2004

Effects of artificial neural network speed-based inputs on heavy-duty vehicle emissions prediction

Nastaran Hashemi
West Virginia University

Effects of Artificial Neural Network Speed-Based Inputs on Heavy-Duty Vehicle Emissions Prediction

Nastaran Hashemi

Thesis submitted to
The College of Engineering and Mineral Resources
at West Virginia University
in partial fulfillment of the requirements
for the degree of

Master of Science
in
Mechanical Engineering

Nigel N. Clark, Ph.D., Chair
Shahab D. Mohaghegh, Ph.D.
Gregory J. Thompson, Ph.D.

Department of Mechanical and Aerospace Engineering

Morgantown, West Virginia
2004

Keywords: Artificial Neural Network, Emissions Prediction, Heavy-Duty Diesel Vehicle

ABSTRACT

Effects of Artificial Neural Network Speed-Based Inputs on Heavy-Duty Vehicle Emissions Prediction

Nastaran Hashemi

The PM split study was performed in Southern California on thirty-four heavy-duty diesel vehicles using the West Virginia University Transportable Heavy-Duty Vehicle Emissions Testing Laboratories to gather emissions data of these vehicles. The data obtained from six vehicles in the 1985-2001 model year and 33,000-80,000 lb weight range exercised through three different cycles were selected in this thesis. To predict the instantaneous levels of oxides of nitrogen (NO_x), carbon dioxide (CO_2), hydrocarbons (HC) and carbon monoxide (CO), an Artificial Neural Network (ANN) was used. Axle speed, torque, their rates of change over different time periods and two other variables as a function of axle speed were defined as the inputs for the neural network. Also, each emissions species was considered individually as the output of the ANN. The ANN was trained on the Highway cycle and applied to the City/Suburban Heavy Vehicle Route (CSHVR) and Urban Dynamometer Driving Schedule (UDDS) with four different sets of inputs to predict the emissions for these vehicles. The research showed an excellent emissions prediction for the neural networks that were trained with only eight inputs (speed, torque, their first and second derivatives, and two variables of Diff. and $\overline{Spd}$ related to the speed pattern over the last 150 seconds). The Diff variable provided a measure of the variability of speed over the last 150 seconds of operation. This variable was able to create a moving speed-dependant window, which was used as an input for the neural networks. The results showed an average accuracy of 0.97 percent for CO_2, 0.89 percent for NO_x, 0.70 for CO and 0.48 percent for HC over the course of the CSHVR, Highway and UDDS.

DEDICATION

To my beloved mom, dad and Niloofar

for their enduring loyalty

ACKNOWLEDGEMENTS

I would like to express my deep appreciation to Dr. Nigel Clark for his remarkable role in helping me to accomplish this thesis. He is not just an outstanding research advisor, but a great human being who provided me with a wonderful definition of living in an absolute different culture.

I would also like to thank Dr. Gary Morris for his supportive character. I thank my committee members for their presence.

My special thanks go to my friends, Reza, Zoa and Azadeh for their emotional encouragements.

I would like to thank the U.S. Department of Energy and National Renewable Energy Laboratory for initiating the PM Split Study that was the source of data for this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

α	Learning rate
γ	Momentum coefficient
θ_i	Dimensionless time at i^{th} increment
μ	Coefficient of rolling resistance
ρ_a	Ambient air density
a	Scalar output
$\mathbf{a}_q$	Network output
A_v	Vehicle frontal area
ADALINE	ADAptive LInear NEuron
ADC	Analog to Digital Conversion
ADVISOR	ADvanced VehIcle SimulatOR
AFR	Air/Fuel Ratio
ANN	Artificial Neural Network
b	Bias
BP	Back-Propagation
BSFC	Brake Specific Fuel Consumption
C_D	Drag coefficient
C_i	Dispersion coefficient at time t_i
CO	Carbon Monoxide
CO_2	Carbon Dioxide
CSHVR	City Suburban Heavy Vehicle Route
Diff	Integral of the difference of instantaneous speed squared and average speed squared over the last t_d seconds
e	Scalar error
EGR	Exhaust Gas Recirculation
EPA	Environmental Protection Agency
f	Transfer function
FTP	Federal Test Procedure
g	Acceleration of gravity

GVWR	Gross Vehicle Weight Rating
HC	Hydrocarbons
LMS	Least Mean Square
M_v	Vehicle mass
MOVES	Multi-scale mOtor Vehicles and equipment Emission System
n	Output
NAAQS	National Ambient Air Quality Standards
NO_x	Oxides of Nitrogen
OTAQ	Office of Transportation and Air Quality
p	Scalar input
$\mathbf{p}_q$	Input vector
P	Power
PM	Particulate Matter
r^2	Linear coefficient
R^2	Coefficient of multiple determination
S	Dispersed speed
Spd	Vehicle speed
$\overline{\text{Spd}}$	Average speed over the last t_d seconds
$\mathbf{t}_q$	Target output
T	Dispersed torque
THDVETL	Transportable Heavy-Duty Vehicle Emissions Testing Laboratory
UDDS	Urban Dynamometer Driving Schedule
VDN	Vessel Dispersion Number
VLBP	Variable Learning rate Back-Propagation
w	Weight
$\mathbf{W}$	Weight matrix

1. INTRODUCTION

The internal combustion engine was introduced to humankind in 1860s by J.J.E. Lenoir. He developed the first coal-gas air mixture at atmospheric pressure. These engines were built between 1860 and 1865 with up to 6 brake horsepower and an efficiency of 5 percent. In 1867, N.A. Otto and E. Langen developed another atmospheric engine with the thermal efficiency of 11 percent. To improve the engine efficiency, in 1876 Otto proposed an engine cycle with four piston strokes: an intake stroke, a compression stroke, an expansion or power stroke and an exhaust stroke. Due to this improvement, this engine was more efficient and there was a substantial reduction in engine volume and weight. Further development followed fast after Otto's four-stroke engine. In 1892, R. Diesel designed a new form of internal combustion engine [1]. He found that by injecting a liquid fuel into compressed air, it is possible to achieve a higher efficiency, compare to other internal combustion engines. Engine developments have continued ever since. The most recent development was the rotary internal combustion engine, which was successfully tested in 1957.

Nowadays, the engine design is significantly affected by other factors such as air pollution and fuel consumption. In 1940s, the automotive air-pollution problem became a matter of concern in Los Angeles and it was demonstrated that automotive emissions are the reasons of smoke problem in that area. As a result, emission standards for automobile were introduced first in California and then in all the United States in the 1960s.

The United States Environmental Protection Agency was established three decades ago to protect the human health and to safeguard the natural environment (air, water, and land). The EPA defined the National Ambient Air Quality Standards (NAAQS) and set pollution limits. Diesel engines are extensively used to power the on-road vehicles such as trucks and buses and also off-road equipment as a result of their outstanding fuel economy. Oxides of nitrogen (NO_x), particulate matter (PM), hydrocarbons (HC) and carbon monoxide (CO) are the main contributions of on-road heavy-duty diesel vehicles to the atmospheric inventory. NO_x reacts with HC and sunlight to form ground-level ozone and also plays an important role in secondary fine particulate matter formation [2]. The EPA estimates that 27% of on-road vehicle emissions of NOx and more than 60% of PM are contributed by heavy-duty vehicles throughout the nation, although these heavy-duty vehicles are only 2% of the total on-road vehicles by number [3]. Heavy-duty vehicles are identified as vehicles of gross vehicle weight rate (GVWR) of above 8,500 lbs in the federal jurisdiction and above 14,000 lbs in California (model year 1995 and later). Considering the growing pollution crisis, emission regulations have been an important topic for environmental researchers. The EPA has provided emission standards for heavy-duty diesel vehicles. Model year 1988-2010 US federal EPA emission standards for the heavy-duty diesel trucks are summarized in Table 1.1.

The PM emission standard must be met by the 2007 heavy-duty vehicle model year. However, the NO_x and HC standards are scheduled to be met between 2007 and 2010. This schedule is based on the percentage of the vehicles sale. 50% of the total

vehicles sale must meet the standard from 2007 to 2009 and 100% of the vehicles sale by 2010.

Year	HC	CO	NOx	PM
1988	1.3	15.5	10.7	0.60
1990	1.3	15.5	6.0	0.60
1994	1.3	15.5	5.0	0.10
1998	1.3	15.5	4.0	0.10
2000	1.3	15.5	4.0	0.10
2004	0.5	15.5	2.0	0.10
2007-2010	0.14	15.5	0.20	0.01

Table 1.1. EPA Emission Standards for Heavy-Duty Diesel trucks, g/bhp-hr [4].

The EPA's regulations consider emissions inventory data and drive continuous efforts to review the diesel engine design and move toward more effective controls for NO_x, PM and HC. For the development of a modeling inventory, it is necessary to have a database of the emissions, also known as the base year inventory [5]. To approach the new analysis needs, the EPA's Office of Transportation and Air Quality (OTAQ) is working on a new system defined as the Multi-scale mOtor Vehicles and equipment Emission System (MOVES) [6]. The MOVES model will estimate emissions for on-road and off-road sources, covering a wide range of pollutants [7]. MOVES is different from the previous MOBILE series of models. The ability to generate emissions for a wide range of pollutants and processes such as emissions from both on-road and off-road vehicles and equipment are combined in this modeling framework. It is also capable of

illustrating model validation and the calculation of uncertainty, easing the model updating, and interfacing with other models. It is possible to perform multiple scale analysis, from fine-scale analysis to national inventory estimation with the MOVES design.

The MOVES design has been developed to satisfy the National inventory development for the EPA reports and regulations, local inventory development, model interaction, policy evaluation, model analysis, project level analysis and model updates. Due to its unique multiple scales and emissions process, this model provides the calculation of emissions from a wide range of emission sources. Therefore, a generic framework has been suggested to calculate the total emissions according to an identified time, location, use type and emission process.

$$\text{Total Emissions}\,i = \text{Total Activity} \times \sum_{n=1}^{j} \text{Emissions Rate}\,i, j \times \text{Bin Distribution}\,i, j$$

Equation 1.1

where i is the use type and j is the bin.

Total activity is the product of population and per-source activity for a specific use type, time, and location. Emissions rate specifies emissions for a given process, source bin, and operating mode bin. It also can be considered for additional effects such as fuel and meteorology. Each discrete bin is dictated by factors that are found to contribute unique emissions based on characteristics of the use type and operating modes. Source bins are defined by fuel type, accumulated use, technology types, standard levels, source weight for on-road sources, and horsepower range for off-road courses.

The Artificial Neural Network (ANN) modeling is able to estimate the instantaneous emissions. It is necessary to understand the vehicles activities characteristics and primarily engines features to analyze their contribution to the emissions. The ANN has been widely employed for modeling the engines characteristics such as combustion and emissions. In this work, the ANN was used to predict the vehicle emissions and identify those specific vehicle parameters that influence the emissions. Further results of this research could be used to identify the relation of the emissions and engine characteristics in order to improve the engine design. This improvement could facilitate the on-road vehicles emissions reduction.

1.1. Objectives

The primary objective of this research was to develop an artificial intelligence method to predict the heavy-duty vehicle emissions. For this purpose the artificial neural networks were chosen due to their ability to solve non-linear problems such as vehicle emissions. In order to achieve the most accurate emissions prediction, the study was concentrated on the contribution of the different inputs and also the techniques that were used to pre-process them. Therefore, three sub-objectives were defined as:

i. Developing a data pre-processing strategy using a dispersion model.
ii. Creating the different inputs
iii. Developing different neural network architectures with different number of inputs.
iv. Evaluating the applied approaches.

1.2. Literature Review

Due to the capability of the ANNs in solving non-linear problems, they have been the center of great attention since the late 80's. The engine emissions research area has not been excluded from the growing neural network applications.

Considering the relationship of the Air/Fuel Ratio (AFR) to the power and consequently the emissions, precise ignition point of the air-fuel mixture is critical to reduce the exhaust emissions and obtain the maximum output power. Bastian [9] used an ANN approach to identify the amount of air for a given engine speed and inlet manifold pressure. His work was one the earliest research efforts in the ANN applications. Frith et al. [10] have also investigated the application of ANNs for adaptive AFR control in gasoline engines. Plenty of research was performed on the application of the artificial neural networks to identify the AFR effects on gasoline engine emissions levels.

ANN application to engine exhaust catalysts was another area of interest. Ramani et al. [11] used neural networks to identify composition-performance relationships in automobile exhaust catalysts. They employed an ANN to undertake a sensitivity analysis of the conversions of pollutant gases as a function of the catalyst composition and the operating conditions. This approach combined the optimum catalyst composition and operating condition in order to produce specific conversions of CO, HC and NO_x to CO_2, water (H_2O) and nitrogen respectively.

Stevens et al. [12] utilized a neural network approach for processing the data and to evaluate relationships between engine emissions and engine state variables.

To predict real-time engine power output along with exhaust emissions, Atkinson et al. [13] developed a neural network-based engine performance, fuel efficiency and emissions prediction system for both spark ignited and compression ignition engines.

Krijnsen et al. [14] proposed the application of ANN as a precise tool to predict the engine's NO_x emission instead of using expensive NO_x analyzers and computer models. Data were collected from a transient operating diesel engine and part of the data was used to train the network, while the other part was used to test the NO_x emission prediction. A single-layer perceptron network with the inputs of engine speed, rack position, intake air pressure, intake air temperature and their rates of change was chosen for this study. The average absolute deviation between the predicted and measured NO_x emission was 6.7%. This work proved that the ANN is an accurate tool to predict the automotive NO_x along with its short computation time qualification.

Quenou et al. [15] proposed an ANN diesel engine exhaust emissions modeling. The main feature of this ANN modeling is the capability to identify the emissions through opacity. They also dealt with the structure of the one hidden layer feedforward neural network and described the Gaussian-Newton training rule based on the general error criterion. Then, they used this model to determine the diesel engine emissions using the standard Levenberg-Marquardt algorithm. The fuel flow, air flow and the engine speed were chosen as inputs to their ANN model.

To control the diesel engine exhaust, Hafner et al. [16] presented a fast neural network to dynamically simulate the different emissions from engine. The ANN was able to calculate a cost function for exhaust versus consumption and determine an optimal

injection angle dependent on the engine's exhaust performance, fuel consumption and the real-time driving condition.

Thompson et al. [17] used the application of a 3-layer neural network to model the continuous output torque and engine emissions from a heavy-duty diesel engine for the FTP (Federal Test Procedure) heavy-duty engine transient cycle. The ANN consisted of 10 input variables including intake air temperature, intake air pressure, injection pressure, injection pulse width, start of injection, acceleration position, engine speed, engine oil temperature, engine coolant temperature, and exhaust temperature. The prediction results were within 5% of their measured values after only 100 minutes of transient dynamometer training.

De Lucas et al. [18] studied the influence of the fuel composition parameters on particulate emissions where the data were fitted along with the engine torque and speed using an ANN. Their model was also able to estimate emissions within 87-90% of confidence for one single operating mode.

To simulate conventional and hybrid vehicles and predict their emissions, Clark et al. [19] used a neural network emission model based on simulated engine speed and torque combined with ADVISOR (ADvanced VehIcle SimulatOR) software package. The emissions predicted by the ANN showed good correlation with the emissions data from the chassis tests using the WVU Transportable Heavy-Duty Vehicle Emissions Testing Laboratory (THDVETL).

Yuanwang et al. [20] examined the application of a backpropagation neural network to predict exhaust emissions including HC, CO, PM and NO_x. Cetane number

was selected as an input and the effects of cetane improver and nitrogen content were also analyzed.

Desantes et al. [21] suggested a mathematical model to correlate NO_x and PM as a function of engine operating parameters. A multi-layer perceptron with a backpropagation learning algorithm was selected for this research. By examining a wide range of inputs, they concluded that exhaust gas recirculation (EGR) rate, fuel mass and start of injection were the most relevant parameters for NO_x, PM emissions and also BSFC (Brake Specific Fuel Consumption). BSFC showed how efficiently the engine used the fuel to produce power. Their model proved to be successful in maintaining the emission values below the obligatory level and minimize BSFC.

Hafner et al. [22] described a modern approach towards a model-based optimization of internal combustion engine control maps. They tried to reduce fuel consumption and exhaust emissions while achieving a good driveability. Therefore, the structure of a torque-oriented engine management system based on control maps was described first. Then in order to develop an optimization tool, an ANN-based model of engine emissions was used to compute the basic control maps for the engine settings. The results in both simulation and measurements proved the quality of the proposed technique.

Tehranian [23] studied the ANN design, architecture, and activation function on the accuracy of emissions predictions. She found that the 3-Layer and Jump Connection neural networks with radial basis functions such as Gaussian were the best architectures to predict five emissions of NO_x, PM, HC, CO, and CO_2. The data were obtained from a 550 hp General Electric DC engine dynamometer-testing unit at the West Virginia

University Alternative Fuels, Engine and Emissions Research Center. The engine was exercised through engine transient test schedules of the ETC, FTP, E-CSHVR, E-Highway and E-WVU-5 Peak schedules. The E-CSHVR, E-Highway and E-WVU-5 were chassis test schedules which were adapted to be performed on a dynamometer.

2. ARTIFICIAL NEURAL NETWORKS

Artificial Neural Networks (ANN) with their outstanding ability to obtain meaning from complicated data can be used to identify patterns that are too complicated to be solved by other computer techniques. The other advantages of ANN can be classified as:

1. Adaptive learning: ANN is able to learn how to perform tasks based on the data presented for training.
2. Real Time Operation: ANN computations may be carried out in parallel. Therefore, special hardware devices are being designed and manufactured which take advantage of this capability.
3. Self-Organisation: An ANN is able to create its own organization or representation of the information it receives during learning time.
4. Fault Tolerance: Partial destruction of a network causes the corresponding degradation of performance but some network capabilities may be maintained even with major network damage [24].

2.1. History

The history of neural networks has developed through conceptual innovations and implementations, two necessary ingredients for the advancement of a technology [25]. Some of the neural network developments occurred in the late 19^{th} and early 20^{th}

centuries in physics, psychology and neurophysiology. Alexander Bain, William James, Hermann von Helmholtz, Ernst Mach and Ivan Pavlov were the pioneers of ANN [26].

The modern view of neural networks began in 1940s by Warren McCulloch and Walter Pitts [27] who showed that networks of artificial neurons could compute any arithmetic or logical function. Their work is often recognized as the origin of the neural network field. In 1949 Hebb [28] formed the basis of 'Hebbian learning,' which is now considered as an important part of ANN theory. The concept of 'Hebbian learning' is based on the principle that every time a neural connection is used, the path is strengthened.

The first practical application of artificial neural networks was introduced in 1958, with the concept of the perceptron network by Rosenblatt [29]. He and his colleagues introduced the perceptron network and showed its ability to perform pattern recognition. Bernard Widrow and his graduate student Marcian Hoff, suggested the ADALINE (ADAptive LInear NEuron) network and LMS (Least Mean Square) algorithm as its learning rule. ADALINE network was very similar to the perceptron, except for its transfer function that was linear.

In 1969 Minsky and Pappert published a book [30], which showed that the perceptron developed by Rosenblatt had serious limitation. They assumed this limitation is a common feature of all neural networks. As a result of this publication, the amount of research work on ANN was reduced for the next 10 years.

However, some important work continued during the 1970s. Kohonen [31] and Anderson [32] independently developed the linear associator neural networks that could act as memories in 1972.

During the 1980s with the aid of powerful computers and new concepts, research in neural network increased dramatically. Two new concepts helped for the rebirth of the neural networks [25]. First was using the statistical mechanics to describe the operation of a specific class of recurrent network that could be used as an associative memory. This concept was introduced by Hopfield [33] in 1982. The other concept was the backpropagation algorithm for training multilayer perceptron networks. This algorithm was the answer to the Minsky's criticism. Rumelhart and McClelland [34] were the pioneers of the backpropagation neural network.

Today there are two classes of ANN patterns, supervised and unsupervised [35]. The backpropagation network is the most popular example of the supervised network. It is very useful to model the nonlinear transfer function between several inputs and one or more outputs. Carpenter and Grossberg [36] proposed the adaptive resonance theory that is an unsupervised network with adaptive architecture. Kohonen's algorithm is another well-known unsupervised neural network [26].

ANNs are being used extensively due to their ability to provide a rapid solution to non-trivial problems. They are not universal solutions to all problems [35]. They are alternative mathematical devices for processing the data.

2.2. The Biological Neuron

The ANN has been developed by inspiration of the characteristics of brain function. The brain consists of approximately 10^{11} highly connected (10^4 connections per element) elements. The neurons have three components: the dendrites, the cell body and the axon. The dendrites are receptive networks of nerve fibers that carry electrical signals

into the cell body. The cell body sums and thresholds these incoming signals. The axon is a single long fiber that carries the signal from the cell body to the other neurons. Synapse is the point of contact between an axon of one cell and a dendrite of another cell. The arrangement of neurons and the strength of the synapses determine the function of the neural network. Figure 2.1 shows the feature of biological neurons and their connections.

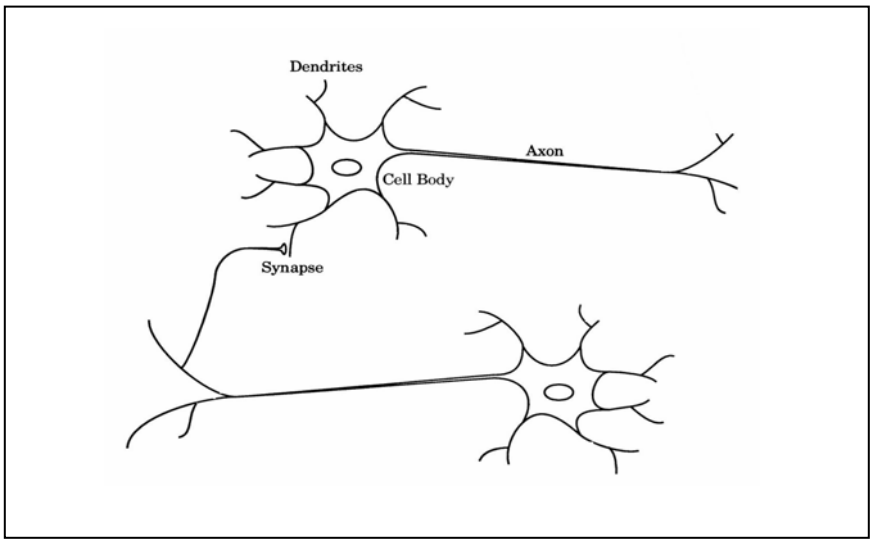


Figure 2.1. Biological Neurons Feature [25]

Part of the neural structures is defined at birth and continues to change during the human life. The similarities between the artificial and biological neural network can be summarized in two main characteristics. First, both networks consist of simple elements, which are connected extensively. Second, the function of the network is determined by the connections of the neurons. Compared to electrical circuits, the biological neurons are very slow. However, the brain performs many tasks much faster than any computer. This is because of parallel structure of the biological neural network and the high number of neural connections.

2.3. The Artificial Neuron

2.3.1. Single-Input Neuron

Figure 2.2 shows a single-input neuron. The scalar input (p) multiplies by weight (w) and sums up with bias (b). The summer output (n) is the net input that goes to the transfer function (f) and produces the scalar output (a).

$$a = f(wp+b)$$ ***Equation 2.1***

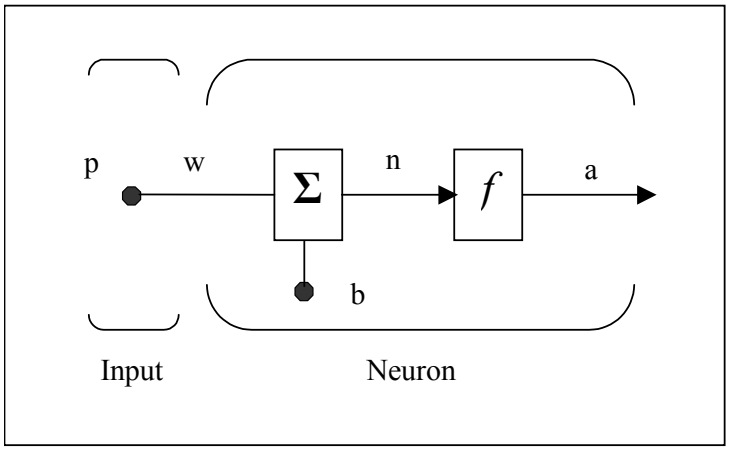


Figure 2.2. Single-Input Neuron

Relating the biological neuron to this model, the w corresponds to the power of the synapse, the summation is representative of the cell body and the transfer function and neuron output represent the axon.

w and b are adjustable scalars that will be varied by the learning rule but the transfer function will be selected by the designer to meet the desired output.

2.3.2. Transfer Functions

The transfer function may be a linear or non-linear function of the net input (*n)*. The way the networks produce their outputs is by applying a transfer function to the sum of the weighted inputs [37].

2.3.2.1. Gaussian

This transfer function maps high values into low ones, and mid-range values into high ones. It helps identifying meaningful characteristics not found at the extreme ends of the sum of weighted inputs. Output of this function is [0,1]. The following transfer functions are summarized in Table 2.1.

$$a = f(n) = \exp(-n^2) \qquad \textbf{Equation 2.2}$$

2.3.2.2. Gaussian Complement

Gaussian complement function reveals meaningful characteristics in the extremes of the data.

$$a = f(n) = 1 - \exp(-n^2) \qquad \textbf{Equation 2.3}$$

2.3.2.3. Sigmoid Logistic

This function is useful for almost all neural network applications and maps the inputs into the (0,1) range. Logistic has always been used for categorized outputs.

$$a = f(n) = \frac{1}{1+\exp(-n)}$$ ***Equation 2.4***

2.3.2.4. Symmetric Logistic

Symmetric logistic function maps the data to (-1,1). This function is mostly been used in hidden and output layers.

$$a = f(n) = \frac{2}{1+\exp(-n)} - 1$$ ***Equation 2.5***

2.3.2.5. Hyperbolic Tangent

Tanh is useful for continuous valued outputs. To use it in the first hidden layer, the inputs must be scaled into [-1,1].

$$a = f(n) = \frac{\exp(n) - \exp(-n)}{\exp(n) + \exp(-n)}$$ ***Equation 2.6***

2.3.2.6. Tanh 1.5n

This function is expressed with the following equation and it is much better than tanh function in some cases.

$$a = f(n) = \frac{\exp(1.5n) - \exp(-1.5n)}{\exp(1.5n) + \exp(-1.5n)}$$ ***Equation 2.7***

2.3.2.7. Sine

The inputs must be scaled into [-1,1] if this function has been used in the first hidden layer.

$a = f(n) = \sin(n)$ ***Equation 2.8***

2.3.2.8. Linear

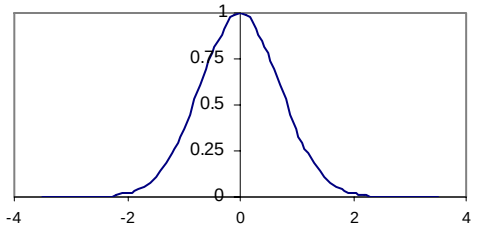
It is useful for problems where the output is a continuous variable. The linear function prevents the network from producing outputs with more error near the min or max of the output scale. It is better to use smaller learning rates, momentums, and initial weight sizes with this transfer function. Otherwise, the network may produce larger errors and weights which consequently does not minimize the error.

$a = f(n) = n$ ***Equation 2.9***

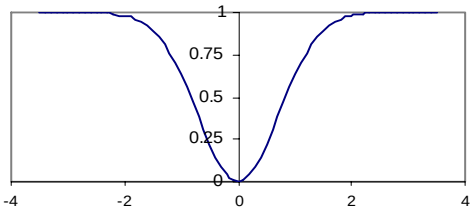
2.3.3. Neural Network Design

The process of designing a neural network is an iterative process and can be achieved through a series of trial and error steps to achieve a desirable network. Designing a neural network includes:

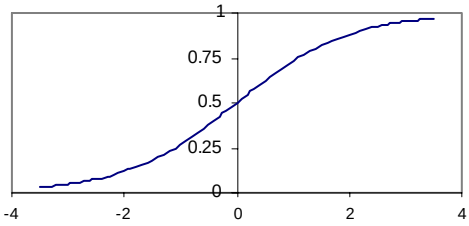
1. Selecting the type of neuron connections for different layers and also the connections among the neurons within a layer.
2. Determining the strength of connection within the network by allowing the network to learn the appropriate values of connection weights by using a training data set [38].
3. Choosing the suitable transfer function or in the other word the way the neuron produces output



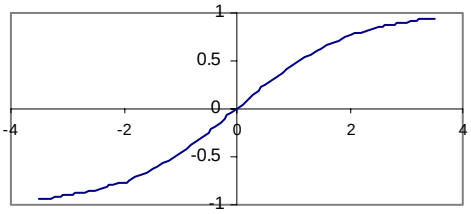
Gaussian



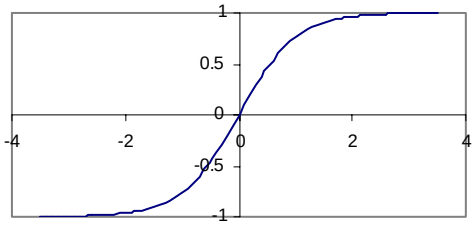
Gaussian Complement



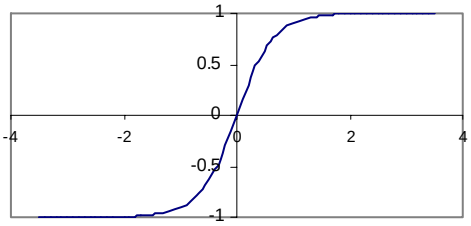
Logistic



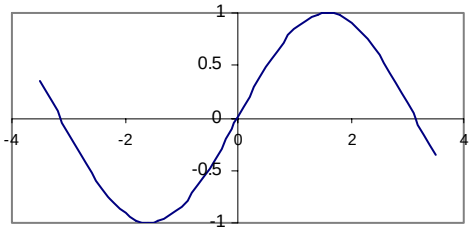
Symmetric Logistic



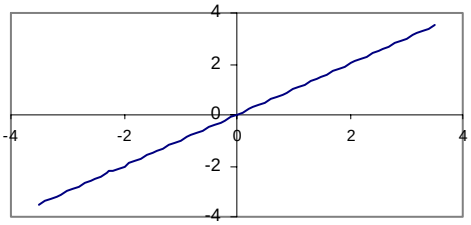
Tanh



Tanh 15



Sin



Linear

Table 2.1. Transfer Functions

2.3.3.1. Network Layers

Usually a neuron has more than one input; however one neuron with even many inputs is not sufficient. Therefore, a layer of neurons working in parallel is needed for complicated problems. The layer includes the weight matrix, the summers, the bias vector, the transfer functions and the output vector. Figure 2.3 shows a layer of S neurons. However, most of the neural networks consist of several layers. The output layer is the layer whose output is the neural network output. The layer of the network inputs is called input layer and the other layers are considered as hidden layers.

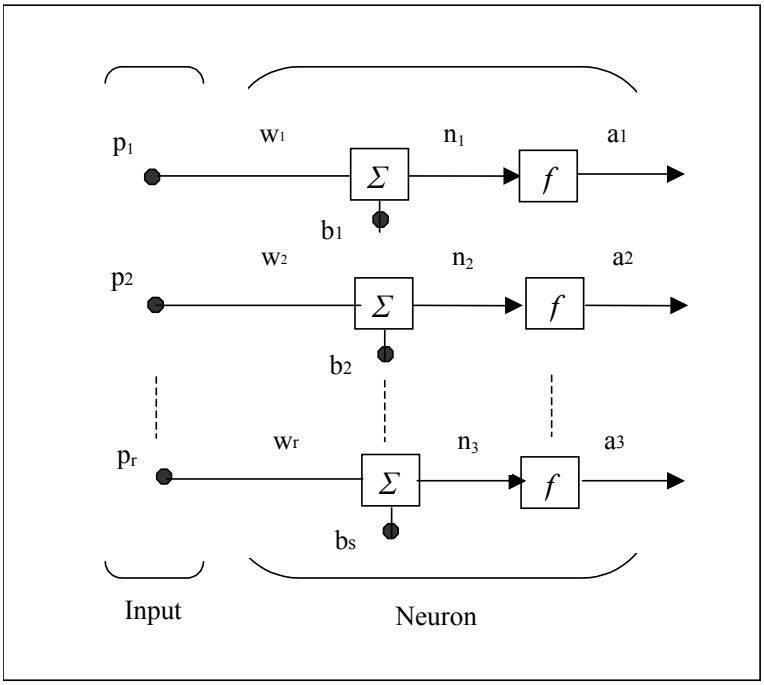


Figure 2.3. Layer of S Neurons

Hidden layers can be more than one layer in the ANN architecture. Figure 2.4 shows the layers of a neural network in which every layer is connected to the next.

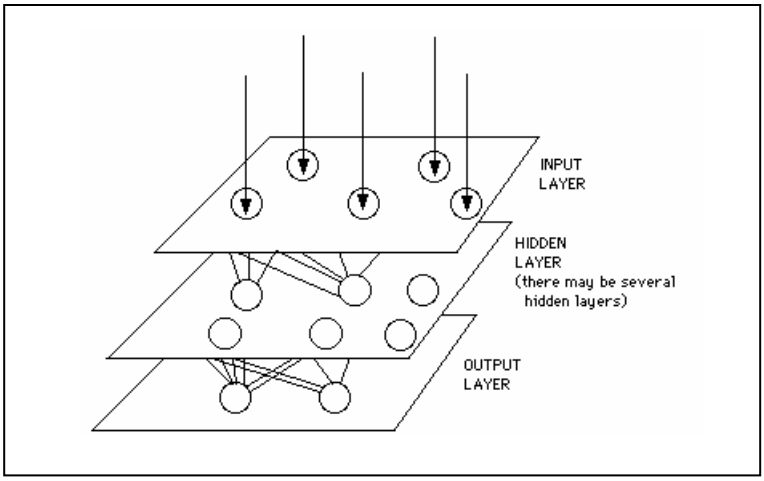


Figure 2.4. Layers of Neural Network

2.3.3.2. Types of Connections

The neurons in a layer can be connected either within the layer or between the layers. The connections of the neurons within the layers are called intra-layer connections however these connections between the layers are identified as inter-layer connections.

There are two types of intra-layer connections: 'Recurrent' and 'On-center/Off surround' [38]. In 'Recurrent' connection, the neurons in one specific layer communicate their outputs with each other within the layer before they send their outputs to another layer. Generally some conditions among the neurons of the layer should be achieved before they communicate their outputs to another layer. However, in 'On-center/Off surround' connections, a neuron within a layer has excitatory connections to itself and its immediate neighbors, and has inhibitory connections to other neurons. In the excitatory connection, the output of one neuron boosts the performance of the neuron that receives the signal but in inhibitory connections of the neurons, the output of the neuron sending a message would reduce the activity of the receiving neuron. Consequently, after a while

the neuron with excitatory connections can update its weight and also weights of its linked neurons.

Different types of inter-layer connections consist of ‘Fully connected,’ ‘Partially connected,’ ‘Feed forward,’ ‘Bi-directional,’ ‘Hierarchical’ and ‘Resonance.’ In a ‘Fully connected’ connection each neuron on the first layer is connected to all neurons of the second layer but in ‘Partially connected’ each neuron on the first layer is connected to the part of the neurons on the second layer. The neurons in ‘Feed forward’ connections send their outputs to the next layer’s neurons and do not receive any input back but in ‘Bi-directional’ another set of connections bring the outputs back from the neurons of the next layer to the previous layer’s. These two types of connections can be ‘Fully or Partially connected.’ ‘Hierarchical’ is a type of structure that the neurons of a lower layer could only communicate with neurons on the next level of layer and ‘Resonance’ presents a kind of bi-directional connection that the layers send signals to each other until a certain condition is achieved.

2.3.3.3. Learning

Like the brain that learns from experience, the ANN learns the solution to a specific problem by adjusting the connection weights (training algorithm). Two features determine the learning ability of a neural network, its architecture and training algorithm. The training algorithm usually consists of:

1. Supervised learning: Supervised neural networks are trained to achieve desired outputs in response to specific inputs. These networks are mostly suitable for

predicting future events, modeling and controlling dynamic systems, and classifying noisy data [39]. The connections among the neurons in the hidden layer are randomly arranged, then a training set of data or an observer identifies how close it is to desired outputs and accordingly rearranges the network connections. One of the most commonly used supervised neural network model is the 'Backpropagation' network. Error information is fed back through the system and is used to adjust the connections between the layers. The followings are examples of supervised networks: 'Radial basis' networks that provide an alternative technique for designing non-linear feed-forward networks, 'Recurrent' networks that use feedback to identify spatial and temporal patterns, and 'Feed-forward' networks which are suitable for pattern recognition, prediction, and nonlinear function fitting.

2. Reinforced learning: This learning is similar to supervised learning, but instead of being provided with correct output for each network input, the algorithm is only given a grade. This grade identifies the network performance. Reinforced learning is not very common and it is most useful for control system applications [40].
3. Unsupervised learning: Unsupervised learning method is not given any target value. A desired output of the network is unknown. During training the network performs some kind of data compression such as dimensionality reduction or clustering. The network learns the distribution of patterns and makes a classification of that pattern where, similar patterns are assigned to the same output cluster. Kohonen network is the best example of unsupervised learning network.

There are a number of mathematical algorithms which are used to update the connections weight and perform network training. Some of them are listed as below.

2.3.3.3.1. Perceptron Learning Rule

This learning rule is an example of supervised training. A set of network behavior is defined as:

$$\{\mathbf{p}_1,\mathbf{t}_1\}, \{\mathbf{p}_2,\mathbf{t}_2\}, \ldots, \{\mathbf{p}_Q,\mathbf{t}_Q\} \qquad \textit{Equation 2.10}$$

where $\mathbf{p}_q$ is an input vector and $\mathbf{t}_q$ is the corresponding target output of the network. The network output ($\mathbf{a}_q$) is compared to the target and then the weights and biases of the network are modified by learning rule to achieve an output close to the target. If the scalar perceptron error is defined as:

$$e = t\text{-}a \qquad \textit{Equation 2.11}$$

The matrix notation of the perceptron learning rule can be written as:

$$\mathbf{W}^{new} = \mathbf{W}^{old} + \mathbf{e}\,\mathbf{p}^T \qquad \textit{Equation 2.12}$$

$$\mathbf{b}^{new} = \mathbf{b}^{old} + \mathbf{e} \qquad \textit{Equation 2.13}$$

where **W** is a weight matrix, **'new'** refers to the computed weight matrix / bias in each step, **'old'** is the weight matrix / bias in the immediate previous step, **e** is an error vector and **b** is network bias.

2.3.3.3.2. Hebbian Learning Rule

The well-known Hebbian rule is the origin of the majority of these learning algorithms. The Hebbian learning rule explains that "When an axon of cell A is close

enough to excite cell B persistently, some growth process occurs in one or both cells that increases the effectiveness of cell A" [25]. To interpret this rule in mathematical form, it should be noticed that the connection between inputs and outputs is the weight matrix. Consequently, if a positive input produces a positive output then the elements of weight matrix should increase.

$$\mathbf{W}^{new} = \mathbf{W}^{old} + \alpha\, \mathbf{t}_q\, \mathbf{p}_q^{T} \qquad \textit{Equation 2.14}$$

Note that α is the learning rate.

2.3.3.3.3. Least Mean Square Learning Rule

The least Mean Square (LMS) or delta rule is an example of supervised training rule that adjusts the weights and biases of the network in order to minimize the mean square error. Existence of a unique minimum point depends on the characteristics of the input vectors. The LMS rule can be identified as:

$$\mathbf{W}\,(k+1) = \mathbf{W}\,(k) + 2\,\alpha\,\mathbf{e}\,(k)\,\mathbf{p}^{T}(k) \qquad \textit{Equation 2.15}$$

$$\mathbf{b}\,(k+1) = \mathbf{b}\,(k) + 2\,\alpha\,\mathbf{e}\,(k) \qquad \textit{Equation 2.16}$$

2.3.3.3.4. Backpropagation Learning Rule

Similar to the LMS rule, backpropagation (BP) rule is an approximate steepest decent algorithm in which the performance index is mean square error. The only difference between LMS and backpropagation is in the way that their derivatives are computed. The backpropagation term comes from the concept of the sensitivities, s^{m}. The sensitivities are propagated backward through the network from the last layer (M) to

the first layer. The sensitivity of transfer function F to changes in the *i*th element of the net input (n) at layer m is defined as:

$$s_i^m = \frac{\partial F}{\partial n_i^m} \qquad \textit{Equation 2.17}$$

$$n = \mathbf{W}p + \mathbf{b} \qquad \textit{Equation 2.18}$$

Therefore

$$s^m = -2\dot{F}^m(n^m)e \qquad \text{for } m = M \qquad \textit{Equation 2.19}$$

$$s^m = \dot{F}^m(n^m)(W^{m+1})^T s^{m+1} \qquad \text{for } m = M\text{-}1, \ldots 2, 1 \qquad \textit{Equation 2.20}$$

where,

$$\dot{F}^m(n^m) = \begin{bmatrix} \dot{f}^m(n_1^m) & 0 & \cdots & 0 \\ 0 & \dot{f}^m(n_2^m) & \cdots & 0 \\ \vdots & \vdots & & 0 \\ 0 & 0 & & \dot{f}^m(n_{S^m}^m) \end{bmatrix} \qquad \textit{Equation 2.21}$$

$n_{S^m}^m$ shows the *s* th input of layer m. Finally, the weights and biases are updated using approximate steepest descent rule [25]:

$$\mathbf{W}^m(k+1) = \mathbf{W}^m(k) - \alpha s^m (a^{m-1})^T \qquad \textit{Equation 2.22}$$

$$\mathbf{b}^m(k+1) = \mathbf{b}^m(k) - \alpha s^m \qquad \textit{Equation 2.23}$$

2.3.3.3.5. Kohonen Learning Rule

This learning is inspired by learning in biological systems and is implemented in the unsupervised methods of learning. The weights of a neuron learn an input vector in Kohonen rule. This learning is suitable for recognition applications. Learning happens when the neuron's index *i* is a member of the set X(*q*). The Kohonen rule can be written in vector notation as:

$$ {}_i\mathbf{w}(q) = {}_i\mathbf{w}(q\text{-}1) + \alpha\,(\,\mathbf{p}(q) - {}_i\mathbf{w}(q\text{-}1)\,) \quad \text{for } i \in X(q) \qquad \textit{Equation 2.24} $$

where ${}_i$w is a vector composed of the elements of the *i*th row of weight matrix.

$$ {}_i\mathrm{W} = \begin{bmatrix} \mathrm{W}_{i,1} \\ \mathrm{W}_{i,2} \\ \vdots \\ \mathrm{W}_{i,R} \end{bmatrix} \qquad \textit{Equation 2.25} $$

This rule also can be used with other identification which is its advantage. It is useful for training networks such as self-organizing neural network [25,38].

2.4. Backpropagation Neural Network

Generalizing the Widrow-Hoff learning rule, backpropagation neural network is created for multiple-layer networks with nonlinear differentiable transfer functions. The target vector and the associated input vector are used to train the network. Generally, it has been showed that 2-layer neural networks with biases, a hidden layer with sigmoid transfer function and the output layer with linear transfer function are able to model any output vector [41]. In BP neural network the weights for the nodes are selected arbitrarily, then the weights are modified according to the gradient of the performance

function. Sometimes, this modification points in the direction to the local minimum. Unfortunately, the local minimum is not always the global minimum, which causes the network to settle in a non-optimal configuration. To prevent the BP neural network from settling in local minima, the number of hidden layer nodes can be increased or decreased or even the network can be tried with different initial conditions. This is because the weights will be reinitialized to a different set of random numbers, which may keep them from falling into a local minimum that is not the global minimum.

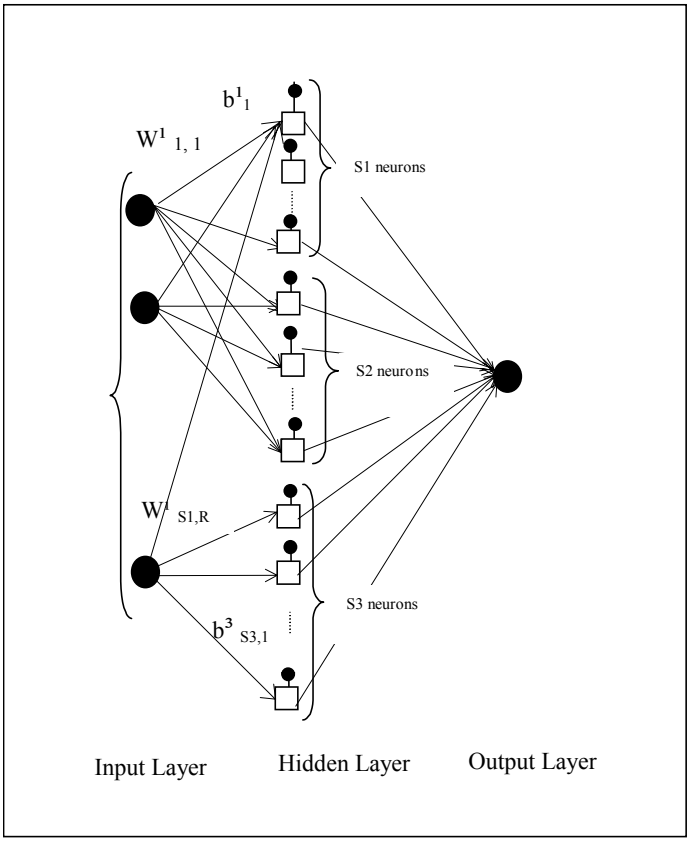


Figure 2.5. Backprobagation neural network. This illustration combines information from Neural Network Design Book and NeuroShell2 Users Manual [25,37].

2.4.1. Convergence

The study of performance function of a multilayer neural network shows that first, the initial parameters should not be set to zero. This is because the origin of the parameter space is inclined to be a saddle point for the performance surface. Second, the initial parameters should not be too large because as we move far away from the global minimum the performance surface would become very flat. Typically, the initial weights and biases are selected as small random values. It is also recommended to try several initial points in order to identify if it is converged to the global minimum or the local minimum [42]. For improvement of the BP neural networks, two methods of variable learning rate and momentum based on numerical optimization algorithms are generally being used [25].

2.4.2. Variable Learning Rate

To speed up the convergence, the learning rate can be increased on the flat area and then be decreased when the slope increase. The error performance of a single-layer linear network is a quadratic function and the maximum stable learning rate for the steepest decent algorithm is two divided by the maximum eigenvalue of the Hessian matrix. But the error surface of the multilayer is not quadratic and its slope is different in each region. By adjusting the learning rate considering the performance function shape, the convergence time can be decreased.

Vogl et al. [43] suggested a straightforward batching approach for adjusting the learning rate according to the performance of the algorithm. The rules of the Variable Learning rate BackPropagation (VLBP) are summarized to:

- If over the training set, the squared error increases by more than a specific percentage (ζ=1-5%) after the weight update, then the weight update is neglected, the learning rate is multiplied by a positive factor less than 1 ($0< \eta<1$) and the momentum coefficient (γ) is set to zero.
- If the squared error decreases after a weight update, then the weight is updated and the learning rate is multiplied by a factor more than 1 ($\eta>1$). If γ has been changed to zero formerly, it is updated to its original value.
- If the squared error increases (less than ζ), then the weight is updated but the learning rate does not change. If γ has been adjusted to zero, it is set to its original value.

This rule suffers from two main disadvantages while using more complex algorithms:

- The modification requires definition of several parameters such as γ, η and ζ. The performance of the algorithm depends on changes in these parameters.
- Sometimes, these modifications to the steepest descent BP cause failing of the convergence.

2.4.3. Momentum

The second method is using the momentum. By smoothing the oscillation of the squared error throughout the iterations, the convergence can be improved. This can be achieved by applying a low-pass filter. It is possible to use a larger learning rate while

maintaining the stability of the algorithm by using the momentum. The other advantage of momentum is that it accelerates the convergence.

$$\Delta W^m(k+1) = \gamma \Delta W^m(k) - (1-\gamma)\alpha s^m (a^{m-1})^T \qquad \textbf{Equation 2.26}$$

$$\Delta b^m(k+1) = \gamma \Delta b^m(k) - (1-\gamma)\alpha s^m \qquad \textbf{Equation 2.27}$$

2.5. Ward Neural Network Architecture

Backpropagation neural networks are proved to be able to generalize well on a wide variety of problems. Typically, training time depends on the number of patterns in BP neural networks and it may be slow but worth it because they are such global algorithms.

The neural network used in this research to predict the emissions was a 2-layer BP network created by Ward Systems Group (NeuroShell2, release 4.0 software). The hidden layer of this neural network was composed of different activation functions. The number of inputs determined the number of neurons in the input layer. The numbers of neurons in the hidden layer were a function of the number of inputs, outputs and training patterns (Equation 2.28).

Number of hidden neurons = 1/2 (Inputs + Outputs) + Square root of the number of the training patterns **Equation 2.28**

The number of hidden neurons determines how well the network can be trained. Too many neurons make the network to memorize the problem and consequently not

generalize well later. Too few neurons make the network generalize well but not learn the patterns competently. Output neurons were equal to the number of outputs.

Applying three different activation functions to the hidden layer identifies different features in each pattern and it offers three ways of viewing the data.

The learning rate and momentum for each link can be adjusted individually or they can be similar for all links.

Figure 2.6 shows the connections of the different layers. Neural network input is presented by Slab 1 but it is not considered as a layer according to ANN terminology. Slab 2,3 and 4 illustrate the hidden layer of the network. Due to applying three different activation functions on hidden layer, three different slabs were shown. Slab 5 refers to output layer and is considered as a layer.

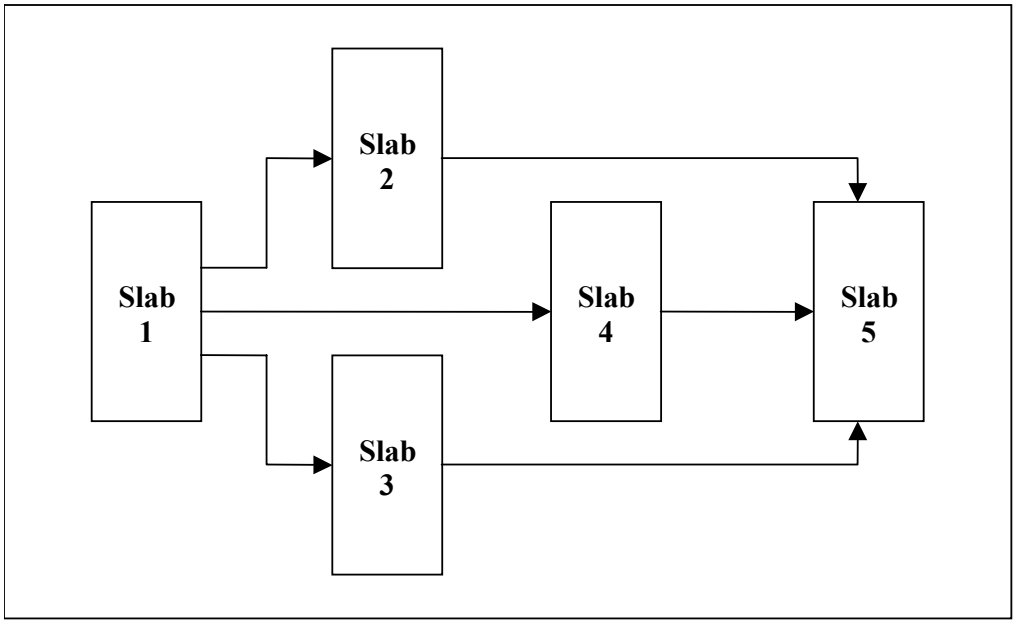


Figure 2.6. Connections of the layers in a 2-layer backpropagation neural network.

3. EXPERIMENATAL EQUIPMENT AND PROCEDURE

The primary objective of this research was to predict second by second heavy-duty vehicle emissions using ANN during the course of transient chassis tests of the vehicles while continuous gaseous emission data were collected. For this purpose, it was first required to train the ANN on actual emissions data in order to use the similar architecture as a predictive tool for the unseen data. All chassis dynamometer emissions data were collected using the West Virginia University (WVU) Transportable Heavy Duty Vehicle Emissions Testing Laboratories. Although the collection of data was not included in this graduate research effort, it is important to describe the acquisition of the data to place the thesis in context.

3.1. WVU Transportable Emissions Testing Laboratory Description

WVU, with cooperation of the Office of Transportation Technologies of the U.S. Department of Energy, designed the Transportable Heavy Duty Vehicle Emissions Testing Laboratories to measure the emissions from heavy-duty vehicles. The laboratories were able to perform transient and steady state chassis dynamometer emissions tests on vehicles and simulate a range of driving cycles to provide performance and fuel efficiency data.

The laboratories can be moved from site to site and cause the reduction in the test timing of the tested vehicles. WVU operated two heavy-duty chassis with twin flywheel sets and twin power absorbers for the vehicles from 20,000 to 80,000 lbs in Gross

Vehicle Weight (GVW). Also there was a medium-duty dynamometer for lighter vehicles (3,000 to 20,000 lbs). This dynamometer had a single-phase power absorber and a single dynamometer.

3.1.1. Heavy-Duty Vehicle Emissions Testing Laboratory

The heavy-duty dynamometer unit was composed of a power absorber for simulation of road load power and a set of flywheels (discs) for simulation of an inertia weight equal to a gross vehicle weight. During the test cycle, torque cell and speed transducers in the power absorber drive train measured the vehicle load and speed. Two tractor-trailers carried the laboratory to the test site. The emission measurement equipment, data acquisition and control were pulled by one of the trailers while the other carried the power absorber unit. At the test site, the dynamometer trailer wheels were removed. This laboratory was supplied by its power generator, which makes it able to operate when there was no power available at the test site. The vehicles were driven up to the dynamometer. The inner wheels were placed on rollers and the outer wheels on each side of the vehicle were removed and substituted by hub adapters. The hub adapters connected the drive axle to the power absorber units. A data acquisition and control system controled the power absorber unit power. To simulate the driving conditions, WVU has developed a wide variety of cycles to simulate various types of in-use vehicle operation [44]. The vehicle to be tested was driven through the test cycles by a human driver.

The vehicle exhaust mixed with air in a full-scale dilution tunnel, which was 18 inches in diameter and 20 feet in length. To guarantee the through mixing of exhaust

before reaching the sampling zone, an orifice was located in 3 feet from the dilution entrance. The flow rate in the tunnel was maintained by a critical flow venturi. There were five venturi setting flow rates of 1000, 1500, 2000, 2500, and 3000 standard cubic feet per minute (scfm).

Stainless steal heated sampling probes and heated sampling lines drew the samples from the dilution tunnel to the gas analyzers.

3.1.2. Particulate Matter Sampling

PM was measured gravimetrically according to the Code of Federal Regulation, Title 40, Part 50. A 3-inches diameter and 30-inches long secondary dilution tunnel was pulled out from the main dilution tunnel to carry out the exhaust to a stainless steel filter holder. The sample passed through a primary and a secondary Pallflex 70-mm fluorocarbon coated glass fiber filter, Model T60A20 with a filtration efficiency of 98% for particles larger than 0.1 micron. The secondary dilution tunnel maintained the double diluted exhaust at 50˚ C or less, instantly before the primary particle filter in the secondary dilution tunnel. The flow rate in secondary dilution tunnel was proportional to the flow rate of the primary dilution tunnel. Prior to being weighed on a Cahn microbalance, the filter media were maintained for at least one hour to 80 hours at 50% relative humidity and 25˚ C in an environmental chamber. The primary dilution air might be passed through a HEPA filtration unit before entering the dilution tunnel to remove ambient particles.

Continuous PM was measured using a Tapered Element Oscillating Microbalance (TEOM), Rupprecht & Patashnick Co., Inc. TEOM Series 1105, Diesel Particulate Mass

Monitor. The TEOM was connected to the secondary dilution tunnel. It was an instrument that drew a sample of diluted exhaust air through a filter, Pallflex TX40, at a constant flow rate. The filter was weighed continuously to find the real-time PM rate. The mass rate of the TEOM filter was measured every 0.42 seconds based on the natural frequency of oscillation [45]. The filter was connected to a hollow tapered element, vibrating at its natural frequency of oscillation. The natural frequency of oscillation changed by the mass change of the filter.

Due to sensitivity of the TEOM to the variables that occurred during the test such as temperature and pressure changes, external vibrations, filter face velocity, filter pressure drop and the water content of the filter, it was not an absolutely accurate real-time PM measuring device [46].

3.1.3. Hydrocarbon Analyzer

Hydrocarbon emissions were measured using a heated Flame Ionization Detector (FID). The HC analyzer consisted of 5 components including flame ionization burner, sample handling system, calibration gas system, oven, and burner-flameout/fuel shutoff circuitry. HC emissions were ionized and the ionization current was proportional to the number of carbon atoms that entered the burner. A pump drew sample to analyzer and through a glass fiber filter that removed the PM and it also helped to minimize the system response time. Non-methane hydrocarbon (NMHC), aldehyde and alcohol emissions were able to be measured by chromatography methods performed at WVU [47].

3.1.4. Oxides of Nitrogen Analyzer

This analyzer used the chemiluminescent method for the determination of NO or NO_x in a sample gas. In the NO mode, the NO was converted to NO_2.10%-15% of these NO_2 molecules were elevated to an electronically excited state followed by reversion to a non-excited state. This occurrence caused the radiation of the photons. A photodiode detector generated DC current proportional to the concentration of NO in the sampled gas. In the NO_x mode, the NO_2 was converted to NO. This analyzer consisted of heated sections, which maintained the sample above 50˚ C. Two separate chemiluminescent analyzers allowed the NO and NO_2 fractions to be determined [48].

3.1.5. Carbon Monoxide (CO) and Carbon Dioxide (CO_2) Analyzers

These analyzers measured the gas concentration based on the principle that each type of gas absorbs unique line spectrum in the infrared region. Carbon monoxide (CO) and carbon dioxide (CO_2) analyzers were Non-Dispersive Infrared (NDIR) analyzers. The CO_2 analyzer contained two cells, a sealed reference cell and flow through sample cell. The difference between the amounts of infrared energy absorbed by the cells showed the concentration of CO_2. The sample passed through a filter and a dryer to remove water and eliminate its absorption issue in the CO_2 and CO measurements. Low CO analyzer measured the low CO concentrations. The CO analyzer shared the same sample line, filter, and dryer as the CO_2 analyzer [49, 50].

3.1.6. Data Acquisition

The data collected by gas analyzers, pressure transducers, and thermocouples was acquired using a RTI-815 digital to analog data acquisition board, a 3-B signal conditioning module and a computer. The data was recorded as analog to digital conversion (ADC) and during the data reduction was converted to conventional units. Data reduction was performed using Microsoft visual basic software according to the CFR, Title 40, Part 86.

3.2. Medium-Duty Vehicle Emissions Testing Laboratory

The medium-duty vehicle emissions laboratory was used to test the vehicles with gross weight of 3,000 to 20,000 pounds. It consisted of a chassis dynamometer test bed and incorporated with the instrumentation trailer of the heavy-duty laboratory as described earlier. This dynamometer was transported to the site using a flatbed tractor truck and a crane mounted on this truck was used to place the dynamometer on the ground. Similar to the heavy-duty dynamometer, the medium-duty dynamometer unit consisted of flywheels and a power absorber. The vehicle to be tested was positioned on the dynamometer so that the drive axle of the vehicle was centered on the rollers. The medium-duty chassis dynamometer, dissimilar to the heavy-duty dynamometer, drew power from the rollers rather than from the drive shafts connected to the axle. The torque cell and speed transducers in-line with the power absorbers measured the vehicle load and speed during the test.

3.3. Transient Cycles and Routes

The test cycles were designed to simulate the real-life vehicle activity and to test the vehicle emissions under various conditions such as acceleration, deceleration, cruise, and idling. Heavy-duty vehicles were tested through engine or chassis dynamometer test cycles for emissions and fuel economy measurements [51]. Engine test cycles were schedules of speed and loads; however chassis test cycles were composed of speed versus time or speed versus distance traces. A cycle was a speed-time schedule however a route was a speed-distance schedule and it was not dependent on time [52]. A diver with visual prompt via computer monitor in the cab was able to incorporate the effects of gear shifting to the desired vehicle speed at each second. A route allowed the vehicle full acceleration, which was possibly representative of real-life driving conditions [53].

In this thesis, the emissions data were gathered from three chassis dynamometer test schedules: the City Suburban Heavy Vehicle Route (CSHVR), Highway, and EPA Urban Dynamometer Driving Schedule (UDDS) or Test D.

3.3.1. City Suburban Heavy Vehicle Route (CSHVR)

The City Suburban Heavy Vehicle Route (CSHVR) chassis dynamometer test was developed by the West Virginia University [54]. The route was achieved by changing the route speed to a function of distance traveled and assigning maximum acceleration rates to portions of the route. The following are CSHVR test schedule features:

- Duration: 1700 seconds with the average speed of 14.15 mph
- Total distance: 6.68 miles
- Maximum speed: 43.84 mph

- Average speed: 14.15 mph

Vehicle speed versus the distance for the CSHVR test schedule is shown in Figure 3.1.

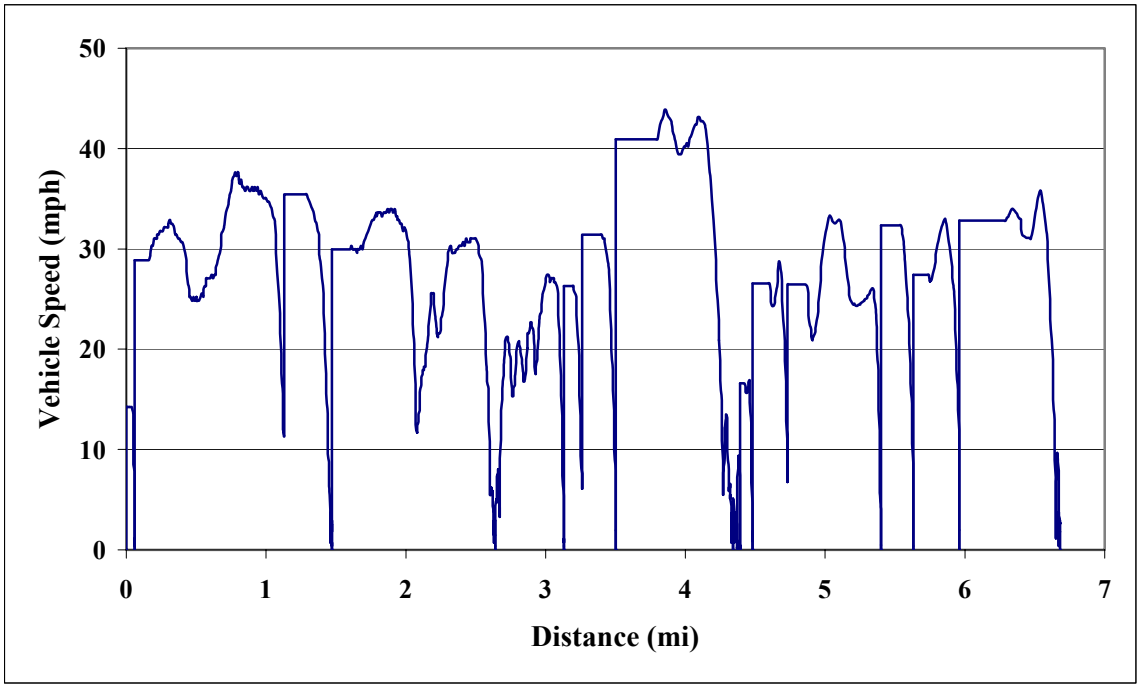


Figure 3.1. Vehicle speed versus distance for the CSHVR test schedule.

3.3.2. Highway Cycle

Based upon the study performed by Clark et al. [54] in which the CSHVR test schedule was developed, the Highway cycle was also developed for the operations of the tractors in highways. The Highway cycle characteristics are:

- Duration: 1640 seconds
- Total distance: 15.51 miles
- Maximum speed: 60.7 mph
- Average speed: 34.05 mph

Figure 3.2 shows the vehicle speed versus time for the Highway test schedule.

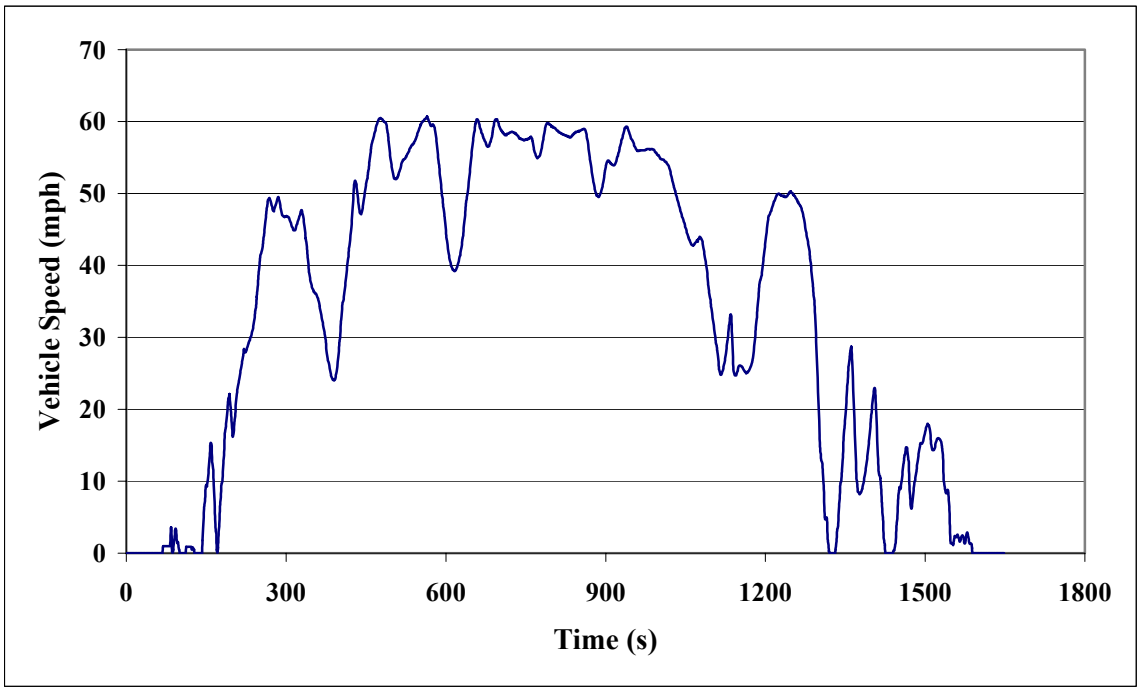


Figure 3.2. Vehicle speed versus time target trace for the Highway test schedule.

3.3.3. Urban Dynamometer Driving Schedule (UDDS)

The Urban Dynamometer Driving Schedule (UDDS) was designed by EPA to simulate the heavy-duty vehicle operation in urban areas (CFR 40, 86, App. I). It was also known as "Test D". Thirty-three percent of the cycle period was idle. The following are UDDS test schedule features:

Duration: 1060 seconds

Total distance: 5.55 miles

Maximum speed: 58 mph

Average speed: 18.86 mph

Vehicle speed versus time for the UDDS test schedule is shown in Figure 3.3.

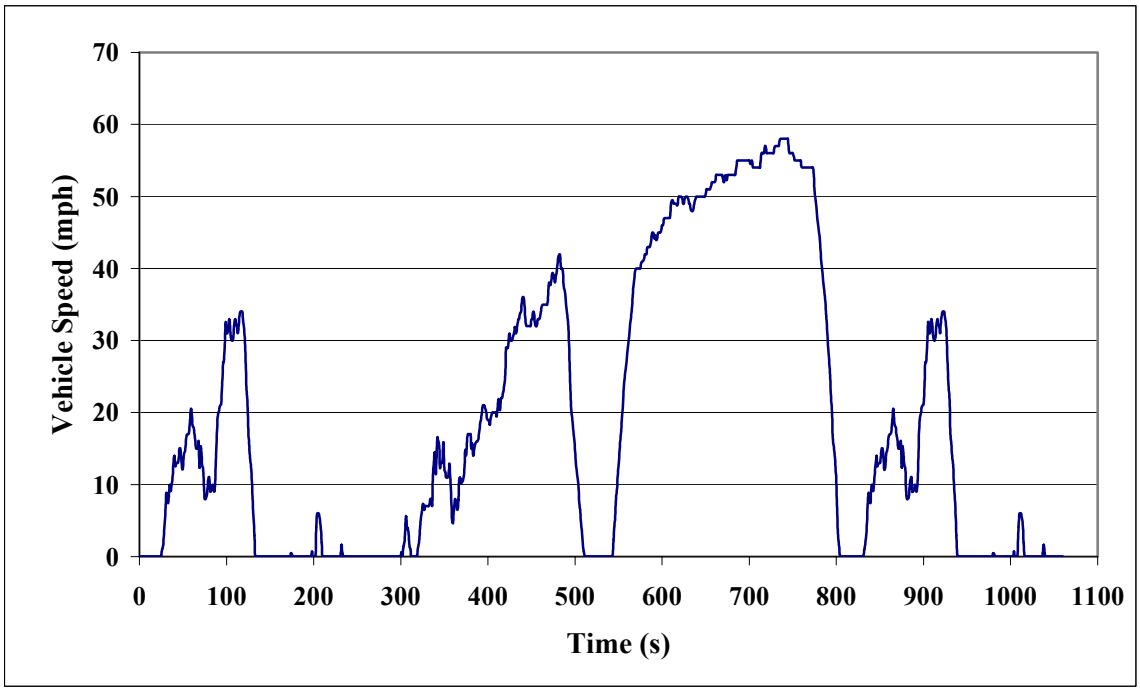


Figure 3.3. Vehicle speed versus time target trace for the UDDS.

4. DATA PROCESSING

4.1. PM Split Study

Due to inaccuracy in the engine-based emissions inventory that was used to represent mobile source emissions, California has turned to chassis dynamometer based emissions data for emissions factor input. Although not a part of this new thrust, the PM split study was performed to gather emissions data from thirty-four heavy-duty diesel vehicles [55]. NO_x, HC, CO, and CO_2 were measured from vehicles operating in southern California. This study [56] was undertaken using the West Virginia University Transportable Heavy-Duty Vehicle Emissions Testing Laboratories described above. The tested vehicles were 2 transit buses, 16 trucks with Gross Vehicle Weight (GVW) over 33000 lbs., 8 trucks with GVW of 14001 to 33000 lbs., and 8 trucks with GVW of less than 14001 lbs. The City/Suburban Heavy Vehicle Route (CSHVR), a Highway cycle, and an Idle period were used to measure emissions from the vehicles. The buses were tested through a Manhattan cycle, an Idle period and CSHVR. Also, a number of these vehicles were tested under cold start idle, cold start CSHVR, and the Urban Dynamometer Driving Schedule (UDDS).

4.2. Data Description

From the PM split study database, the group of vehicles in the 33,000-80,000 lb weight range was selected for examining in this thesis. There were six tractor trucks in this group, and their characteristics are listed in Table 4.1. To perform the power

dispersion modeling, three of them (Vehicles 1, 2 and 6) with different engine manufacturers were considered.

Vehicle	Type	Year	Manufacturer	Odometer Mileage	GVW (lb)	Vehicle Tested Weight (lb)	Engine	Displacement (liter)
Vehicle 1	Tractor Truck	1985	Freightliner	769413	80,000	42,000	Caterpillar 3406B	14.6
Vehicle 2	Tractor Truck	1994	Freightliner	639105	80,000	42,000	Detroit Diesel Series 60	12.7
Vehicle 3	Tractor Truck	1998	Sterling	327300	80,000	42,000	Detroit Diesel Series 60	12.7
Vehicle 4	Tractor Truck	1999	Sterling	272307	80,000	42,000	Caterpillar C-12	12
Vehicle 5	Tractor Truck	2000	Sterling	255880	80,000	42,000	Caterpillar C-12	12
Vehicle 6	Tractor Truck	2001	Volvo	327300	80,000	42,000	Cummins N14-370	14

Table 4.1. Vehicle and engine descriptions

Vehicles 1 and 2 were exercised through 3 different driving test schedules, the City/Suburban Heavy Vehicle Route (CSHVR), Highway cycle and Heavy-Duty Urban Dynamometer Driving Schedule (UDDS or Test D). Vehicle 6 was exercised only through CSHVR and Highway cycles. Appendix A presents the detailed information on the vehicles.

4.3. Cross-Correlation

In order to find the time delay between the actual emission production instance and emissions measuring time, the cross-correlation was performed for the power and emissions data series as a function of time. The power was selected due to the relationship of this parameter to the engine emissions [57]. Figure 4.1 shows the relationship between the power and the vehicle emissions along with the time delay. This time delay includes transport time in the dilution tunnel, sampling line and also the analyzers' response time.

Cross-correlation is a standard method of estimating the degree to which two data series are correlated. The cross-correlation product (represented by $\otimes$) of two complicated functions *f(t)* and *g(t)* of a real variable *t* can be identified by several possible functions. The sum of squared differences method can be used when the data series are of the same units [58].

$$f \otimes g = \int_{-\infty}^{+\infty} (f(\tau) - g(t-\tau))^2 d\tau \qquad \textbf{Equation 4.1}$$

The other method is termed the sum of products:

$$f \otimes g = \int_{-\infty}^{+\infty} f(\tau) g(t-\tau) d\tau \qquad \textbf{Equation 4.2}$$

These methods may be used to align two signals in time when one signal suffers a time offset relative to the other during data logging. The first method describes a measure of difference between the two functions, which is supposed to be minimized to find the best correlating point. However, the second principle uses a quantity that has to be maximized in order to find the target point.

Cross-correlation has been performed by both methods (Equations 4.1 and 4.2) using MATLAB software. In order to use the sum of squared differences method, the

power (hp or kW) and its relevant emissions data (g/sec) at each second must be dimensionless since they are in different units. Therefore, the data have been made dimensionless by dividing them by their average value over the entire test schedules. The two methods agreed very well and confirmed one another. The average shifting time was 8-9 seconds for NO_x, 12-14 seconds for CO and 13-15 seconds for CO_2. HC shifting time varied in different cycles. Since HC may not increase monotonically with power, its time alignment can be difficult to find.

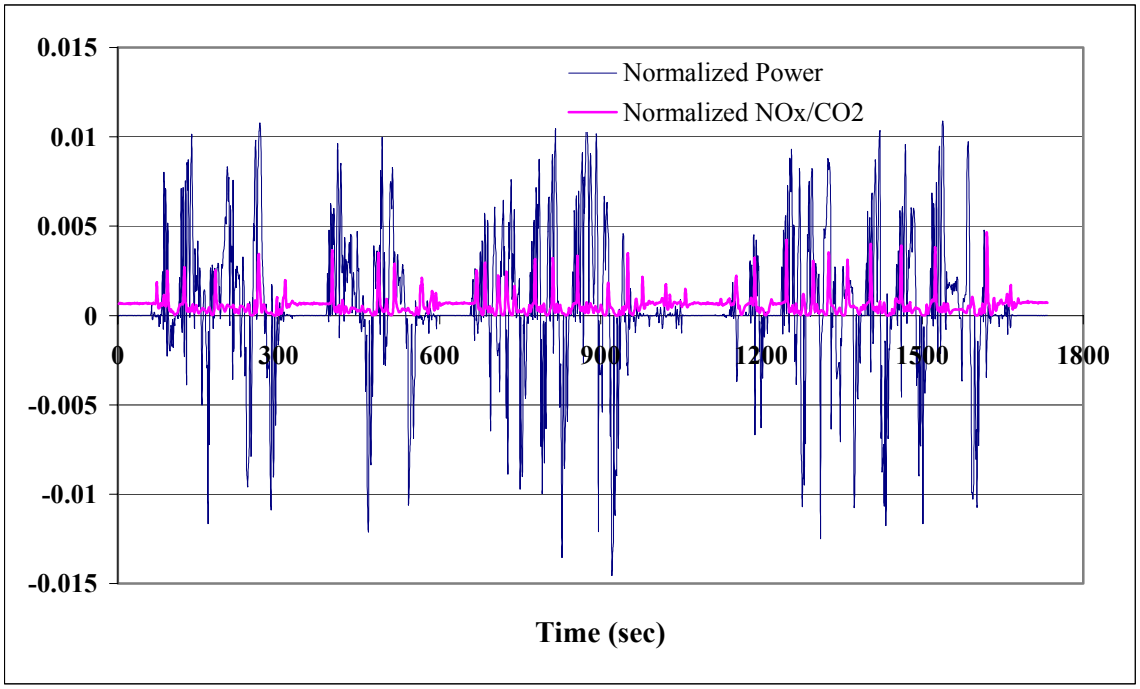


Figure 4.1. Normalized values of NO_x/CO_2 and power versus time.

4.4. Dispersion

To measure emissions, a standard full-scale dilution tunnel and emissions analyzers are required. When the exhaust gas passes through the measurement system, gas elements take different routes through the dilution tunnel and sampling lines. This

phenomenon is called axial dispersion and it causes the delay for the emissions to pass through the system and be measured. In addition, the time taken by the analyzer to respond and measure the data also contributes to this delay. Considering the known relationship between the power and emissions, Clark et al. [59] applied a dispersion model to the axle power to find the emissions dispersion model. Therefore, the power was also assumed to be dispersed to correlate the power and the emissions appropriately.

Levenspiel [60] suggested a dispersion model that is applicable to turbulent flow in pipes and laminar flow in long tubes or channels. The dispersion coefficient, D, represents the spreading process of the flow; the large quantity of D shows rapid spreading in time and vice versa. To characterize the distribution, the centroid of the distribution, $\bar{t}$, must be described.

$$\bar{t} = \frac{\int_0^\infty t\, C\, dt}{\int_0^\infty C\, dt} \qquad \textit{Equation 4.3}$$

For small quantities of dispersions (D/UL<0.01), the spreading curve is symmetric and belongs to a family of gaussian curves and represents by the Equation 4.4.

$$C_i = \frac{1}{2\sqrt{\pi(D/UL)}} \exp\left(\frac{-(1-\theta_i)^2}{4(D/UL)}\right) \qquad \textit{Equation 4.4}$$

where θ_i is $t_i/\bar{t}$ and D/UL is the Vessel Dispersion Number (VDN) that is a dimensionless group characterizing the spread in the tunnel. U is the average velocity of the fluid in a tunnel of length L. If D/UL>0.01 there is a nonsymmetrical curve. In this case, regarding the boundary condition situation two possible models exist: Undisturbed flow as it passes the entrance and exit boundaries and plug flow outside the vessel up to the boundaries.

The first condition represents a commonly used experimental device and is defined as [60]:

$$C_i = \frac{1}{2\sqrt{\pi\theta_i (D/UL)}} \exp\left(\frac{-(1-\theta_i)^2}{4\theta_i (D/UL)}\right) \qquad \textit{Equation 4.5}$$

Baskaran et al. [61], by injecting CO_2 into the tunnel for a period of 4 seconds, showed that the analyzer response had a delay to detect the injection. It also illustrated the diffusion of the input signal. Their dispersion model presented the starting dead time and a gamma distribution shape that represented the response of well-mixed tanks in series. The values of C_i were normalized in their model.

Jarrett [53] added two variables of 'a' and 'b' to Equation 4.4 and tried to minimize the actual measured emissions and the suggested emission dispersion model by modifying the variables of 'a,' 'b' and VDN. However in his model the sum of the C_i was not equal to 1 and its associated error was about 4%.

4.5. Power Dispersion Modeling

In this research, to find the best power dispersion model, the linear correlation between dispersed power and measured CO_2 was set to be optimum in relation to different values of D/UL. In most cases, the cross-correlation of the power and CO_2 estimated the time delay to be 15 seconds, therefore the centroid of distribution was assumed to be located at 15 seconds. t_i was defined over a wide period of 31 seconds to show the absolute nature of delay which consisted of onset time and dispersion time and

$C_i \in [1,31]$ relevant to t_is. Table 4.2 shows the values of $C_i \in [1,31]$ relevant to t_i which were normalized to 1.

In this work, it was assumed that the dispersion phenomenon has been started at the moment of the point injection but the mathematical model (Figure 4.2) showed that the C_is of the first 10 seconds after injection were extremely low (Table 4.2) which can be easily neglected. Inspection of the C_i values revealed that the beginning of the dispersion was actually occurred at t=11 and lasted for 13 seconds. The C_is relevant to the last 8 seconds of the studied period were also very small. The actual values of the C_i versus t_i are shown in Figure 4.2. The first 10 and the last 8 dispersion coefficients are so small that can not be seen in figure.

C_1	7.4e-130	**C_{12}**	6.4e-02	**C_{23}**	2.1e-03
C_2	1.2e-56	**C_{13}**	1.2e-01	**C_{24}**	8.5e-04
C_3	1.0e-32	**C_{14}**	1.6e-01	**C_{25}**	3.2e-04
C_4	4.3e-21	**C_{15}**	1.7e-01	**C_{26}**	1.1e-04
C_5	2.1e-14	**C_{16}**	1.5e-01	**C_{27}**	4.1e-05
C_6	3.7e-10	**C_{17}**	1.1e-01	**C_{28}**	1.4e-05
C_7	2.5e-07	**C_{18}**	7.6e-02	**C_{29}**	4.6e-06
C_8	2.2e-05	**C_{19}**	4.4e-02	**C_{30}**	1.4e-06
C_9	5.4e-04	**C_{20}**	2.3e-02	**C_{31}**	4.5e-07
C_{10}	4.9e-03	**C_{21}**	1.1e-02		
C_{11}	2.3e-02	**C_{22}**	5.0e-03		

Table 4.2. Normalized values of C_i relevant to t_i summed to 1.

The power was dispersed for different values of D/UL (0.085, 0.02 and 0.011). CO_2 versus dispersed power was mapped to find the best dispersion model. Tables 4.3 through 4.5 show the correlation coefficient (r^2) values. The results demonstrated the best

linear correlation for D/UL=0.011 which represented the onset time of 10 seconds and dispersion time of 13 seconds. In this case, the C_i curve followed the larger dispersion formula (D/UL>0.01) but it was very close to a Gaussian symmetric curve. Figure 4.3 shows the continuous C_i curves for different D/UL versus θ. Neglecting the C_i <0.001, θ $\in$[0.333 , 2.067] for D/UL=0.085 that shows the 5 seconds onset time with 26 seconds dispersion period, θ $\in$[0.533 , 1.800] for D/UL=0.02 which corresponds to 8 seconds onset time and 19 seconds dispersion period, and θ $\in$[0.667 , 1.533] for D/UL=0.011 which was described earlier as the best power dispersion model.

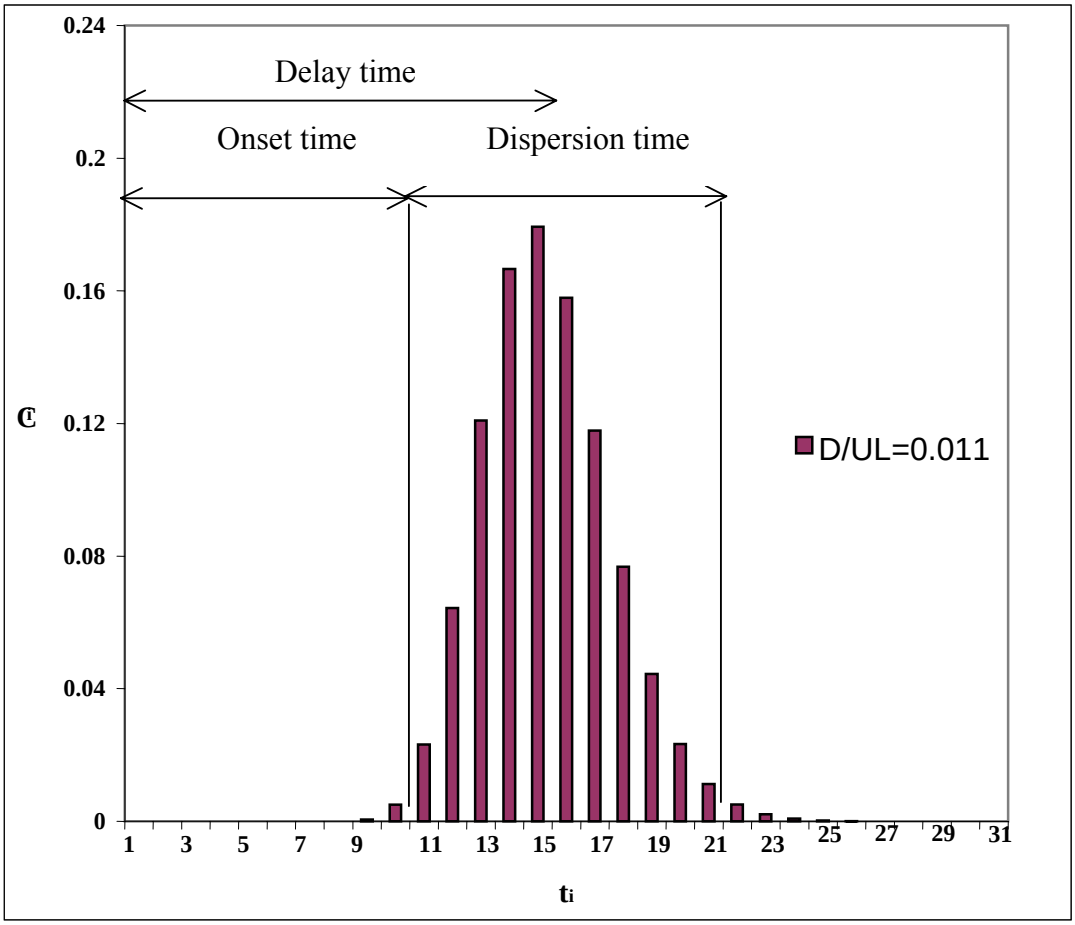


Figure 4.2. Normalized C_i relevant to t_i over a 31 second period shows an onset time of 10 seconds and dispersion period of 13 seconds.

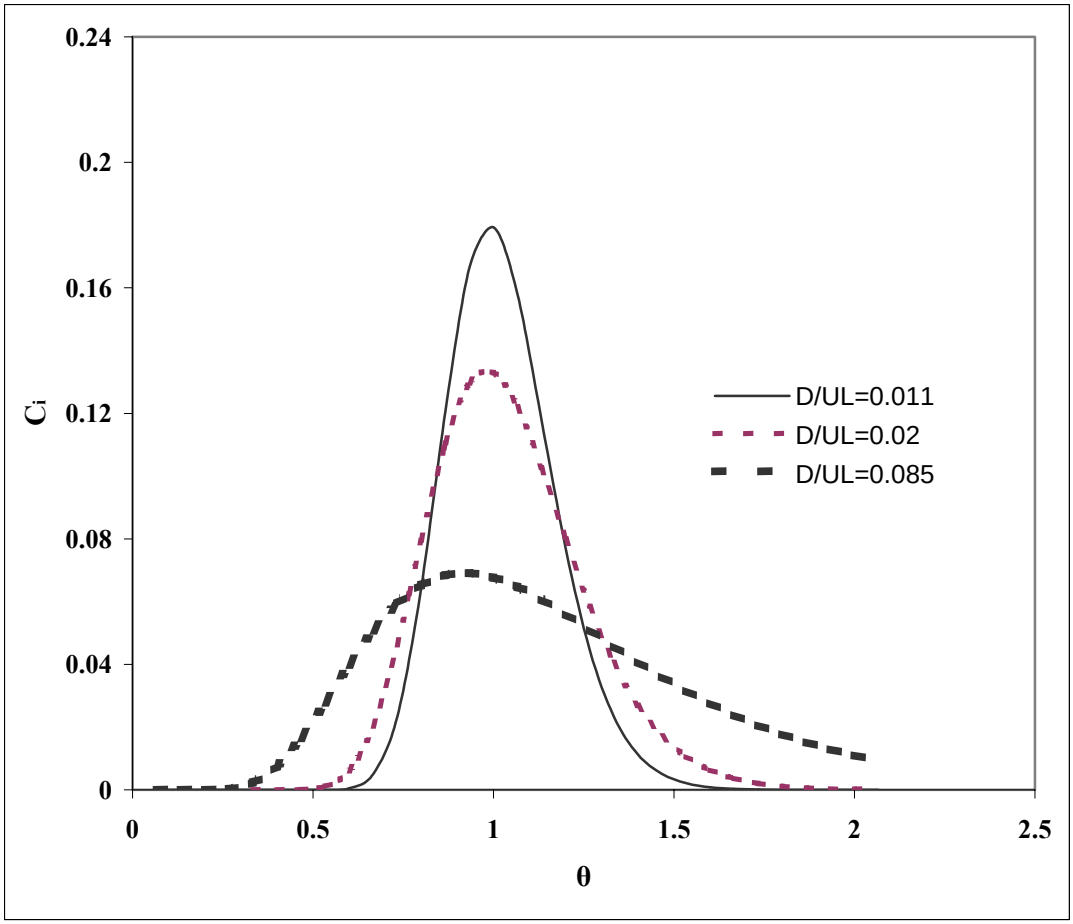


Figure 4.3. Dispersion of the point injection for three different values of D/UL (0.085, 0.02, 0.011) corresponding to Equation 4.5.

	D/UL=0.085	D/UL=0.02	D/UL=0.011
Vehicle 1	0.7956	0.8507	0.8507
Vehicle 2	0.7732	0.8552	0.8668
Vehicle 6	0.6640	0.8265	0.8684

Table 4.3. r^2 results of linear fit for CO_2 versus dispersed power through CSHVR.

	D/UL=0.085	D/UL=0.02	D/UL=0.011
Vehicle 1	0.8868	0.9262	0.9274
Vehicle 2	0.7885	0.8302	0.8131
Vehicle 6	0.8435	0.9002	0.9067

Table 4.4. r^2 results of linear fit for CO_2 versus dispersed power through Highway cycle.

	D/UL=0.085	D/UL=0.02	D/UL=0.011
Vehicle 1	0.8625	0.9040	0.9119
Vehicle 2	0.7848	0.8593	0.8908

Table 4.5. r^2 results of linear fit for CO_2 versus dispersed power through UDDS.

CO_2 arises in proportion to fuel used. Heat, auxiliary and axle power are the production of the fuel consumption. With axle power as the biggest part of the fuel consumption productions, the relationship between the CO_2 emissions and the axle power is apparent.

The instantaneous power was correlated with the vehicle emissions and then was dispersed using the above non-symmetric dispersion model. The dispersed power was linearly mapped versus CO_2 continuous data gathered by the analyzer, which had a significant contribution to the CO_2 dispersion. The results also confirmed the relationship between the measured CO_2 and the dispersed power distinctly (Tables 4.3 to 4.5). Figure 4.4 shows the linear fit of r^2=0.9067 with the D/UL=0.011 dispersion model for vehicle 6 exercised through the Highway cycle. Some scatter is due to varying the engine efficiency over the operating envelope and the consumption of fuel to run auxiliary devices such as the compressor and cooling fan. For instance, 1,000 rpm at 1,000 ft-lb of

torque may consume fuel at a different rate that 2,000 rpm at 500 ft-lb, but power is the same.

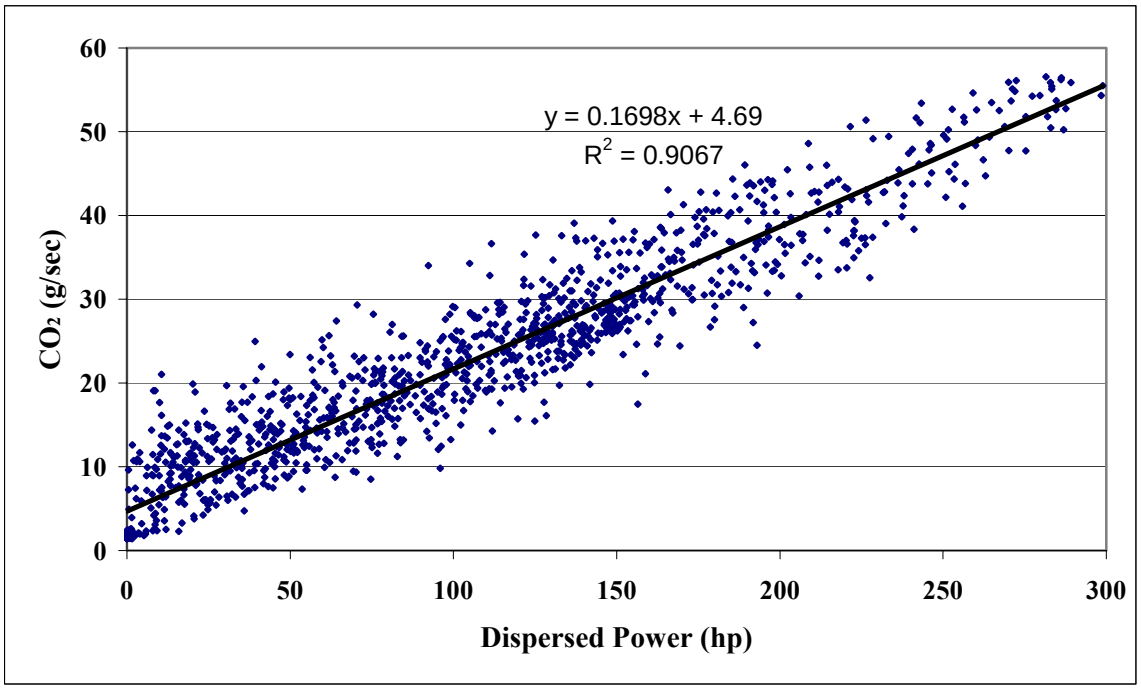


Figure 4.4. Linear fit of CO_2 versus dispersed power with D/UL=0.011 for vehicle 6 exercised through Highway cycle.

This model is capable of representing the onset time and the dispersion period at the same time. The delay time was defined as the duration of the onset time along with the time taken for the C_i curve to reach to its peak. This delay time consisted of transport time in the dilution tunnel, the sampling lines and the analyzers' response.

Due to the known correlation of the NO_x and CO_2 to power, their delay time could be found by this model easily. The CO and PM are considered to be affected by the rates of the power change along with power which represent the transient effects.

In this research, the power has been used to find the time delay for all of the considered emissions (HC, NO_x, CO_2 and CO). The emissions data were cross-correlated by the mentioned methods and dispersed before being applied to the ANN.

4.6. Neural Network Inputs

In this research, the ANN was trained individually for each emission species (NO_x, CO_2, CO and HC) as the network output. Considering the known relationship [62, 63] between the vehicle emissions and the power, the hub speed and torque were chosen for training the ANN as inputs.

$$P = (\mu M_v g + \tfrac{1}{2} \rho_a C_D A_v Spd^2) Spd \qquad \textit{Equation 4.6}$$

where μ is the coefficient of rolling resistance, M_v is the mass of the vehicle, ρ_a is the ambient air density, C_D is drag coefficient, A_v is frontal area of the vehicle and Spd is the vehicle speed.

The neural network was trained with three different sets of inputs (8, 14 and 20 inputs) derived from the hub speed and torque. The 20 input neural network consisted of the two variables of hub speed and torque and their first and second rates of change over three different time ranges (1,5,10 seconds) in addition to two new variables of speed defined as:

$$Diff = \sum_{\text{last } t_d \text{ sec.}} S^2 - \bar{Spd}^2 \qquad \textit{Equation 4.7}$$

where,

$$\bar{Spd} = \frac{\sum_{\text{last } t_d \text{ sec.}} S}{t_d} \qquad \textit{Equation 4.8}$$

Prediction of "off-cycle" emissions is confusing for ANNs. Variables such as Equation 4.7 and 4.8 that measure the steadiness of the speed through the test schedule are able to demonstrate if there is not a remarkable change in vehicle speed how well the ANN can predict the emissions. The above variables were calculated over three different

time periods (50, 100 and 150 seconds). Table 4.6 presents the inputs calculation technique of the inputs.

Input 1	Dispersed speed (S)	**Input 11**	$\Delta S/\Delta t$ (10s)
Input 2	Dispersed torque (T)	**Input 12**	$\Delta T/\Delta t$ (10s)
Input 3	$\Delta S/\Delta t$ (1s)	**Input 13**	$\Delta^2 S/\Delta t^2$ (10s)
Input 4	$\Delta T/\Delta t$ (1s)	**Input 14**	$\Delta^2 T/\Delta t^2$ (10s)
Input 5	$\Delta^2 S/\Delta t^2$ (1s)	**Input 15**	S^2 - Ave. S^2 (50s)
Input 6	$\Delta^2 T/\Delta t^2$ (1s)	**Input 16**	S^2 - Ave. S^2 (100s)
Input 7	$\Delta S/\Delta t$ (5s)	**Input 17**	S^2 - Ave. S^2 (150s)
Input 8	$\Delta T/\Delta t$ (5s)	**Input 18**	Sum (S^2 - Ave. S^2 (50s)) over the last 50s
Input 9	$\Delta^2 S/\Delta t^2$ (5s)	**Input 19**	Sum (S^2 - Ave. S^2 (100s)) over the last 100s
Input 10	$\Delta^2 T/\Delta t^2$ (5s)	**Input 20**	Sum (S^2 - Ave. S^2 (150s)) over the last 150s

Table 4.6. S and T present dispersed speed and torque consequently. dS /dt (1s) shows first derivative of dispersed speed over 1 second. d^2S/dt^2 (1s) shows second derivative of dispersed speed over 1 second. Inputs 18, 19 and 20 refer to Equation 4.7

The fourteen inputs of the hub speed and torque and their first and second rates of change for three different time periods (1,5,10 seconds) were applied to the 14 input neural network.

Also the 8 input neural network employed the speed and torque, their first and second rates of change at one second and Diff and $\bar{S}pd$ in 150 seconds as the network inputs. Table 4.7 summarizes three different neural network architectures that have been used in this research.

	20 Input Network	14 Input Network	8 Input Network
Input 1*	X	X	X
Input 2	X	X	X
Input 3	X	X	X
Input 4	X	X	X
Input 5	X	X	X
Input 6	X	X	X
Input 7	X	X	
Input 8	X	X	
Input 9	X	X	
Input 10	X	X	
Input 11	X	X	
Input 12	X	X	
Input 13	X	X	
Input 14	X	X	
Input 15	X		
Input 16	X		
Input 17	X		X
Input 18	X		
Input 19	X		
Input 20	X		X

Table 4.7. Applied neural network architectures with three different sets of inputs. * See Table 4.6.

4.7. Relative Contribution Factors

To show the effect of a particular input in neural network training, the strength of each input relevant to the other inputs was considered. Research was initiated with the 20 input neural network. The architecture of this network consisted of one hidden layer with different activation functions in each hidden slab. The linear transfer function was chosen for the input. Tanh15, tanh and symmetric logistic were selected for the Slabs 2, 3 and 4 of the hidden layer respectively. Logistic transfer functions have been chosen for slab 5.

Each slab in the hidden layer had 15 neurons and every input connected to each neuron by a weight. The output layer was a one-neuron layer relevant to the desired emission species. In further steps, the number of inputs was decreased to those specific inputs that contributed more to the prediction results. Each slab of the hidden layer of the 14 input neural network was composed of 14 neurons. However, the 8 input neural network had 13 neurons in each hidden layer slab. The study performed on all the vehicles for NO_x emissions (Table 4.8). The results showed that the dispersed speed was the most controlling input in the neural network training for all the emissions prediction cases. The second derivative of torque, first derivatives of speed and torque, and the Diff variable were the next most powerful inputs respectively. Figure 4.5 shows the average relative contribution of each input to predict NO_x emissions of the vehicles using the 20 input neural networks. Since there may be an overlap of information between some of these variables such as T and $\Delta S/\Delta t$ (1s), $\Delta S/\Delta t$ (5s) and $\Delta S/\Delta t$ (10s), Figure 4.5 may not ultimately classify the relative importance of each variable.

	Vehicle 1	Vehicle 2	Vehicle 3	Vehicle 4	Vehicle 5	Vehicle 6	Average
Input 1	0.087	0.104	0.105	0.104	0.089	0.091	0.0973
Input 2	0.053	0.045	0.057	0.046	0.041	0.042	0.0477
Input 3	0.070	0.055	0.062	0.063	0.052	0.047	0.0587
Input 4	0.053	0.046	0.041	0.055	0.079	0.072	0.0583
Input 5	0.052	0.046	0.043	0.030	0.046	0.040	0.0433
Input 6	0.082	0.050	0.057	0.090	0.068	0.059	0.0680
Input 7	0.051	0.053	0.044	0.043	0.046	0.041	0.0468
Input 8	0.037	0.038	0.050	0.040	0.052	0.040	0.0433
Input 9	0.058	0.059	0.044	0.043	0.051	0.048	0.0510
Input 10	0.049	0.047	0.033	0.057	0.059	0.055	0.0505
Input 11	0.053	0.053	0.052	0.054	0.052	0.050	0.0529
Input 12	0.028	0.033	0.038	0.033	0.034	0.029	0.0328
Input 13	0.039	0.033	0.043	0.029	0.030	0.032	0.0348
Input 14	0.038	0.034	0.035	0.040	0.036	0.033	0.0363
Input 15	0.036	0.044	0.042	0.033	0.031	0.045	0.0388
Input 16	0.038	0.039	0.044	0.041	0.035	0.036	0.0395
Input 17	0.035	0.048	0.046	0.046	0.043	0.052	0.0452
Input 18	0.043	0.052	0.048	0.053	0.049	0.055	0.0503
Input 19	0.049	0.056	0.042	0.053	0.052	0.062	0.0529
Input 20	0.039	0.056	0.064	0.037	0.046	0.060	0.0507

Table 4.8. Relative contribution of the 20 input neural network to predict NO_x emissions of the vehicle. * See Table 4.6.

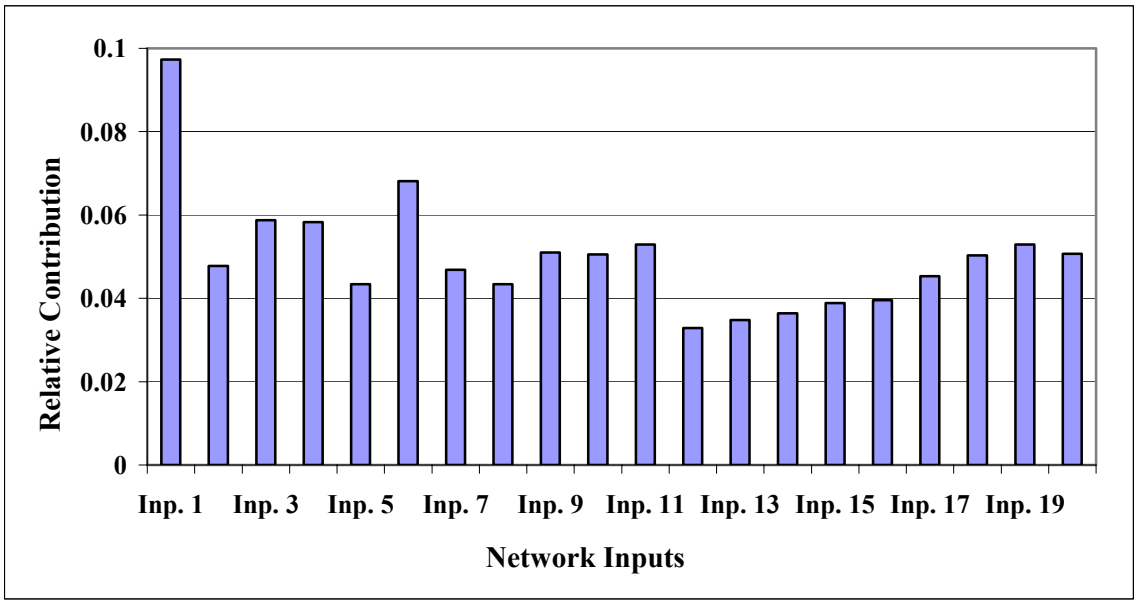


Figure 4.5. Average relative contribution of each input to predict NO_x emissions using the 20 input neural networks.

5. RESULTS AND DISCUSSION

Based on the association of the vehicle power (separated into speed and torque) and the emissions, the test schedules with a wide range of speed and torque were able to predict the emissions more effectively than the test schedules with a lower range of speed and torque [64]. Therefore among the available test schedules, the neural network was trained on the Highway cycle using different inputs and the trained network was used to predict the CSHVR and UDDS emissions. The neural network was also trained on CSHVR and UDDS test schedules and applied to the other two cycles. The ANNs that were trained on Highway cycle yielded better results. These results also confirmed the assumption of using the test schedules with a wide range of speed and torque. The networked was trained for every vehicle independently and the trained network was applied to the other schedules of the same vehicle. The results showed outstanding improvement in emission prediction in respect to the previous study at WVU performed by Clark et al. [64]. The new defined variables of Diff and $\overline{Spd}$ maintained a record of the constancy of the speed additional inputs to the neural network and helped to improve the results.

5.1. NO_x prediction

It was important to find an effective NO_x prediction neural network using a limited number of inputs to model the measured NO_x. Therefore, the first step was to find the inputs, which had larger contribution in neural network training and consequently

achieving the desired output. The desired output in this case was the measured NO_x (which was already naturally dispersed). Accordingly the neural network which have never seen the actual high rates of torque and speed change might not be able to predict real-life emissions.

The neural network was trained on Highway cycle of each vehicle and applied to the other test schedules related to the same vehicle. D/UL=0.011 yielded the best power dispersion results in chapter 4. Therefore the network inputs, continuous speed and torque, were dispersed using this dispersion number. Their first and second derivatives at 1, 5 and 10 seconds along with Diff and $\overline{Spd}$ variables at time ranges of 50, 100 and 150 seconds were computed. In 20 the input neural network, the linear fits of the predicted versus measured NO_x yielded an average R^2 of 0.97 for the Highway cycles. Applying the trained network on the other test schedules showed an average R^2 of 0.82 and 0.85 for the CSHVR and UDDS respectively.

The other concern in this work was the contribution of the Diff and $\overline{Spd}$ variables to network training. Therefore, the 14 input neural network was used to examine the prediction accuracy while these variables were excluded from the previous neural network (20 input neural network). The results are shown in Table 5.1 through Table 5.3.

At the same time, to investigate the effect of using the suggested dispersion model in Chapter 4, the same 14 input neural network with the same initial learning rate, momentum and weights was employed to predict NO_x. The only difference was that the inputs of this neural network were preprocessed by the other dispersion model studied by Jarrett [53].

Knowing the relationship between the vehicle steady speed [62] and NO_x emissions, to obtain an accurate neural network prediction with few inputs, the research was directed to use mostly the speed-dependant inputs. Therefore, the research was performed with the 8 input neural network including the Diff and $\bar{Spd}$ variables in the last 150 seconds. The motivation to choose the 150 seconds was due to its excellent contribution to the ANN emissions prediction compared to the 50 and 100 seconds alternatives. The Diff variable worked as a momentum and made the fluctuations of the speed smoothed throughout the test schedule. This variable was able to create a steady speed-dependant input for the neural networks. The momentum conception of the Diff variable also showed that choosing a wider time frame for this variable can improve the ANN prediction because the emissions were directly relevant to the steady speed. However if the time frame was selected too wide, it was impossible to follow the speed changes. Therefore, it was very critical to find an appropriate time frame for this variable.

	NO_x			
	20 inputs	**14 inputs (preprocessed by Jarrett [53] dispersion model)**	**14 inputs**	**8 inputs**
Vehicle 1	0.810	0.790	0.811	0.830
Vehicle 2	0.754	0.807	0.825	0.827
Vehicle 3	0.687	0.730	0.732	0.643
Vehicle 4	0.926	0.944	0.933	0.918
Vehicle 5	0.902	0.928	0.907	0.887
Vehicle 6	0.854	0.924	0.888	0.854

Table 5.1. ANN NO_x emissions prediction results using different numbers of inputs for all 6 vehicles exercised through the CSHVR. The neural network was trained on the Highway cycle.

	NO_x			
	20 inputs	**14 inputs (preprocessed by Jarrett [53] dispersion model)**	**14 inputs**	**8 inputs**
Vehicle 1	0.970	0.971	0.962	0.963
Vehicle 2	0.967	0.937	0.935	0.946
Vehicle 3	0.986	0.971	0.972	0.982
Vehicle 4	0.974	0.952	0.934	0.958
Vehicle 5	0.978	0.939	0.935	0.943
Vehicle 6	0.975	0.870	0.855	0.931

Table 5.2. ANN NO_x emissions prediction results using different numbers of inputs for all 6 vehicles exercised through the Highway cycle. The neural network was trained on the same cycle.

	NO_x			
	20 inputs	**14 inputs (preprocessed by Jarrett [53] dispersion model)**	**14 inputs**	**8 inputs**
Vehicle 1	0.891	0.918	0.918	0.912
Vehicle 2	0.814	0.907	0.900	0.847

Table 5.3. ANN NO_x emissions prediction results using different numbers of inputs for all the vehicles exercised through the UDDS. The neural network was trained on the Highway cycle.

Tables 5.1 to 5.3 show that the predictions were significantly improved for the CSHVR and UDDS test schedules using 8 inputs with respect to 20 and 14 input networks. The 20 input network yielded better prediction while it was trained on the Highway cycle. However, the prediction was not remarkable for the UDDS and the CSHVR because of the overtraining on the Highway cycle. Table 5.4 shows the integration of the actual and ANN predicted instantaneous NO_x over the CSHVR,

Highway and UDDS test schedules along with their percentage error. Figures 5.1 to 5.3 show the neural network NO_x prediction for the Highway cycle and the CSHVR. Investigating of the Figure 5.2 shows the warm-up effect of the engine on the neural network predictions. There is a significant difference between the ANN predicted and actual NO_x at the beginning of the test schedule which decreases throughout the schedule. The UDDS prediction is presented in Appendix B.

		Vehicle 1	Vehicle 2	Vehicle 3	Vehicle 4	Vehicle 5	Vehicle 6
CSHVR	**Actual**	247	239	323	94.9	97.1	137
	NN	232	189	235	93.6	88.2	132
	Error (%)	5.98	20.9	26.9	1.39	9.15	3.33
Highway	**Actual**	496	402	733	201	197	253
	NN	497	408	733	200	196	244
	Error (%)	-0.084	-1.37	-0.092	0.288	0.723	3.59
UDDS	**Actual**	203	199	_	_	_	_
	NN	200	170	_	_	_	_
	Error (%)	1.13	14.4	_	_	_	_

Table 5.4. Integration of the actual and ANN predicted instantaneous NO_x over the CSHVR, Highway and UDDS test schedules along with their percentage error. Positive and negative percentage errors represent over- and under-prediction respectively. The 8 input neural network was used.

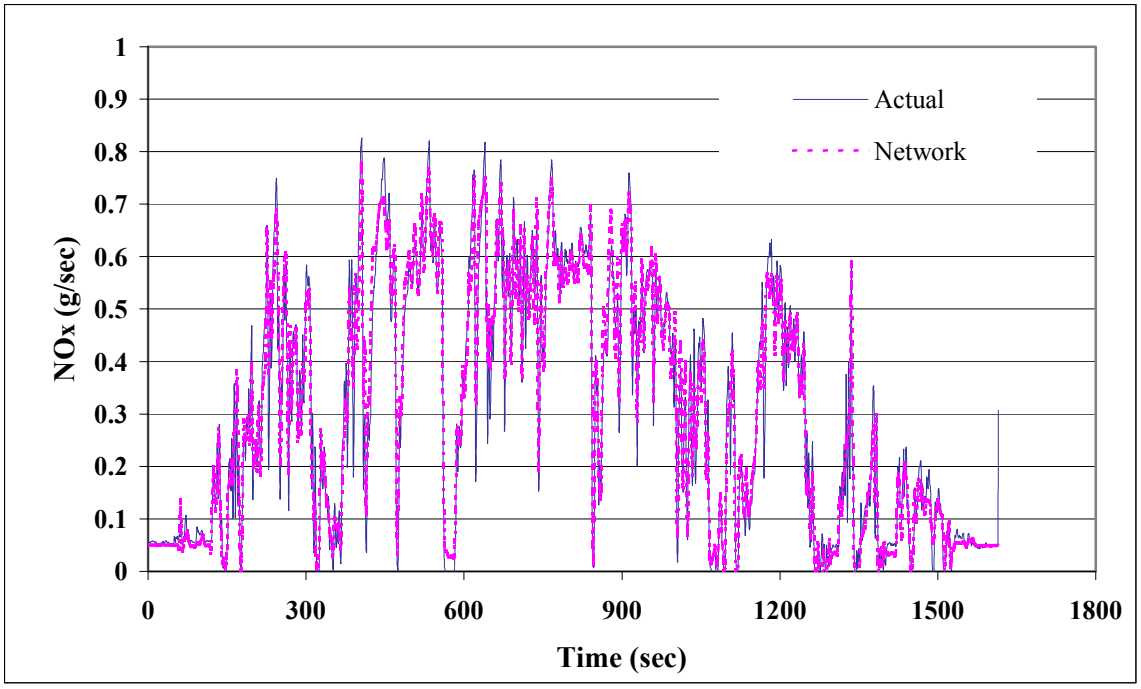


Figure 5.1. Actual and ANN prediction NO_x emissions using 8 inputs for Vehicle 1 exercised through the Highway cycle. The neural network was trained on the same cycle.

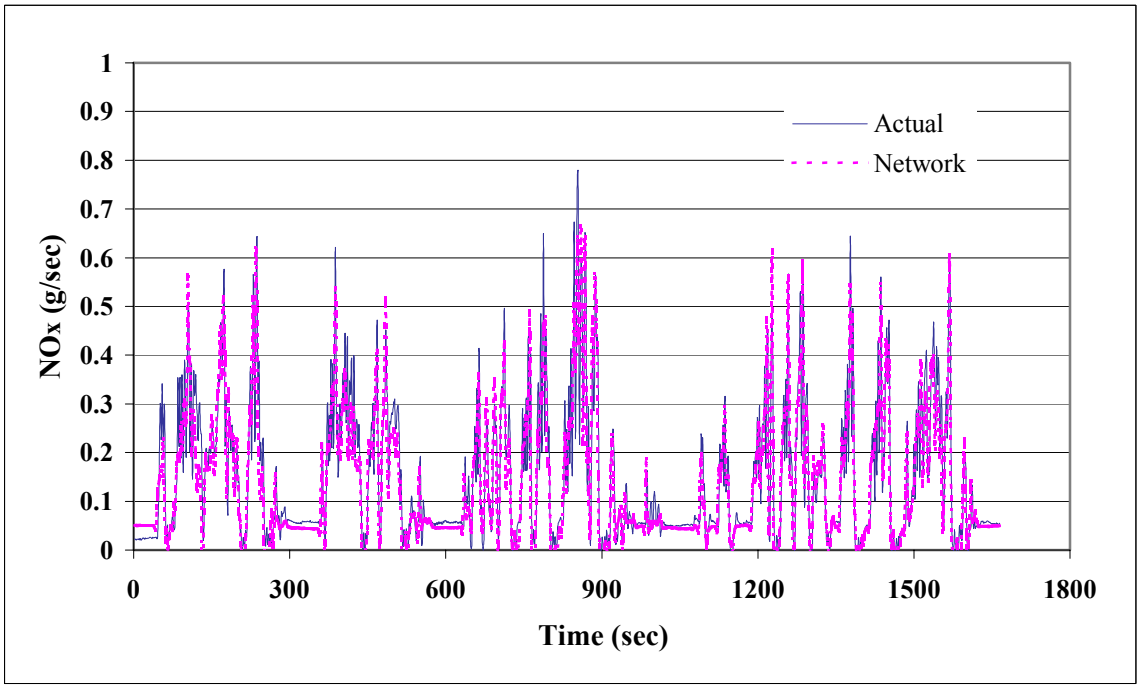


Figure 5.2. Actual and ANN prediction NO_x emissions using 8 inputs for Vehicle 1 exercised through the CSHVR. The neural network was trained on the Highway cycle.

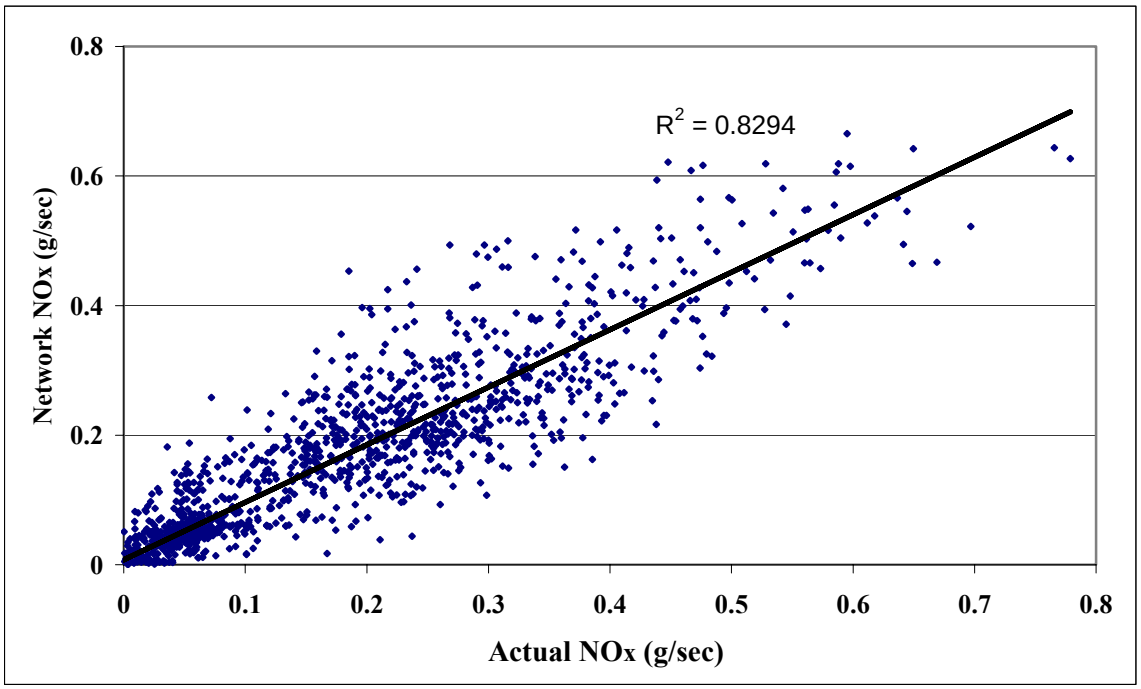


Figure 5.3. ANN prediction versus actual NO_x emissions related to the Figure 5.2 shows correlation coefficient of r^2=0.829.

5.2. CO_2 prediction

The same neural network architectures were applied to the data to predict CO_2 emissions. The average prediction accuracy of the 20 input neural network trained on the Highway cycle was R^2= 0.99. The same ANN was used to predict CO_2 emissions of the other test schedules. The results showed an average R^2 of 0.96 for both the CSHVR and UDDS.

Both 14 input neural networks yielded almost the same predictions for the Highway cycle and UDDS test schedule. However the Jarrett [53] dispersion model showed better prediction for the CSHVR (Table 5.5).

Table 5.5 clarified that the 8 input ANN was the most accurate architecture for the CO_2 prediction of the CSHVR and UDDS in all cases. Therefore, the author highly

recommend using this neural network architecture due to its capabilities while reducing the computing time.

	CO_2			
	20 inputs	**14 inputs (preprocessed by Jarrett [53] dispersion model)**	**14 inputs**	**8 inputs**
Vehicle 1	0.963	0.964	0.971	0.972
Vehicle 2	0.928	0.931	0.928	0.937
Vehicle 3	0.964	0.975	0.971	0.975
Vehicle 4	0.972	0.980	0.978	0.979
Vehicle 5	0.969	0.990	0.971	0.973
Vehicle 6	0.953	0.944	0.943	0.950

Table 5.5. ANN CO_2 emissions prediction results using different numbers of inputs for all the 6 vehicles exercised through the CSHVR. The neural network was trained on the Highway cycle.

	CO_2			
	20 inputs	**14 inputs (preprocessed by Jarrett [53] dispersion model)**	**14 inputs**	**8 inputs**
Vehicle 1	0.993	0.992	0.992	0.991
Vehicle 2	0.993	0.990	0.990	0.988
Vehicle 3	0.994	0.992	0.991	0.992
Vehicle 4	0.994	0.985	0.986	0.989
Vehicle 5	0.994	0.990	0.987	0.990
Vehicle 6	0.992	0.981	0.982	0.986

Table 5.6. ANN CO_2 emissions prediction results using different numbers of inputs for all the 6 vehicles exercised through the Highway cycle.

	CO_2			
	20 inputs	**14 inputs (preprocessed by Jarrett [53] dispersion model)**	**14 inputs**	**8 inputs**
Vehicle 1	0.973	0.977	0.979	0.980
Vehicle 2	0.940	0.939	0.946	0.939

Table 5.7. ANN CO_2 emissions prediction results using different numbers of inputs for the vehicles exercised through the UDDS. The neural network was trained on the Highway cycle.

		Vehicle 1	**Vehicle 2**	**Vehicle 3**	**Vehicle 4**	**Vehicle 5**	**Vehicle 6**
CSHVR	**Actual**	1.89e04	1.45e04	1.50e04	1.49e04	1.54e04	1.81e04
	NN	1.77e04	1.46e04	1.47e04	1.60e04	1.57e04	1.81e04
	Error (%)	6.11	-1.04	2.02	-7.73	-1.71	0.089
Highway	**Actual**	3.24e04	2.05e04	2.54e04	2.77e04	2.73e04	2.82e04
	NN	3.22e04	2.06e04	2.54e04	2.78e04	2.73e04	2.79e04
	Error (%)	0.608	-0.453	-0.004	-0.165	-0.072	1.07
UDDS	**Actual**	1.43e04	1.08e04	_	_	_	_
	NN	1.34e04	0.97e04	_	_	_	_
	Error (%)	5.68	10.2	_	_	_	_

Table 5.8. Integration of the actual and ANN predicted instantaneous CO_2 over the CSHVR, Highway and UDDS test schedules along with their percentage error. Positive and negative percentage errors represent over- and under-prediction respectively. The 8 input neural network was used.

Figures 5.5 and 5.6 show the measured and predicted CO_2 for the 1998 Sterling tractor truck exercised through CSHVR and UDDS test schedules, respectively. The neural network was trained on the Highway cycle (Figure 5.4).

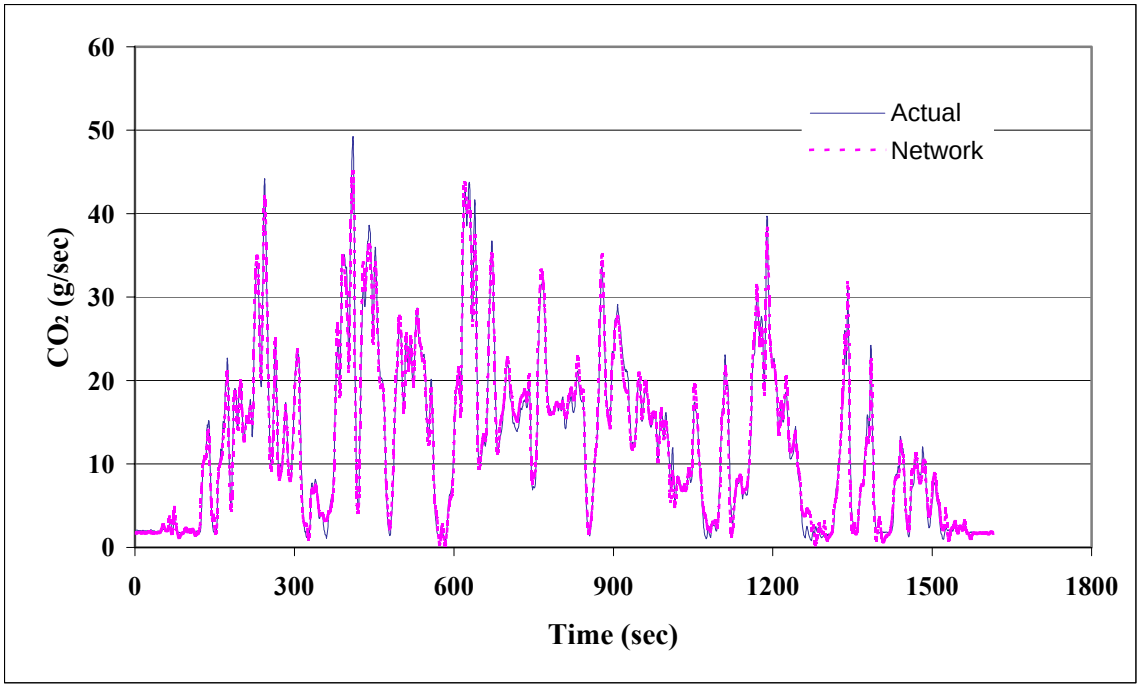


Figure 5.4. Actual and ANN prediction CO_2 emissions using 8 inputs for Vehicle 2 exercised through the Highway cycle. The neural network was trained on the same cycle.

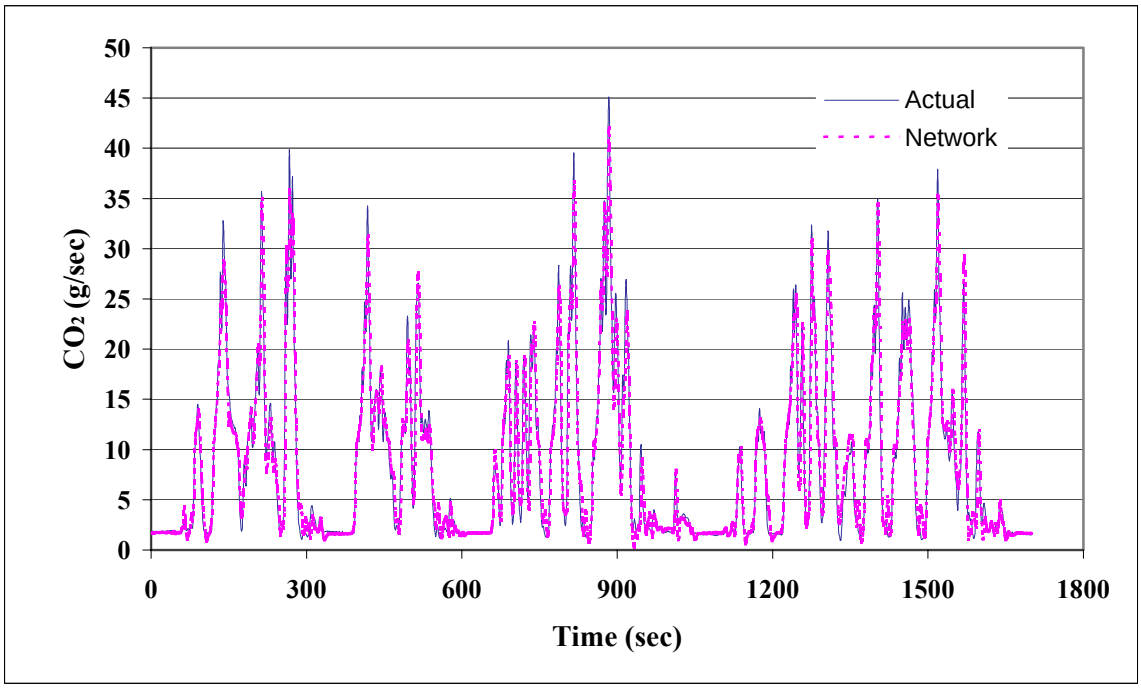


Figure 5.5. Actual and ANN prediction CO_2 emissions using 8 inputs for Vehicle 2 exercised through the CSHVR. The neural network was trained on the Highway cycle.

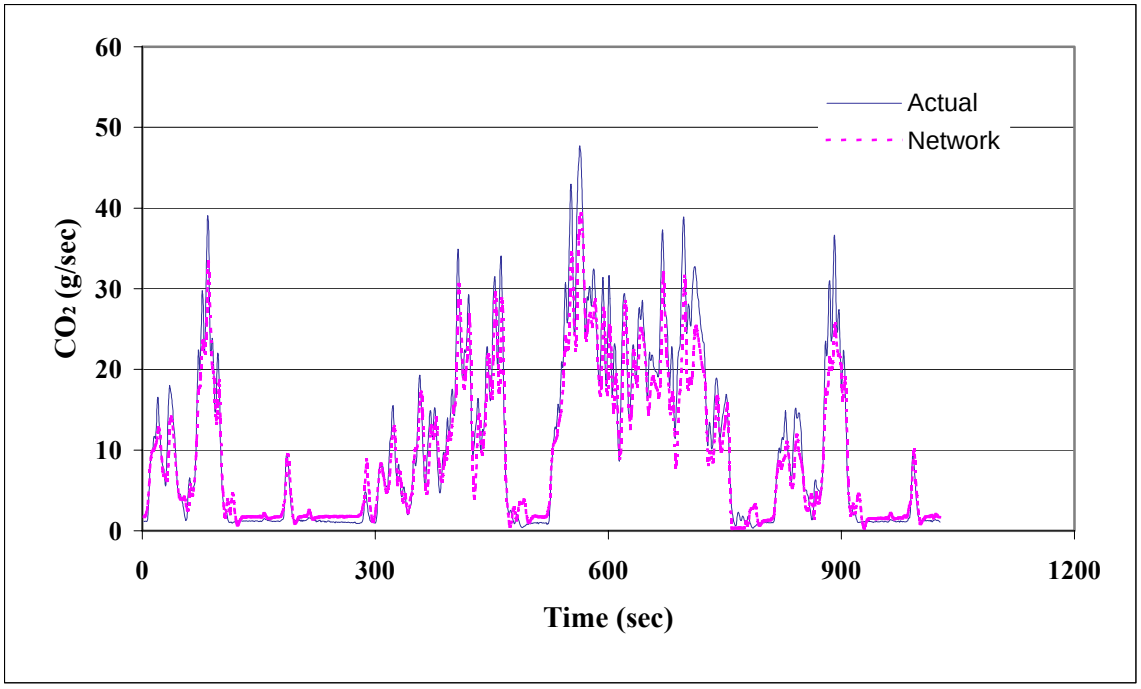


Figure 5.6. Actual and ANN prediction CO_2 emissions using 8 inputs for Vehicle 2 exercised through the UDDS. The neural network was trained on the Highway cycle.

5.3. HC prediction

The 20 input neural network yielded the average R^2 of 0.85 for HC prediction of the Highway cycle. Applying the trained network on other cycles provided an average R^2 of 0.31 for the CSHVR and R^2 of 0.22 for the UDDS test schedules. Research was extended to examine the accuracy of HC prediction using the 8 and 14 input neural network.

Comparing the 14 input ANNs predictions showed that using the Equation 4.5 dispersion model has improved the prediction results significantly for the self-trained neural network (Table 5.10) and when it has been applied on the CSHVR (Table 5.9). For the UDDS prediction, they are almost the same.

The 8 input ANN was the most accurate architecture for the HC prediction of the UDDS. Compare to the 20 input neural network, vehicles 2 and 4 yielded better results

while using less neurons in hidden layer and consequently decreasing the training process time.

However considering the results illustrated that HC prediction are poor in relation to the other emissions predictions. Generally, the HC level emitted from the vehicle engines through the application of the test schedules is very low compared to the atmospheric levels in concentration units. HC emissions are even difficult to measure by the HC analyzers. This also reinforced the difficulties of the HC prediction using neural networks.

	HC			
	20 inputs	**14 inputs (preprocessed by Jarrett [53] dispersion model)**	**14 inputs**	**8 inputs**
Vehicle 1	0.175	0.178	0.180	0.151
Vehicle 2	0.089	0.181	0.117	0.194
Vehicle 3	0.084	0.006	0.007	0.008
Vehicle 4	0.395	0.455	0.716	0.429
Vehicle 5	0.459	0.412	0.450	0.420
Vehicle 6	0.681	0.602	0.610	0.647

Table 5.9. ANN HC emissions prediction results using different numbers of inputs for all the 6 vehicles exercised through the CSHVR. The neural network was trained on the Highway cycle.

	HC			
	20 inputs	**14 inputs (preprocessed by Jarrett [53] dispersion model)**	**14 inputs**	**8 inputs**
Vehicle 1	0.865	0.734	0.644	0.866
Vehicle 2	0.873	0.311	0.311	0.262
Vehicle 3	0.886	0.382	0.371	0.659
Vehicle 4	0.618	0.335	0.636	0.452
Vehicle 5	0.907	0.612	0.568	0.822
Vehicle 6	0.954	0.829	0.870	0.930

Table 5.10. ANN HC emissions prediction results using different numbers of inputs for all the 6 vehicles exercised through the Highway cycle. The neural network was trained on the same cycle.

	HC			
	20 inputs	**14 inputs (preprocessed by Jarrett [53] dispersion model)**	**14 inputs**	**8 inputs**
Vehicle 1	0.449	0.445	0.448	0.486
Vehicle 2	0.008	0.311	0.265	0.334

Table 5.11. ANN HC emissions prediction results using different numbers of inputs for the vehicles exercised through the UDDS. The neural network was trained on the Highway cycle.

		Vehicle 1	Vehicle 2	Vehicle 3	Vehicle 4	Vehicle 5	Vehicle 6
CSHVR	**Actual**	9.60	2.07	10.0	6.27	7.65	14.7
	NN	6.22	1.96	9.62	8.10	7.90	13.8
	Error (%)	35.2	5.31	4.08	-29.1	-3.33	6.20
Highway	**Actual**	5.86	2.10	10.0	7.13	6.60	14.8
	NN	5.89	2.00	10.0	6.86	6.53	14.5
	Error (%)	-0.511	4.76	-0.19	3.81	1.06	1.62
UDDS	**Actual**	2.58	1.31	_	_	_	_
	NN	3.52	1.22	_	_	_	_
	Error (%)	-36.4	6.87	_	_	_	_

Table 5.12. Integration of the actual and ANN predicted instantaneous HC over the CSHVR, Highway and UDDS test schedules along with their percentage error. Positive and negative percentage errors represent over- and under-prediction respectively. The 8 input neural network was used.

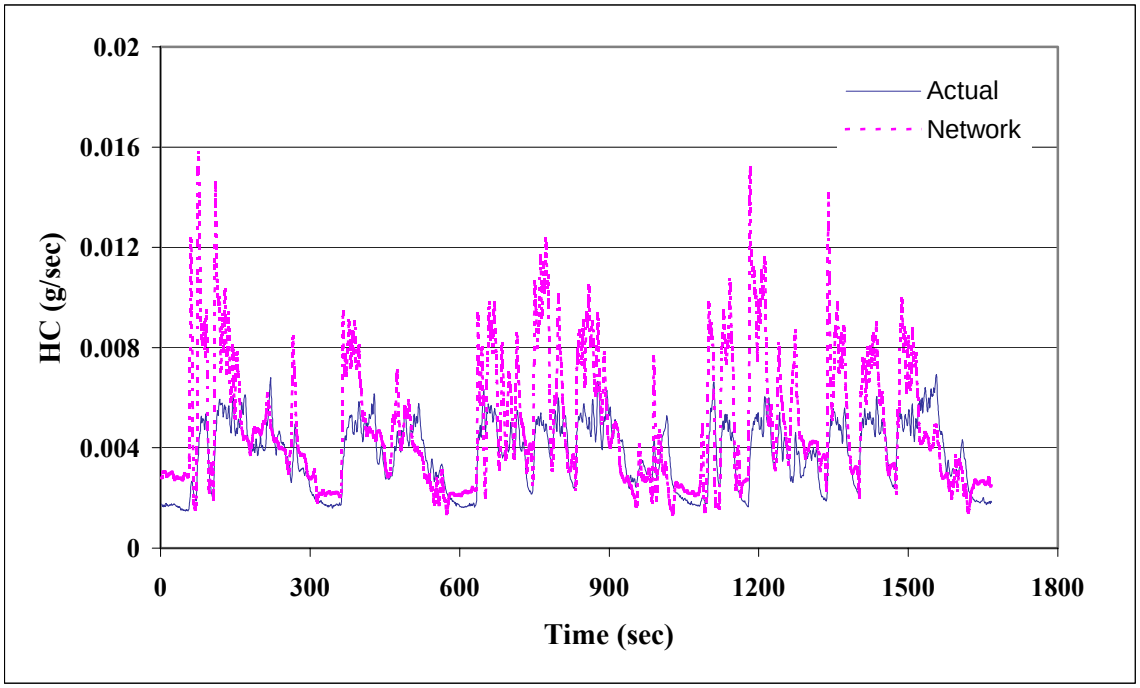


Figure 5.7. Actual and ANN prediction HC emissions using 8 inputs for Vehicle 4 exercised through the CSHVR. The neural network was trained on the Highway cycle

5.4. CO prediction

Tables 5.13 to 5.15 show the actual versus predicted CO results. The prediction results revealed that the neural network prediction while it has been trained using 8 inputs is more accurate than the 14 and 20 input neural networks even on the self-trained cycle. Results for a few vehicles on the CSHVR were acceptable while for the UDDS the results were uniformly good. Inspection of Figure 5.8 shows that the ANN generally predicted 'spikes' of CO at the appropriate points in the schedule, but that the magnitude of many spikes was poorly modeled. However, the transient 'spikes' that were shorter than smoothing dispersion model were seen accurately. The first 700 seconds of the actual and ANN prediction CO emissions versus time for vehicle 2 exercised through the CSHVR is presented in Figure 8.3 (Appendix B).

	CO			
	20 inputs	**14 inputs (preprocessed by Jarrett [53] dispersion model)**	**14 inputs**	**8 inputs**
Vehicle 1	0.595	0.614	0.603	0.660
Vehicle 2	0.605	0.529	0.541	0.674
Vehicle 3	0.391	0.672	0.662	0.685
Vehicle 4	0.680	0.736	0.780	0.732
Vehicle 5	0.298	0.371	0.313	0.448
Vehicle 6	0.720	0.787	0.796	0.791

Table 5.13. ANN CO emissions prediction results using different numbers of inputs for all the 6 vehicles exercised through the CSHVR. The neural network was trained on the Highway cycle.

	CO			
	20 inputs	**14 inputs (preprocessed by Jarrett [53] dispersion model)**	**14 inputs**	**8 inputs**
Vehicle 1	0.857	0.779	0.774	0.767
Vehicle 2	0.935	0.795	0.823	0.712
Vehicle 3	0.948	0.678	0.683	0.810
Vehicle 4	0.914	0.710	0.717	0.832
Vehicle 5	0.843	0.615	0.643	0.666
Vehicle 6	0.954	0.844	0.836	0.925

Table 5.14. ANN CO emissions prediction results using different numbers of inputs for all the 6 vehicles exercised through the Highway cycle. The neural network was trained on the same cycle.

	CO			
	20 inputs	**14 inputs (preprocessed by Jarrett [53] dispersion model)**	**14 inputs**	**8 inputs**
Vehicle 1	0.493	0.591	0.662	0.559
Vehicle 2	0.195	0.302	0.303	0.314

Table 5.15. ANN CO emissions prediction results using different numbers of inputs for the vehicles exercised through the UDDS. The neural network was trained on the Highway cycle

		Vehicle 1	Vehicle 2	Vehicle 3	Vehicle 4	Vehicle 5	Vehicle 6
CSHVR	**Actual**	115	104	128	63.9	60.1	43.9
	NN	114	97.1	105	66.4	55.6	48.1
	Error (%)	0.286	6.28	17.8	-3.87	7.35	-9.46
Highway	**Actual**	128	72.7	86.9	68.6	65.2	44.6
	NN	136	75.1	90.6	69.4	67.2	44.4
	Error (%)	-5.73	-3.31	-4.25	-1.07	-3.17	0.494
UDDS	**Actual**	84.5	43.0	_	_	_	_
	NN	69.9	48.4	_	_	_	_
	Error (%)	17.3	-12.5	_	_	_	_

Table 5.16. Integration of the actual and ANN predicted instantaneous CO over the CSHVR, Highway and UDDS test schedules along with their percentage error. Positive and negative percentage errors represent over- and under-prediction respectively. The 8 input neural network was used.

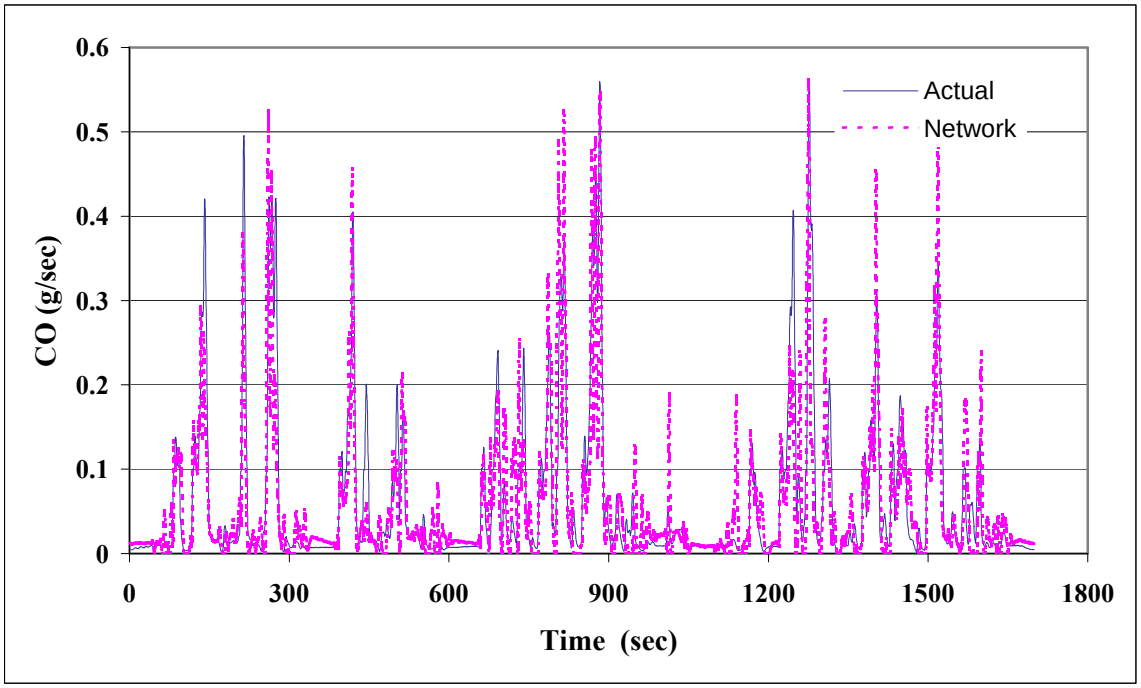


Figure 5.8. Actual and ANN prediction CO emissions using 8 inputs for Vehicle 2 exercised through the CSHVR. The neural network was trained on the Highway cycle.

5.5. Discussions

In each cycle, to investigate the areas of divergence of the predicted emissions species from the measured emissions, the dimensionless difference between predicted and actual was considered and plotted against time. Figure 5.9 shows the difference between predicted and measured NOx at each second. To compute the dimensionless difference, the difference between the measured and the predicted emissions at each second was divided by the average of the measured emissions over the entire test schedule. The dispersed speed pattern was illustrated as a solid line (Figures 5.9 and 5.10).

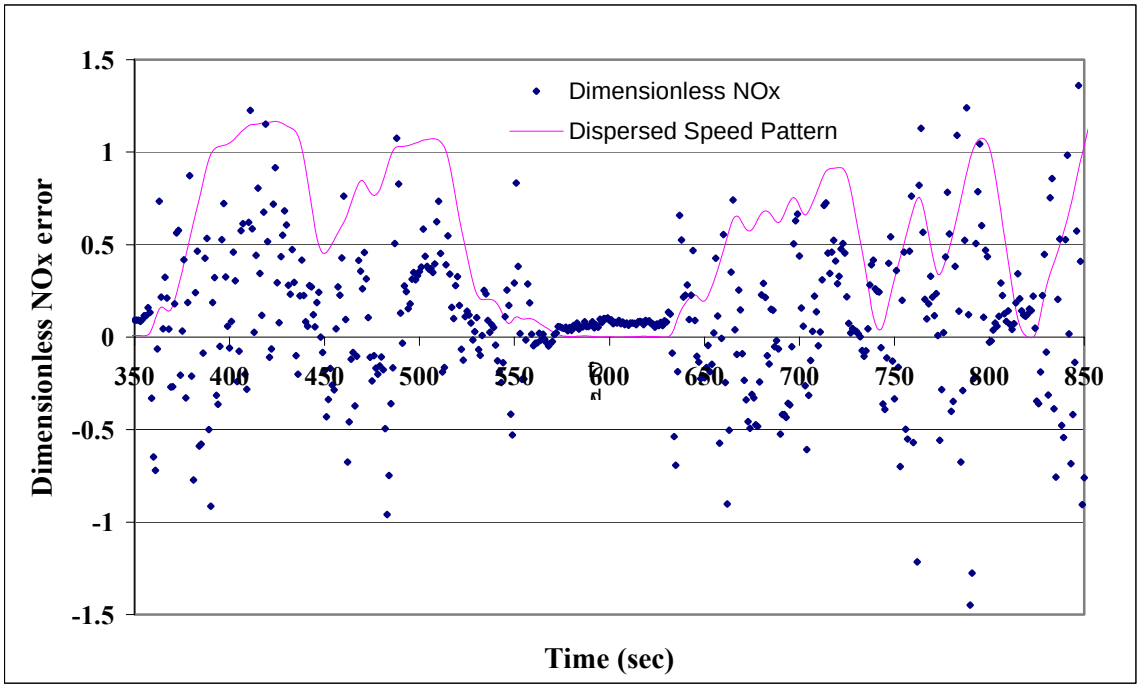


Figure 5.9. Difference of ANN predicted and actual real-time NO_x for vehicle 1 exercised through the City/Suburban Heavy Vehicle Route. The neural network was trained on the Highway cycle. This illustration shows only 500 seconds of the test schedule.

The predicted NOx showed more divergence for the area that the test schedule experienced a large change in the speed. However, while the speed was constant over a period of the time, the difference was very small. This occurred possibly because the "off-cycle" NO_x is confusing to the neural networks. Emissions of NO_x are strongly affected by injection timing changes while emissions of CO_2 are affected slightly. During the 1990's, some manufacturers implemented timing changes in the engine control strategy once the vehicle was detected to be cruising rather than pursuing transient activity. This "off-cycle" timing behavior was curtailed following a consent decree between manufacturers and the U.S. government [65]. Functions to detect steady behavior (Diff and $\overline{Spd}$) offer the neural network a method for predicting timing changes, with consequent rise in NO_x, although these functions can not hope to mimic manufacturers' strategies faithfully.

CO_2 difference was not uniformly distributed at each second for the CSHVR but it followed the speed pattern for the UDDS. On the other hand, the rate of speed change was correlated to the magnitude of predicted and measured emissions difference.

Neural network HC prediction differences did not follow the speed pattern for the UDDS and CSHVR and were spread irregularly throughout the test schedules.

The CO prediction differences were also spread irregularly with no correlation to speed.

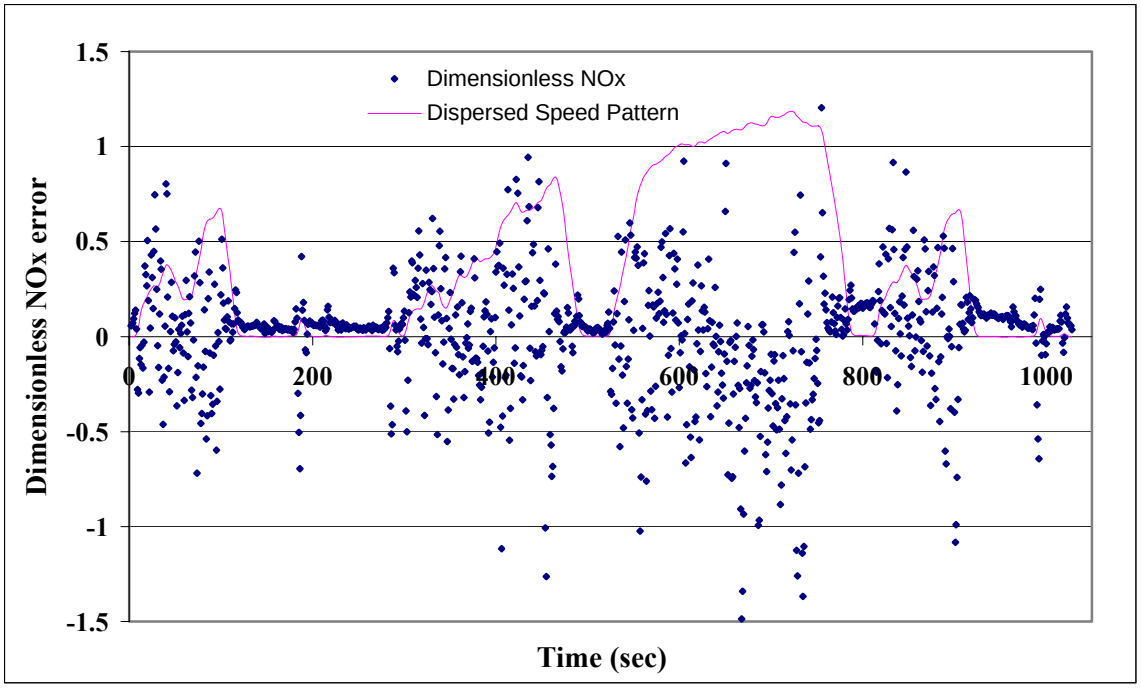


Figure 5.10. Difference of ANN predicted and actual real-time NO_x for vehicle 1 exercised through the Heavy-Duty Urban Dynamometer Driving Schedule (UDDS).

6. CONCLUSIONS

A Levenspiel [60] dispersion model was used to mimic the response of analyzers to instantaneous emissions. Linear correlation of the dispersed power and CO_2 proved that the best dispersion model is achievable using a D/UL value of 0.011. This VDN yielded a non-symmetric dispersion model. CO_2 was chosen for this purpose because of the known relation of this emissions species to the power. The suggested model is capable of illustrating the onset time and the dispersion period at the same time. To find the delay time, a C_i value of 0.001 was set as a criterion. Therefore, the dispersion coefficients, C_is, were classified into three groups. The first region with values of C_i less than 0.001 (10e-130 to 10e-4) was considered as the onset time. The next region was the dispersion time period with values of the C_i bigger than 0.001 corresponding to 13 seconds. The third region (10e-4 to 10e-7) was the indicator of the damping time of the spike. Using the wide range of dispersion coefficients made it possible to model the entire track of an emission spike. The delay time was defined as the duration of the onset time along with the time taken for the C_i curve to reach to its peak. This delay time consist of transport time in the dilution tunnel, the sampling lines and the analyzer's response and it was found to be 16.5 seconds for the D/UL=0.011. NO_x and CO_2 were perfectly correlated to this dispersion model. However CO was known to be dependant on power rate as well as power.

The preprocessing of the vehicle speed and torque was followed by feeding these data as inputs to the neural network. Among the applied neural network architectures, the

one with 8 inputs of axle speed, torque, their first and second derivatives at one second and Diff and $\bar{S}pd$ in 150 seconds yielded the most accurate emissions prediction.

NO_x prediction was improved while the trained network applied to the CSHVR and UDDS compared to the 20 input neural network. This improvement also can be seen in the 8 input self-trained network (for Highway cycle). However, the 14 input neural network yielded 2.5% improvement for the CSHVR and 2.7% for the UDDS compared to the 8 input ANN.

The 8 input neural network proved to predict the CSHVR and UDDS CO_2 and CO emissions much better compared to the other networks.

However, the HC prediction was poor due to its low emissions level and difficulty to measure them, but the 8 input neural network predicted HC emissions very well also.

Reviewing the results showed that the dispersed speed and the speed-based inputs such as the Diff variable were the most controlling input in the neural network training for all the emissions prediction cases. Smoothing the oscillations of the speed using the Diff variable and feeding it as an input to ANN increased the accuracy of the prediction. In some cases this might account for "off-cycle" timing behavior. Overall, ANN would be well suited to inventory prediction and would allow access to the real life emissions of a vehicle that is operated over a roadway route that is very different from the existing tests cycles.

The author's recommendation for future steps is to concentrate on the optimization of the "off-cycle" timing modeling as an input to the neural network. The other area

would be considering a smaller number of the neural network inputs while maintaining the prediction accuracy.

7. REFERNCES

1. Heywood, J. B., "Internal Combustion Engine Fundamentals," McGraw-Hill, Inc., New York, 1988.
2. U.S. Environmental Protection Agency, "Emission Control Potential for Heavy-duty Diesel Engines," EPA 420-F-97-015, 1997.
3. Yanowitz, J., McCormick, R. L., and Graboski, M. S., "In-Use Emission from Heavy-Duty Diesel Vehicle," Environmental Science & Technology, Vol. 34, No. 5, pp. 729-740, 2000.
4. http://www.dieselnet.com/standards/us/hd.html, accessed 10/22/2003.
5. U.S. Environmental Protection Agency, "Technology Transfer Network Clearinghouse for Inventories and Emission Factors - Emissions Modeling: Inventories." http://www.epa.gov/ttn/chief/emch/invent/ accessed 1/13/2004.
6. Koupal, J., Cumberworth, M., Michaels, H., Beardsley, M. and Brzezinski, D., "Draft Design and Implementation Plan for EPA's Multi-Scale Motor Vehicle and Equipment Emission System (MOVES)," EPA420-P-02-006, 2002. http://www.epa.gov/otaq/models/ngm/p02006.pdf accessed 3/2/2004.
7. Koupal, J., "Emission Rate Development, EPA Office of Transportation & Air Quality," FACA Modeling Workgroup Meeting, December 2003. http://www.epa.gov/otaq/models/ngm/mwg1203e.pdf, accessed 4/20/2004.
8. Clark, N. N., Gajendran, P. and Kern, J. M., "A Predictive Tool for Emissions from Heavy-Duty Diesel Vehicles," Environmental Science & Technology, Vol. 37, No. 1, pp. 7-15, 2003.
9. Bastian, A., "Modeling Fuel Injection Control Maps Using Fuzzy Logic," IEEE International Conference on Fuzzy Systems, Vol. 2, pp. 740-743, 1994.
10. Frith, A., M., Gent, C. R., Beaumont, A. J., "Adaptive Control of Gasoline Engine Air-Fuel Ratio Using Artificial Neural Networks," IEEE Conference Publication, Proceedings of the 4th International Conference on Artificial Neural Networks, pp. 274-278, 1995.

11. Ramani, S., Miranda, R., "Neural-Network-Aided Design of Automobile Exhaust Catalysts," Chemical Engineering Communications, Vol. 156, pp. 147-160, 1996.
12. Stevens, S. P., Shayler, P. J., Ma, T. H., "Experimental Data Processing Techniques to Map the Performance of a Spark Ignition Engine," Journal of Automobile Engineering, Vol. 209, pp. 297-306, 1995.
13. Hanzevack, E. L., Long, T. W., Atkinson, C. M., Traver, M. L., "Virtual Sensors for Spark Ignition Engines Using Neural Networks," Proceedings of the American Control Conference, Vol. 1, pp. 669-673, 1997.
14. Krijnsen, H. C., Van Kooten, W. J., Calis, H. A., Verbeek, R. P., Van den Bleek, C. M., "Prediction of NO_x Emissions from a Transiently Operating Diesel Engine Using an Artificial Neural Network," Chemical Engineering and Technology, Vol. 22, pp. 601-607, 1999.
15. Quenou Gamo, S., Ouladsine, M., Rachid, A., "Diesel Engine Exhaust Emission Modeling Using Artificial Neural Networks," SAE paper 1999-01-1163.
16. Hafner, M., Schueler, M., Nelles, O., "Dynamical Identification and Control of Combustion Engine Exhaust," Proceedings of the American Control Conference, Vol. 1, pp. 222-226, 1999.
17. Thompson, G. J., Atkinson, C. M., Clark, N. N., Long, T. W., Hanzevack, E., "Neural Network Modeling of the Emissions and Performance of a Heavy-Duty Diesel Engine," Proc. Inst. Mech. Engrs., Vol. 214, Part D, pp. 111-126, 2000.
18. De Lucas, A., Duran, A. Carmona, M., Lapuerta M., "Modeling Diesel Particulate Emissions with Neural Networks," Fuel, Vol. 80, pp. 539-548, 2001.
19. Clark, N. N., Conley, J., Jarrett, R. P., Nennelli, A., Toth-Nagy, C., "Emissions Modeling of Heavy-Duty Conventional and Hybrid Electric Vehicles," SAE paper 2001-01-3675.
20. Yuanwang, D., Meilin, Z., Dong, X., Xiaobei, C., "An Analysis for Effect of Cetane Number on Exhaust Emissions from Engine with the Neural Network," Fuel, Vol. 81, pp. 1963-1970, 2002.
21. Desantes, J., Lopez, J., Garcia, J., Hernandez, L., "Application of Neural Networks for Prediction and Optimization of Exhaust Emissions in a H.D. Diesel Engine," SAE paper 2002-01-1144.

22. Hafner, M., Isermann, R., "Multiobjective Optimization of Feedforward Control Maps in Engine Management Systems Towards Low Consumption and Low Emissions," Transactions of the Institute of Measurement and Control, Vol. 25, pp. 57-74, 2003.
23. Tehranian, A., "Effects of Artificial Neural Networks Characterization on Prediction of Diesel Engine Emissions," M.S. Thesis, West Virginia University, 2004.
24. Stergiou, C., Siganos, D., "Neural Networks" http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html, accessed 5/11/04.
25. Hagan, M. T., Demuth, H. B., Beale, M, "Neural Network Design," PWS publishing Co., Boston, MA, 1996.
26. Olmsted, D. D., "History and Principles of Neural Networks to 1960," http://www.neurocomputing.org/History/body_history.html, accessed 8/31/03.
27. McCulloch, W. and Pitts, W., "A Logical Calculus of the Ideas Immanent in Nervous Activity," Bulletin of Mathematical Biophysics, Vol. 5, pp. 115-133, 1943.
28. Hebb, D. O., "The Organization of Behaviors," John Wiley & Sons, New York, 1949.
29. Rosenblatt, F., "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain," Psychological Review, Vol. 65, pp. 386-408, 1958.
30. Minsky, M. L., Papert, S., "Perceptrons," MIT Press, Cambridge, MA, 1969.
31. Kohonen, T., "Correlation Matrix Memories," IEEE Transactions on Computers, Vol. 21, pp. 353-359, 1972.
32. Anderson, J. A., "A Simple Neural Network Generating an Interactive Memory," Mathematical Bioscience, Vol. 14, pp. 197-220, 1972.
33. Hopfield, J. J., "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," Proceedings of the National Academy of Sciences, Vol. 79, pp. 2554-2558, 1982.

34. Rumelhart, D.E., McClelland, J. L., "Parallel Distributed Processing: Explorations in the Microstructure of Cognition," Vol. 1, MIT Press, Cambridge, MA, 1986.
35. Jain, L. C., Martin, N. M., "Fusion of Neural Networks, Fuzzy Sets, and Genetic Algorithms," CRC Press LLC, Boca Raton, Florida, 1999.
36. Carpenter, G. A., Grossberg, S., "ART2 Self-Organisation of Stable Category Recognition Codes for Analog Input Patterns," Applied Optics, 26, pp. 4919-4930, 1987.
37. NeuroShell2 User's Manual, Ward Systems Group, Inc., Fredrick, MD, 1996.
38. http://hem.hj.se/~de96klda/NeuralNetworks.htm, accessed 5/11/04.
39. https://tagteamdbserver.mathworks.com/ttserverroot/Download/807_8511v04_NeuralTbX.pdf, accessed 5/12/04.
40. White, D., Sofge, D., "Handbook of Intelligent Control," Van Nostrand Reinhold, New York, 1992.
41. http://www.mathworks.com/access/helpdesk/help/toolbox/nnet/backpr52.html#12955, accessed 6/9/2004.
42. http://www.xs4all.nl/~dpsol/data-machine/nmtutorial/improvementsinbackpropagationtraining.htm, accessed 6/9/2004.
43. Vogl, T., P., Mangis, J., K., Zigler, A., K., Zink, W., T., Alkon, D., L., "Accelerating the Convergence of the Backpropagation Method," Biological Cybernetics, Vol. 59, pp. 256-264, 1988.
44. Clark, N. N., Nine, R. D., Daley J. J., and Atkinson, C. M., "Development of a Heavy-Duty Chassis Dynamometer Driving Route," Proceedings of Institution of Mechanical Engineers. Part D., Journal of Automobile Engineering, Vol. 213, pp. 561-574.
45. Rupprecht & Patashnick Co., Inc., TEOM, Series 1105, Diesel Particulate Mass Monitor, Operating Manual, September 1997.
46. Okrent, D. A., "Optimization of a Third Generation TEOM Monitor for Measuring Diesel Particulate in Real-Time," SAE paper No. 980409, 1998.
47. Rosemount Analytical Inc., Model 402 Hydrocarbon Analyzer, Instruction Manual, 1993.

48. Rosemount Analytical Inc., Model 955 NO/NOX Analyzer, Instruction Manual, Literature No. 748190-D,1992.
49. Rosemount Analytical Inc., Model 868 Non-Dispersive Infrared Analyzer, Instruction Manual, Instructions 015-748003-J, 1991.
50. Horiba Instrument Incorporated, AIA-210/220 Infrared Analyzer, Instruction Manual, Horiba Manual No. 091215, 1995.
51. http://www.dieselnet.com/standards/cycles/index.html, accessed 03/25/04
52. Nine, R. D., Clark, N. N., Norton, P., "Effects on Emissions of Multiple Driving Test Schedules Performed on Two Heavy-Duty Vehicles," SAE paper 2000-01-2818.
53. Jarrett, R. P., "Evaluation of Opacity Particulate Matter, and Carbon Monoxide From Heavy Duty Diesel Transient Chassis Tests," M.S. Thesis, West Virginia University, 2000.
54. Clark, N. N., Daley, J. J., Nine, R. D., Atkinson, C. M., "Application of the New City-Suburban Heavy Vehicle Route (CSHVR) to Truck Emissions Characterizations," SAE paper 99011467.
55. Clark, N. N., Wayne, W. S., Lyons, W. L. and Thompson G., "Gasoline-Diesel PM Split Study: Heavy-Duty Vehicle Exhaust Collection Phase," Submitted to: National Renewable Energy Laboratory by West Virginia University Research Corporation, 2002.
56. Clark, N. N., Wayne, W. S., Nine, R. D., Buffamonte, T. M., Hall, T., Rapp, B. L., Thompson, G., and Lyons, D. W., "Emissions from Diesel-Fueled Heavy-Duty Vehicles in Southern California," SAE/JSAE Spring Fuels & Lubricants Meeting, Yokohama, Japan 2003, JSAE paper 20030232.
57. Ramamurthy, R. and Clark N. N., "Atmospheric Emissions Inventory Data for Heavy Duty Vehicles," Environmental Science & Technology, Vol. 33, pp. 52-62, 1999.
58. http://www.eso.org/projects/dfs/papers/jitter99/node10.html, accessed 4/6/2004.
59. Clark, N. N., Jarrett, R. P., Atkinson, C. M., "Field Measurements of Particulate Matter Emissions and Exhaust Opacity from Heavy Duty Vehicles," Journal of the Air and Waste Management Association, Vol. 49, 1999.

60. Levenspiel, O., "Chemical Reaction Engineering," 3rd Edition, John Wiley & Sons, Inc., New York, NY, 1999.
61. Baskaran, G. and Clark N. N., "Relationships Between Instantaneous and Measured Emissions in Heavy Duty Applications," SAE paper 2001-01-3536.
62. Kern, J., Clark, N. N., Nine, R. and Atkinson, C. M., "Factors Affecting Heavy-Duty Diesel Vehicle Emissions," Journal of the Air and Waste Management Association, Vol. 52, pp. 84-94, 2002.
63. Brodrick, C. J., Dwyer, H. A., Farshchi, M., Harris, D. B., King, F. G., "Effects of Engine Speed and Accessory Load on Idling Emissions from Heavy-Duty Diesel Truck Engines," Journal of the Air and Waste Management Association, Vol. 52, No. 9, 2002.
64. Clark, N. N., Tehranian, T., Jarrett R. P., and Nine R. D., "Translation of Distance-Specific Emissions Rates between Different Heavy Duty Vehicle Chassis Test Schedules," SAE paper 2002-01-1754.
65. "Title 13. California Air Resources Board: Notice of Public Hearing to Consider Amendments to Adopt Not-to-Exceed and Euro III European Stationary Cycle Emission Test Procedures for the 2005 and Subsequent Model Year Heavy-Duty Diesel Engines" www.arb.ca.gov/regact/ntetest/notice.pdf, accessed 4/18/2004.

APPENDIX A – Vehicle Description

VEHICLE 1

WVU Test Reference Number: WVU-TEDDY2-CARB

Fleet Owner Full Name	West Virginia University
Vehicle Type	Tractor Truck
Vehicle ID Number (VIN)	1FUPYSYB7FH258124
Vehicle Manufacturer	Freightliner
Vehicle Model Year	1985
Gross Vehicle Weight (GVW) (lb.)	80000
Vehicle Total Curb Weight (lb.)	12000
Vehicle Tested Weight (lb.)	42000
Odometer Reading (mile)	769413
Transmission Type	Manual
Transmission Configuration	9-Speed
Number of Axles	3
Engine Type	Caterpillar 3406B
Engine ID Number	7FB30060
Engine Displacement (Liter)	14
Number of Cylinders	6
Engine Rated Power (hp)	350
Primary Fuel	CARB

VEHICLE 2

WVU Test Reference Number: RDO-94 Freight-CARB

Fleet Owner Full Name	RDO Truck Sales
Vehicle Type	Tractor Truck
Vehicle ID Number (VIN)	1FUYDZYB85H706048
Vehicle Manufacturer	Freightliner
Vehicle Model Year	1994
Gross Vehicle Weight (GVW) (lb.)	52000
Vehicle Total Curb Weight (lb.)	17120
Vehicle Tested Weight (lb.)	42000
Odometer Reading (mile)	639105
Transmission Type	Manual
Transmission Configuration	10 Speed
Number of Axles	3
Engine Type	Detroit Diesel Series 60
Engine ID Number	l14139
Engine Displacement (Liter)	12.7
Number of Cylinders	6
Engine Rated Power (hp)	470
Primary Fuel	CARB

VEHICLE 3

WVU Test Reference Number: RAGRO-5939-CARB

Fleet Owner Full Name	Ralph's Grocery
Vehicle Type	Tractor Truck
Vehicle ID Number (VIN)	2FWYJFEB3XAB21287
Vehicle Manufacturer	Sterling
Vehicle Model Year	1998
Gross Vehicle Weight (GVW) (lb.)	52000
Vehicle Total Curb Weight (lb.)	16560
Vehicle Tested Weight (lb.)	42000
Odometer Reading (mile)	327300
Transmission Type	Manual
Transmission Configuration	10 speed
Number of Axles	3
Engine Type	Detroit Diesel Series 60
Engine ID Number	06R0461436
Engine Displacement (Liter)	12.7
Number of Cylinders	6
Engine Rated Power (hp)	470
Primary Fuel	CARB

VEHICLE 4

WVU Test Reference Number: RAGRO-5009-CARB

Fleet Owner Full Name	Ralph's Grocery
Vehicle Type	Tractor Truck
Vehicle ID Number (VIN)	2FWYHWEB5YAF83127
Vehicle Manufacturer	Sterling
Vehicle Model Year	1999
Gross Vehicle Weight (GVW) (lb.)	52000
Vehicle Total Curb Weight (lb.)	16240
Vehicle Tested Weight (lb.)	42000
Odometer Reading (mile)	272307
Transmission Type	Manual
Transmission Configuration	10 speed
Number of Axles	3
Engine Type	Caterpillar C-12
Engine ID Number	2KS15562
Engine Displacement (Liter)	12
Number of Cylinders	6
Engine Rated Power (hp)	425
Primary Fuel	CARB

VEHICLE 5

WVU Test Reference Number: RAGRO-5007-CARB

Fleet Owner Full Name	Ralph's Grocery
Vehicle Type	Tractor Truck
Vehicle ID Number (VIN)	2FWYHWEB1YAF83125
Vehicle Manufacturer	Sterling
Vehicle Model Year	2000
Gross Vehicle Weight (GVW) (lb.)	52000
Vehicle Total Curb Weight (lb.)	16460
Vehicle Tested Weight (lb.)	42000
Odometer Reading (mile)	255880
Transmission Type	Manual
Transmission Configuration	10 Speed
Number of Axles	3
Engine Type	Caterpillar C-12
Engine ID Number	12KS15574
Engine Displacement (Liter)	12
Number of Cylinders	6
Engine Rated Power (hp)	425
Primary Fuel	CARB

VEHICLE 6

WVU Test Reference Number: RDO-01 Volvo-CARB

Fleet Owner Full Name	RDO Truck Sales
Vehicle Type	Tractor Truck
Vehicle ID Number (VIN)	4V4NC9JH91N255387
Vehicle Manufacturer	Volvo
Vehicle Model Year	2001
Gross Vehicle Weight (GVW) (lb.)	52000
Vehicle Total Curb Weight (lb.)	17680
Vehicle Tested Weight (lb.)	42000
Odometer Reading (mile)	327300
Transmission Type	Manual
Transmission Configuration	10 speed
Number of Axles	3
Engine Type	Cummins 2N14-370
Engine ID Number	12002625
Engine Displacement (Liter)	14
Number of Cylinders	6
Engine Rated Power (hp)	370
Primary Fuel	CARB

APPENDIX B – Supplementary Figures

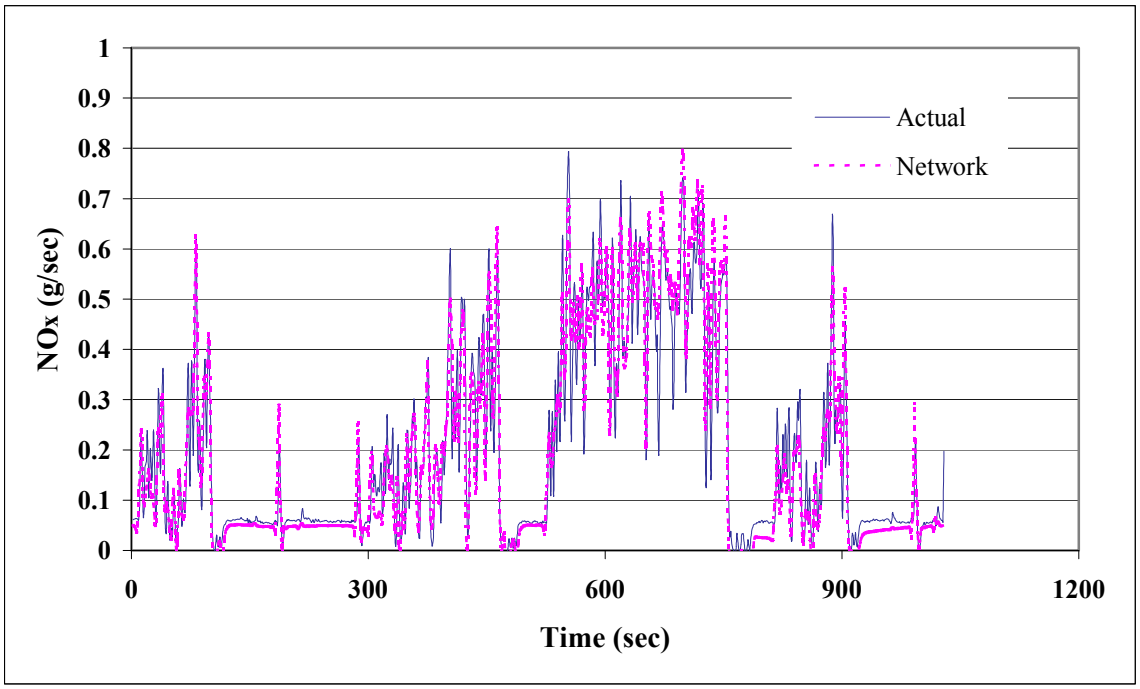


Figure 8.1. Actual and ANN prediction NO_x emissions using 8 inputs for Vehicle 1 exercised through the UDDS. The neural network was trained on the Highway cycle.

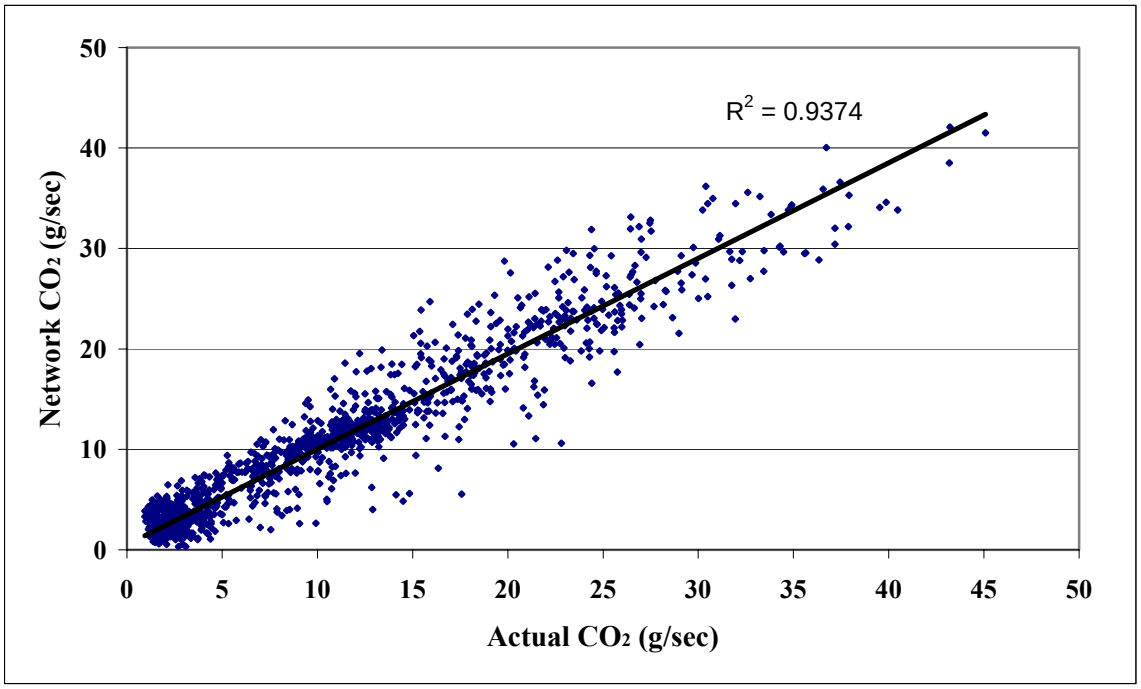


Figure 8.2. ANN prediction versus actual CO_2 emissions related to the Figure 5.5 shows correlation coefficient of r^2=0.937.

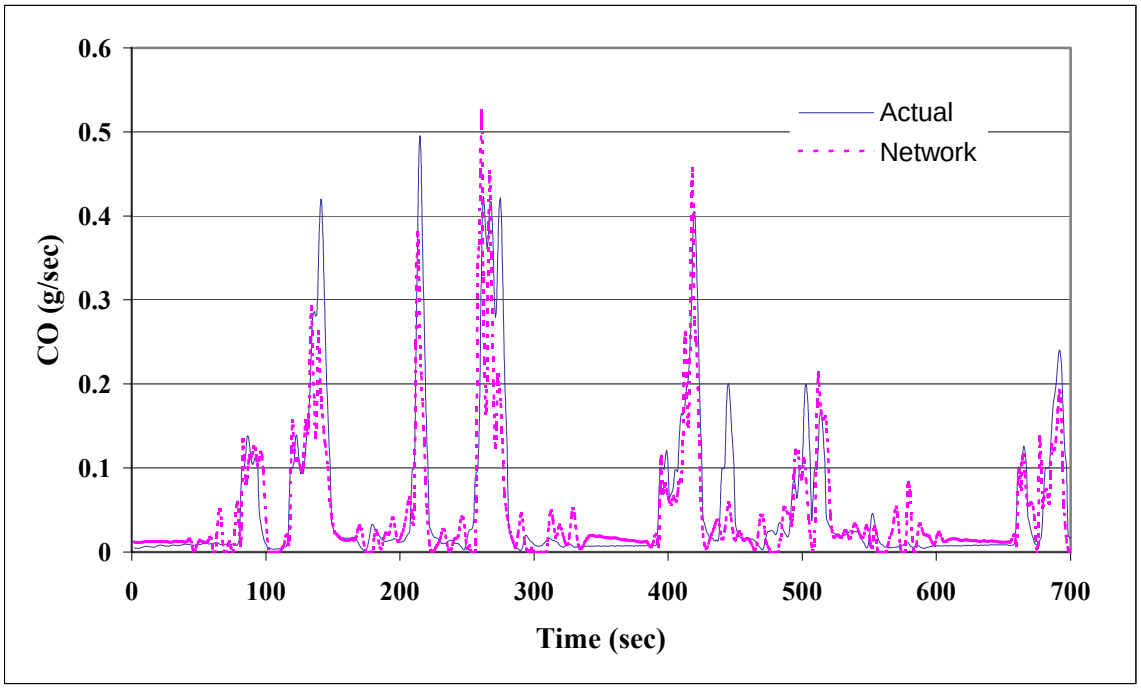


Figure 8.3. First 700 seconds of the actual and ANN prediction CO emissions using 8 inputs for Vehicle 2 exercised through the CSHVR. The neural network was trained on the Highway cycle.