Nelles, and R. Isermann, "Fast neural networks for diesel engine control design," *Control Engineering Practice*, vol. 8, no. 11, pp. 1211–1221, 2000.

[51] O. Nelles and R. Isermann, "Basis function networks for interpolation of local linear models," in *Proceedings of 35th IEEE Conference on Decision and Control*, vol. 1, pp. 470–475, IEEE, 1996.

[52] M. Grahn, K. Johansson, and T. McKelvey, "Data-driven emission model structures for diesel engine management system development," *International Journal of Engine Research*, vol. 15, no. 8, pp. 906–917, 2014.

[53] C. Atkinson and G. Mott, "Dynamic model-based calibration optimization: An introduction and application to diesel engines," tech. rep., SAE Technical Paper, 2005.

[54] G. Thompson, C. Atkinson, N. Clark, T. Long, and E. Hanzevack, "Neural network modelling of the emissions and performance of a heavy-duty diesel engine," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 214, no. 2, pp. 111–126, 2000.

[55] U. Karrenberg, *An interactive multimedia introduction to signal processing*. Springer Science & Business Media, 2007.

[56] Y. Zhu, *Multivariable system identification for process control*. Elsevier, 2001.

[57] J. Yang and V. Honavar, "Feature subset selection using a genetic algorithm," in *Feature extraction, construction and selection*, pp. 117–136, Springer, 1998.

[58] R. Bro and A. K. Smilde, "Principal component analysis," *Analytical Methods*, vol. 6, no. 9, pp. 2812–2831, 2014.

[59] R. Wasserstein, "George box: A model statistician," *Significance*, vol. 7, no. 3, pp. 134–135, 2010.

[60] E.-W. Bai, "Prediction error adjusted gaussian process for nonlinear non-parametric system identification," *IFAC Proceedings Volumes*, vol. 45, no. 16, pp. 101–106, 2012.

[61] O. Nelles, *Nonlinear system identification: from classical approaches to neural networks and fuzzy models*. Springer Science & Business Media, 2013.

[62] R. Haber and L. Keviczky, "Nonlinear system identification-input-output modeling approach," 1999.

[63] S. Billings, "Identification of nonlinear systems using parameter estimation techniques," in *Institute of Electrical Engineers Conference*, pp. 183–190, 1981.

[64] S. Roy, R. Banerjee, A. K. Das, and P. K. Bose, "Development of an ann based system identification tool to estimate the performance-emission characteristics of a crdi assisted cng dual fuel diesel engine," *Journal of Natural Gas Science and Engineering*, vol. 21, pp. 147–158, 2014.

[65] H. T. Siegelmann, B. G. Horne, and C. L. Giles, "Computational capabilities of recurrent narx neural networks," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 27, no. 2, pp. 208–215, 1997.

[66] J. L. Elman, "Finding structure in time," *Cognitive science*, vol. 14, no. 2, pp. 179–211, 1990.

[67] J. L. Elman, "Distributed representations, simple recurrent networks, and grammatical structure," *Machine learning*, vol. 7, no. 2-3, pp. 195–225, 1991.

[68] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[69] J. A. Pérez-Ortiz, F. A. Gers, D. Eck, and J. Schmidhuber, "Kalman filters improve lstm network performance in problems unsolvable by traditional recurrent nets," *Neural Networks*, vol. 16, no. 2, pp. 241–250, 2003.

[70] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.

[71] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.

[72] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Gated feedback recurrent neural networks," in *International Conference on Machine Learning*, pp. 2067–2075, 2015.

[73] F. Visin, K. Kastner, K. Cho, M. Matteucci, A. Courville, and Y. Bengio, "Renet: A recurrent neural network based alternative to convolutional networks," *arXiv preprint arXiv:1505.00393*, 2015.

[74] J. Mao, W. Xu, Y. Yang, J. Wang, Z. Huang, and A. Yuille, "Deep captioning with multimodal recurrent neural networks (m-rnn)," *arXiv preprint arXiv:1412.6632*, 2014.

[75] M. Al-Smadi, O. Qawasmeh, M. Al-Ayyoub, Y. Jararweh, and B. Gupta, "Deep recurrent neural network vs. support vector machine for aspect-based sentiment analysis of arabic hotels' reviews," *Journal of computational science*, vol. 27, pp. 386–393, 2018.

[76] T. Mikolov, M. Karafiát, L. Burget, J. Černockỳ, and S. Khudanpur, "Recurrent neural network based language model," in *Eleventh annual conference of the international speech communication association*, 2010.

[77] S. Yeung, O. Russakovsky, N. Jin, M. Andriluka, G. Mori, and L. Fei-Fei, "Every moment counts: Dense detailed labeling of actions in complex videos," *International Journal of Computer Vision*, vol. 126, no. 2-4, pp. 375–389, 2018.

[78] H. Stuhler, T. Kruse, A. Stuber, K. Gschweitl, W. Piock, H. Pfluegl, and P. Lick, "Automated model-based gdi engine calibration adaptive online doe approach," tech. rep., SAE Technical Paper, 2002.

[79] Y. Nozaki, T. Fukuma, and K. Tanaka, "Development of a rule-based calibration method for diesel engines," tech. rep., SAE Technical Paper, 2005.

[80] E. Rask and M. Sellnau, "Simulation-based engine calibration: tools, techniques, and applications," *SAE transactions*, pp. 821–832, 2004.

[81] P. Gnanam, "Mobeo: Model based engine development and calibration," *Future Powertrains Conference*.

[82] I. Brahma and J. N. Chi, "Development of a model-based transient calibration process for diesel engine electronic control module tables–part 2: modelling and optimization," *International Journal of Engine Research*, vol. 13, no. 2, pp. 147–168, 2012.

[83] T. C. Zannis, D. T. Hountalas, E. A. Yfantis, R. G. Papagiannakis, and Y. A. Levendis, "Intake-air oxygen-enrichment of diesel engines as a power enhancement method and implications on pollutant emissions," in *ASME 2009 Internal Combustion Engine Division Fall Technical Conference*, pp. 295–308, American Society of Mechanical Engineers, 2009.

[84] S. Billings, M. Korenberg, and S. Chen, "Identification of non-linear output-affine systems using an orthogonal least-squares algorithm," *International Journal of Systems Science*, vol. 19, no. 8, pp. 1559–1568, 1988.

[85] S. Chen, S. A. Billings, and W. Luo, "Orthogonal least squares methods and their application to non-linear system identification," *International Journal of control*, vol. 50, no. 5, pp. 1873–1896, 1989.

[86] H.-L. Wei, S. A. Billings, and J. Liu, "Term and variable selection for non-linear system identification," *International Journal of Control*, vol. 77, no. 1, pp. 86–110, 2004.

[87] C. Hametner, M. Stadlbauer, M. Deregnaucourt, S. Jakubek, and T. Winsel, "Optimal experiment design based on local model networks and multilayer perceptron networks," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 1, pp. 251–261, 2013.

[88] L. Pronzato, "Optimal experimental design and some related control problems," *Automatica*, vol. 44, no. 2, pp. 303–325, 2008.

[89] T. J. Santner, B. J. Williams, W. Notz, and B. J. Williams, *The design and analysis of computer experiments*, vol. 1. Springer, 2003.

[90] L. Ljung, *System Identification: Theory for the User*. Prentice Hall, 1999.

[91] R. Pintelon and J. Schoukens, *System identification: a frequency domain approach*. John Wiley & Sons, 2012.

[92] H. Steven, "Development of a worldwide harmonised heavy-duty engine emissions test cycle," *Final Report, April*, 2001.

[93] T. J. Barlow, S. Latham, I. McCrae, and P. Boulter, "A reference book of driving cycles for use in the measurement of road vehicle emissions," *TRL Published Project Report*, 2009.

[94] P. Falcone, M. C. De Gennaro, G. Fiengo, L. Glielmo, S. Santini, and P. Langthaler, "Torque generation model for diesel engine," in *42nd IEEE International Conference on Decision and Control (IEEE Cat. No. 03CH37475)*, vol. 2, pp. 1771–1776, IEEE, 2003.

[95] L. Ljung, *System identification toolbox 7: Getting started guide*. The MathWorks, 2008.

[96] J. Ko, D. Jin, W. Jang, C.-L. Myung, S. Kwon, and S. Park, "Comparative investigation of nox emission characteristics from a euro 6-compliant diesel passenger car over the nedc and wltc at various ambient temperatures," *Applied energy*, vol. 187, pp. 652–662, 2017.

[97] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[98] W. Schindler, C. Haisch, H. A. Beck, R. Niessner, E. Jacob, and D. Rothe, "A photoacoustic sensor system for time resolved quantification of diesel soot emissions," *SAE transactions*, pp. 483–490, 2004.

[99] W. Schindler, M. Arndt, A. Painsi, and B. Giechalskiel, "Avl micro soot sensor applications," in *AVL Tech Days*, 2012.