THESIS

ARTIFICIAL NEURAL NETWORKS FOR FUEL CONSUMPTION AND EMISSIONS MODELING IN LIGHT DUTY VEHICLES

Submitted by

Shiva Tarun Chenna

Department of Mechanical Engineering

In partial fulfillment of the requirements

For the degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2019

Master's Committee:

Advisor: Shantanu Jathar

Thomas Bradley
Chuck Anderson

ABSTRACT

ARTIFICIAL NEURAL NETWORKS FOR FUEL CONSUMPTION AND EMISSIONS MODELING IN LIGHT DUTY VEHICLES

There is growing evidence that real world, on-road emissions from mobile sources exceed emissions determined during laboratory tests and that the air quality, climate, and human health impacts from mobile sources might be substantially different than initially thought. Hence, there is an immediate need to measure and model these exceedances if we are to better understand and mitigate the environmental impacts of mobile sources. In this work, we used a portable emissions monitoring system (PEMS) and artificial neural networks (ANNs) to measure and model on-road fuel consumption and tailpipe emissions from Tier-2 light-duty gasoline and diesel vehicle.

Tests were performed on at least five separate days for each vehicle and each test included a cold start and operation over a hot phase. Routes were deliberately picked to mimic certain features (e.g., distance, time duration) of driving cycles used for emissions certification (e.g., FTP-75). Data were gathered for a total of 49 miles and 145 minutes for the gasoline vehicle and 52 miles and 165 minutes for the diesel vehicle. Fuel consumption and emissions data were calculated at 1 Hz using information gathered from the vehicle using the onboard diagnostics port and the PEMS measurements. Route-integrated tailpipe emissions did not exceed the Tier-2 emissions standard for CO, NO_x, and non-methane organic gases (NMOG) for either vehicle but did exceed so for PM for the diesel vehicle.

We trained ANN models on part of the data to predict fuel consumption and tailpipe emissions at 1 Hz for both vehicles and evaluated these models against the rest of the data. The ANN models performed best when the training iterations (or epochs) were set to larger than 25 and the number of neurons in the hidden layer was between 7 and 9, although we did not see any specific advantage in increasing the

number of hidden layers beyond 1. The trained ANN model predicted the fuel consumption over test routes within 5.5% of the measured value for both gasoline and diesel vehicles. The ANN performance varied significantly with pollutant type for the two vehicles and we were able to develop satisfactory models only for unburned hydrocarbons (HC) and NO_x for diesel vehicles. Over independent test routes, the trained ANN models predicted HC within 12.5% of the measured value for the gasoline vehicle and predicted NO_x emissions within 3% of the measured values for the diesel vehicle. The ANN performed better than, and hence could be used in lieu of, multivariable regression models such as those used in mobile source emissions models (e.g., EMFAC). In an 'environmental-routing' case study performed over three origin-destination pairs, the ANNs were able to successfully pick routes that minimized fuel consumption.

Our work demonstrates the use of artificial neural networks to model fuel consumption and tailpipe emissions from light-duty passenger vehicles, with applications ranging from environmental routing to emissions inventory modeling.

ACKNOWLEDGEMENTS

I would first like to thank my advisor Dr. Shantanu Jathar for always guiding me in the right direction and helping me in accomplishing this degree. I must express my very profound gratitude to my family and to my friends for showing me unconditional love and encouragement throughout my life. This work would not have been possible without them.

TABLE OF CONTENTS

1. Introduction

The vast majority of anthropogenic greenhouse gases come from the combustion of fossil fuels used for energy production and transportation (Kampa and Castanas, 2008). According to 2016 United States Environmental Protection Agency (EPA) report, in the United States, the transportation and electricity generation sectors contribute to the majority of total greenhouse gas emissions, with each contributing 28 percent each. Within the transportation sector, light-duty vehicles (LDV) which include passenger cars below 8500 lbs contribute to 60% of total emissions followed by heavy-duty trucks (HDT) which account for 23%. Greenhouse gases from transportation include carbon dioxide (CO_2), methane (CH_4), nitrous Oxides (NO_X) which are a result of the combustion of fossil fuels. Passenger vehicles have a huge environmental impact and contribute to 38.1% of total CO emissions, 34.7% of total NO_X emissions, 7.1% of total $PM_{2.5}$ and 10.8% of total PM10 emissions (U.S. EPA Office of Air and Radiation, n.d.). These pollutants not only have significant climate impact but also have adverse health effects (Kampa and Castanas, 2008) (Brugge et al., 2007) (Laumbach and Kipen, 2012) (Zhang and Batterman, 2013). For example, breathing elevated levels of CO reduces the amount of oxygen reaching the body's organs and tissues. This can result in chest pain and other serious symptoms like a coma. On the climate front, CO emissions, contribute to the formation of CO_2 and ozone, greenhouse gases that warm the atmosphere. NO_X also enhances the production of ozone (O_3) causing greenhouse effects. Since vehicular pollution has a significant role in global climate change it is important to reduce the emissions from vehicles in the most cost-efficient way possible. One forthright approach to reducing greenhouse gas emissions is by improving fuel economy which results in a lesser amount of fuel burnt leading to reduced emissions.

To improve fuel economy the US EPA imposed Corporate Average Fuel Economy (CAFE) standards in 1982. These are a fleet-wide average fuel economy standard in a given model year, expressed in miles per gallon that the manufacturers must attain to be compliant and avoid fines. These standards indirectly contribute to better air quality through reduced emissions of CO, NO_X, and $PM_{2.5}$. To improve fleet wide

averages, manufacturers have also been encouraged to increase the production of electric and hybrid vehicles. While there are many methods to improve fuel economy in vehicles, in our study, we try to improve fuel economy by Eco-routing. Eco-routing is the ability for a vehicle to acknowledge all possible routes to get to a destination and identify the most fuel efficient and/or emissions efficient way to help the driver reduce the environmental impact of their journey. To do this we need a better understanding of individual vehicles on-road vehicular emissions for a representable drive cycle.

All vehicles regardless of size and model year must undergo emissions test procedures and comply with the certification standards. The emissions test procedures take place on a chassis dynamometer for standard emission test cycles. Chassis dynamometer drive cycles are defined as vehicle velocity as a function of time intended to capture various driving conditions such as rural, urban, low-speed city driving and aggressive highway driving. Common drive cycles used for chassis dynamometer tests are the Federal Test Procedure-75 (FTP-75), New York City Cycle (NYCC), California Unified Cycle (UC, LA92) and, Inspection and Maintenance Driving Cycle (IM240). The existing models to predict on-road emissions in the United States are EMFAC (EMission FACtor) by California Air Resources Board and MOtor Vehicle Emission Simulator (MOVES) which are quite extensive and are built using data from these emissions testing facilities that use chassis dynamometer tests. These models are only accurate for on-road results if the data with which they are built are accurate. Chassis dynamometer tests often fail to capture the effects of real-world driving such as driver behavior, weather, traffic and, rash driving habits which can result in under-predicting on-road emissions (Frey et al., 2008). Many studies have pointed out that the on-road emissions of vehicles are often much higher than certified limits especially during suburban/rural driving conditions (Pelkmans and Debal, 2006)), (Kumar Pathak et al., 2016). Anenberg et al., (2017) compiled information from 11 major markets (USA, China, Japan, EU) to show that over half of on-road light-duty diesel vehicles were exceeding certification limits. Figure 1.1 shows that in the United States, on-road NO_X emissions for tier 2 light-duty diesel vehicles were found at 0.35 $g.mile^{-1}$

while the certified limit was 0.7 g.mile^{-1}. These excess emissions (totaling 4.6 million tons) were linked to 38,000 premature deaths globally in 2015.

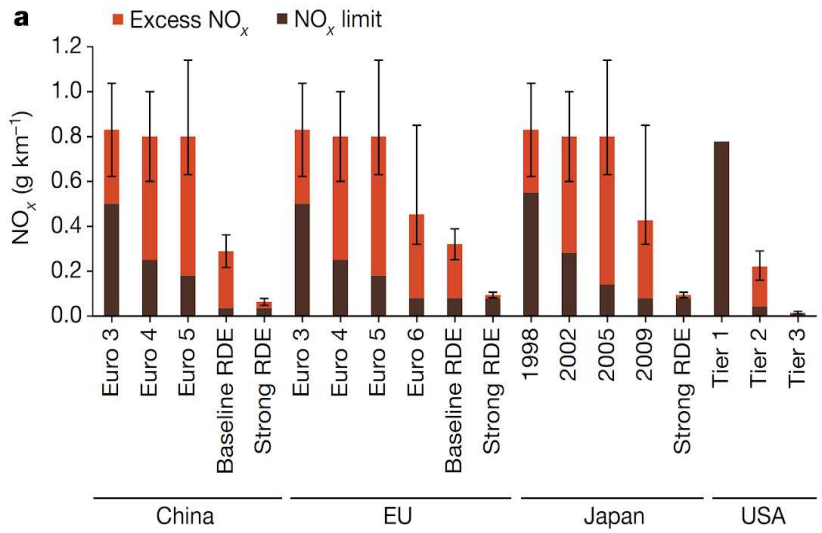

Fig 1.1 Real world NO_X emission factors by vehicle emissions standards in different regions (Anenberg et al., 2017)

Real-World drive cycles help in capturing the difference between the chassis dynamometer emissions and real driving emissions. Over the past decade, portable emissions monitoring systems (PEMSs) have been used to measure tailpipe emissions from on- and off-road vehicles for in-use tests ((Frey et al., 2003); (O'Driscoll et al., 2016); (Kwon et al., 2017);). While their use earlier was limited to studying differences between laboratory and on-road use, they are now increasingly used for regulatory purposes. For instance, federal (Environmental Protection Agency) and state (California Air Resources Board) emissions certification for heavy-duty vehicles in the United States requires compliance with Not To Exceed (NTE) standards measured via PEMS devices (https://www.dieselnet.com/standards/cycles/nte.php). The European Commission now requires the use of PEMS-based testing to meet real driving emissions (RDE) standards in all of Europe and the United Kingdom. PEMS devices have evolved significantly in terms of the quality, performance, and cost and systematic tests against laboratory-grade reference instruments show very little difference (Durbin et al., 2007). Using PEMS devices to acquire in-use emissions data for

individual vehicles and developing vehicle specific emissions models using this data can deliver better estimates of real-world emissions from vehicles.

The goal of our study is a step to model fuel consumption and on-road emissions in real time based on parameters that can be calculated ahead of driving and give the user suggestions about the route that could be used which results in the least fuel consumed or least amount of emissions. We aim to develop models that can better predict on-road emissions from passenger vehicles to achieve more accurate emission estimates from the transportation sector. The advent of smart and connected vehicles helps us to have a better sense of the route to be driven in advance. Developing models for each individual vehicle and identifying the possible routes between the origin and destination pair will help us in predicting the fuel consumption and tailpipe emissions in advance. In our work, we construct models using the data collected from on-road vehicles under real driving conditions using a Portable Emissions Measurement Device (PEMS). Artificial Neural Network (ANN) models and multivariable linear regression models were developed using data from the PEMS to predict on-road fuel consumption and tailpipe emissions from LDV's. The primary objectives of this work are (i) to determine whether ANN models perform better when compared to linear multivariable models in predicting on-road fuel consumption and emissions from light duty vehicle, (ii) determine efficient ways to train and test an ANN model to predict fuel consumption and tailpipe emissions and (iii) to understand how the performance of ANN model varies by the type of pollutant, and (iv) to determine if ANNs can be used to improve on-road fuel economy.

2. Methods

In the sections below, we describe the experimental (Section 2.1) and numerical methods (Section 2.2) used in this work. The experimental methods describe the use of the Portable Emissions Monitoring System (PEMS) to measure fuel consumption and tailpipe emissions from two light-duty passenger vehicles. The numerical methods describe the multivariable regression and artificial neural network models used to model fuel consumption and tailpipe emissions.

2.1 Experimental Methods

2.1.1 Portable Emissions Monitoring System (PEMS)

In this work, we used a PEMS AxionR/S manufactured by Global MRV (Buffalo, NY)(RDE PEMS, n.d.). This is a low to mid-range device that costs approximately $100,000 and was loaned through a no-cost contract with Lightning Systems (Loveland, CO). The PEMS (shown in Figure 2.1(a)) consists of a central unit in a briefcase that includes the sensors and the integrated system to acquire, process, record, and display data in real-time. It has a small form factor, weighs approximately 17 kg, and can be easily placed on the front passenger seat during tests. The PEMS measures concentrations of CO_2, O_2, CO, NO_X, HC (hydrocarbons), and PM (particulate matter) at 1 Hz; details for the different gas and particle sensors are provided in Table 2.1.

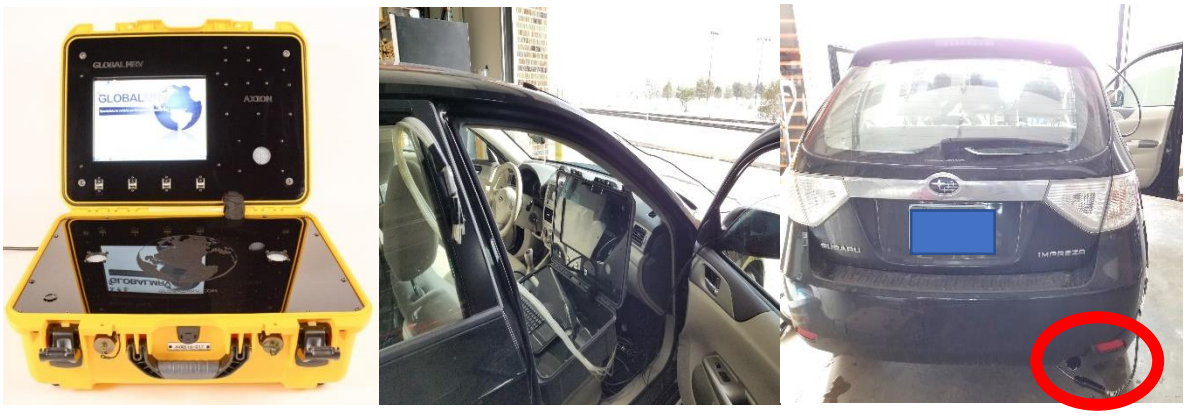

Figure 2.1: (a) PEMS device, (b)PEMS in the passenger seat during experiment(c)Exhaust sampling line

Table 2.1: Sensors and sensor features used in the PEMS AxionR/S. Values reflect those listed in the AxionR/S manual (Global MRV, 2016).

Species	***Measurement Type***	***Measurement Range***	***Accuracy***	***Resolution***	***Sensor Flow Rate***
CO_2	NDIR	0.01 – 16%	± 0.3% abs.	0.01 vol. %	1 liter/min
O_2	Electrochemical	0.01 – 25 %	± 0.1% abs	0.01 vol. %	1 liter/min
CO	NDIR	0.001 – 10%	± 0.02% abs.	0.001 vol. %	1 liter/min
NO	Electrochemical	1 – 4000 ppm	± 25 ppm abs.	1 ppm	1 liter/min
HC	NDIR	1 – 4000 ppm	± 8 ppm abs.	1 ppm	1 liter/min
PM	Light scattering	0.01 – 300 mg/m^3	± 2%	0.01 mg/m^3	4 liters/min

Tailpipe emissions were sampled through a stainless-steel probe (2.25 mm OD) inserted ~25 cm inside the tailpipe and secured using a steel hose clamp. The probe was connected to two conductive tubes that ran from the tailpipe to the PEMS device placed inside the vehicle (~8 m). One of the tubes sampled undiluted exhaust at 5 liters per minute (lpm) and delivered the sample to the gas sensors while the other tube sampled undiluted exhaust at 5 lpm and delivered the sample to the particle sensor. Both sample streams were run through a water trap to remove the condensing water and the gas stream was also filtered for particles using a particle filter. We did not measure losses of the species through the relatively long sample tubing but expect the HC and PM to be substantially affected based on the tube material (Deming et al., 2019). Some modern PEMS devices are located in the trunk of the vehicle or on a skid above the tailpipe outside the vehicle, presumably to shorten the sampling lines and maintain the sample integrity (e.g., Figure 2.2).

Figure 2.2: Examples of on-road sampling from light-duty vehicles with the PEMS placed in the trunk (left) and on a skid above the tailpipe (right). Courtesy: (Portable Emissions Measurement System)

2.1.2 Vehicles, Routes, and Experimental Details

The PEMS was used to measure tailpipe emissions from in-use gasoline and a diesel vehicle. Both vehicles were solicited from researchers working at the Powerhouse Energy Campus and were road legal (i.e., valid registration and emissions certificate provided under Colorado's inspection and maintenance program (AirCare Colorado, n.d.). We deliberately chose older vehicles to make sure we could make robust measurements with this low- to medium-range PEMS and not have to worry about signals near or below the limit of detection. We note that the goal of this research was to study the potential of artificial neural networks to model tailpipe emissions and the choice of vehicle should have little to no influence on the findings from this work. Details for both vehicles are provided in Table 2.2. Both vehicles were driven on commercially available fuel from local gas stations.

Table 2.2: Vehicles and vehicle specifications used for this study.

Attribute	***Gasoline***	***Diesel***
EPA Tier standard	Tier 2	Tier 1
Model year	2008	2003
Make and model	Subaru Impreza	Volkswagen Jetta
Engine displacement	2.5 L	1.9 L
Gross vehicle weight	4040 lbs.	4023 lbs.

Vehicle miles	88,000	120,000
Emissions control	3-way catalytic converter (CO, THC, NO_X)	Exhaust gas recirculation for NO_X

We performed a total of eight experiments with four tests each for the gasoline and diesel vehicle. Each test was performed on a different day since to ensure that our test, like those performed on chassis dynamometers, included a cold start. Gasoline tests were conducted in spring 2018 (ambient temperatures of 0 to 6 ℃) and the diesel tests conducted in spring and summer of 2018 (ambient temperatures of 7 to 22 ℃). Tests were performed on the same urban-suburban route through Fort Collins, CO that was ~10 miles long and required less than 30 minutes to navigate. The test route with the velocity superimposed on a map of Fort Collins, CO is shown in Figure 2.2(a) and the timeseries for the vehicle velocity is shown in Figure 2.2(b). The developed drive cycle starts at Fort Collins downtown to capture the urban/downtown driving and later routes through the suburbs of the city to mimic cruising/highway speeds. We can clearly see the low-velocity patterns near the starting point (A) and higher speeds as we move away from downtown Fort Collins. Since these were on-road tests, the vehicle operation during each of these tests was a little different. In addition to the test route mentioned above, we drove one random route with the gasoline vehicle on one of its four experiment days and drove four random routes with the diesel vehicle on each of its four experiment days. These random route tests do not include a cold start.

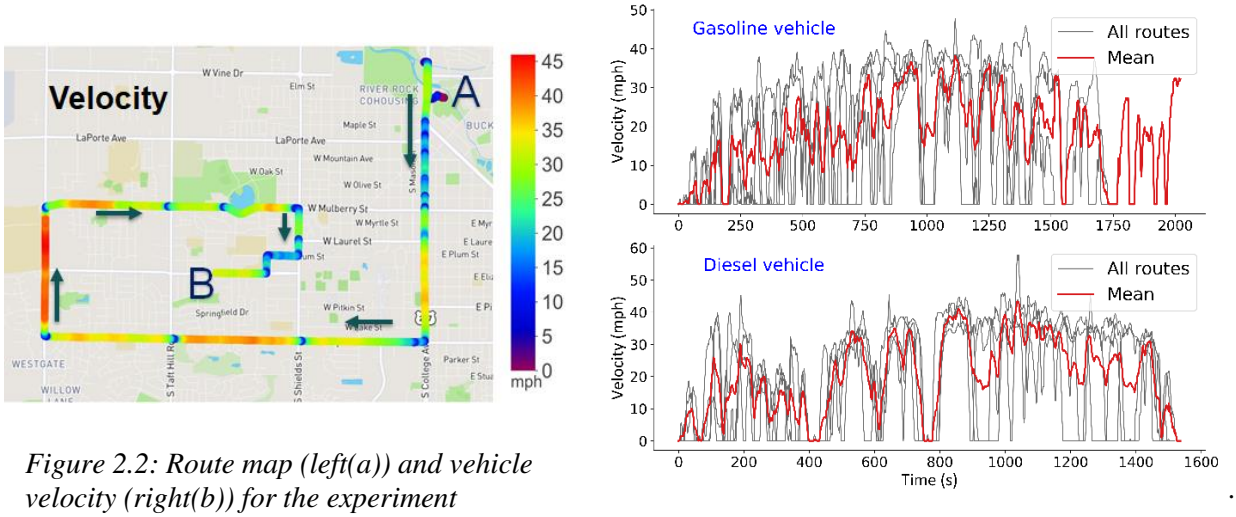

Figure 2.2: Route map (left(a)) and vehicle velocity (right(b)) for the experiment performed on January 3rd, 2018 with the gasoline vehicle

At the beginning of each test day, the PEMS was placed on the front passenger seat, following which the PEMS was powered and warmed up for 45 minutes. A calibration was then performed with zero air and a low and high concentration gas mixture (low: 6% CO_2, 0.5 ppmv CO, 300 ppmv NO, 200 ppmv of propane and high: 12% CO_2, 8 ppmv CO, 3000 ppmv NO, 3200 ppmv propane). The manufacturer recommends calibration for every 10 hours of PEMS operation and hence a calibration was performed for every test day. The PEMS was powered using a dedicated 12V, 80 Ah lead acid battery that was placed in the passenger row of the vehicle. On a full charge, the lead-acid battery can provide 10 hours of uninterrupted power to the PEMS and was enough for the single day tests performed in this work. All lines to and from the PEMS (sample lines, zero air line to measure the background air, exhaust lines) were taped to the exterior of the vehicle for safety purposes. The PEMS was interfaced with the On-Board Diagnostics-II (OBD-II) port on the vehicle to read vehicle parameters that included, velocity, intake air temperature, engine speed, and manifold air pressure; a complete list can be found in the appendix. Finally, a GPS (global positioning system) module connected to the PEMS was affixed to the top of the vehicle to record latitude, longitude, and altitude. The PEMS was made to sample ambient air for 45 minutes at the end of each test day to flush the sample and exhaust lines.

2.1.3 PEMS Data Management

The PEMS recorded and stored raw and processed data on a local hard drive in the ASCII format. These data included, but were not limited, to local time (hh:mm:ss), tailpipe concentrations(%, ppmv, or μg m^{-3}) and emission factors (g $mile^{-1}$) and rates ($mg.s^{-1}$) for CO_2, O_2, CO, NO_X, THC, and PM, air intake ($g.s^{-1}$), exhaust flow ($g.s^{-1}$), fuel consumption ($g.s^{-1}$), latitude, longitude, altitude (m), and vehicle speed (km h^{-1}). We should note that tailpipe concentrations and emissions were corrected for the residence time in the sampling line. The ASCII data from the PEMS for each experiment was organized as a structure array and stored as a .mat file for further processing in MATLAB (MathWorks, MA). The raw ASCII files from the PEMS along with MATLAB structure arrays and codes are currently stored on a network drive but will eventually be archived with CSU Libraries.

2.2 Numerical Methods

In this study, Linear regression models and Artificial Neural Network (ANN) models were the two different predictive modeling approaches used to model on-road fuel consumption and tailpipe emissions from the two light-duty vehicles. Both these models are useful to solve tasks that are difficult to solve with fixed programs. The sections below briefly describe the theory and our process for model development and application.

2.2.1 Linear Regression Modeling

A linear regression model is a form of predictive modeling technique where the model can take a set of parameters as input and can predict a scalar output corresponding to those inputs. As the name suggests, the output of the linear regression model is a linear function of the inputs. The three different types of linear regression models based on the number of inputs and outputs of the model are simple regression, multivariable linear regression, and multivariate linear regression. Simple regression is when a single output is predicted using one dependent variable (input). In multivariable regression, a single output is

predicted using two or more dependent variables and in multivariate regression, one or more dependent variables are used to predict multiple output parameters at once.

- Simple Regression $Y = \beta_0 + \beta_1 * X$
- Multivariable Linear Regression $Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \cdots \beta_n * X_n$
- Multivariate Regression $\hat{Y} = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \cdots \beta_n * X_n$

Y is the response variable, X_1 through X_n are the dependent variable or predictors, β_0 is the bias or intercept, and β_1 through β_n are the coefficients associated with respective predictors and, $\hat{Y}$ is set of response variables that are dependent on each of the dependent variables.

In this study, we used Multivariable (MV) linear regression model to predict fuel consumption and tailpipe emissions individually using multiple predictors such as vehicle velocity, acceleration, intake air temperature, vehicle rotations per minute (rpm), vehicle specific power and, time since the start of the experiment. In the training phase, each set of inputs has an associated target output value which results in a set of weights that are calculated using the method of least squares (Geladi and Kowalski, 1986), the greater the weight associated with an input parameter, the greater is its impact on the response or output variable.

2.2.2 Artificial Neural Networks

Artificial Neural Networks (ANN) or Machine Learning (ML) models are learning algorithms that were inspired to be computational models of biological learning. As defined in (Mitchell, 1997) " A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance of tasks in T, as measured by P, improves with experience E". ANNs can be used to perform multiple tasks and some common tasks solved using ANNs are classification, denoising and, regression. A common example of classification is object detection where the input is an image and the output are the identification of the object in the image if any. Denoising is when the neural networks are required to denoise a corrupted input signal with the help of a model built using uncorrupted input

signals. Regression is when the ANN model is trained on an existing set of observations and predicts a numerical output based on a previously unobserved input or set of inputs. An example of this is temperature forecasting based on previous trends and various factors contributing to the change in temperature. ANN models are non-linear models, primarily used when the underlying relationship within the data is unknown. These models can identify and learn correlated patterns between the inputs and the associated target output even if the data is non-linear, complex and, noisy and can predict the output of new independent data (Lek and Guégan, 1999) .

ANN algorithms are broadly classified as supervised and unsupervised learning (LeCun et al., 2015). Unsupervised learning is when neural network models learn by understanding the multiple features of the input dataset that does not have a corresponding observed target value, this is mostly associated with the classifier algorithms such as image recognition (Weber et al., 2000) and clustering analysis (Erman et al., 2006) . In our work, we have used supervised ANN models that learn from not only the features of the input values but also have an associated target value for each input. Multiple studies have previously used ANN models for various emissions modeling (Kukkonen et al., 2003) (Khoshnevisan et al., 2013) (Nagendra and Khare, 2006) (Mohamed Ismail et al., 2012) (Thompson et al., 2000). A multi-layer feedforward neural network also known as multi-layer perceptron (MLP) is a popular supervised neural network architecture where the model is constructed based on experimental data with known output. In this architecture, the neurons are arranged in successive layers from the input layer to the output layer through hidden layers where information flows unidirectionally (Lek and Guégan, 1999). The number of neurons in each of these layers varies according to the complexity of the data. The most important phase in model development is the training phase where the ANN model learns using a set algorithm to perform the designated task. In this phase, the network is presented with a training dataset consisting of input/output pairs obtained from the real-world experimentation. After the training phase, the performance of the model is assessed using a different dataset called test dataset. The primary challenges associated with training the model are overfitting, underfitting, and generalization of models. Underfitting

is when the trained model has high errors when validated against the same dataset with which it is trained with while overfitting happens when the model performs well when validated against the training dataset, but the performance drops when validated against a different dataset. Generalization refers to the model's ability to perform well on previously unobserved inputs during training. We can say that a model performs well on generalized data if the difference in error between train validation and test validation is minimal.

In our work, we used neural network toolbox available in MATLAB which has various training algorithms to train the neural network models (NN). In this study, we used a two-layer feed-forward network trained with the Levenberg-Marquardt algorithm and a sigmoid activation function.

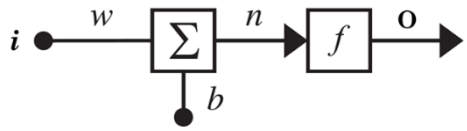

Figure 2.3: Basic neuron block in an artificial neural network model

Figure 2.3 depicts the basic block of all NN models which is a single input neuron where a scalar input (i) (e.g., Velocity) is multiplied with a scalar weight (w) and this weighted input is then added to a bias (b) term which becomes the net input that is passed through a transfer/activation function (f) to produce a scalar output (o) (e.g., fuel consumption or emissions) . The neural network element computes a linear combination of its input signals (net input) and applies a sigmoid function $f(x) = \frac{1}{1+e^{-x}}$ to the result to introduce nonlinearity in the model by mapping the net input that can vary between minus infinity to plus infinity to 0 and 1. The output from the hidden layer is then fed as an input to the output layer where a similar procedure with a linear activation function is applied to approximate the function value. A typical neural network architecture consists of multiple neurons in different layers. Figure 2.4 illustrates an example of the NN architecture in the MATLAB toolbox used in this study where there are 7 independent inputs to the network, one hidden layer with 8 neurons and an output layer to predict one response at any

given time. In MATLAB, the training data set is split into three segments for training, validation, and verification. The amount of data allocated for each of these can be user-specified and in our work, we chose it to be 70%, 15%, and 15% respectively. The process of training a neural network model involves changing the weights and biases of the network to optimize the performance of the network. The common performance measure used in training feedforward neural networks is mean square error (MSE) between the predicted output of the 15% validation dataset and the respective observed values.

MSE = $\frac{1}{N}\sum_{i=0}^{N}(M_i - O_i)^2$ where N is the total number of data points, M_i is the model output and, O_i is the observed/target value.

There are many standard numerical optimization functions for training multilayer feedforward neural network models. The optimization methods mainly use either the gradient of network performance with respect to network weights or the Jacobian of network errors with respect to network weights (https://www2.cs.siu.edu/~rahimi/cs437/slides/nnet.pdf). Jacobian uses the gradient method to compute the final weights and biases. The Levenberg-Marquardt (LM) algorithm uses the Jacobian method to optimize network performance and is also the fastest algorithm available in MATLAB toolbox (Levenberg-Marquardt backpropagation, n.d.). The LM method was found to be more efficient and to have a high convergence rate when compared to the gradient algorithm (Hagan and Menhaj, 1994)

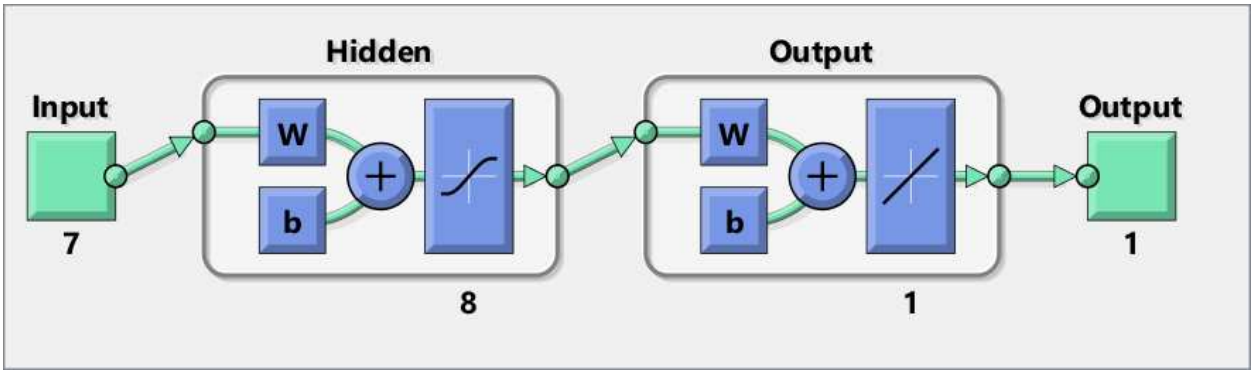

Figure 2.4. Example of an artificial neural network architecture from MATLAB 2018 toolbox

The performance of a Neural network model also depends on the number of neurons in the hidden layers, the number of hidden layers and the number of epochs and the initial weights assigned. Typically, these parameters are set beforehand. Epochs refer to the number of times the network weights and bias terms are updated to obtain the best performing model. To capture the effects of these parameters on the model performance multiple simulations were run while varying all the above-mentioned parameters. A general consensus is that by increasing the number on any of these parameters will eventually lead to overfitting of the model which trickles down to poor performance when validated against a new dataset (Sheela and Deepa, 2013). The effect of initial weights on the model performance was minimized by training each NN model with a set number of epochs, hidden layers and number of neurons, multiple times and picking the model with the best performance on the validation dataset.

In our work, we focus on both linear and ANN regression models to predict fuel consumption and emissions from on-road LDV's given a certain set of inputs. The models are developed/trained using the on-road experimental data obtained in this study. The performance of the model was validated using training dataset (train validation) to understand the model performance against the data it was trained with. Models were also validated with data from a different experiment (test validation) which were independent to the training dataset, but also identically distributed meaning performed on the same vehicle on the same route with the same driver.

2.2.3 Processing PEMS Data for Modeling

For all experiments, we performed quality assurance by checking for data gaps (e.g. since both vehicles had a manual transmission, gaps caused due to poor gear shifts leading to engine off) and outliers (e.g. NaN values caused when the engine shuts off during the experiment). PEMS software automatically corrects for any signal delays between the parameters and outputs the data as a second by second ASCII comma delimited text (csv). These files were read using MATLAB software as multiple row vectors with each vector representing one particular parameter from that experiment. All the row vectors from one

experiment were saved in a structure data-type (struct) that can store vectors of different data types in it. Each experiment had one struct file with 19 different row vectors such as velocity, rpm, acceleration, fuel consumption, CO, CO_2, NO_X, HC, PM, manifold air pressure, intake air temperature, time in seconds from start of the experiment, latitude, longitude, catalyst temperature, manifold air flow, altitude, flow in and flow out. One other parameter calculated was the limit of detection (LoD) of the PEMS device for different pollutants at every second since the device does not account for measurements performed at or below the LoD. LoD of the PEMS device depends on the lowest measurement range for each pollutant in milligrams per meter cubed, density of air at the experimental location and the intake air flow rate i.e. $LoD\ (mg\ s^{-1}) = X\ (mg\ m^{-3}) * \frac{flow\ rate\ (g\ s^{-1})}{\rho\ (g\ m^{-3})}$ where X is the pollutant being measured and ρ is the density of air. Tailpipe emissions from vehicles were a combination of highly nonlinear and time-dependent data, to develop models that can capture these trends struct files were also created by applying different transformations on data from all the experiments. Four different transformations were used on the experimental data and the performance of the models was assessed. The following are the transformations used in this study

- T1 - No transformation
- T2 - 7 and 13 second time averaged transformation - to capture the causality
- T3 - Logarithmic transformation - to capture the non-linearity in emissions
- T4 - Normalized transformation - to restrict the input data range between 0 and 1

2.2.4 Model Selection and Validation

The most important factor that drives the model performance in our modeling approach is the choice of independent input variables selected. The other factors along that drive the model performance are the choice of transformation on the training dataset and NN model architecture, i.e., the number of epochs, number of neurons and number of hidden layers. In our study, a variety of model types that varied in the number of input variables and types of input variables were used on all transformations of data and the

performance of each model was assessed to pick the best model for predicting fuel consumption and emissions independently.

Table 2.3 : Type of inputs used for different modeling schemes

Types	***Input Parameters***
M1	V
M2	VSP
M3	VSP, t
M4	V, VSP
M5	V, a
M6	V, VSP, t
M7	V, RPM, VSP
M8	V, a, t
M9	V, V^3, t
M10	V, V^3, a*V, t
M11	V, V^3, a*V, RPM, t
M12	V, V^3, a*V, V*RPM, t
M13	V, RPM, a, a*V, IAT, t
M14*	V, RPM, a, a*V, IAT, VSP, t

V=velocity, a=acceleration, V^3= velocity cubed, RPM=revolutions per minute, VSP=vehicle specific power, t=time since ignition, IAT=intake air temperature.

In table 2.3, **VSP** is the measure of load on a vehicle and is defined as power per unit mass to overcome inertial acceleration, rolling resistance, road grade, and aerodynamic drag (Frey et al. 2010).

$VSP = V\{a * (1 + \epsilon) + gr + gC_r\} + 0.5 * \rho * V^3 * (\frac{C_D A}{M})$ where ϵ is mass factor for rotational mass, g is acceleration due to gravity (m s^{-2}) , r is road grade , ρ is ambient air density (kg m^{-3}), C_r is rolling resistance (dimensionless), C_D is aerodynamic drag coefficient, A is vehicle frontal area (m^2) and m is vehicle mass (metric tons)

The different model input parameters were selected based on how each independent variable is physically related to the output variable and based on previous literature. These modeling types were used for both MV and NN models. The models were used to predict changes in fuel consumption or tailpipe emissions based on the changes in input parameters. A generic NN architecture was first set with 10 neurons, one hidden layer and 50 epochs and models were developed with the same architecture but with different transformations on the training dataset

- For no transformation, the training dataset was directly fed into the NN without any transformation
- For 7 and 13-second time-averaged transformation, training dataset was time averaged at 7 seconds and 13-second intervals before the modeling.
- If the logarithmic transformation was used, a natural log was applied to the training dataset. Once the model was built, the output was transformed back into the actual values using an inverse log transformation
- For normalized transformation, a z-transformation was used to convert the training dataset to have a mean of 0 and a standard deviation of 1 using the function $Z_i = \frac{(X_i - \bar{X})}{\sigma}$where Z_i is the respective Z score value for the respective observed X_i, $\bar{X}$ is the average of all observed X and σ is the standard deviation of the observed X values. The developed model produces a z-score output (Z_o) of the response variable which needs to be transformed back into the absolute output(Y_o) value using $Y_o = Z_o * \sigma + \bar{X}$.

Multiple models each having a different number and type of input variables were developed and the models with the highest performance were picked to predict both fuel consumption and emissions.

2.2.5 Error Metrics

Models were developed to predict on-road fuel consumption and emissions for a given set of independent variables. The performance of all the developed MV and NN models were measured using Relative Error (RE) and coefficient of determination (R^2).

1. Absolute Relative Error (RE)

$$RE = \frac{1}{N}\sum_{i=1}^{N} \left| \frac{M_i - O_i}{O_i} \right|$$

2. The coefficient of determination (R^2)

$$R^2 = 1 - \frac{SSE}{SSE + SSR}$$

$$SSE = \sum_{i=0}^{N}(O - M)^2 \; SSR = \sum_{i=0}^{N}(M - \bar{O})^2$$

M - Model output, O - Observed values, $\bar{O}$ - Mean of observed values, SSE - Sum of squared errors, SSR - Sum of squared residuals

RE helps in measuring model performance on a route integrated basis, i.e., it is the error percentage by which the model differs from the observed value in predicting either the total fuel consumed for the experimental route or total mass emissions of a particular pollutant for the given route. R^2 measures model performance in capturing the fuel consumption/emissions time trace on a per second basis for the given route. Data from one experiment was used as training dataset and data from a different experiment performed on the same vehicle on a different day was used as validation dataset. Two types of validation methods were used in this study, one is internal validation and the other is cross-validation. Internal validation is when the model performance is measured in terms of its prediction capability of the output variable from the same dataset as it was trained/built on, meaning the model is trained and validated using

the same dataset from the same experiment. Cross-validation is when the model is built using the data from one specific experiment but validated against a dataset from a different experiment. In our study, internal validation is referred to as training data and cross-validation is referred to as test data. Both these validations give different insights into our model performance. If a model performs well against the training data but performs poorly against test dataset it means that the model is overfitted. Our goal is to get a high performance against the test dataset which would make the model more versatile in predicting results from any experiments.

2.3 Case study

In our work, three routes between two origin-destination pairs each were identified and driven on a 2004 gasoline LDV with a PEMS device onboard that was used to collect vehicles' engine data through the OBD-II port. Routes used for the case study comprised of routes with highway driving and urban driving and varied widely in maximum and average velocity, total distance and total time when compared to the experimental route. Models developed using the data from gasoline LDV experiments were used to predict fuel consumption of our case study routes. By comparing the fuel consumed on different routes for each origin-destination pair we can identify the route that results in the least fuel consumed and the associated time penalty if there was any. The primary purpose of this case study was to understand whether there is a significant difference in fuel consumed based on the choice of route.

3. Results

3.1 Experimental Results

Raw data from PEMS were processed and compared against standard emissions test cycles and against EPA emission factor for Tier-1 and Tier-2 standards. We describe the experimental results and how they vary between gasoline and diesel vehicle experiments before going into the modeling results.

3.1.1 Drive Cycle Comparisons

In our study, we chose the EPA Federal Test procedure - 75 (FTP-75) as a base to develop our experimental route. Figure 3.1 shows the metric comparison between FTP-75 and the experimental data for all the experiments. Both gasoline and diesel vehicle experimental data compare well for maximum velocity, average velocity, the standard deviation of velocity, total distance, stop time and, vehicle specific power not only within the experiments but also against the FTP-75 cycle. While the maximum acceleration from the FTP-75 cycle is lower than that of all the experimental routes, the average acceleration compares well with the FTP-75 cycle showing that our developed drive cycle is realistic to the existing emissions test cycles.

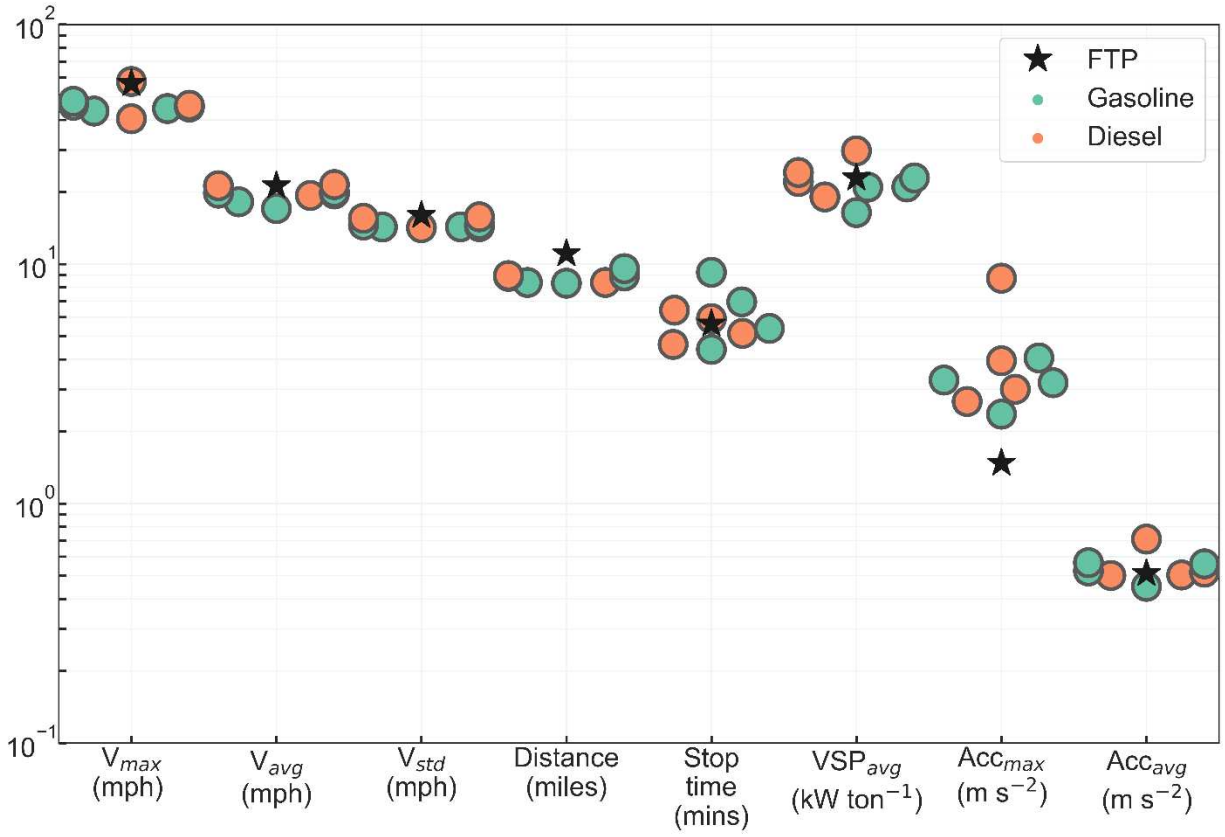

Figure 3.1: Route metrics comparison between FTP-75 and experimental data

3.1.2 Emission Factor Comparisons

Tailpipe emissions from on-road vehicles are subject to stringent standards based on grams of pollutant emitted per mile driven. Tier-2 emission standards imposed by EPA regulate the following tailpipe pollutants: carbon monoxide (CO), oxides of nitrogen (NO_X), hydrocarbons (HC) and particulate matter (PM). These pollutants emitted during our experiments on diesel and gasoline vehicles that were driven on the developed drive cycle were compared with the EPA standards. Figure 3.2 compares on-road CO, NO_X, and HC emissions from all experiments with their respective standards. The PM measurements from all gasoline vehicle experiments were below the limit of detection of PEMS and hence were not used in our study. Tier-2 standards do not regulate CO_2 emissions from tailpipe directly (Epa and OAR 2016) and hence the plot only shows the real-world emissions of CO_2 from both the test vehicles.

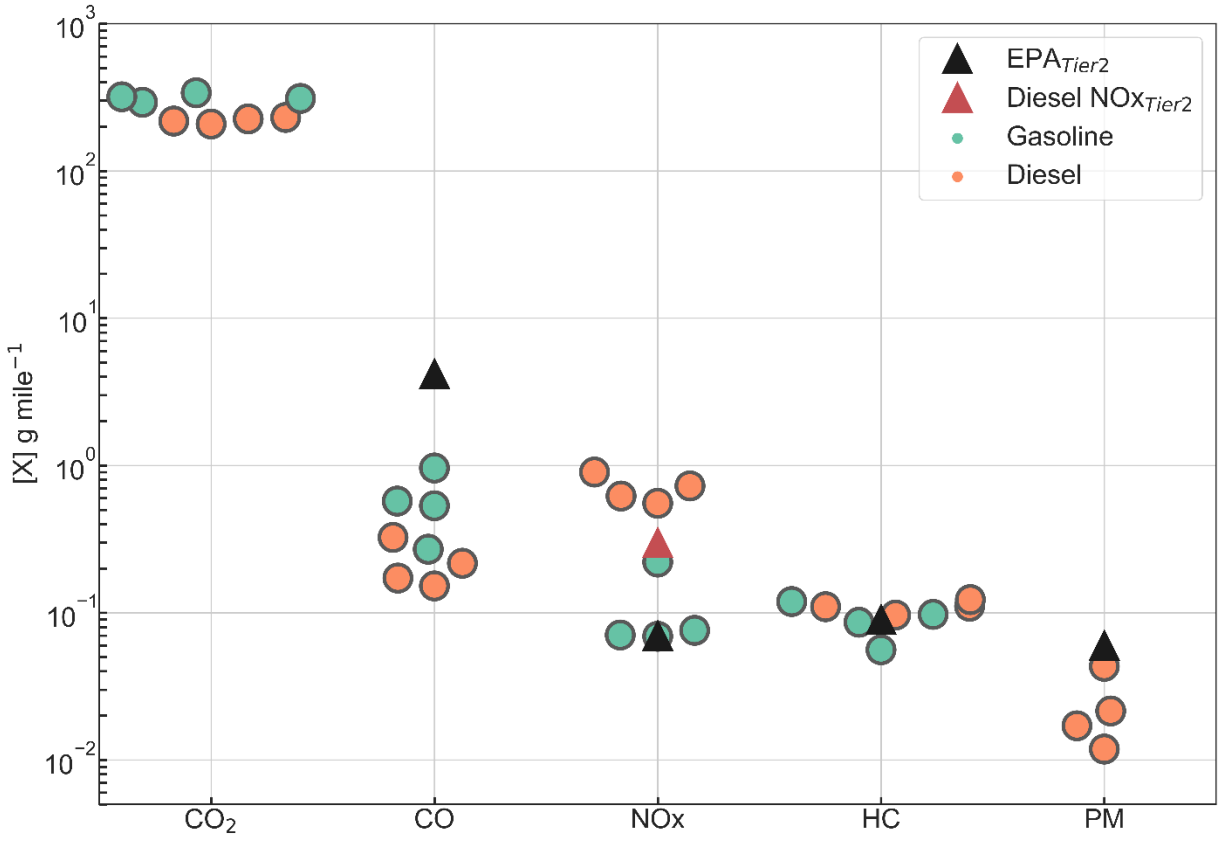

Figure 3.2: Comparison of real-world emission factors obtained from this study with the US EPA emission standards

CO_2 emissions from gasoline vehicles have an average of 315 g.mile^{-1} and are slightly higher than emissions from diesel vehicles which have an average of 220 g.mile^{-1}. CO from gasoline vehicles have an average value that is two times higher than that of diesel vehicles, but both these vehicles fall well within the permissible standards. Tier-2 standards for NO_X emissions from diesel vehicles illustrated with a red marker in the plot are less stringent at 0.3 g.mile^{-1} (bin 9) when compared to that of gasoline vehicles which are at 0.07 g.mile^{-1} (bin 3)(Emission Standards: USA: Cars) . NO_X from gasoline vehicles compared closely with the imposed standards except for one outlier. NO_X emissions from diesel vehicles had a much higher average value of 0.7 g.mile^{-1} making on-road emissions greater than a factor of two when compared to the standard. HC emissions from both the vehicles on an average were only slightly higher when compared with the imposed standards and PM emissions from diesel vehicles were within

the Tier-2 standard for diesel vehicles. On-road emissions from all the gasoline experiments were comparable to the Tier-2 emission standards for all pollutants except for one outlier in NO_X emissions for one experiment. On-road emissions from diesel engines were within the standards for all pollutants except NO_X.

3.1.3 Cold phase and Hot phase

In our study, emissions from the gasoline vehicle where much higher within the first eight minutes of operation. Cold phase in vehicles primarily refers to when there is a difference in temperature from regular operating conditions (Reiter and Kockelman, 2016). Emissions from the LDV's during the cold phase are an important source of NO_X, HC, PM and CO. During the first minutes of vehicle operation when the engine block and coolant temperature are low, incomplete combustion paired with low catalyst temperature results in significantly higher emissions than at nominal operating conditions (Cao, 2007). Low ambient temperatures also result in higher cold start emissions (Reiter and Kockelman, 2016) (Weilenmann et al., 2005). The emissions from diesel and gasoline vehicles are greatly reduced by catalysts during the hot phase when the catalyst reaches its normal operating temperature. In our work, we observe the effect of cold phase and hot phase emissions for various pollutants for both gasoline and diesel vehicle experiments and arbitrarily picked the first 8 minutes of vehicle operation to be in the cold phase. Figure 3.3 illustrates the normalized cumulative emissions of CO_2, CO, NO_X, HC and, PM (diesel) from engine start until the end of the experiment for a representative gasoline and diesel vehicle experiment. For gasoline vehicle experiments the effect of cold phase is very prominent with more than 80% of CO, HC and NO_X emissions being limited to the cold phase. We also observed that CO_2 emissions from both gasoline and diesel vehicle experiments were linear across the experiment. In diesel vehicle experiments, the cold phase has a significant effect in CO emissions with more than 60% of emissions within the cold phase while the other pollutants are not affected by the phase of the vehicle. A study by (Weilenmann et al., 2005) also shows that cold phase emissions are significantly lower for diesel vehicles than gasoline vehicles.

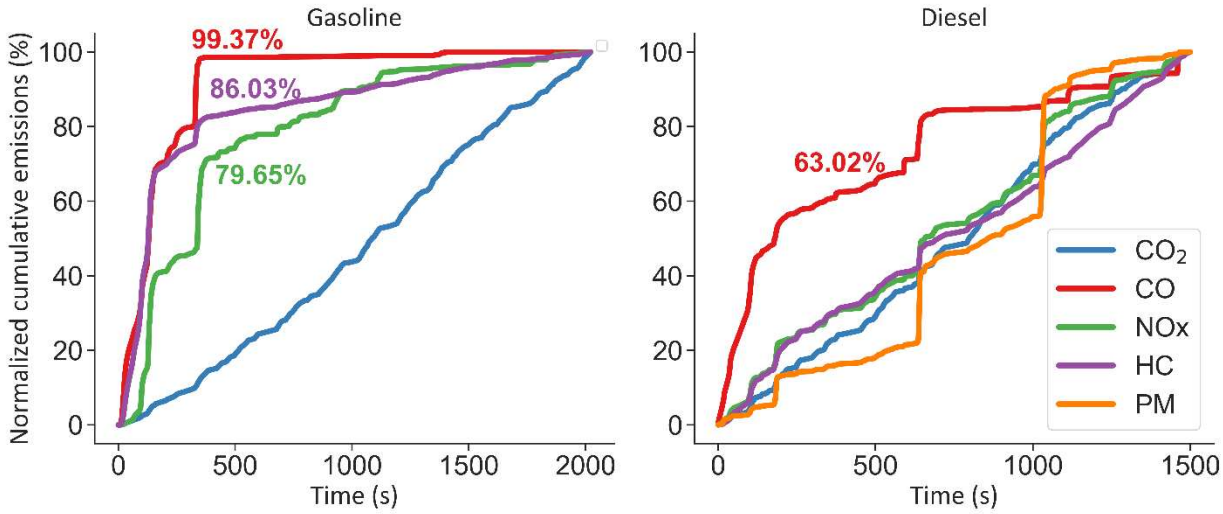

Figure 3.3: Normalized cumulative emissions from the start of experiment for a representative gasoline and diesel vehicle experiment illustrating the effects of cold phase.

3.1.4 EMFAC Model comparison

In our work, vehicle specific ANN models were developed using real-world emissions and fuel consumption data. Unlike the ANN models developed in this study, existing emissions inventory models such as EMFAC are developed based on emissions from chassis dynamometer tests. Figure 3.4 compares the results of median grams per mile emissions with respect to binned velocity from all gasoline experiments in our study to the values generated by the EMFAC model for the pollutants NO_X, HC, and CO. PM emissions were not measured for the gasoline vehicle experiments since typical PM values were below the PEMS detection limit. EMFAC model results were generated as California statewide emission rates for the annual calendar year 2017 for a 2008 model year light-duty gasoline vehicle for velocities 0 to 50 mph. For both NO_X and HC emissions, we can clearly see that for lower velocity bins, on-road emissions were much higher than those of the EMFAC models and at the higher velocity bins, both observed and EMFAC model results compared well. This trend suggests that the EMFAC model may underestimate the on-road vehicular NO_X and HC emissions at lower speeds. On the other hand, observed CO emissions from on-road vehicles are lower than that of the EMFAC model predictions except at the lowest velocity bin. Higher emissions in the lower velocity bins are also a result of higher cold phase emissions which were significant during lower vehicle velocities.

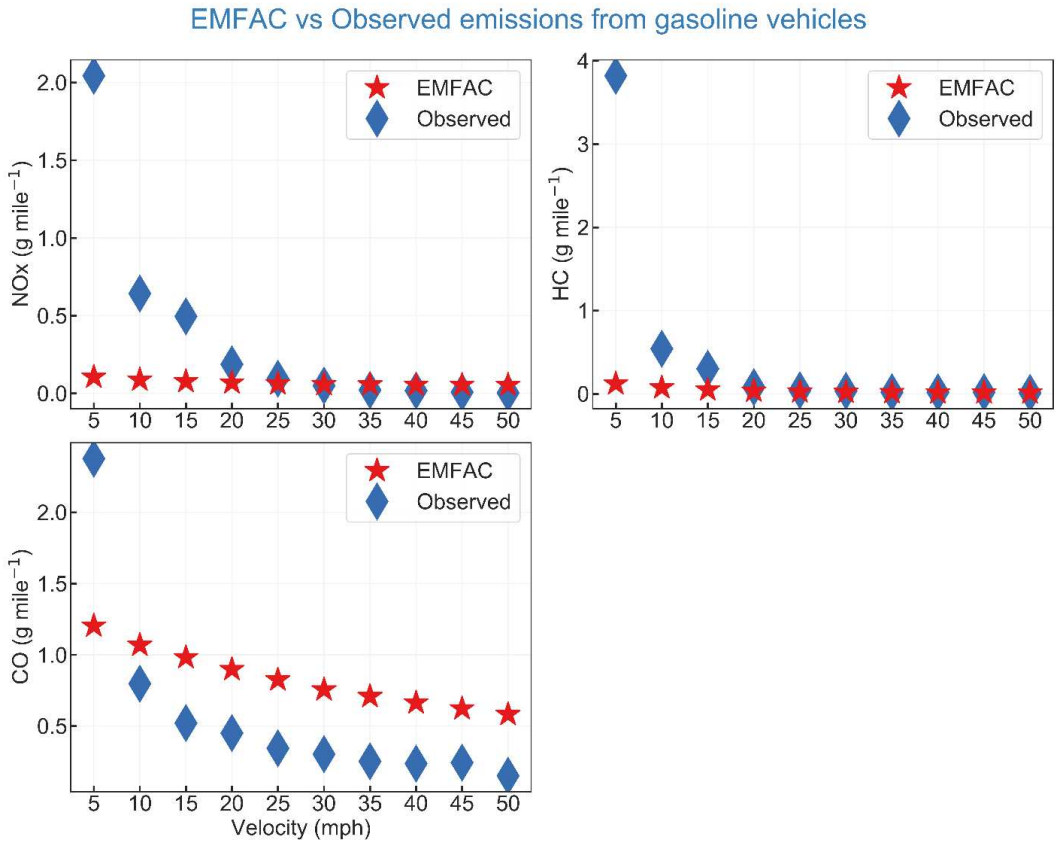

Figure 3.4: Comparison of velocity binned emission factors generated by EMFAC model for calendar year 2017 for a 2008 model year light-duty gasoline vehicle with the median observed emissions from gasoline vehicle experiments in our study.

Like gasoline vehicles, velocity binned median NO_X, HC, CO, and PM emission factors from all the diesel vehicle experiments were also compared to the EMFAC model results in Figure 3.5. EMFAC model results were generated as California statewide emission rates for the annual calendar year 2017 for a 2003 model year light-duty diesel vehicle for velocities 0 to 50 mph. We can see that diesel NO_X emission rates are much higher than that of EMFAC emission rates at lower bin velocities of 5,10 and 15 mph.

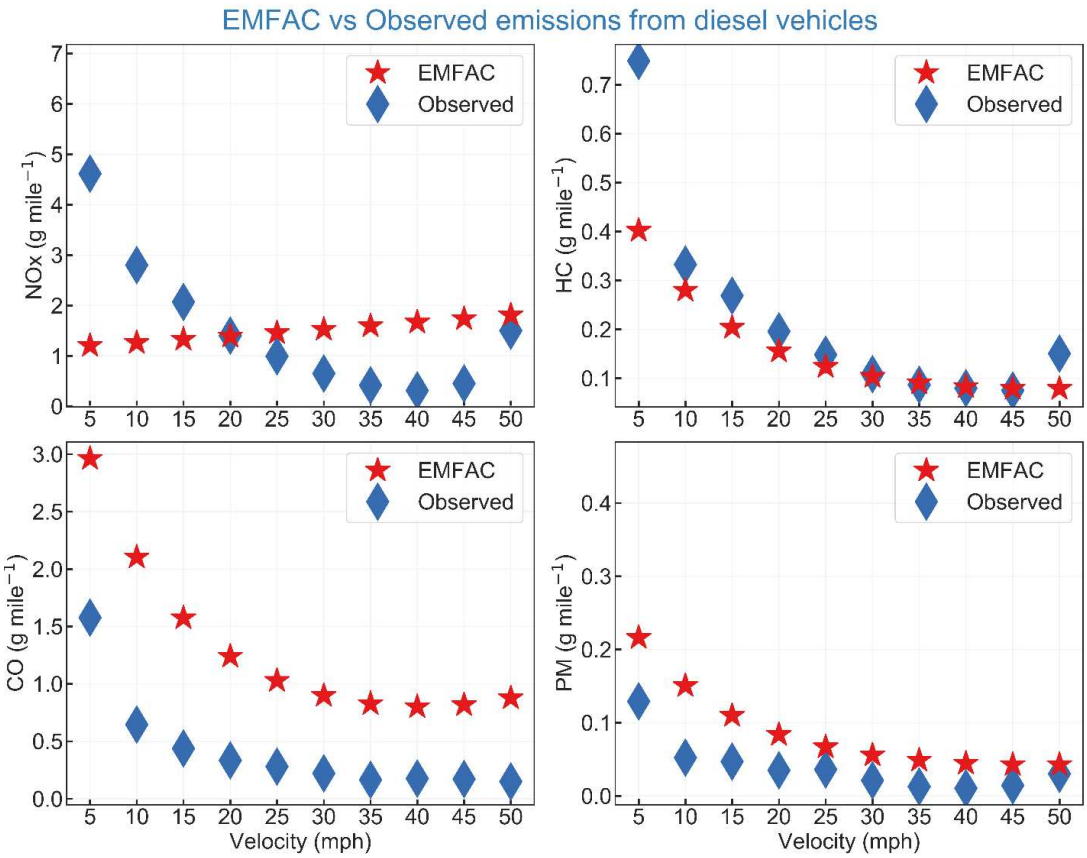

Figure 3.4: Comparison of velocity binned emission factors generated by EMFAC model for calendar year 2017 for a 2003 model year light-duty diesel vehicle with the median observed emissions from gasoline vehicle experiments in our study.

At higher velocities, the median of observed NO_X emission rates is lower than that of EMFAC emission rates from which we can say that in urban driving conditions where the vehicle velocities are within 20 mph the EMFAC model underpredicts on-road NO_X emissions from diesel (Chossière, 2017), (Wang et al., 2016). On the other hand, HC emission rates from both the EMFAC model and observed experimental values compare well across all velocity bins except at the lowest velocity bin. Finally, for both CO and PM, EMFAC model over-predicts the on-road emission rates from diesel vehicles across all velocities. Developing vehicle specific fuel consumption and emissions models can provide better estimates of on-road emissions.

3.2 Model performance

Performance of the model was highly dependent on the type of modeling used and, on the transformation, type used. In our study, we first tested out different modeling types that resulted in the best performance and later this model was used to test the effects of different transformations.

3.2.1 Parameter selection for models

In our work, data from multiple experiments were available to develop models to predict FC and emissions from both gasoline and diesel vehicles. Although multiple datasets existed to train models, only one experiment was used to develop a model at any given time. For example, models built to predict gasoline FC used a dataset from one experiment from gasoline experiment set to train the model and data from a different experiment from the same gasoline set was used to validate the model. A similar process was used to build diesel FC, NO_X and HC. One important aspect to look into is the effect of the choice of training dataset on the model performance. Although all experiments were conducted on one single set experimental route, the day to day traffic and the timing of the experiments result in a highly diverge time series of both fuel consumption and emissions, these diverse time series will lead to change in model performance based on the type of training dataset used to develop the model in the first place. For example, a model that is built using data from an experiment that had many stops in the first few hundred seconds from the start of the experiment will have frequent stops and acceleration events in that time period. If this model is used to validate fuel consumption or emissions from an experiment that have a similar behavior, it would result in a higher performance and similarly if this model is validated against a different experiment which has very few stops in the initial phase of the experiment, it might provide a very different result in terms of performance. To understand the effects of training/validation dataset on model performance, models were developed and validated using all possible combinations of training and validation datasets for each of the output variables. The change in model performance was studied for both MV and NN models for all modeling types defined in table 2.3 for each independent output variable with no transformation on the training dataset.

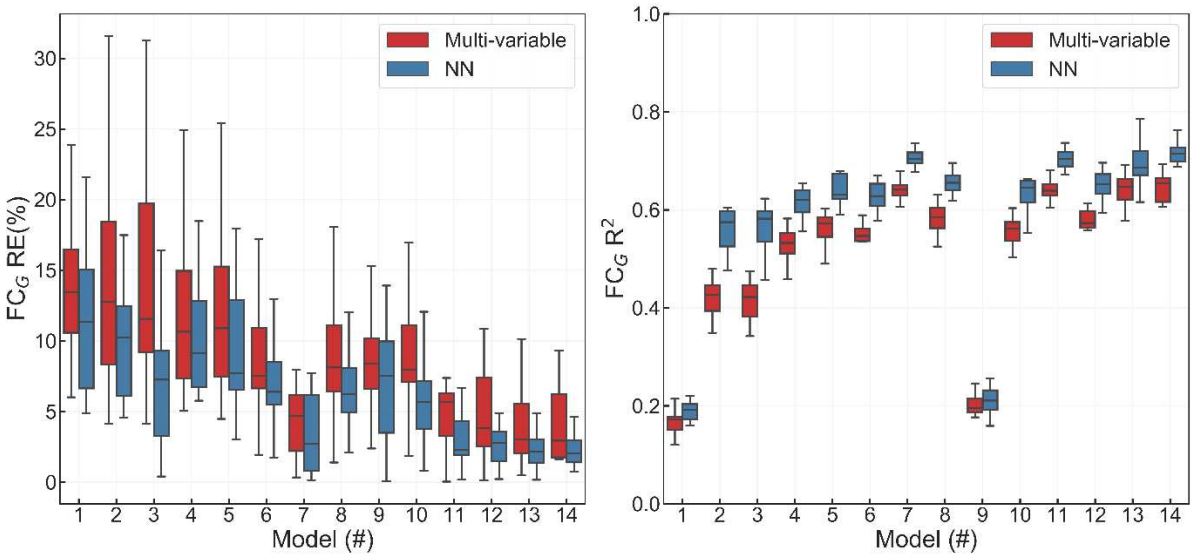

Figure 3.5: Performance effect due to the choice of input parameters on the Multi-Variable and Neural-Network models developed to predict fuel consumption in gasoline vehicles. The boxplots show the variability in RE and R^2 for each model due to the choice of training and validation datasets.

Figure 3.5 shows the trends of RE and R^2 values for both MV and NN models developed to predict FC from gasoline vehicles. The boxplots represent the variability in RE and R^2 based on the choice of training and validation dataset used in modeling. The route integrated model performance is measured using RE and the per second prediction capability of the model is measured using the R^2 .(models are trained and validated using different datasets). In the FC models, apart from few outliers, the RE decreases as we go from model 1 through 14, with model 14 having the least median relative error in predicting fuel consumption. Model 14 parameters, when used with NN models without any transformation, has a median R^2 of 0.72 and median relative error of ~2% meaning if the route integrated fuel consumption of a vehicle was 1 gallon on a particular drive cycle then the model would predict fuel consumed within 0.98 to 1.02 gallons.

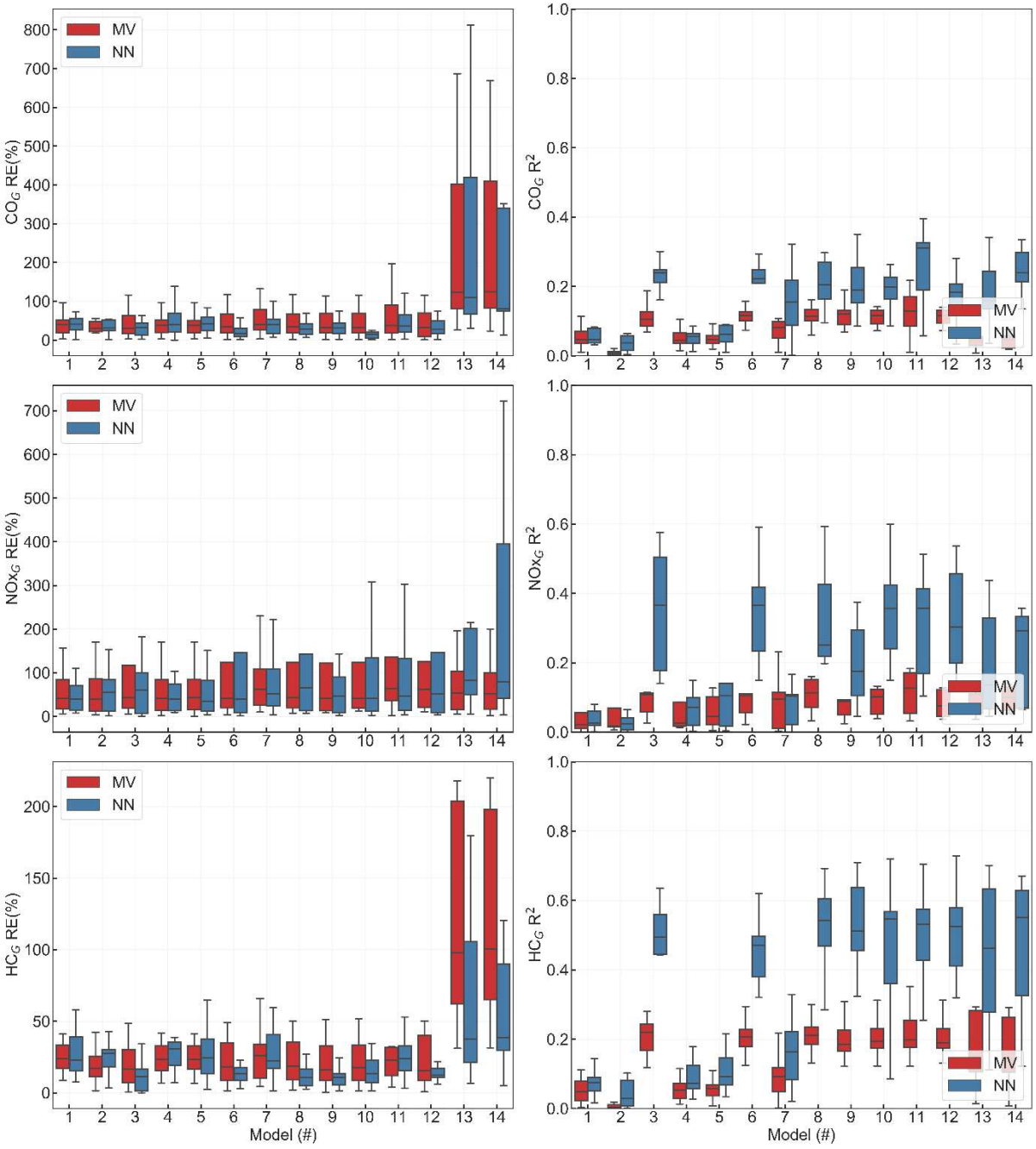

Figure 3.6: Performance effect due to the choice of input parameters on the Multi-Variable and Neural-Network models developed to predict emissions from gasoline vehicles.

R^2 values reflect the model performance in predicting the time series of FC, these values vary from 0 to 1 and a higher value corresponds to a better match in time series predictions. For all MV and NN models, the median R^2 rapidly increase as we go from model 1 through 5 but do not increase significantly between

models 11 through 14 indicating that although there is a decreased RE value associated with model 14, it doesn't necessarily indicate better performance in time-series predictions. Figure 3.6 represents the trends of RE and R^2 values for both MV and NN models developed to predict on-road emissions from gasoline vehicles. Unlike the models developed for predicting fuel consumption, the models developed to predict emissions from gasoline vehicles had high relative errors and poor R^2 values. In the models developed to predict CO emissions, both MV and NN models had similar RE for each modeling type and unlike FC models were RE had a decreasing trend as we went from model 1 through 14, here model 13 and 14 have the highest median RE values at 103% and 76% respectively suggesting that the model with more number of inputs does not necessarily improve the model performance and that the type of inputs used in the model need to be correlated to the response variable. While model 10 has a modest median RE at 16%, it has a corresponding R^2 value of 0.2 suggesting that although the models route integrated prediction capability is high, it fails in capturing the time series of CO emissions. The models developed predict NO_X emissions from gasoline vehicles have no significant improvement as we go from model 1 through 14 with all the models having high median RE of >35% and R^2 of <0.4 resulting in a poor route integrated and time series prediction capabilities. In the models developed to predict HC emissions from gasoline vehicles, we did not find any visual trends in performance change as we go from model 1 through 14. Among all the models, model 12 had a least median RE of ~12% but with a corresponding median R^2 of 0.5 which results in good route integrated prediction but a poor time series prediction. Overall, for all the models developed to predict emissions from on-road gasoline vehicles, the performance was poor with high median RE values and corresponding low median R^2 values. One primary reason for this is due to the huge difference in magnitude in the emissions during the cold phase and hot phase of vehicle operation. Since we know that more than 80% of total emissions from gasoline vehicles were associated with the cold phase, only this data (i.e. first 8 minutes of every experiment) was later modeled and it was found that the model performance remained poor with no significant improvement in predictions, this could be attributed to the reduced sample size that is involved in training the model. For cold phase NO_X emissions from gasoline vehicles, all models resulted in a relative error of

more than 40%. RE for NN models developed to predict cold phase CO and HC emissions were lower when compared with that of NO_X at an average RE of 34% and 22% respectively, but this alone does not guarantee better performance. Respective R^2 values should also be higher to ensure that this performance is not just attributed to the choice of training and test dataset.

Comparing the modeling results between MV and NN models, we see that NN models in most cases have a lower RE and a higher R^2 value when compared to the MV models regardless of the response variable being modeled. For FC models, where NN model 14 had the highest performance with median RE at ~2% and R^2 at 0.72, the corresponding MV model has a RE of 3% and a lower R^2 of 0.65 suggesting that MV model has a poor correlation with time series when compared to the developed NN model. Unlike gasoline FC models, in the models developed to predict emissions from gasoline vehicles both NN and MV models had high REs and low R^2 values suggesting poor prediction capabilities of the developed models regardless of the type of inputs used.

For diesel vehicle experiments, FC, NO_X, HC and, PM was modeled. The pollutants from diesel vehicles in our experiments had no influence of cold starts from vehicles. Like gasoline, all input types with no transformation to the training data were used to develop models for each of the response variables using both MV and NN. Figure 3.7 represents the trends in RE and R^2 for both MV and NN models when developed using different model configurations from table 2.3 to predict FC and NO_X from diesel vehicles. Models were trained with different training datasets and the effect of these various training and validation dataset is illustrated as boxplots. We can see that for models developed to predict diesel FC, NN model 14 has the least median RE of 0.7% with a very high R^2 of 0.81 suggesting a very high route integrated performance as well as highly correlated time series prediction. Although RE values from both MV and NN models do not show a visual trend as we go from model 1 through 14, R^2 value has a slowly increasing trend as we go from model 1 through 14 except for one outlier. We can see a similar trend in models developed to predict NO_X emissions from diesel vehicle where NN model 14 had a low median

RE of 1% and a high median R^2 of 0.8. A median RE of 1% indicates that if the route integrated NO_X emissions consumed on a particular drive cycle was 1000 grams then the model would output a route integrated NO_X value between 990 and 1010 grams.

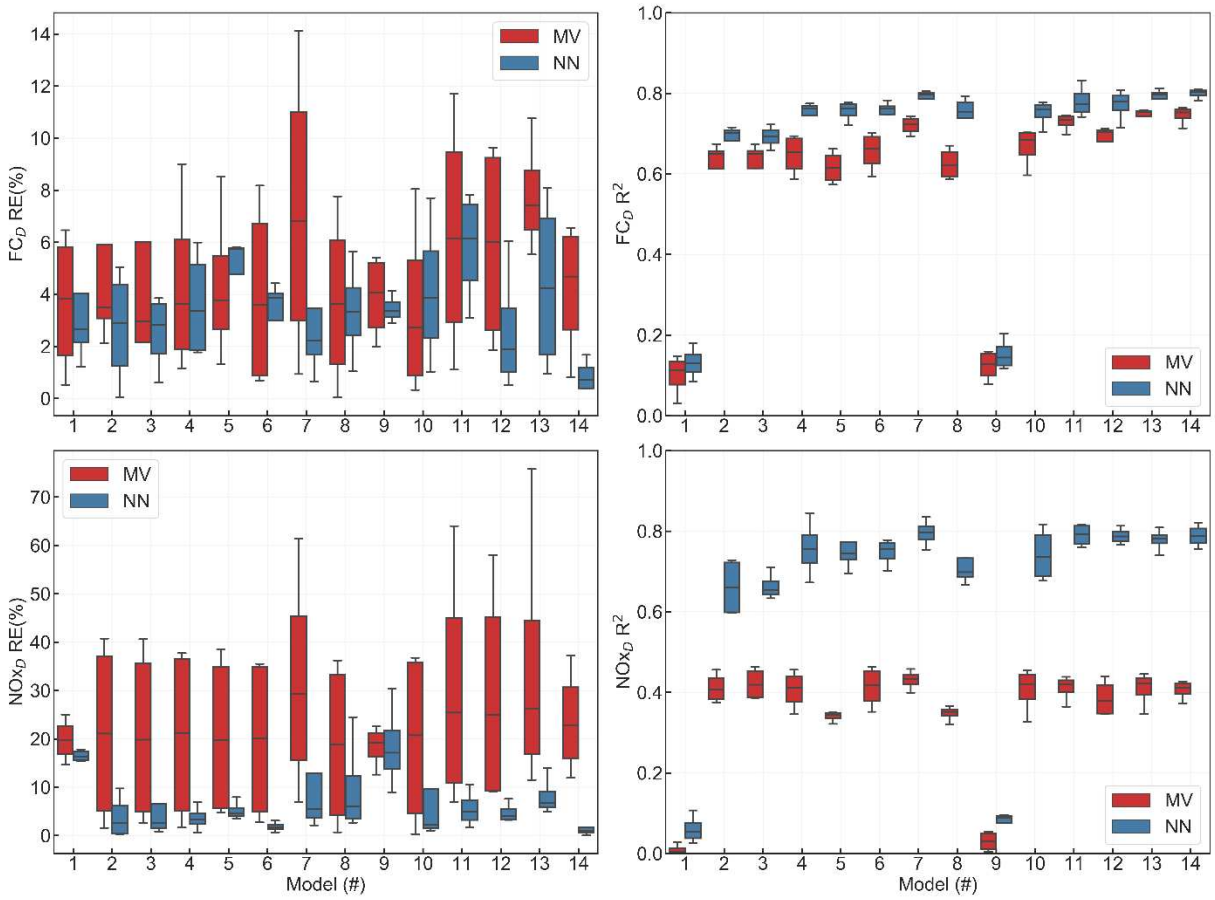

Figure 3.7: Performance effect due to the choice of input parameters on the Multi-Variable and Neural-Network models developed to predict fuel consumption and NO_X *emissions from diesel vehicles. The boxplots show the variability in RE and* R^2 *for each model due to the choice of training and validation datasets.*

The performance trends of different models developed to predict diesel CO, HC and NO_X emissions are shown as boxplots of RE and R^2 values obtained by using different training and validations datasets for all the modeling types in figure 3.8. Similar to gasoline vehicle CO models, all models developed to predict CO form diesel vehicles had high RE values and very low R^2 values. In the figure, we can see that NN model 6 has a low relative error of 2.8% but the corresponding R^2 value for the same mode was less than

0.3 suggesting a poor time series correlation. The highest median R^2 value among the CO models for NN model 11 at 0.4 with a corresponding median RE of 17%, combined these will result in a model with poor prediction capabilities. The poor performance in predicting CO emissions from diesel vehicles can be attributed to the failure in capturing the cold phase CO emissions which were dominant in the diesel vehicles as well as gasoline vehicles. Unlike CO models, models developed to predict HC emissions from diesel vehicles had lower RE and higher R^2 values. The most comprehensive NN model with the highest number of inputs which was model 14, had the least median RE of 16% and a high median R^2 value of 0.67. Although NN model 6 has a lower median RE of 4.2%, it is associated with a lower median R^2 of 0.49 which will result in a less correlated time series when compared to model 14. For the diesel PM models, all models developed using both MV and NN model had high REs and very poor R^2 values, with a maximum R^2 of 0.3. The low R^2 value will result in a poor time series correlation and will result in poor prediction when applied on random drive cycles. The poor performance of models in predicting PM can be attributed to the high number of PM emissions from diesel vehicles which were below the limit of detection of PEMS and were assumed to be at the limit of detection, which can result in reducing the correlation of the inputs with the PM emitted.

Comparing the model performance within MV and NN models for diesel vehicle models we can see similar trends to that from gasoline vehicle models. NN model 14 developed to predict FC has a median RE of 0.7% and an R^2 of 0.81 while the MV model had a RE of 4.9% and an R^2 of 0.75. Similarly, NN model 14 developed to predict NO_X had a higher performance (RE - 1 %, R^2 - 0.8) when compared to the MV model 14 (RE - 22 %, R^2 - 0.4) and a similar trend is observed for the best HC model as well. In all the best models picked for predicting diesel FC, NO_X and, HC emissions, NN models had lower median RE and higher median R^2 values when compared to the MV models.

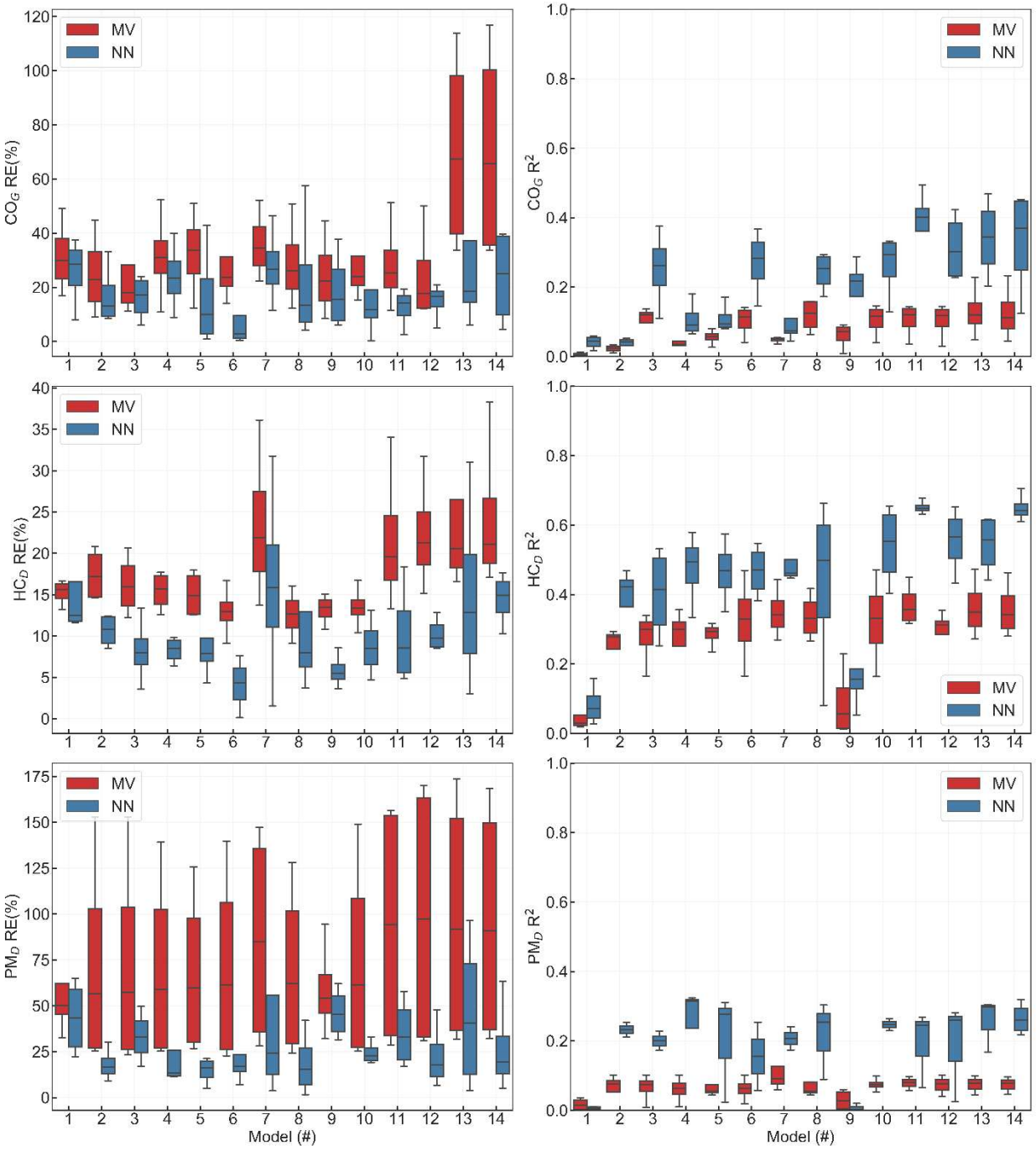

Figure 3.8: Performance effect due to the choice of input parameters on the Multi-Variable and Neural-Network models developed to predict CO, HC and PM from diesel vehicles. The boxplots show the variability in RE and R^2 for each model due to the choice of training and validation datasets.

Overall, the models developed to predict CO, NO_X and, HC emissions from gasoline vehicles, and CO and PM emissions from diesel vehicle experiments had poor performance. Hence, only the models

developed to predict FC from gasoline vehicles and FC, NO_X, and HC from diesel vehicles were more closely looked at in our study.

3.2.2 Effect of input data transformation

Model performance depends on the relationship of inputs and the output variable. The input dataset was transformed in different ways in an effort to improve model performance by targeting a few key issues that can improve the correlation between the input variable and the output variable. Different transformations used in our study were time averaging, log transformation and normal transformation to reduce causality, nonlinearity and restrict the span of all variables during modeling between -1 and 1. Transformations were only used on the NN model as these models performed better than MV and Hybrid models

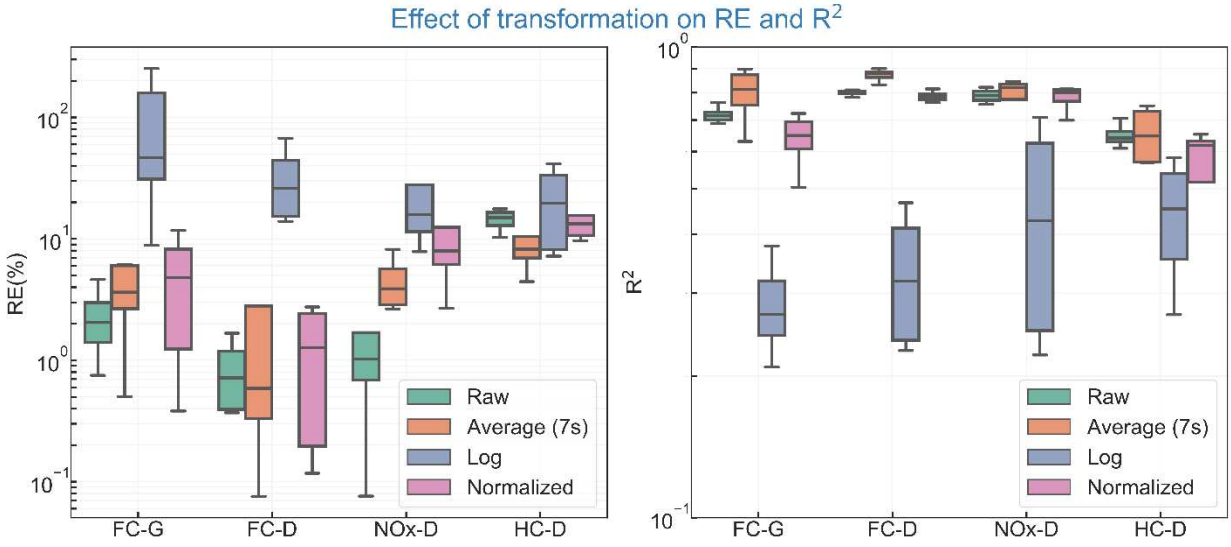

Figure 3.9: Effect of transformation on the performance of Neural-Network models generated to predict FC for gasoline engines, FC for diesel engines, NO_X *for diesel engines and HC for diesel engines*

For NN models the RE and R^2 varied based on the type of transformation used on input parameters. Different transformations were used to look at the effect of model performance in predicting FC from gasoline and diesel engines as well as NO_X and HC from diesel engines. For gasoline FC, RE is the least without any transformation on the dataset followed by normalization and averaging. Log transformation

of the data resulted in the highest error. R^2 values are comparable for raw data and normalization but the averaged dataset has the highest R^2 value among all of them.

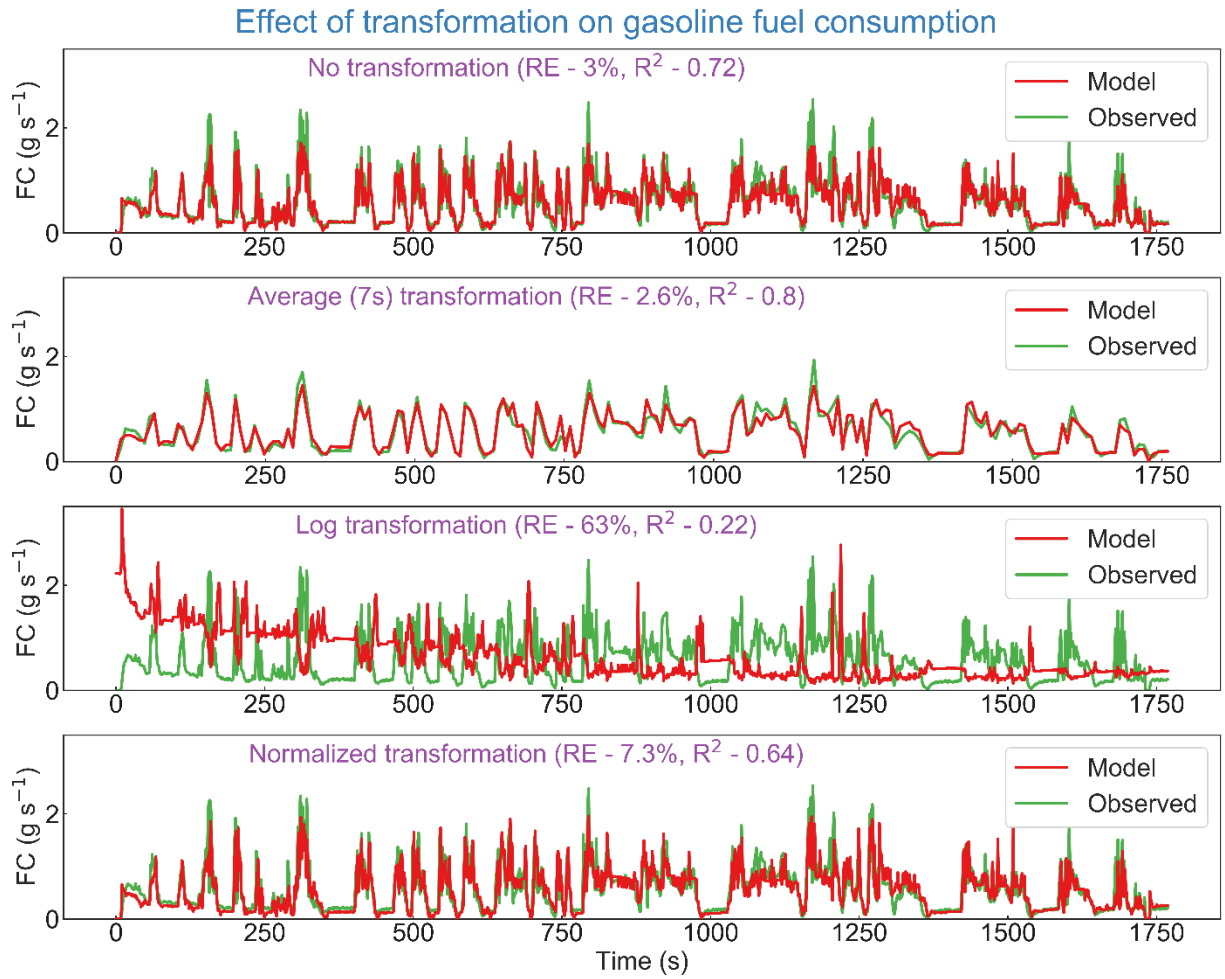

Figure 3.10: Effect of training dataset transformation in predicting gasoline fuel consumption.

Although averaged transformation yields an excellent median R^2 value of 0.9, it is at a cost of lower time resolution. Figure 3.10 illustrates the model predicted gasoline FC time series as compared to observed values as a result of different transformations. The model was trained using a dataset from one experiment and validated against dataset from a different experiment. Models with no transformation and Normalized transformation predict the time-series of fuel consumption with great accuracy whereas log transformation of Input data gives us a rather skewed result and overpredicts when compared to the observed values. In our work, averaged transformation is when the model training dataset is arbitrarily

averaged for every 7 seconds and is used to predict output variables as a 7-second averaged time series. This approach results in a very high correlation between the model outputs and the observed values but also comes at a cost of reduced time series resolution.

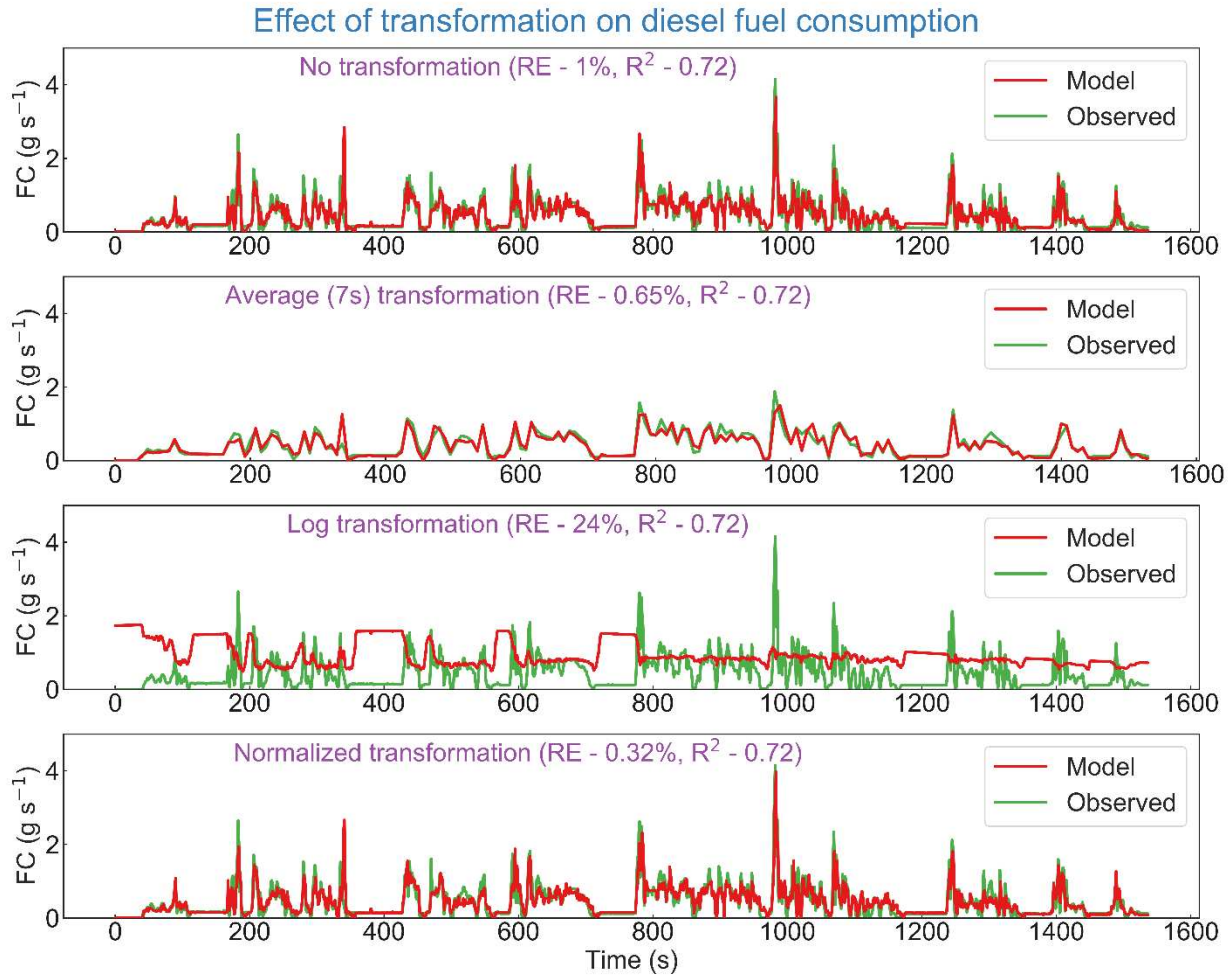

Figure 3.11: Effect of training dataset transformation in predicting diesel fuel consumption

Diesel FC predictions have a slightly different trend when compared to gasoline FC. When a normalized transformation, which transforms each variable in the dataset with a mean of 0 and standard deviation of 1, was applied on the training dataset and validation datasets resulted in a low RE and a slightly poor R^2 in predicting FC from diesel vehicle. However, normalization has a high dependence on the type of training dataset used as the mean and standard deviation of FC from the training dataset are used to

transform the normalized output of the NN model to respective fuel consumed per second. R^2 for averaged transformation is significantly higher for all output pollutants in diesel vehicle experiments but they lack the time resolution provided by other models. Log transformation is inefficient for FC for both gasoline and diesel which can be a result of linear trends in FC for both vehicle types. Figure 3.11 shows the time series trends in predicting diesel fuel consumption based on different transformations. Log transformation has visually the least correlation between the model predicted values and the observed values. Training dataset with no transformation and normalized training dataset yield similar results in terms of time-series with little to no visual difference. Similar to gasoline FC, diesel FC when modeled using averaged training dataset to predict average FC results in a very high correlation between observed and modeled results with low RE but has a lower resolution in time series.

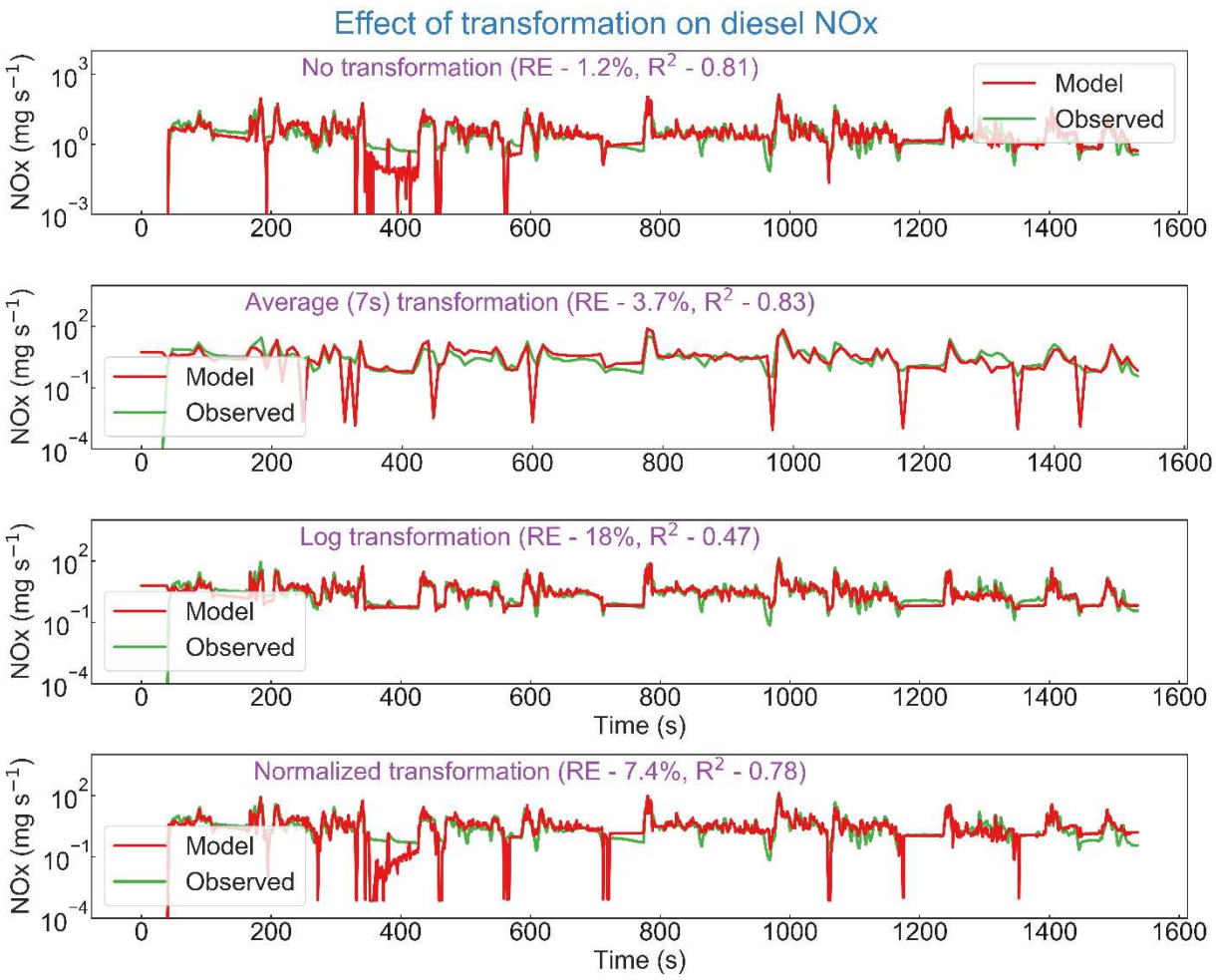

Figure 3.12: Effect of training dataset transformation in predicting diesel NO_X

Emissions from diesel vehicles have a wide range in magnitude and vary non-linearly throughout the experiment when compared to linear fuel consumption trends. They are heavily impacted by the vehicles instantaneous velocity and acceleration and result in higher grams of emission per second when there is a higher load on the vehicle. One main difference between FC and NO_X modeling based on the type of transformation is the significant improvement in performance when a log transformation is used on the training dataset to predict NO_X emissions. In figure 3.12 we can see that model built by log transformation of the training dataset to predict NO_X visually compares indifferently to those models developed by either normalization or no transformation and both the RE and R^2 values obtained using log transformation are also comparable to all other transformations' types used. Among all the

transformations, normalization has the least RE of 0.15% and an R^2 of 0.73 while the model with no transformation has a RE of 2% and a slightly higher R^2 of 0.76. Since normalization transformations' performance is highly dependent on the training dataset used to build the model, the model built using no transformation were used to model NO_X emissions from diesel vehicles.

Similar to NO_X emissions, HC from diesel vehicles have a high nonlinearity associated with the rapid acceleration and velocity events of the drive cycle. Figure 3.13 compares the variation in performance of M14 when used to predict HC from diesel vehicles based on the type of transformation used. When log transformation is used on the training dataset to predict HC results, we can see that the model performance is very similar to that obtained with any other type of transformation. Models with no transformation on training dataset tend to perform better in both RE and R^2 domain. Overall, for artificial neural network models, the transformation of model training dataset did not have any significant improvement in model performance. We observed that NN models perform really well as long as the training data or the model inputs used are a good physical representation of the dependent variable being predicted. An appropriate sample size, a suitable number of input variables and dataset that is free of outliers and signal delay, resulted in best performing models for predicting both fuel consumption and emissions from the test vehicles. Since transformation did not lead to consistent improvement in performance, all models were built using only the training dataset without any transformation.

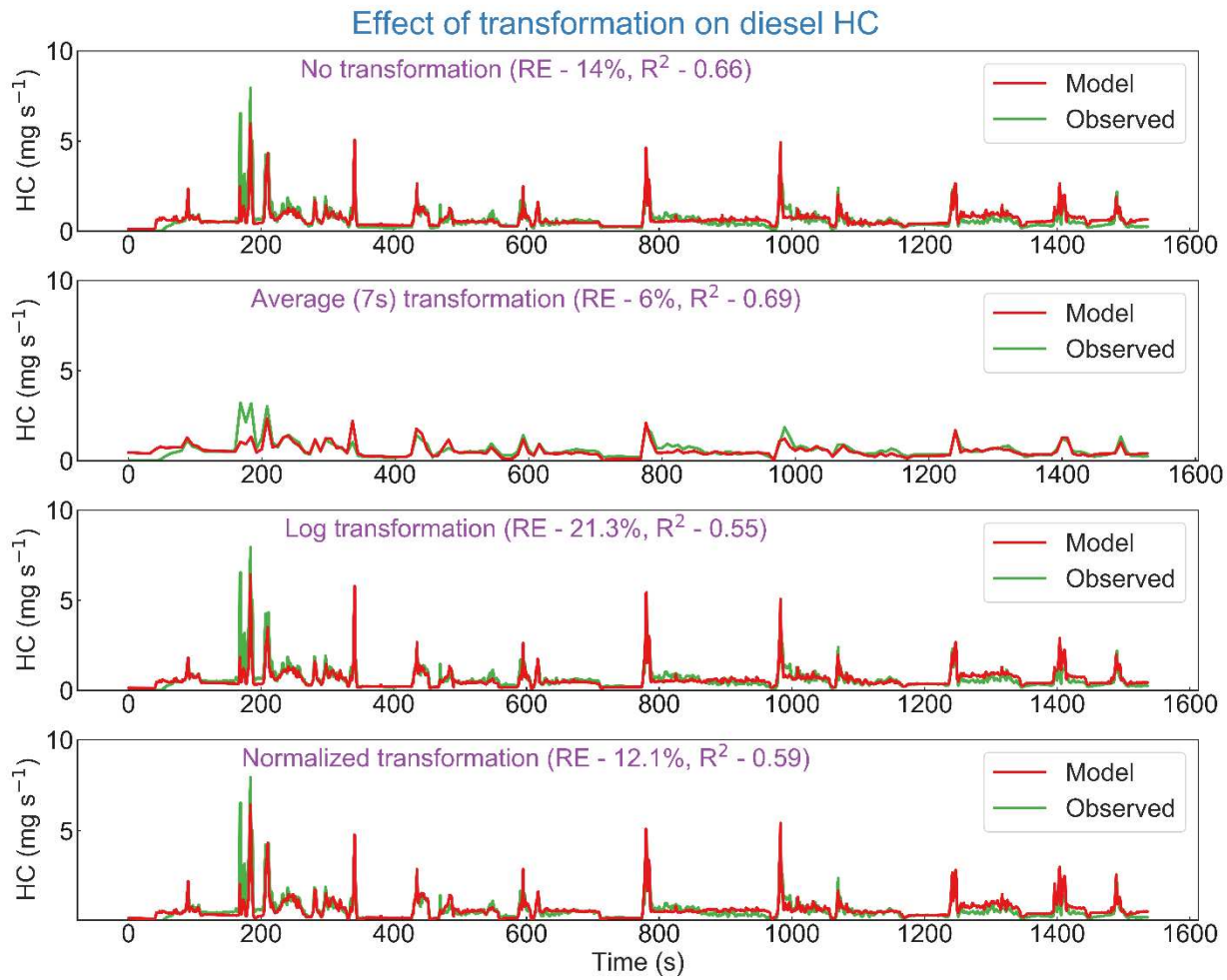

Figure 3.13: Effect of training dataset transformation in predicting diesel HC

3.2.3 NN model parameter selections

Apart from the choice of inputs and transformation used on the training dataset, the performance of NN models is dependent on the model architecture settings that include the number of neurons in a hidden layer, the number of epochs and the number of hidden layers used. In our work, we did a comprehensive study on how changing each of those parameters affected model performance. Once the model inputs, type of transformation and the training validation dataset were set, each one of the model parameters was selected using a series of simulations. For example, to decide on the number of neurons, epochs were set arbitrarily at 30 and the hidden layer at 1, the number of neurons was varied from 1 to 20 to see the change in performance as we increase the number of neurons in a hidden layer. Table 3.1 illustrates all the simulations performed to narrow the number of neurons, number of hidden layers and the number of

epochs for a model. The results presented here are for the models developed to predict diesel FC and NO_X, when trained using a dataset from one diesel experiment and validated using data from a different diesel vehicle experiment performed on the same route on a different day.

Table 3.1: Different simulations used in this study to determine the neural network architecture.

Simulation #	**Neurons (N)**	**Hidden Layers (HL)**	**Epochs (Ep)**
1	N=1:20	1	Set at 30
2	[N, N-1] where N=2:20	2	Set at 30
3	[N, N-1, N-2] where N=3:20	3	Set at 30
4	Set at 8	1	Ep=1:30
5	Set at [8,7]	2	Ep=1:30
6	Set at [8,7,6]	3	Ep=1:30

For example, in simulation 2, epochs were set constant at 30 while the number of hidden layers was set at 2. The second hidden layer had one neuron less than the first hidden layer and the number of neurons varied from 2 to 20. Similar simulations were performed for understanding the effects of number of epochs. In simulation 1, NN models were developed by setting the epochs at 30 while varying number of neurons from 1 to 20. Each model was trained on a specific representative dataset and validated against data from a different experiment. Figure 3.14 illustrates the change in model performance on validation dataset in terms of RE and R^2 in predicting FC and NO_X emissions from diesel vehicles during simulation 1. For models developed to predict FC, an increase in the number of neurons from 1 to 12 leads to a decrease in RE and increasing the number of neurons any further resulted in increasing the error. On the other hand, R^2 values have little to no effect due to the increasing number of neurons. Using these results the number of neurons was selected using the model that performed well in both RE and R^2. Setting the number of neurons at 9 for models that predicted FC from diesel vehicles was sufficient to get good overall performance.

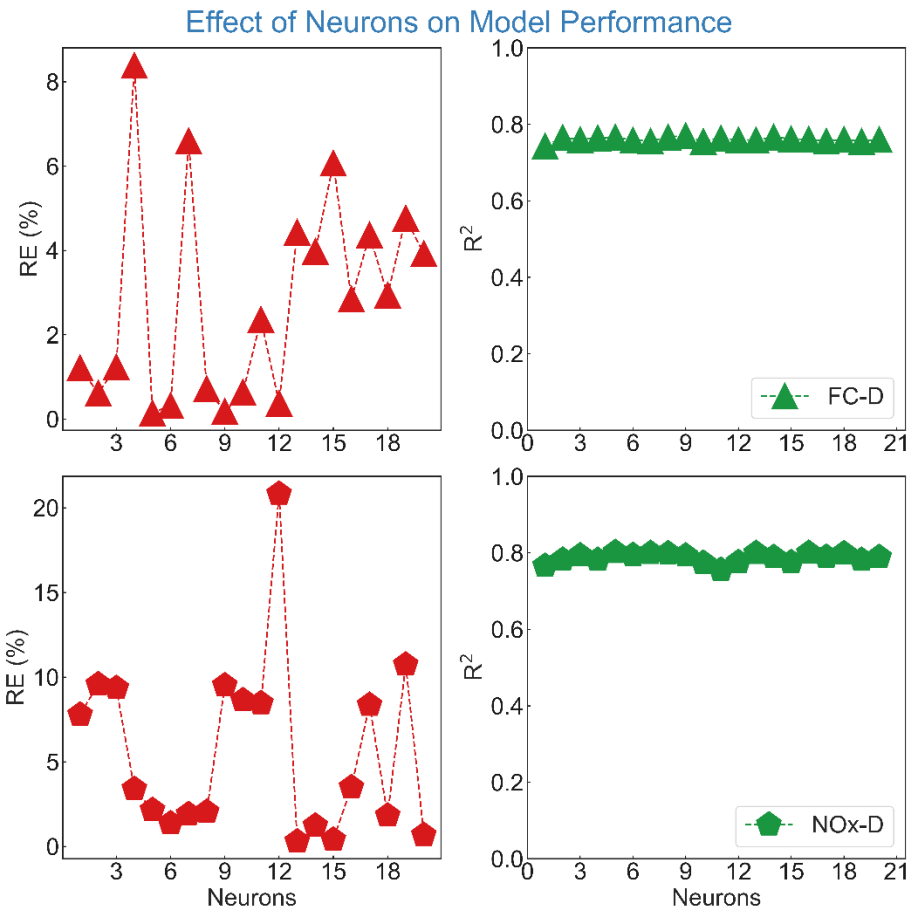

Figure 3.14: Simulation-1, Effect of number of neurons on model performance in predicting fuel consumption and NO_X *emissions for diesel vehicles*

For the models developed to predict NO_X emissions, as the number of neurons was increased in the model from 1-9, RE decreased significantly and a further increase in the number of neurons resulted in a random increase and decrease in the error. As the number of neurons varied, R^2 values fluctuated between 0.75 and 0.81 with no identifiable pattern. The number of neurons for NO_X models was set at 8 which had a low RE of ~3 and a corresponding R^2 value of ~0.795. Increasing the number of neurons for any model will result in a model that is overfitted against the training data. This yields a high performance when the model is validated against the same dataset with which it was trained but will result in poor performance when validated against any other experimental data.

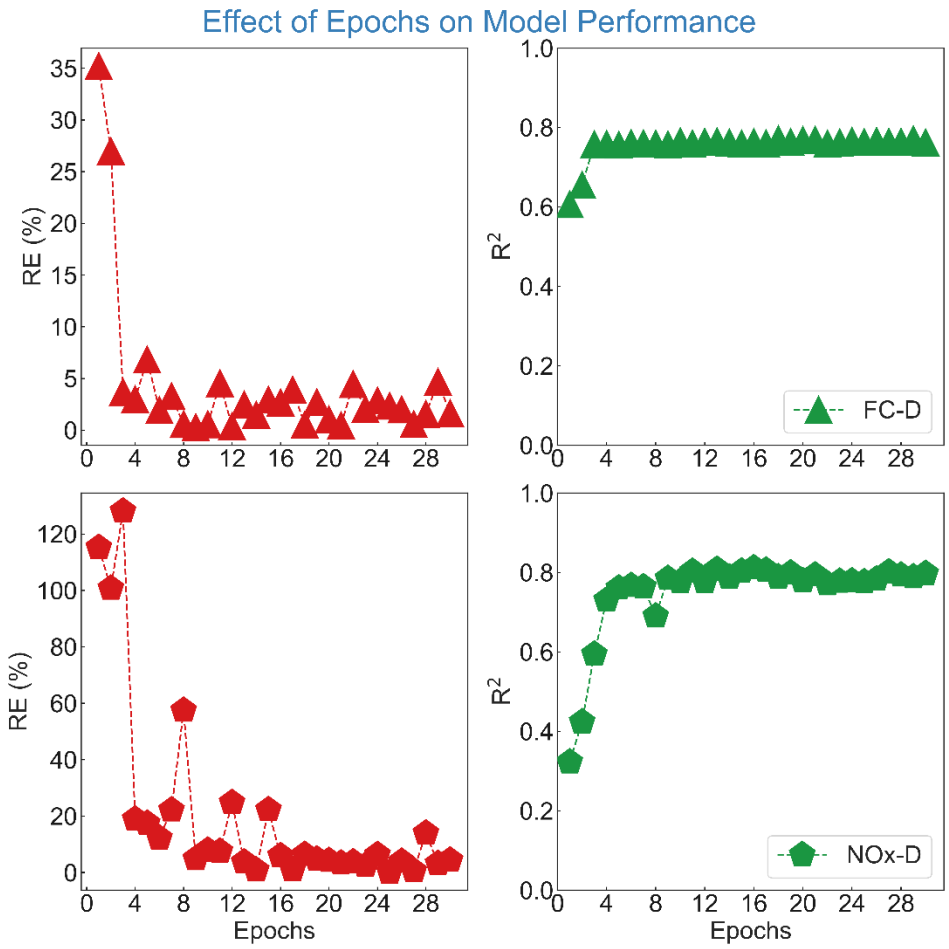

Figure 3.15: Simulation - 4, Effect of number of epochs on model performance in predicting fuel consumption for diesel vehicles

In simulation 4, NN models were developed by setting the number of neurons based on the results from simulation and varying the number of epochs from 1 to 30. Also, the number of maximum epochs can be set for any NN model, but the actual number of epochs used by the model can be less than the number of maximum epochs set, and this happens when the NN model performance worsens when the model overfits the data which is indicated by an increased error on validation dataset. Figure 3.15 illustrates the changes in performance when the number of maximum epochs varied between 1 and 30 in the models that were built to predict FC and NO_X emissions. For both FC and NO_X emissions, an increase in epochs from 1 to 8 resulted in a sharp decrease in error and increasing the number of epochs any further did not result in a significant improvement in performance. R^2 values for FC models increased sharply and

reached a steady value with as low as 3 epochs and the R^2 values for NO_X models increased and obtained a steady value within 8 epochs. These results indicate that having a high maximum number of epochs does not affect the performance significantly as long the number of epochs is greater than a minimum value. This is because if the number of maximum epochs is much higher than what is required by the model, MATLAB automatically stops training the model when the minimum mean squared error is obtained on the 15% of the training dataset that was assigned for validation. This is to avoid overfitting as well as to reduce the computational time in training the model.

3.3 Fuel consumption modeling results

In the previous sections, we have established that NN models outperformed MV models for all parameters of interest. NN model 14, with no transformation on the training dataset, was set as a base model for all parameters.

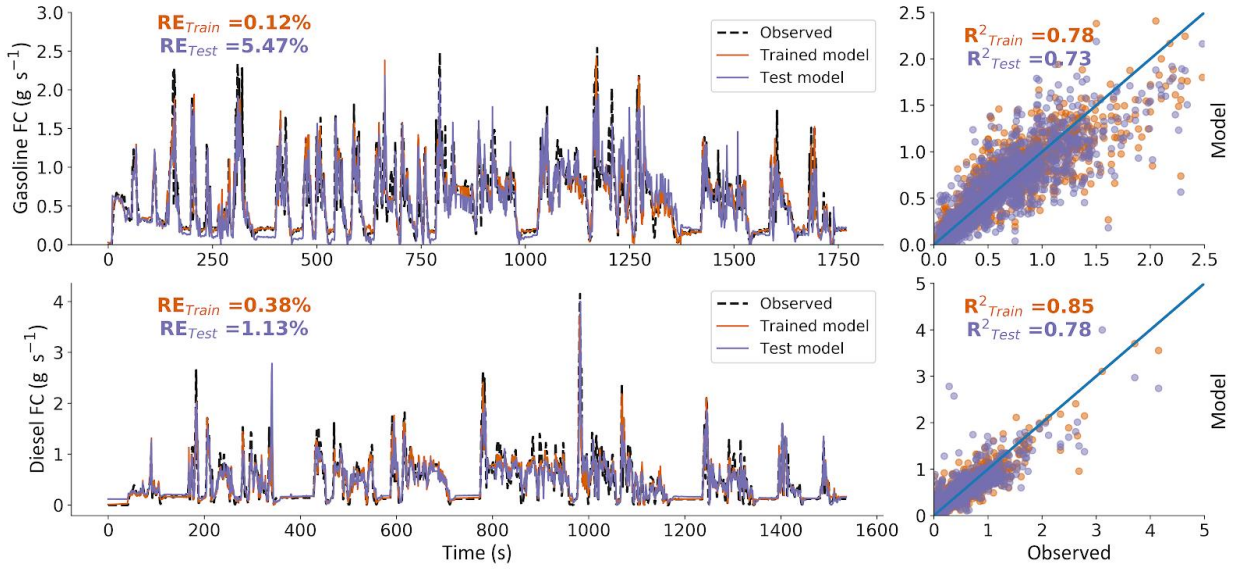

Figure 3.16: Time series comparison of observed fuel consumption from PEMS device for diesel and gasoline vehicle for a representative route with model results when validated against a trained dataset and when validated against a test dataset (left). Scatter plot of observed values against the trained and test model results (right).

NN model with no transformation on the input set M14 with 8 neurons, 1 hidden layer and, the maximum number of epochs set at 30 was used to model fuel consumption from both gasoline and diesel vehicles.

Figure 3.16 is the time series representation as well as the scatter plots of FC modeling results compared to the observed values from gasoline and diesel experiments for a representative experiment respectively. FC is illustrated as grams of fuel consumed per second from the start of the experiment. Observed fuel consumption is the experimental data obtained using the PEMS, trained model data is an output of the NN model which is trained and validated using the data from the same experiment, test model data is the output of the NN model that was trained using the data from one experiment and validated against data from a different experiment. In gasoline FC time series, we can see that the trained model compares well on a per second basis with the observed values and has an R^2 of 0.78 and RE of only 0.12 % but it does not represent the models' prediction capabilities when validated against a random dataset from a different drive cycle. Test model gives a generalized sense on how the model performs for any given route regardless of the model training dataset. We can see that for gasoline FC the test model does not capture the FC during idling and few cases where there is a high FC. The test model, in this case, has an R^2 of 0.73 and a RE 5.47%. The scatter plots provide good insight into how well the NN model performs when compared to changes in the observed values. The scatter plot consists of observed FC values in grams per second on the x-axis and both modeling results test and trained in grams per second on the y-axis. For gasoline FC we can see that there are many data points from test model with a small deviation from the 1-1 line at lower FC values suggesting that the model either under-predicts or over-predicts when the FC is low and as the FC increases the scatter between the model and observed results increases. The scatter at higher FC does not affect the overall relative error as the number of points that deviate from the observed value are less compared to the total data points. We also investigated the changes in the performance of the model based on the choice of training and validation dataset used to develop the NN model, which we will further discuss in the next section.

In diesel FC time series, we can see that the trained model has an error of only 0.38% with an R^2 value of 0.85. Test model developed for the diesel FC perform better than the models developed for gasoline FC. We can see that the test model's time series has a high correlation with the observed values with a high R^2

value of 0.78 and with a RE as low as 1.1%. Unlike gasoline FC, test model results from diesel FC aligns well when the FC values are lower. The scatter plot shows a lower scatter from the 1-1 line when compared to gasoline FC. Overall, we have found that the fuel consumption from diesel vehicles was more predictable than fuel consumed by gasoline vehicles.

3.4 Emission modeling

3.4.1 Gasoline emissions modeling

Exhaust emissions from gasoline experiment vehicle were highly dependent on the phase of the vehicle (cold/hot phase) and on the effective operation of the 3-way catalytic converter.

As discussed earlier, >80% of all emissions from gasoline vehicles occurred within the first 500 seconds of the experiment resulting in a disparity between the emissions before 500 seconds and after. Figure 3.17 illustrates the time series predictions of NN models for various emissions from gasoline vehicle and compares the model results to the observed values obtained from the PEMS. The time series for all pollutants is illustrated as milligrams of pollutant emitted per second from the start of the experiment whereas the scatter plot compares the per second observed values against the respective predictions form both the test and trained NN models. The first row in figure 3.17 consists of the NN model time series comparison for CO emissions from gasoline vehicles when validated against the same dataset with which the model was trained (training dataset) and when validated against data from a different experiment (test dataset) and the corresponding scatter plot. We can clearly see the high variance in the observed values of CO emissions where most emissions are concentrated within 500 seconds and the emissions after that are very close to zero. Trained NN model predictions compared well with the observed values with a RE of 9% and a high R^2 of 0.76, which was expected as the weights of NN model were optimized to give the least error against the training dataset.

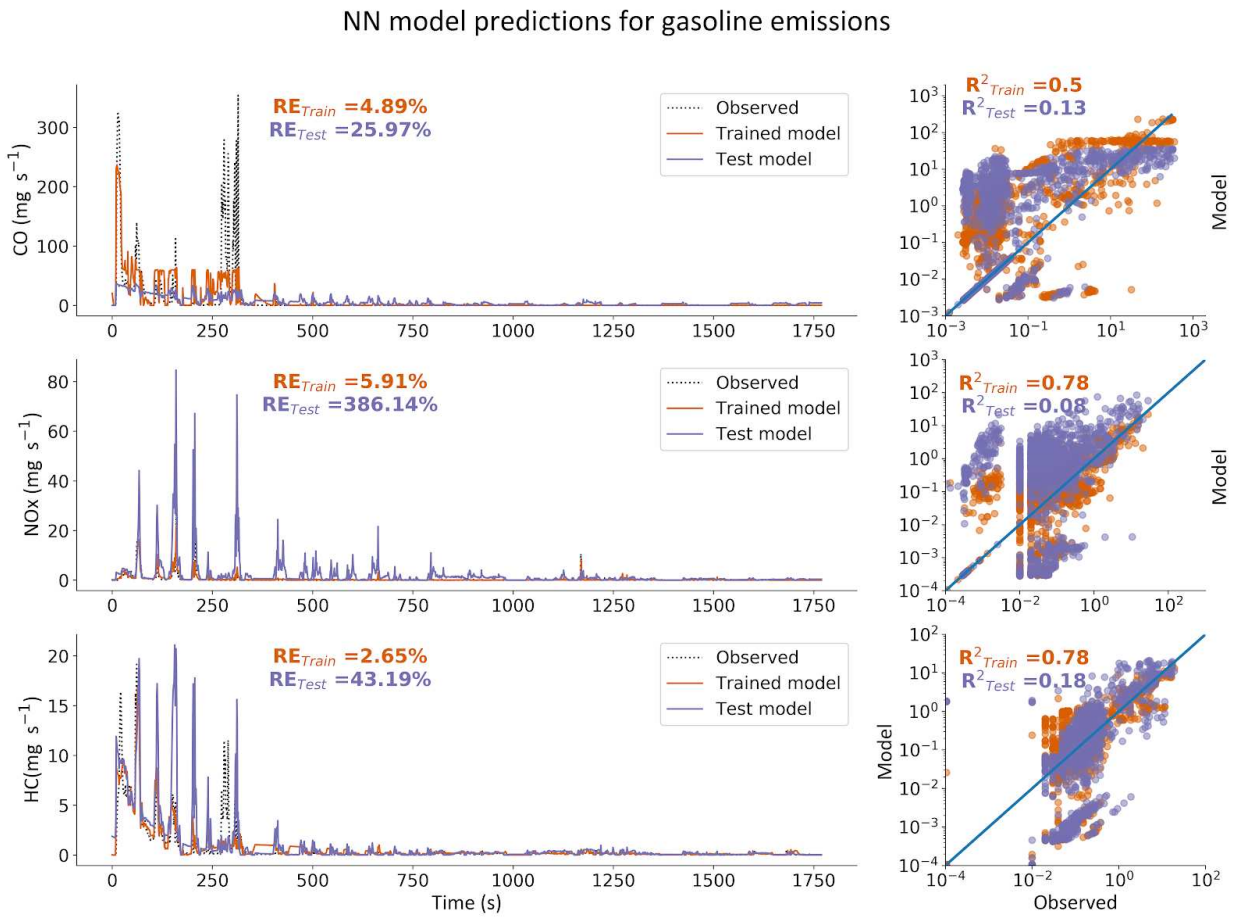

Figure 3.17: Time series comparison of observed CO, NO_X and HC emissions from gasoline vehicle for a representative route with model results when validated against a trained dataset and when validated against a test dataset (left). Scatter plot of observed values against the trained and test model results (right).

Test NN model performance was poor because of high RE of 25.97% and a very low R^2 of 0.13. Scatter plot on the right compares both trained and test model results to the observed values and since the values vary in orders of magnitude, a log-log scatter plot was used where both the axis are logarithmic. We can see that most points of the trained model are scattered around the 1-1 line while the test model results have no identifiable trend and are either underpredicted or overpredicted in most cases. Both the time series and the scatter plot of the test model against the observed values indicate that the developed model cannot predict CO emissions reliably from a gasoline vehicle. The second row of figure 3.17 illustrates the model performances in predicting NO_X emissions from gasoline vehicles. NO_X emissions like CO were confined mostly in the cold phase and occasional high values in the hot phase. The trained model

has a very low RE of 5.9% with a high correlation in time series with an R^2 of 0.78 while the test model has a RE of 386% with a very low R^2 of 0.08. The log-log scatter plot compares the test and trained model results with the observed values and indicates that the training model data compares well with the observed data and the test model in most cases overpredicts the NO_X emissions from gasoline vehicles. Test model performance also varies from iteration to iteration but never approaches a reliable performance metric in predicting NO_X emissions from gasoline vehicles. Similar to CO and NO_X models, models developed to predict HC emissions from gasoline vehicles also had poor performance metrics for the test datasets with a RE of 43.19% and R^2 of 0.18.

Clearly, all models developed to predict emissions from the gasoline vehicle performed well only when validated against training data and in all other cases, the models had very high overall error and a very poor time series comparison. Since >80% of emissions were within the cold phase for gasoline vehicles, developing a successful model to predict just cold phase emissions from these vehicles will help in understanding and reducing these emissions for each individual vehicle. To develop these models, only the cold phase data from an experiment was used to train the NN model and consequently, these models were validated against the cold phase experimental data used to train the model and also cold phase data from different experiments. Figure 3.18 illustrates the NN model predictions for cold phase gasoline CO, NO_X, and HC emissions. The first row of figure 3.18 is split into two parts one that illustrates the time series of observed cold phase CO emissions, trained model and test model predictions of CO all as milligrams of CO emitted per second and the second part consists of a log-log scatter plot that depicts the observed CO values at any particular second on the x-axis and the corresponding trained and test model predictions on the y-axis.

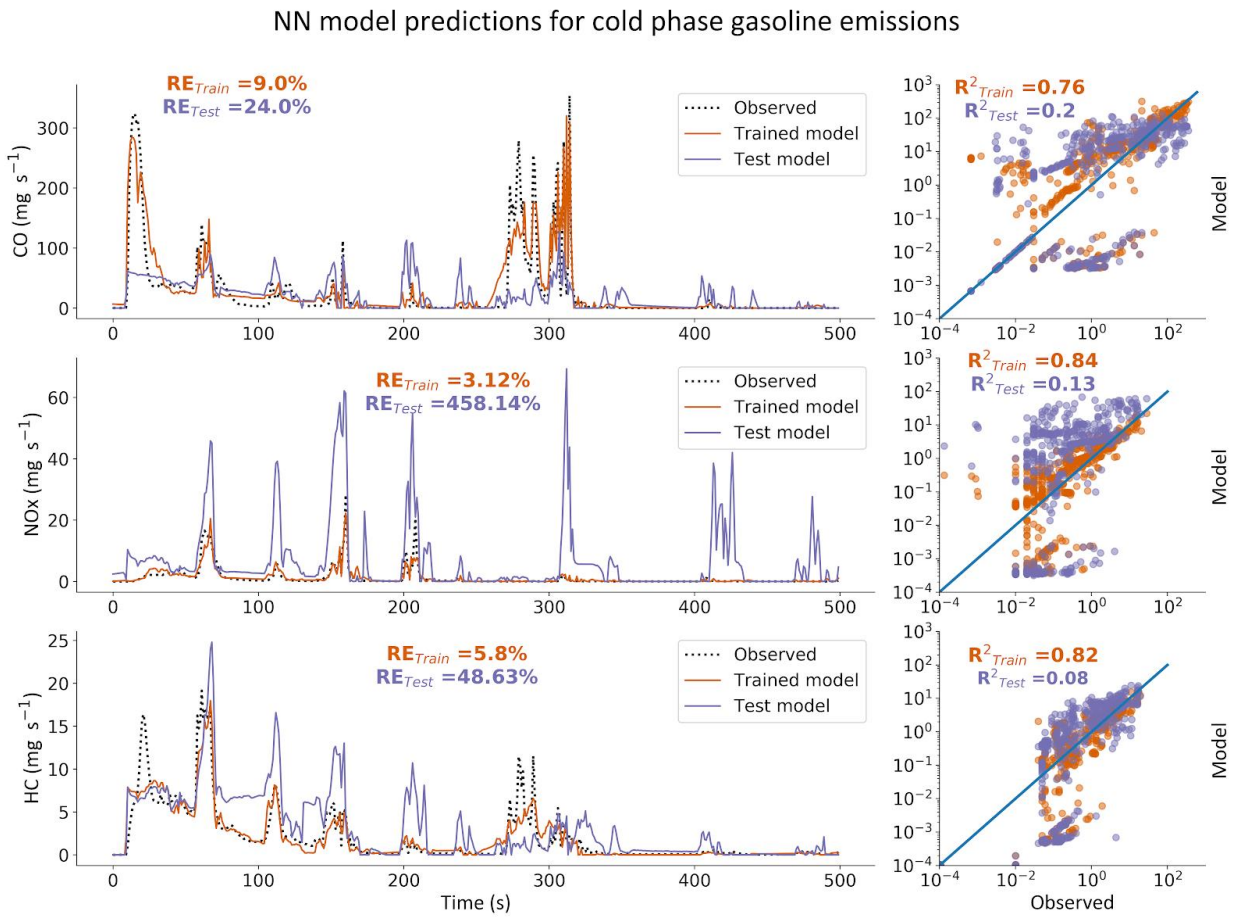

Figure 3.18: Time series comparison of observed cold phase CO, NO_X and HC emissions from gasoline vehicle for a representative route with model results when validated against a trained dataset and when validated against a test dataset (left). Scatter plot of observed values against the trained and test model results (right).

CO emissions from the trained model compare well with the observed values and in most cases capture both the peaks as well as the lower values in CO, the RE error of the trained model was 9% with a fairly high R^2 of 0.76 but as mentioned earlier high performance in the trained model is expected as the model built was optimized to perform well against the training dataset. Test model performance was subpar with RE of 24% and a poor R^2 of 0.2. Modeling the cold phase for CO emissions did not help in improving the model performance significantly. The second row of figure 3.18 illustrates observed and modeled cold phase NO_X emissions from the gasoline vehicle for a representative route as both time series and a log-log scatter plot. Although the trained model has very good performance metrics (RE-3%, R^2-0.84) the test model performed poorly with a very high RE of 458% and an R^2 of 0.13. It is evident from the time series

plot that the test model drastically overpredicts the NO_X emissions which also reflects in the scatter plot where most of the test model data is above the 1-1 line. Results from models developed to predict HC emissions and the observed values are represented in the third row of Figure 3.18 along with the corresponding scatter plot. Similar to models developed to predict cold phase CO and NO_X emissions, models developed to predict HC emissions did not show a significant improvement in performance (RE-5.8%, R^2-0.82) when validated against a test dataset. Observing the time series and the scatterplot we can clearly state that the test model overpredicts the HC emissions in most cases.

All models developed to predict cold phase emissions from gasoline vehicles performed poorly and few reasons that lead to the poor performance are the lower number of data points available for the NN model to train, highly nonlinear nature of the emissions from gasoline vehicles, and also having significant number of data points that are below the limit of detection of PEMS device used. Having a larger dataset for the cold phase emissions along with a more sensitive PEMS device can improve the prediction capabilities of the models developed to predict tailpipe emissions from gasoline vehicles. In our study, despite of trying different modeling schemes, transformation types and modeling architectures, we were not able to successfully develop a model to simulate tailpipe emissions from a gasoline vehicle.

3.4.2 Diesel emissions modeling

In our work, emissions from diesel vehicles were highly dependent on vehicle velocity and acceleration events and were sporadic in nature meaning the emissions were distributed across the entire length of the experiment, unlike gasoline vehicle experiments where >80% of emissions are in the cold phase. Models developed to predict these emissions performed better when compared to models developed to predict emissions from gasoline vehicles. Within the tailpipe pollutants from diesel vehicles, models developed to predict NO_X and HC emissions had a very low RE and high R^2 values while the models developed to predict CO and PM emissions had comparatively higher RE and lower R^2 values.

NN models built to predict tailpipe emissions from diesel vehicles are illustrated in figure 3.19 in the form of both time series of observed, trained and test model results in milligrams of pollutant emitted per second, as well as a log-log scatter plot, comparing the observed values against the test and trained model values.

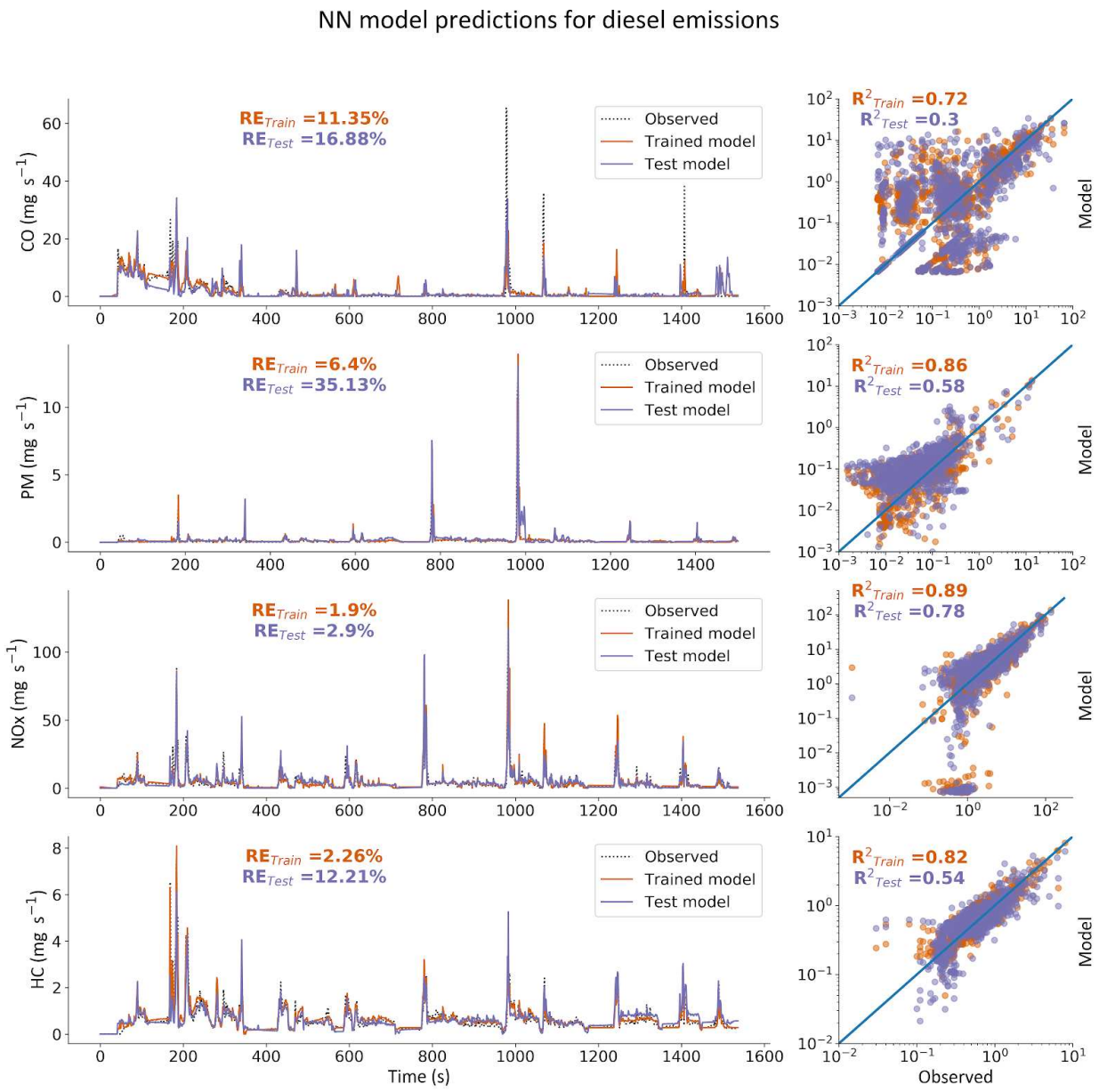

Figure 3.19: Time series comparison of observed CO, PM, NO_X and HC emissions from diesel vehicle for a representative route with model results when validated against a trained dataset and when validated against a test dataset (left). Scatter plot of observed values against the trained and test model results (right).

The ANN model developed to predict CO emissions performed well when validated against the dataset used to build the model with a low RE of 11.35% and an R^2 of 0.72, however, the test model in spite of having a low RE of 16.88% has a very poor correlation with time series with an R^2 of only 0.3. The low RE value is a result of the test model capturing the high CO emission events over the drive cycle and the low R^2 is due to poor performance in capturing low CO emissions. This trend can also be observed in the scatter plot where there is a higher scatter of CO emissions at lower values and as the emissions increase the scatter decreases and most data points are close to the 1:1 line.

Another reason for this can be due to the numerous data points that were found to be below the limit of detection of the PEMS device. Since all data points that were below LoD of PEMS were set to be at the limit of detection, this can affect the model performance as the value at these particular instances reduce the correlation of the input parameters and the CO emissions being modeled. NN models developed to predict PM emissions had a very high R^2 of 0.86 and with a RE of 6.4% when validated against the trained experimental data. The model, when validated against the test data, had a decent R^2 of 0.58 with and a RE of 35.13%. When compared to the models developed to predict CO, the test model to predict PM has a lower scatter at lower PM values and the scatter decreased as the value of PM increased leading to a higher R^2 when compared to those of CO models. The lower R^2 in predicting CO and higher RE in predicting PM make both these models not suitable for applying for a more generic case. If the NN model predictions are poor for a different experiment performed on the same vehicle and in similar conditions by the same driver then the predictions for other vehicles of the same category will invariably be poor.

In all the emission models developed in our study, models developed to predict diesel NO_X and HC emissions had the best performance. NO_X and HC emissions from the test diesel vehicle were more sporadic and were highly dependent on vehicle parameters like velocity, acceleration, and engine speed. The time series of observed NO_X emissions and predicted emissions from trained and test model in the form of milligrams of pollutant emitted per second and scatter plot that compares how the model

predictions vary based on the observed values are illustrated in the third row of Figure 3.19. NO_X models, when validated against the experimental data that were used to train the model, had a low RE of 1.9% and a high R^2 of 0.89. The model when validated against data from a different representative experiment performed with a low RE of only 2.9% and a very R^2 of 0.78. Scatter plot suggests that the difference in model predictions to the observed value is higher when the NO_X emissions are low and as the instantaneous NO_X emission increases the scatter between the observed and trained and test models is decreased. This can be attributed to the limit of detection assumptions and a more sensitive PEMS can yield better results in NO_X predictions.

The last row of Figure 3.19 illustrates the time series of observed, trained and test model results of HC emissions from a representative route as milligrams of HC emitted per second along the route along with the corresponding scatter plot of observed vs trained and test model results. Similar to models developed to predict NO_X, HC models also perform well when validated against a trained dataset (RE-2.26%, R^2 - 0.82). The models also perform moderately well in terms of both RE and R^2 when validated against data from a different experiment (RE - 12.21%, R^2-0.54). Unlike for NO_X models, there is no sign of the effect of the magnitude of HC emissions on the model prediction capabilities and the scatter plot is uniformly distributed along the 1:1 line throughout. In the models developed to predict emissions from diesel vehicles, NO_X and HC models had high performance metrics with a low RE of ~3% and ~12% with a corresponding R^2 of 0.78 and 0.54 respectively and hence can be used on more generic routes.

3.5 Multivariable and NN model comparison

A primary focus of our work was to establish whether the NN models are invariably better than the linear multivariable models which are commonly used in predicting both fuel consumption and emissions from on-road vehicles. Since we have established that models developed to predict fuel consumption and NO_X emissions from diesel vehicles had low RE and high R^2 values, we can compare the differences in MV and NN models developed to predict these two parameters. Figure 3.20 illustrates how the MV model

compares with the NN model in predicting FC and NO_X from diesel vehicle in the form of both time series and a scatter plot. The first row of figure 3.20 contains the observed time series of fuel consumed in grams per second along with the respective MV and NN model predictions for FC.

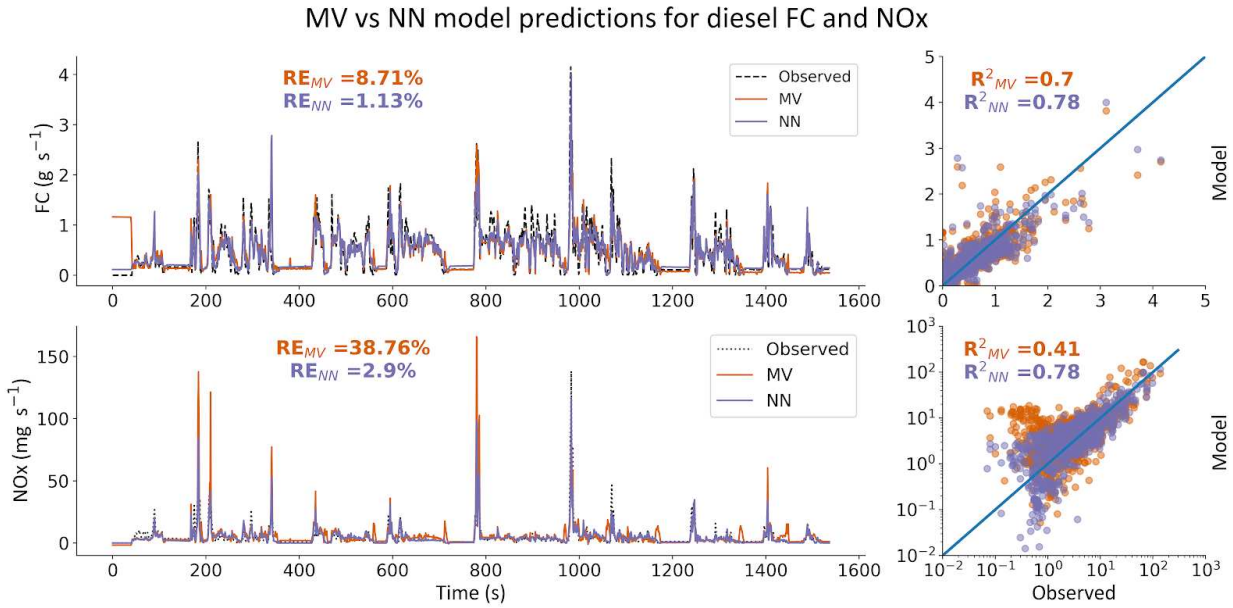

Figure 3.20: Time series comparison of observed FC and NO_X emissions from diesel vehicle for a representative route with MV and NN model results when validated against a test dataset (left). Scatter plot of observed values against the trained and test model results (right).

The scatter plot on the right compares the observed FC in grams on the x axis to the respective MV and NN model FC output in grams on the y axis. NN model has a better performance in predicting both instantaneous FC (R^2- 0.78) and route-integrated FC (RE-1.13%) when compared to the MV model that has an R^2 of 0.7 and RE of 8.71%. One reason for the higher error in MV model for this case is that the model has a higher error in the initial few seconds of the experiment where the observed FC value is at 0 (g s^{-1}) while the MV model prediction is at a constant value of 1.1 (g s^{-1}). Although both MV and NN models have a similar scatter, the correlation between observed and NN model is slightly higher than that of observed and MV model output.

Difference between MV and NN model in predicting NO_X emissions from diesel vehicles is illustrated in the second row of figure 3.20. Linear MV model performs better when the output of the model changes linearly with respect to the inputs used whereas for NN model it is not necessary to have linear relationship between the inputs and output of the model to obtain high performance. In predicting NO_X emissions from diesel vehicles, NN models significantly outperform the MV model. NN models had a very low route integrated error of 2.9% and a high R^2 of 0.78 suggesting a high instantaneous prediction capability. On the other hand, we can see that the MV model has a very high RE of 38.76% and a poor R^2 of 0.41. From the time series and scatter plot it is evident that when the emitted NO_X value is higher the MV model overpredicts the NO_X emissions while the NN model underpredicts.

The performance trends between MV and NN models are similar for all other parameters modeled in this study where the NN model outperforms the MV models regardless of the output variable.

3.6 Choice of training dataset for the model

It is also important to validate the developed models on a random route to understand the applicability of the model across different routes on the same vehicle. Figure 3.21 illustrates the observed values of FC and NO_X emissions of a diesel vehicle when driven on a random route and compares them to the output generated by the NN model in predicting these parameters. The first row of figure 3.21 depicts the timeseries of observed FC and the respective NN model prediction in grams per second, as well as the scatter plot comparing the correlation of observed values and model output. The model predicts the route integrated FC with a low error of 8.18% and the instantaneous FC with a high correlation of R^2=0.78. The second row of figure 3.21 depicts the timeseries of observed NO_X and the respective NN model prediction in milligrams per second, as well as a logarithmic scatter plot comparing the observed NO_X values and model output. NN model developed predicts the route integrated NO_X with an error of 9.45% and the instantaneous NO_X with a correlation of R^2=0.69, where we can see that at lower NO_X values model deviates from the observed values resulting in a lower overall R^2. Although the performance is slightly

lower when compared to models validated against experimental data from the same route, the current model has considerably high performance when validated against a random route for the same vehicle for diesel FC and NO_X with a RE<10% and R^2>0.68 in both cases.

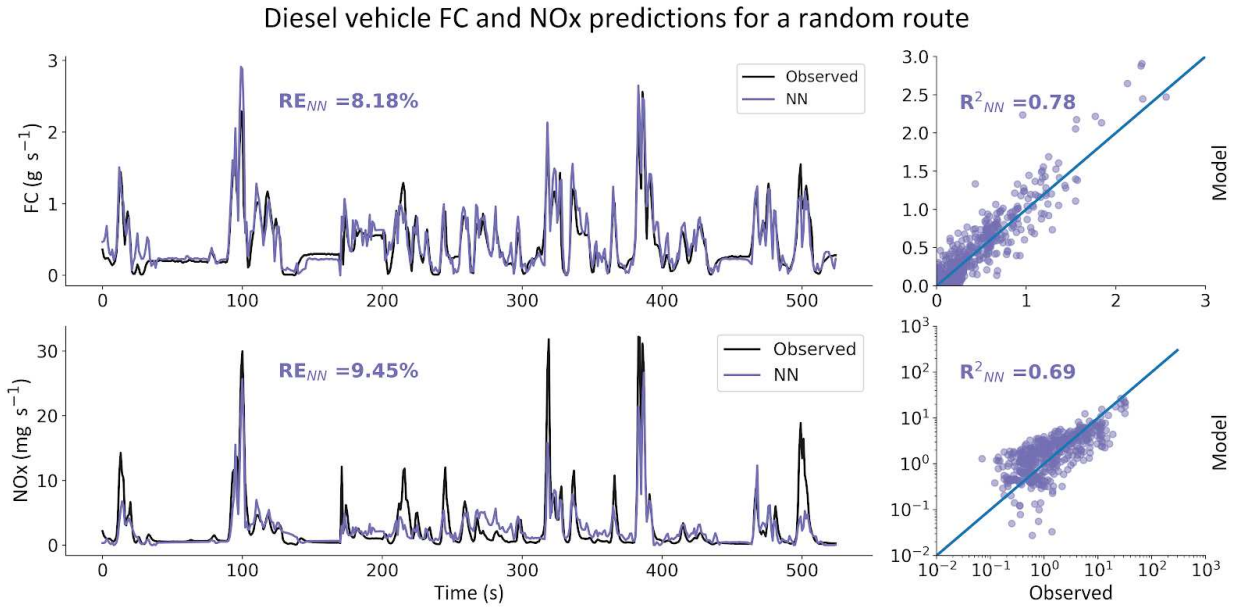

Figure 3.21: Time series comparison of observed FC and NO_X emissions from diesel vehicle with NN model results when validated against a random dataset (left). Scatter plot of observed values against the model results (right).

Within gasoline vehicle experiments, NN models built to predict FC had the highest performance. Similar to NN models built to predict diesel vehicle fuel consumption and emissions, the NN model developed to predict FC from gasoline vehicle was trained on data from an experiment performed on a specific day on the designed experimental route and validates using the data from an experiment performed on a different day but on the same route. To successfully apply this model for any given route driven on the same vehicle it is important to understand the performance of the model when validated against a random route. Figure 3.22 represents the timeseries of observed values of FC and the corresponding NN model prediction for the gasoline vehicle driven on the test route. It also represents the scatter plot comparing the correlation of observed FC values and NN model output. The model predicts the route integrated FC with a very low error of 9.51% but has a very poor R^2 value of 0.53. In figure 3.22, we can clearly observe that

there is a high scatter between the observed values and model predicted value when the observed FC is lower, and we can also see that at higher FC values the NN model underpredicts which is also reflected in the timeseries plot where the NN model fails to capture the peaks in the observed FC values. The model performance in predicting instantaneous FC does not affect the route integrated FC prediction since the model both underpredicts and overpredicts at lower FC values which results in a decreased relative error over the entire route.

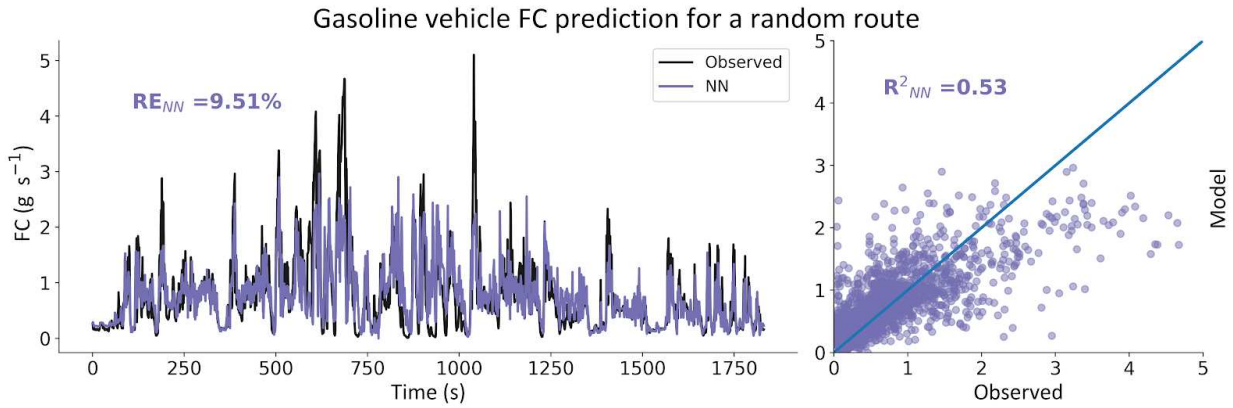

Figure 3.22: Time series comparison of observed FC of gasoline vehicle with NN model results when validated against a random dataset (left). Scatter plot of observed values against the model results (right).

3.7 Case Study

Developing a model that can predict FC and emissions from parameters that are readily available through onboard sensors can help in providing a real-time estimation of the FC and emissions based on user choice of route to arrive at a destination. Given the possibility of smart connected cars over the horizon, if a person needs to travel from one point to another and a real-time velocity time pattern is available for all possible routes through a platform like Google Maps, we can incorporate our NN models to calculate emissions and FC for each individual route before driving through them and can estimate the route that results in the lowest emissions or least amount of fuel consumed. This vehicle navigation method that aims to minimize fuel or energy consumption is known as Eco-routing (M Kubička et al., 2016 IEEE). Using the models developed, we tested the amount of fuel required to travel between two origin-

destination pairs on 3 different routes each. Figure 3.23(a) depicts the three different routes in the city of Fort Collins having the same start and end points. The routes were driven on the gasoline experimental vehicle and all the input parameters necessary for model 14 from table 2.3 were readily available through the vehicle's onboard sensors. The model developed in our study was used to predict FC from the three routes and we found that the route that had the highest fuel consumption of 0.55 gallons which took ~30 minutes was 27% more than the route that had the least fuel consumption of 0.431 gallons which only took 17.3 minutes.

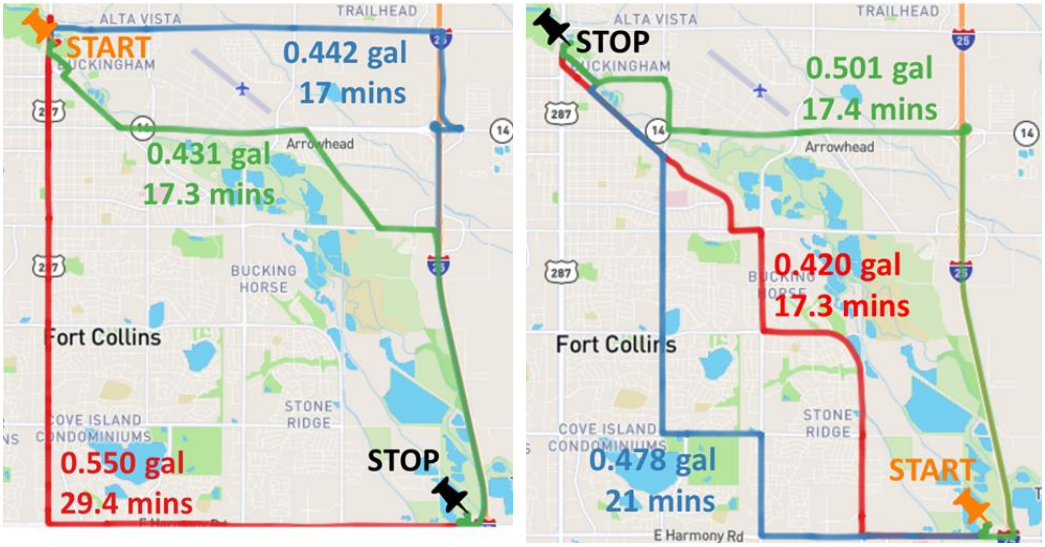

Figure 3.23: Three different routes with same origin and destination point in the city of fort Collins (left(a)), three different routes with origin and destination point interchanged (right(b))

In figure 3.23(b) gasoline experimental vehicle was driven on 3 new routes with the start and stop points interchanged from figure 3.23(a) and the ANN model was used to predict gasoline fuel consumption. We can observe that the red route has the least fuel consumption of 0.42 gallons while the highest fuel consumption was observed on the green route with 0.501 gallons. In this, we can also see that both routes took a similar amount of time to drive but one route had 25% more fuel consumption than the other route. With this case study, we can infer that by having route information between origin-destination pairs and vehicle specific ANN models we can optimize our route selection leading to lower fuel consumption and emissions from vehicles.

4. Summary and conclusions

On-road vehicle emissions are often higher than those of laboratory type-approval methods. On-road diesel vehicles contribute approximately to 20% of the global anthropogenic NO_X emissions (Stohl, A. et al.,2015). A study by Anenberg et al., 2017 estimated that over half of on-road light-duty diesel vehicle NO_X emissions among 80% of total global diesel vehicles sales are in excess of certification limits. These excess emissions from vehicles have significant health and environmental impacts ((Holland et al., 2016), (Oldenkamp et al., 2016)). To better understand and regulate these excess emissions there is a necessity to develop accurate models to predict both on-road fuel consumption and emissions. The existing US emission models EMFAC for California and MOVES for rest of US predict on-road emissions by having a fixed emission factor (EF) along with vehicle activity data like velocity paired with temperature and relative humidity amongst other, for estimating on-road emissions. These models are built based on emissions observed during chassis dynamometer tests that do not represent on-road emissions accurately, which results in imprecise predictions. In our work, we develop models based on the on-road fuel consumption and emissions data collected using a PEMS device on a representative test route to better predict real-world fuel consumption and emissions from gasoline and diesel light-duty passenger vehicles.

In our work, we developed linear Multivariable (MV) and Artificial Neural Networks (ANNs) to predict fuel consumption and emissions from LDV's. Performance of the model was assessed using absolute relative error (RE) and coefficient of correlation (R^2) where the former estimates the route integrated prediction capability while the later gives an idea of instantaneous prediction at any given second during the drive cycle. All models were developed with multiple configurations to maximize performance for each individual output variable for both diesel and gasoline vehicle. The performance of the NN model varied based on the variable being predicted and also had a dependency on the fuel type of the vehicle.

NN model route integrated fuel consumption predictions for both gasoline and diesel vehicles for any given route were within 11% and 5% respectively. On the other hand, instantaneous fuel predictions measured in terms of R^2 were also high with a median value of ~0.7 and ~0.8 for gasoline and diesel vehicles respectively. One key finding of this work is that NN models developed to predict fuel consumption and emissions consistently performed better than MV models regardless of the vehicle fuel type.

Models built to predict emissions from gasoline vehicles performed poorly when compared to all other models. The primary reason for poor performance was due to the emission activity from a gasoline vehicle where >80% of all emissions were within the cold phase (~8 mins) of the experiment leading to a highly nonlinear trend. The reason for this could be due to the catalytic converter not achieving the operating temperatures because of the cold start. Another reason for poor performance is due to a number of data points being below the limit of detection of the PEMS that was used in this study, leading to a decreased correlation between the model inputs and the output variable. On the contrary, emissions from diesel vehicles are much more sporadic over the entire drive cycle leading to more efficient models in predicting emissions. Model predictions for route integrated NO_X emissions from diesel vehicle were within 3% and 17% respectively regardless of the experimental data or the type of transformation used during training the model and the instantaneous time series predictions for NO_X and HC emissions were also high with a median R^2 of 0.79 and 0.64 resulting in models with high prediction capabilities. Models developed to predict CO and PM emissions had low performance metrics with high median RE of 26% and 19% and low R^2 of 0.38 and 0.24 respectively. Overall, among all the models developed, NN models outperformed MV models regardless of the type of transformation or modeling schemes used and the NN models developed to predict fuel consumption from gasoline vehicles and fuel consumption, NO_X, and, HC from diesel vehicles had the highest performance.

Bibliography

1) Anenberg, S. C., Miller, J., Minjares, R., Du, L., Henze, D. K., Lacey, F., Malley, C. S., Emberson, L., Franco, V., Klimont, Z. and Heyes, C.: Impacts and mitigation of excess diesel-related NOx emissions in 11 major vehicle markets, Nature, 545(7655), 467–471, doi:10.1038/nature22086, 2017.

2) Emission Standards: USA: Cars and Light-Duty Trucks—Tier 2, [online] Available from: https://www.dieselnet.com/standards/us/ld_t2.php (Accessed 20 May 2019a), n.d.

3) Home :: AirCare Colorado, [online] Available from: https://aircarecolorado.com/ (Accessed 15 May 2019b), n.d.

4) Levenberg-Marquardt backpropagation - MATLAB trainlm, [online] Available from: https://www.mathworks.com/help/deeplearning/ref/trainlm.html (Accessed 15 May 2019c), n.d.

5) Portable Emissions Measurement System | Center for Alternative Fuels Engines and Emissions | West Virginia University, [online] Available from: https://www.cafee.wvu.edu/laboratories/portable-emissions-measurement-system (Accessed 20 May 2019d), n.d.

6) In-Use Emissions Testing of Light-Duty Diesel Vehicles in the United States, Publications | Center for Alternative Fuels Engines and Emissions | West Virginia University, [online] Available from: https://www.cafee.wvu.edu/publications (Accessed 17 May 2019e), n.d.

7) RDE PEMS Real Driving Emissions Portable Emissions Measurement Systems - micro-PEMS, nano-PEMS | Global MRV, [online] Available from: https://www.globalmrv.com/pems-axionrs-2/ (Accessed 15 May 2019f), n.d.

8) Brugge, D., Durant, J. L. and Rioux, C.: Near-highway pollutants in motor vehicle exhaust: a review of epidemiologic evidence of cardiac and pulmonary health risks, Environ. Health, 6, 23, doi:10.1186/1476-069X-6-23, 2007.

9) Cao, Y.: Operation and Cold Start Mechanisms of Internal Combustion Engines with Alternative Fuels, in SAE Technical Paper Series, vol. 1, SAE International, 400 Commonwealth Drive, Warrendale, PA, United States., 2007.

10) Chossière, G. P.: Public Health Impacts of Excess NOx Emissions from Volkswagen Diesel Passenger Vehicles in Germany, Massachusetts Institute of Technology, Department of Aeronautics and Astronautics. [online] Available from: https://market.android.com/details?id=book-QTvPtAEACAAJ, 2017.

11) Durbin, T. D., Johnson, K., Cocker, D. R., 3rd, Miller, J. W., Maldonado, H., Shah, A., Ensfield, C., Weaver, C., Akard, M., Harvey, N., Symon, J., Lanni, T., Bachalo, W. D., Payne, G., Smallwood, G. and Linke, M.: Evaluation and comparison of portable emissions measurement systems and federal reference methods for emissions from a back-up generator and a diesel truck operated on a chassis dynamometer, Environ. Sci. Technol., 41(17), 6199–6204, doi:10.1021/es0622251, 2007.

12) Epa, U. S. and OAR: Final Rule for Model Year 2017 and Later Light-Duty Vehicle Greenhouse Gas Emissions and Corporate Average Fuel Economy Standards, [online] Available from:

https://www.epa.gov/regulations-emissions-vehicles-and-engines/final-rule-model-year-2017-and-later-light-duty-vehicle (Accessed 20 May 2019), 2016.

13) Erman, J., Arlitt, M. and Mahanti, A.: Traffic classification using clustering algorithms, in Proceedings of the 2006 SIGCOMM workshop on Mining network data, pp. 281–286, ACM., 2006.

14) Frey, H. C., Unal, A., Rouphail, N. M. and Colyar, J. D.: On-road measurement of vehicle tailpipe emissions using a portable instrument, J. Air Waste Manag. Assoc., 53(8), 992–1002, doi:10.1080/10473289.2003.10466245, 2003.

15) Frey, H. C., Zhang, K. and Rouphail, N. M.: Fuel use and emissions comparisons for alternative routes, time of day, road grade, and vehicles based on in-use measurements, Environ. Sci. Technol., 42(7), 2483–2489, doi:10.1021/es702493v, 2008.

16) Geladi, P. and Kowalski, B. R.: Partial least-squares regression: a tutorial, Anal. Chim. Acta, 185, 1–17, doi:10.1016/0003-2670(86)80028-9, 1986.

17) Hagan, M. T. and Menhaj, M. B.: Training feedforward networks with the Marquardt algorithm, IEEE Trans. Neural Netw., 5(6), 989–993, doi:10.1109/72.329697, 1994.

18) Holland, S. P., Mansur, E. T., Muller, N. Z. and Yates, A. J.: Damages and Expected Deaths Due to Excess NOx Emissions from 2009 to 2015 Volkswagen Diesel Vehicles, Environ. Sci. Technol., 50(3), 1111–1117, doi:10.1021/acs.est.5b05190, 2016.

19) Kampa, M. and Castanas, E.: Human health effects of air pollution, Environ. Pollut., 151(2), 362–367, doi:10.1016/j.envpol.2007.06.012, 2008.

20) Khoshnevisan, B., Rafiee, S., Omid, M., Yousefi, M. and Movahedi, M.: Modeling of energy consumption and GHG (greenhouse gas) emissions in wheat production in Esfahan province of Iran using artificial neural networks, Energy, 52, 333–338, doi:10.1016/j.energy.2013.01.028, 2013.

21) Kukkonen, J., Partanen, L., Karppinen, A., Ruuskanen, J., Junninen, H., Kolehmainen, M., Niska, H., Dorling, S., Chatterton, T., Foxall, R. and Cawley, G.: Extensive evaluation of neural network models for the prediction of NO2 and PM10 concentrations, compared with a deterministic modelling system and measurements in central Helsinki, Atmos. Environ., 37(32), 4539–4550, doi:10.1016/S1352-2310(03)00583-1, 2003.

22) Kumar Pathak, S., Sood, V., Singh, Y. and Channiwala, S. A.: Real world vehicle emissions: Their correlation with driving parameters, Transp. Res. Part D: Trans. Environ., 44, 157–176, doi:10.1016/j.trd.2016.02.001, 2016.

23) Kwon, S., Park, Y., Park, J., Kim, J., Choi, K.-H. and Cha, J.-S.: Characteristics of on-road NOx emissions from Euro 6 light-duty diesel vehicles using a portable emissions measurement system, Sci. Total Environ., 576, 70–77, doi:10.1016/j.scitotenv.2016.10.101, 2017.

24) Laumbach, R. J. and Kipen, H. M.: Respiratory health effects of air pollution: update on biomass smoke and traffic pollution, J. Allergy Clin. Immunol., 129(1), 3–11; quiz 12–3, doi:10.1016/j.jaci.2011.11.021, 2012.

25) LeCun, Y., Bengio, Y. and Hinton, G.: Deep learning, Nature, 521(7553), 436–444, doi:10.1038/nature14539, 2015.

26) Lek, S. and Guégan, J. F.: Artificial neural networks as a tool in ecological modelling, an introduction, Ecol. Modell., 120(2), 65–73, doi:10.1016/S0304-3800(99)00092-7, 1999.

27) Mitchell, T. M.: Does machine learning really work?, AI magazine [online] Available from: https://www.aaai.org/ojs/index.php/aimagazine/article/view/1303, 1997.

28) Mohamed Ismail, H., Ng, H. K., Queck, C. W. and Gan, S.: Artificial neural networks modelling of engine-out responses for a light-duty diesel engine fuelled with biodiesel blends, Appl. Energy, 92, 769–777, doi:10.1016/j.apenergy.2011.08.027, 2012.

29) Nagendra, S. M. S. and Khare, M.: Artificial neural network approach for modelling nitrogen dioxide dispersion from vehicular exhaust emissions, Ecol. Modell., 190(1), 99–115, doi:10.1016/j.ecolmodel.2005.01.062, 2006.

30) O'Driscoll, R., ApSimon, H. M., Oxley, T., Molden, N., Stettler, M. E. J. and Thiyagarajah, A.: A Portable Emissions Measurement System (PEMS) study of NOx and primary NO2 emissions from Euro 6 diesel passenger cars and comparison with COPERT emission factors, Atmos. Environ., 145, 81–91, doi:10.1016/j.atmosenv.2016.09.021, 2016.

31) Oldenkamp, R., van Zelm, R. and Huijbregts, M. A. J.: Valuing the human health damage caused by the fraud of Volkswagen, Environ. Pollut., 212, 121–127, doi:10.1016/j.envpol.2016.01.053, 2016.

32) Pelkmans, L. and Debal, P.: Comparison of on-road emissions with emissions measured on chassis dynamometer test cycles, Transp. Res. Part D: Trans. Environ., 11(4), 233–241, doi:10.1016/j.trd.2006.04.001, 2006.

33) Reiter, M. S. and Kockelman, K. M.: The problem of cold starts: A closer look at mobile source emissions levels, Transp. Res. Part D: Trans. Environ., 43, 123–132, doi:10.1016/j.trd.2015.12.012, 2016.

34) Sheela, K. G. and Deepa, S. N.: Review on Methods to Fix Number of Hidden Neurons in Neural Networks, Math. Probl. Eng., 2013, doi:10.1155/2013/425740, 2013.

35) Thompson, G. J., Atkinson, C. M., Clark, N. N., Long, T. W. and Hanzevack, E.: Technical Note: Neural network modelling of the emissions and performance of a heavy-duty diesel engine, Proc. Inst. Mech. Eng. Pt. D: J. Automobile Eng., 214(2), 111–126, doi:10.1177/095440700021400201, 2000.

36) U.S. EPA Office of Air and Radiation: Air Quality Improves as America Grows, [online] Available from: https://gispub.epa.gov/air/trendsreport/2018/ (Accessed 15 May 2019), n.d.

37) Wang, T., Jerrett, M., Sinsheimer, P. and Zhu, Y.: Estimating PM2.5-associated mortality increase in California due to the Volkswagen emission control defeat device, Atmos. Environ., 144, 168–174, doi:10.1016/j.atmosenv.2016.08.074, 2016.

38) Weber, M., Welling, M. and Perona, P.: Unsupervised Learning of Models for Recognition, in Computer Vision - ECCV 2000, pp. 18–32, Springer Berlin Heidelberg., 2000.

39) Weilenmann, M., Soltic, P., Saxer, C., Forss, A.-M. and Heeb, N.: Regulated and nonregulated diesel and gasoline cold start emissions at different temperatures, Atmos. Environ., 39(13), 2433–2441, doi:10.1016/j.atmosenv.2004.03.081, 2005.

40) Zhang, K. and Batterman, S.: Air pollution and health risks due to vehicle traffic, Sci. Total Environ., 450-451, 307–316, doi:10.1016/j.scitotenv.2013.01.074, 2013.