
ISSN: 2321-2152
IJMECE
International Journal of modern
electronics and communication engineering
E-Mail
editor.ijmece@gmail.com
editor@ijmece.com
www.ijmece.com

ISSN2321-2152 www.ijmece

Vol 8, Issuse.2 June 2020

A Machine Learning Model for Average Fuel Consumption in Heavy Vehicles

S.Raju[1],E.Muralidhar Reddy[2],Firdose Fathima[3],Dr.M.Bal Raju[4],

Abstract—Data summary for personalised machine learning models for fuel usage should take into account distance rather than time, according to this article. This strategy is combined with seven predictors obtained from vehicle speed and road grade to develop a highly predictive neural network model for the average fuel consumption in large trucks. This method. A fleet's whole fuel consumption may be reduced by using the suggested model, which can be devised and implemented for each vehicle individually with ease. Distance travelled is pooled into predictors for the model. Fuel consumption may be predicted with a 0.91 coefficient of determination and a mean absolute peak-to-peak percent error of less than 4% for routes that contain both city and highway duty cycle segments using a 1 km window, according to the findings of the study. Index Neural networks; vehicle modelling; fleet management; average fuel usage; data summarization.

INTRODUCTION

Manufacturers, regulators, and customers all have an interest in fuel consumption models for automobiles. Every stage of a vehicle's lifecycle necessitates its use. During the operation and maintenance phase of heavy trucks, we simulate the average fuel consumption of these vehicles. Fuel consumption models may be divided into three primary groups, namely: Understanding the physical system is the basis for the development of physics-based models. These models use mathematical equations to explain the dynamics of the vehicle's components at each time step [1], [2]. Models that are data-driven and reflect an abstract mapping from a group of predictors to the objective outcome, in this instance average fuel consumption [3], [4], are called machine learning models. It is possible to create statistical models that are based on data and that establish a relationship between an indicator's probability distribution and the desired result. Both [5] and [6] may be found. When it comes to cost and precision, you'll have to choose one over the other depending on your specific needs. We provide a concept for large fleets of heavy vehicles that can be readily customised for each vehicle's particular needs in this article. An efficient fleet manager may use realistic models for each vehicle to improve route planning for all of the vehicles in a fleet, which ensures that all route assignments are aligned to reduce total fleet fuel consumption. Road transportation of commodities [7], public transit [3], construction trucks [8], and waste trucks [9] are all examples of these fleets. Methodology has to be applicable and adaptable for any fleet, regardless of the wide variety of vehicle technology (current and future) and configurations. Because of these criteria, machine learning is the preferred method when weighing the precision sought against the expense of developing and adapting a customised model for each vehicle in the fleet individually.

Professor[1,2,3,4], Assistant Professor[1,2,3,4], ,Associate Professor[1,2,3,4],
Department of CSE Engineering,
Pallavi Engineering College,
Mail ID:srajunayak@gmail.com, Mail ID:krishna81.reddy@gmail.com,
Kuntloor(V),Hayathnagar(M),Hyderabad,R.R.Dist.-501505.

There have been a number of earlier models for both immediate and average fuel usage. Instantaneous fuel consumption is best predicted using physics-based models [1], [2], which can capture the dynamics of the system's behaviour at various time steps. Because of the difficulties in recognising patterns in real-time data, machine learning algorithms are unable to accurately estimate immediate fuel use [3]. These models, on the other hand, can accurately detect and learn patterns in average fuel usage [4]. Machine learning models for average fuel consumption have previously used a series of predictors gathered over a period of time to forecast the associated fuel consumption in terms of either gallons per mile or litres per kilometre, depending on the vehicle type. The input space of the predictors is quantized with regard to a given distance rather than a constant time period, but our strategy is still focused on average fuel use. An aggregate of all predictors based on a fixed window that reflects vehicle travel distance provides a superior mapping of input space to output space in the proposed model.... However, prior machine learning models had to execute a time-to-distance translation from the input domain to its output domain in order to discover patterns in that data (i.e., average fuel consumption). Several advantages come from keeping the model's input and output areas at the same scale: We gather data as fast as it has an effect on the result. A vehicle that is stopped may gather the same amount of data as a vehicle that is moving when the input space is sampled with regard to time. Vehicle fuel consumption may be predicted based on factors such as duty cycle and environmental conditions (e.g., the number of stops in an urban traffic over a given distance). Fewer predictors with fewer storage and transmission bandwidth needs may be derived from raw sensor data on-board. Data summarization is best done on-board near the source of the data, given the rise in processing capability of modern vehicles. Additional fuel efficiency improvement may be achieved at the vehicle, route, and time of day levels using new technologies such as V2I and dynamic traffic management [10–12]. In the following sections, you will find: Chapters II and III provide background information, Chapter IV presents a description of the data collection and analysis procedures, Chapter V presents the results of applying the proposed machine learning model in various configurations, and Chapter VI offers conclusions and recommendations for future research.

II. RELATED WORK

Average fuel usage has been modelled using physics-based, machine learning, and statistical models, as previously indicated. These models are based on physics-based, comprehensive vehicle simulations developed by the EPA and the European Commission ([1]). When compared to actual flowmeter results, these models have an accuracy of 3% in forecasting average fuel usage [2]. An extensive development effort was required to achieve this degree of precision. Statistics processes used under rigorous testing settings are at the opposite extreme of the modelling spectrum, and they are used to guarantee that the reported findings are consistent and repeatable. Fuel consumption for new cars may be estimated using well-defined statistical techniques for particular duty cycles constructed from segments of real-world journeys, such as those provided by the Code of Federal Regulations [5]. After market revisions or changes in operating circumstances, the SAE J1321 [6] standard is used to estimate fuel consumption for trucks and buses. This standard compares the performance of identical vehicles on the same route under similar operating circumstances based on real-world data. For example, in [13], the standard was used to compare the fuel consumption of a control car to that of two test vehicles after changing lubricating fluids in the engine, gearbox, and axle, respectively. Additionally, the standard was used in [8] to evaluate the performance of three fuel technologies in two coal mine trucks. In several studies, the generalizability of machine learning models to diverse vehicles and operating situations drew researchers to this modelling tool for fuel consumption predictions. These models will be discussed further in terms of the underlying machine learning approach, as well as representations for input space and output space, in the next sections of this paper. For the objective of simulating fuel use, a variety of machine learning algorithms have been employed and contrasted. Consider the following examples: in [3], [4], and [7], neural networks are pitted against each other, whereas neural networks are pitted against multivariate regression splines. These research have identified a preferred method based on their findings. As mentioned in [7] and [14], the procedures are equivalent. However, the differences between these techniques are minor. Differences in data collecting and data summarising approaches may be at blame, we suspect. Because neural networks operate best with models that have continuous input and output variables, that's the approach we used in our article. To top it all off, neural networks are less vulnerable to noisy data[15]. There is also a wide range in the input of previously suggested fuel consumption models. Driver behaviour, vehicle dynamics, and how the

surrounding environment interact could all be included in a comprehensive model. Combinations of first, second, third and fourth order vehicle acceleration and speed are used in [4] as predictors of vehicle performance. Vehicle speed, distance travelled, elevation, longitude/latitude and day of the week are some of the variables in [3]. Road characteristics (e.g., gradient, curvature, and roughness) and operational circumstances of the vehicle (e.g. speed vehicle, acceleration, gear, and percent torque) are taken into consideration in [7]. [8] Acceleration, percent torque, and gradient were determined to be the most significant predictors in this earlier investigation. During data collection, vehicle speed was not a factor since it was kept almost constant. There were more than 30 factors examined in [15], with the wind speed, platooning, engine strength, and braking rate being the most significant. The road slope, speed, and weight of the vehicle were also shown to be significant predictors. In [15], the weight of the vehicle was determined using its suspension, which is not a conventional sensor. Vehicle speed and road slope are also used as predictors of the proposed model in this article. Non-invasive, inexpensive, and generally accessible telematics devices can collect these characteristics. Different sensor readings that are taken at regular intervals are often used as predictors in models [3] and [4]. When comparing the accuracy of suggested fuel consumption models with regard to input data obtained at 1 minute and 10 minute intervals, [15] author concluded the 10 minute period generates more accurate models. There is a minute-by-minute collection of measurements in [7]. Because the trucks were travelling at a consistent pace, this amounted to gathering input data across a set distance of about one mile. For fuel consumption modelling, both [7] and [15] appear to indicate that data collection over a longer distance is preferable. The values of the predictors are aggregated across specified windows of journey distance in this article. Additionally, we examine how the length of the window affects model performance in terms of real-world duty cycles with differing vehicle speeds. You may get either the fuel rate (liters/hour) or the average fuel consumption (liters/100km) from fuel consumption models. Fuel rate models are used to anticipate how much fuel a vehicle will use at any given time. Because of their lack of fine detail, these models are notoriously inaccurate (i.e., at the level of each sample). To provide a more realistic picture of typical fuel consumption, the models take a long-term or long-distance average of the expected fuel rates. Accuracy ranged from 3.73 percent to 6.83 percent from the recorded average fuel rate during a whole real-world duty cycle when using this method in [4]. [3] utilised a similar method and was able to estimate gasoline consumption throughout a 365-kilometer journey with an accuracy of 2 percent using the model's expected fuel rates. In this investigation, the correlation between point-wise predicted fuel rates and observed fuel rates was found to be poor (0.3). A model's input and output are both in the time domain in each of these investigations. The result of the models described in [7] is the ratio of fuel burned to the distance travelled. Because the cars in this research were all travelling at the same speed at the same time, this study was able to accurately sample both input and output data. The anticipated fuel usage had an average absolute inaccuracy of 3%. An R2 value over 0.8 indicates that the model described in [7] is better equipped to generate point-wise predictions than the two previous investigations. Some of the difficulties that machine learning models experience when their input and output are from distinct domains are highlighted in the models provided in [15]. Over a period of 10 minutes, the input is aggregated and the result is fuel usage for the distance travelled during that time period. For a typical fuel consumption of 30 l/100km, the root mean square error of the anticipated fuel consumption was 7.4 l/100km for the whole duty cycle. Compared to models that compute the error throughout the whole journey, this error, which is calculated point-wise for each 10 minute time period, is quite significant.

III. A DIGITAL MODEL USING NEURAL NETWORKS

ANNs, as previously indicated, are often used in the development of digital models for complex systems (e.g., [4], [16], [17]). A transfer function F(p) = o may be used to depict a complicated system, with F() standing for the system, p for the input predictors, and o for the system response or output. As a result, an approximation of the system is provided.

$$\hat{o} = f(\vec{p})$$

F(o) is the prediction model, and o is an estimate of the actual output. Model f (p) in (1) is not confined to artificial neural networks. Examples include employing differential equations [18] or a weight matrix generated by support vector machines to represent f (p) [19]. Feed Forward Neural Networks are the ANNs used in this study (FNN). For example, the matrix W, the vectors w, b, and c represent adjustable weights, and p is an input vector consisting of a collection of predictors that capture the system state and are thus considered to be able to reliably

forecast o. The nonlinear properties of a complex system may be represented using a nonlinear activation function () in FNNs rather than linear models (e.g., linear regression) (3).

$$\hat{o} = \sigma\left(\Sigma_k w_k \cdot \sigma\left(\Sigma_l w_{lk} \cdot p_l + b_l\right) + \right.$$

$$\sigma(\tau) = \frac{1}{1+e^{-\tau}}.$$

To define the model, we use the weight coefficients of the following formulas: Wlk = W, Wk = W, Bl = B, and Ck = C. In this equation, Wlk = W, Wk = W, Bl = B, and Ck = C define the model where Wk = W, Wk = W, and Bl = B define the hidden layer. Weights are selected such that the discrepancy between the system response (o) and the model output (o) is as little as possible. In this model, the weights W and w determine the connection between a neuron's input and output, while the biases Wb and c allow for some degree of flexibility in the model. As part of the training phase, the network is given a set of parameters to operate with. Predictors (p) and outputs (o) must be known before to training (o).

$$E = \frac{1}{2}\left(o - \hat{o}\right)^2.$$

Ftr and Fts are feature sets (F) that generally include two disconnected subsets of input predictors (I) and output (O). For the FNN, the weights and biases are calculated using the first batch of data, Ftr. Validation of the model takes place using the second dataset, Fts. That way, even if the input values are different from what was seen during the training, the model may still generalise. Particle swarm optimization [20] and back propagation [21] are two methods for training that may be used in conjunction with each other. Both approaches are used in this study. Various more ways will be investigated in the future in order to see whether they can enhance the model's capacity to accurately forecast the future. A random pair of (input, output) features from Ftr is randomly selected for each iteration of training and the network's weights are updated accordingly. Using the model's anticipated output value and the actual output value, the error is calculated (4). This mistake is then retransmitted from the network's output layer back to its input layer. According to W, w, b, and c, the gradient of this mistake is determined as illustrated in W. (5).

$$\frac{\partial E}{\partial w_{l,k}} = \frac{\partial\left(\frac{1}{2}\left(o - \hat{o}\right)^2\right)}{\partial w_{l,k}}.$$

The new weights in W(s) are updated for each training iteration step s by adding w (s) l,k as stated by (6) to the prior weight value at step s1 according to (7). When substituting w (s) l,k with w (s) k, b (s) l, and c (s) k, the corresponding values for w (s), b (s) l, and c (s) k are obtained in the same way. A model's learning rate is controlled by the constants and in (6), which are the control parameters. With regard to prior training iterations and the direction of gradient (5), the constant determines how much weights should be added to weights in the direction of that gradient.

$$\Delta w_{l,k}^{(s)} = -\epsilon \cdot \frac{\partial E}{\partial w_{l,k}} + \mu \cdot \Delta w_{l,k}^{(s-1)}$$

$$w_{l,k}^{(s)} = w_{l,k}^{(s-1)} + \Delta w_{l,k}^{(s)}$$

After the model has been trained, it must be validated to see whether it can generalise to (input, output) characteristics from Fts that were not utilised in the training phase.

$$C_D = 1 - \frac{\sum_i \left(o_i - \hat{o}_i\right)^2}{\sum_i \left(o_i - \bar{o}\right)^2}$$

This performance may be assessed using a variety of measures applied to the test data. With this numerator as the numerator, we get the Coefficient of Determination (8), which measures how well our model predicts system behaviour in relation to actual system behaviour. In order to get the denominator, subtract the actual responses from the system's mean responses. The machine learning model's projected reaction and the system's actual point-wise response are both represented by the coefficient of determination (CD). Accurate point-by-point predictions are indicated by CD values near to 1. Point-wise forecasts are less accurate when the value is near to zero or negative. The CD may have a value ranging from -infinity to one.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|o_i - \hat{o}_i|$$

$$MAPE_{pk} = \frac{1}{N}\sum_{i=1}^{N}\frac{|o_i - \hat{o}_i|}{o_{pk}} \cdot 100$$

$$RMSE = \frac{1}{N}\sqrt{\sum_{i=1}^{N}(o_i - \hat{o}_i)^2}$$

Furthermore, the mean absolute error (MAE) as indicated in (9) [23] and the mean absolute peak-to-peak percent error (MAP Epk) as shown in (10) are used to assess the model's accuracy in addition the coefficient of determination.

$$o_{pk} = max(\vec{o}) - min(\vec{o})$$

In Fts, N is the total number of test points, It was found that MAP Epk was preferable than MAP E (Mean Absolute Percent Error) because it avoids biassing and unsymmetrical outcomes when oi is near to oi. Model prediction errors are measured using MAE, whereas MAP Epk presents these errors as a percentage of the greatest potential prediction errors. Additionally, the RMSE (11) is supplied where it is appropriate. All of the chosen measures assess the model's performance on a point-by-point basis rather than cumulatively over the course of a journey. Predictors' relative relevance is determined by comparing their weight (IW) with other model inputs, as specified in (13) [22].

$$IW_l = \frac{\sum_j |w_{lk}|}{\sum_l \sum_k |w_{lk}|} \cdot 100$$

It was necessary to alter the measure above to accommodate for input characteristics using a broad range of methods. The paper's corrected IW metric is provided by

$$AIW_l = \frac{IW_l}{\sum_k IW_k} \cdot \bar{p}_l$$

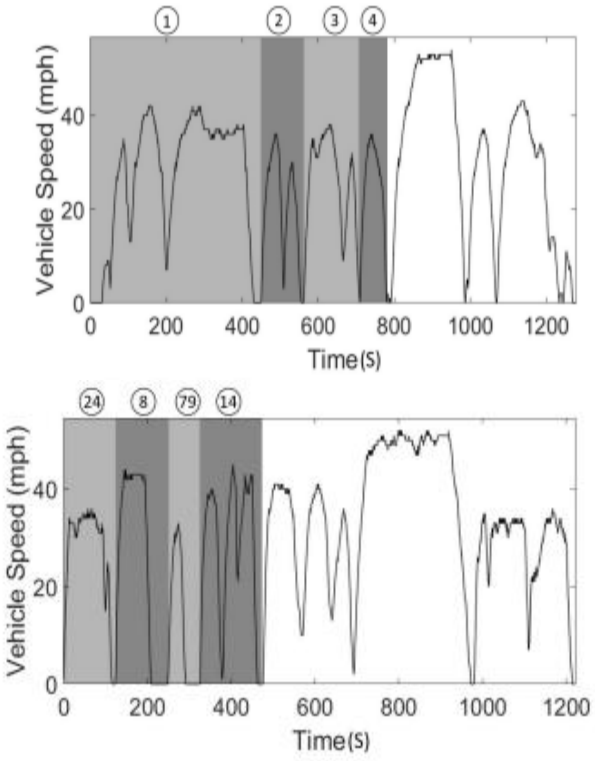
in where p l is the mean value of the input feature pl throughout the test data set Fts; AIWl is a measure of how much excitation the input feature (pl) provides the network's first layer in comparison to other input features. An estimate of the feature's importance may be found here. Excitation of low AIW features may have a considerable impact on output of the model as it propagates through its buried layer.

DATA COLLECTION AND SUMMARIZATION

An 8-ton truck with an estimated mass of 8,700 kg is used to create the model, which is based on the duty cycles of that vehicle's shifts in Indianapolis. Serial control and communications in heavy-duty vehicle networks were implemented in accordance with the SAE J1939 standard [24]. Good or poor conduct was questioned of 12 drivers on two separate routes. If you're driving well, you anticipate the need to brake and let the car do the rest. The distribution of drivers and routes is not homogeneous throughout the data set since some drivers participated more than others. 3 302 890 data points sampled at 50 Hz from the vehicle CAN bus were produced during this field test, which included 56 excursions of varied lengths totaling 778.89 km in total distance travelled. The majority of the journeys were between 10 and 15 kilometres long. From the field duty cycles picked at random, synthetic duty cycles were assembled to enhance the number of points of data collection. As a result, distinct sets of data are generated for training (Ftr) and testing (Fts) segments, with one set of drivers assigned to each.

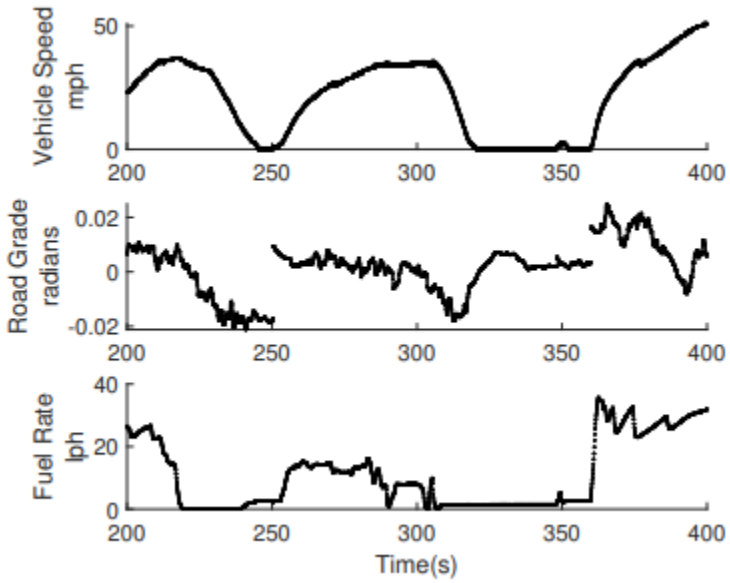
Model Predictors

The model's predictors required a number of processing steps. Road grade and transmission output speed are used to create these predictions. Downsampling of the road grade and calculating vehicle speed from the trans mission output speed were the initial steps in the processing. An on-board inclinometer was used to measure the road gradient, which was then down-sampled to 1 Hz. The data also demonstrated that the vehicle speed and transmission output speed have a linear connection, as shown in the following equation: A moving average low pass filter was used to minimise noise in the variable, and the variable was down-sampled from 50 Hz to 1 Hz to further reduce noise. Creating the synthetic duty cycles was the goal of the second processing phase. It was for this reason that data from actual vehicle stops was used to partition the duty cycles collected (Figure 1). All of the study's twelve drivers provided a total of 455 actual data segments. The training data set (Ftr) was derived from 358 segments received from nine drivers, while the testing data set (Fts) was derived from 97 segments obtained from the remaining three drivers in the research.



Sample genuine duty cycles are shown in Fig. 1. (top). Real data segments 24, 8, 79, and 14 were concatenated to construct this synthetic duty cycle (bottom).

To construct a synthetic duty cycle, genuine data segments are sampled and concatenated until a total distance of 15 kilometres is attained. The 15-kilometer route was chosen to reflect the actual paths used during data collecting in the field. An average of five segments is required to generate 15 kilometres of data, according to a study. As you can see in Figure 1, we were able to produce a synthetic duty cycle via the use of this method. When segments were combined in the manner described above, the resulting vehicle speed was constant. However, as illustrated in Figure 2, road grade discontinuities were found between segments. Based on the selected window, these duty cycles are then averaged over a predetermined trip distance (x). For each data set and window size studied in this research, the total number of data points (i.e. windows) and total distance are shown in Table I. Predictors for the proposed model are generated as the last stage in the input data processing. Using the speed and gradient of the road, these predictors are computed for each window. • amount of stops, • time spent in a halting state



Experiment with a three-segment duty cycle, as seen in Figure 2. The duty cycle was designed to maintain a constant pace for the vehicle. It's possible that road conditions or gasoline prices may change.

For the training (F(x)tr) and testing (F(x)ts) data sets, the number of data points (I.E., WINDOWS), as well as the total distance between the data points, is shown in Table I. (I.E., 1, 2, AND 5 km.)

Window size	$\mathbb{F}(x)_{tr}$		$\mathbb{F}(x)_{ts}$	
	Number of Points	Distance (km)	Number of Points	Distance (km)
$x = 1\,km$	20,000	20,000	32,089	32,089
$x = 2\,km$	20,000	40,000	23,106	46,212
$x = 5\,km$	20,000	100,000	6,061	30,305

Average speed, characteristic acceleration, and aerodynamic speed squared are all terms used to describe this phenomenon.

• kinetic energy shift and

Potential energy shifts as a result ofSince these variables are thought to reflect both vehicle dynamics and the driver's behavioural patterns and the route's influence on goal output, they were chosen for inclusion in this study (i.e., fuel consumption). It has previously been shown that the characteristic acceleration (16) and aerodynamic speed squared (18) may be used to predict a vehicle's total fuel

consumption for a certain duty cycle. According to this research, the effect of aerodynamics on fuel consumption is best captured by calculating aerodynamic speed squared from the vehicle's characteristic acceleration and the inertia work required to accelerate the vehicle [25].

$$\widetilde{a} = \frac{1}{x} \cdot \sum_{i=1}^{N-1} \left| \frac{\left(v_{i+1}^2 - v_i^2\right)}{2} + g \cdot \Delta h_{i,i+1} \right|$$

This is the total distance between a vehicle and the window and which represents the vehicle's speed at successive intervals ti and ti+1, respectively, where g is the gravitational constant. According to this equation, the elevation difference between two points on the same road is equal to the distance travelled between the two points, divided by the road grade at each point in time (in this case, the road grade at time ti+1).

$$\Delta h_{i,i+1} = \Delta d_{i,i+1} \cdot sin(\alpha_i)$$

$$v_{aero}^2 = \frac{1}{x} \cdot \sum_{i=1}^{N-1} \frac{v_{i+1}^3 + v_{i+1}^2 v_i + v_{i+1} v_i^2 + v_i^3}{4} \cdot \Delta t_{i,i+1}$$

where ti,i+1 represents the study's set sampling interval of one second. On the basis of this fact, the starting expression (18) presented in [25] was reduced to the approximation displayed in order to decrease the amount of on-board computing required for v 2 aero (19).

$$v_{aero}^2 \approx \frac{1}{x} \cdot \sum_{i=1}^{N-1} \left(\frac{v_{i+1} + v_i}{2} \right)^3$$

To the aforementioned predictors in [25], which aim to capture the vehicle's transient dynamics, we added two more predictors that track changes in the vehicle's average dynamics over the course of a given window. According to (20) and (21), these predictors are the change in kinetic and potential energy (CKE and CP E), respectively.

$$CKE = \frac{1}{2} \cdot m \cdot (v_N - v_1)^2$$

There is a constant in this research, m, the mass of the vehicle, and v1 and vN are the vehicle speeds at which the vehicle is travelling at the beginning and end of this window.

$$CPE = m \cdot g \cdot \sum_{i=1}^{N-1} \Delta h_{i,i+1}$$

Due to the very small amount of total energy spent by the vehicle, it is critical to monitor the vehicle's kinetic and potential energy changes over the duty cycle. The amount of fuel energy that is converted to kinetic and/or potential energy decreases as the distance travelled increases.

Model Output

Each window's average fuel consumption in l/100km represents the model's output. The CAN bus is used to get fuel consumption data for the average consumption. Discontinuities in the fuel rate are noticed from one segment to the next, as is the case with road grade and synthetic duty cycles produced from a random selection of genuine duty cycle segments (Figure 2). Because the fuel rates are averaged throughout the whole window, these discontinuities have no effect on the model's result (i.e., average fuel consumption). Data gathered on the road has shown that average fuel usage for all journeys varies by section. There was a 20% variation in fuel usage between excellent and poor drivers on en-tire journeys. The average fuel usage varies widely depending on the size of the window, as well. For the 1, 2, and 5 km windows, Table II displays the mean and standard deviation of the average fuel usage for all journeys. Standard deviation decreases as window size rises, despite the fact that the mean fuel consumption across all windows remains roughly constant. All of the proposed model's input characteristics are generated from the vehicle speed and the road grade, which are recorded at a rate of 1 Hz. A telematics device may provide the data needed to calculate this information. These findings are the result of this research.

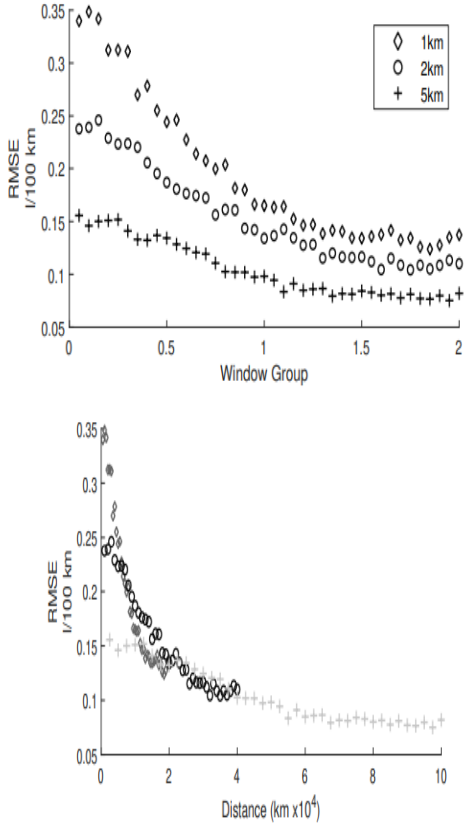
TABLE II MEAN AND STANDARD DEVIATION OF AVERAGE FUEL CONSUMPTION OVER 1, 2, AND 5 km WINDOWS IN THE REAL DATA. THE LAST ROW OF THE TABLE SHOWS THE AVERAGE FUEL CONSUMPTION FOR ALL THE TRIPS

	mean l/100km	*std* l/100km
1 *km*	20.42	9.17
2 *km*	20.69	5.94
5 *km*	20.98	3.97
trips	20.88	3.96

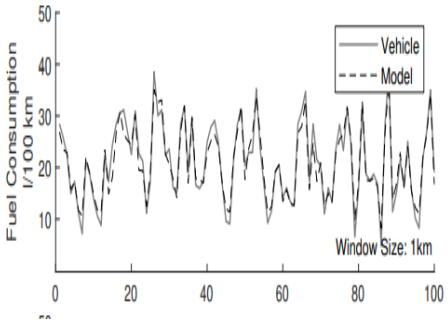
CAN bus sensor readings were used to create variables. As the data source and sample frequency change, so does the model's accuracy. The model's accuracy is also dependent on the output feature's precision. According to the Might bus, fuel consumption can be inaccurate by as much as 5 percent [26]. There is a greater degree of precision that may be achieved using flowmeters. Flowmeters, on the other hand, are more costly. In [7] and [15], as well as in this article, gasoline consumption values obtained from the CAN bus were used, as were high precision fuel sensors. It will be investigated in the future how the data sources are accurate.

MODEL VALIDATION

The neural network model is fed by the seven predictors given in Section IV. The initial layer of the network is made up of this. An additional hidden layer of five neurons receives the input from the top layer. A single neuron feeds the output player from the hidden layer. Three models with window widths of 1, 2 and 5 km are shown in Figure 3 with the RMSE (11) during training. The RMSE values for each data point in the top plot are based on the model's performance after training on 500 windows. Using this figure, we can see that the RMSE of all models is less than 0.2 l/100km. The models' rates of convergence, on the other hand, vary. Although at first it seems to start out at 0.16l/100km after 500 training windows, the model's RMSE drops to 0.08l/100km after it has reached its convergence point. For the 1 km model, the values are 0.34 l/100km and 0.14 l/100km, respectively. As shown in Table II, aggregating input and output data over 5 km provides a stable profile for the fuel consumption of the vehicle over the routes, and this profile does not necessitate extensive learning. This trend is further supported by a difference in standard deviation of the average fuel consumption for the 1 km and 5 km windows This conclusion is in line with other research. When it comes to estimating fuel usage, extended route segments for small cars in metropolitan locations are more accurate than shorter ones, according to [14]. In a prior survey, 64 percent of the excursions were less than 5 kilometres in length. It was shown that data collection over 10 minute intervals was more accurate than data collection over 1 minute intervals in [15]. The longer the data collecting period, the more likely it is that fuel usage and distance travelled will have a linear connection. Despite the fact that this method provides a decent estimate of typical fuel use,



For window widths of 1, 2 and 5 km, there was a significant decrease in the RMSE throughout the training procedure. Long-distance fuel usage may not match point-wise expected values when RMSE is plotted (top) against the number of window groups with 500 consecutive windows per group (bottom).



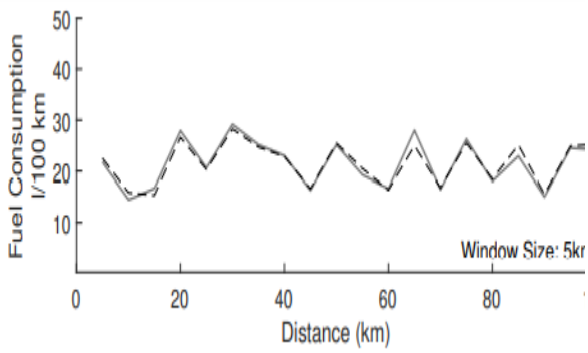


FIGURE 4: Fuel usage throughout the first 100 kilometres from Fts, predicted vs real, broken down into one-kilometer and five-kilometer windows (bottom).

For the first 100 kilometres of the test data set (Fts), fuel consumption for window widths of 1 km and 5 km is shown in Figure 4 to demonstrate this tendency. Per-window tracking is easier using the 1 km model, which is more accurate than the 2 km model. When comparing anticipated and actual fuel use, the same conclusion was reached, as shown in Figure 5. Figures 4 and 5 show the findings.Models trained using the training data set (Ftr) in Table I are listed here. As a comparison of the point-wise prediction performance for changing window widths, five alternative models are trained

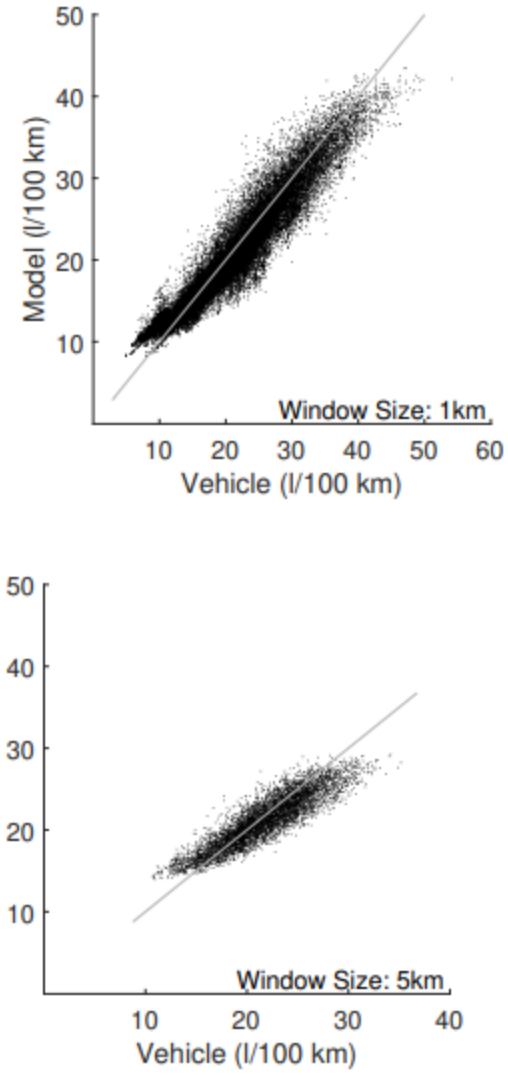


Figure 5 compares the predicted fuel consumption (left) with the actual fuel consumption (right) for the whole test data set during the 1 km and 5 km windows (right).

neural network with random starting values for the weights. Five training/testing cycles for each window size yielded these five performance measures (Ta ble III). Table III shows the standard deviation for each of the five runs in parentheses underneath the metric's performance value. Stable models may be inferred from the standard deviations shown here.

TABLE III PREDICTIVE ACCURACY OF THE FUEL CONSUMPTION MODELS FOR 1, 2 AND 5 km AGGREGATION WINDOWS

Window	1 km	2 km	5 km
CD	0.91 (0.0066)	0.87 (0.0085)	0.79 (0.0136)
$RMSE$ (l/100km)	0.0132 (0.0005)	0.0142 (0.0005)	0.0234 (0.0008)
MAE (l/100km)	1.88 (0.0626)	1.69 (0.0515)	1.43 (0.0466)
$MAPE_{pk}$	3.74% (0.12%)	4.20% (0.13%)	5.83% (0.19%)
$Points$	32,089	23,106	6,061

Overall, the 1 km model performs better than the other two window widths in all measures, according to Table III of the results It has been previously stated that the performance metrics measure the model's performance in terms of points. An accuracy of 0.91 in the 1 km model's coefficient of determination (CD) shows that the model can accurately predict fuel usage over that distance. The CD lowers in size as the window size grows. The suggested model outperforms [3], [7] in terms of MAE and CD, despite the inclusion of high-precision fuel sensors in [3]. Modelling errors are also smaller than in [15], with an average difference of of 0.025 l/100km. However, the test distance employed in this article is greater than that used in [15]. The smaller the RMSE, the better. Fuel consumption accuracy for the models is likewise within the MAP Epk values stated in [4]. This study compares the actual vs. projected error at the window level, while [4] reports the error for the full journey. According to Table III's performance measures, the suggested models seem to be utilising highly predictive input characteristics and that these features are properly translated to the output space of the model. Table IV summarises the AIW values of the predictors in order to better appreciate their contributions.

TABLE IV ADJUSTED INFLUENCE OF WEIGHTS (AIW) FOR TR THE PREDICTORS IN THE PROPOSED MODEL

Window	1 km	2 km	5 km
No. of Stops	1.49	2.29	4.63
Stop Time	0.62	1.24	3.44
Avg. Moving Speed	13.73	10.78	8.98
$\widetilde{a}$	12.47	14.32	12.98
v_{aero}^2	11.73	11.64	10.30
CKE	17.04	16.13	12.26
CPE	13.73	11.45	9.38
Bias	29.21	32.15	38.03

As the window size rises, the relevance of the number of stops and the stop duration increases. Because there are fewer pauses in the 1 km window compared to the 2 or 5 km windows, this is to be anticipated. Other than those, all of the predictors show strong AIW across all window widths. Predictive accuracy was reduced by removing any of these predictors. As window widths rise, the model depends less on vehicle dynamics and more on events linked to the distance travelled to estimate fuel consumption, as seen by the increase in AIWs for number of stops and stop duration, as well as the drop in AIWs for the remaining predictors. The two novel predictors suggested in this article, according to Ta ble IV, contribute to fuel consumption prediction in a similar way as typical acceleration and aerodynamic speed [15], as well as the average moving speed [25].

CONCLUSION

Using machine learning, this article constructed a model for every heavy vehicle in a fleet. There are seven variables that make up the model: number of stops, the duration of each stop, the average speed, the typical acceleration of the vehicle, and the change in kinetic and potential energy, respectively. This research introduces two more predictors to better represent the vehicle's average dynamic behaviour. Speed and grade are used to determine the model's predictors. Telematics devices, which are increasingly becoming an intrinsic feature of automobiles with internet connectivity, may provide quick access to these data. In addition, these two variables make it simple to calculate the predictors while travelling. Predictions from the model are averaged across a predetermined distance travelled (i.e., window) rather than a predetermined time period. Machine learning models for fuel consumption with an RMSE of less than 0.01l/100km were created by mapping the input space to the distance domain, which matches with the goal output. Models with 1, 2, and 5 km windows were tested. For precision, the 1 kilometre window is the best option. This model has a CD of 0.91 and can accurately forecast fuel usage on a per-kilometer basis. For complete long-distance excursions, this model's performance is more in line with physics-based models, and it outperforms earlier machine learning models. The cost of the model in terms of data collection and on-board computation should be taken into account when selecting a window size. Furthermore, the size of the window is likely to vary depending on the programme. It is advised to use the 1 km window for fleets with short journeys (e.g., construction trucks inside a site or urban traffic routes). A 5 kilometre window may be enough for

long-haul fleets. It was found that the 1 km window was better suited to the duty cycles in this research since they included both highway and city traffic. Identifying these distinct elements and determining the optimal window size will be the focus of future effort. The model is being extended to include other types of automobiles, such as those with changing mass or older models. It will be necessary to include predictors for these features in order to account for the effect on fuel consumption that changes in vehicle mass and wear have. Aside from this, future research will investigate the minimum distance necessary for training each model, as well as the frequency of online training in order to ensure that a model's prediction accuracy is maintained.

REFERENCES

[1] B. Lee, L. Quinones, and J. Sanchez, "Development of greenhouse gas emissions model for 2014-2017 heavy-and medium-duty vehicle compliance," SAE Technical Paper, Tech. Rep., 2011.

[2] G. Fontaras, R. Luz, K. Anagnostopoulus, D. Savvidis, S. Hausberger, and M. Rexeis, "Monitoring co2 emissions from hdv in europe-an experimental proof of concept of the proposed methodolgical approach," in 20th International Transport and Air Pollution Conference, 2014.

[3] S. Wickramanayake and H. D. Bandara, "Fuel consumption prediction of fleet vehicles using machine learning: A comparative study," in Moratuwa Engineering Research Conference (MERCon), 2016. IEEE, 2016, pp. 90–95.

[4] L. Wang, A. Duran, J. Gonder, and K. Kelly, "Modeling heavy/mediumduty fuel consumption based on drive cycle properties," SAE Technical Paper, Tech. Rep., 2015.

[5] Fuel Economy and Greenhouse gas exhaust emissions of motor vehicles Subpart B - Fuel Economy and Carbon-Related Exhaust Emission Test Procedures, Code of Federal Regulations Std. 600.111-08, Apr 2014.

[6] SAE International Surface Vehicle Recommended Practice, Fuel Consumption Test Procedure - Type II, Society of Automotive Engineers Std., 2012.

[7] F. Perrotta, T. Parry, and L. C. Neves, "Application of machine learning for fuel consumption modelling of trucks," in Big Data (Big Data), 2017 IEEE International Conference on. IEEE, 2017, pp. 3810–3815.

[8] S. F. Haggis, T. A. Hansen, K. D. Hicks, R. G. Richards, and R. Marx, "In-use evaluation of fuel economy and emissions from coal haul trucks using modified sae j1321 procedures and pems," SAE International Journal of Commercial Vehicles, vol. 1, no. 2008-01-1302, pp. 210–221, 2008.

[9] A. Ivanco, R. Johri, and Z. Filipi, "Assessing the regeneration potential for a refuse truck over a real-world duty cycle," SAE International Journal of Commercial Vehicles, vol. 5, no. 2012-01-1030, pp. 364–370, 2012.

[10] A. A. Zaidi, B. Kulcsr, and H. Wymeersch, "Back-pressure traffic signal control with fixed and adaptive routing for urban vehicular networks," IEEE Transactions on Intelligent Transportation Systems, vol. 17, no. 8, pp. 2134–2143, Aug 2016.

[11] J. Zhao, W. Li, J. Wang, and X. Ban, "Dynamic traffic signal timing optimization strategy incorporating various vehicle fuel consumption characteristics," IEEE Transactions on Vehicular Technology, vol. 65, no. 6, pp. 3874–3887, June 2016.

[12] G. Ma, M. Ghasemi, and X. Song, "Integrated powertrain energy management and vehicle coordination for multiple connected hybrid electric vehicles," IEEE Transactions on Vehicular Technology, vol. 67, no. 4, pp. 2893–2899, April 2018.

[13] A. C. Brandt, E. A. Frame, and R. W. Warden, "Sae j1321 testing using m1083a1 fmtvs," SOUTHWEST RESEARCH INST SAN ANTONIO TX TARDEC FUELS AND LUBRICANTS RESEARCH FACILITY, Tech. Rep., 2010.

[14] W. Zeng, T. Miwa, and T. Morikawa, "Exploring trip fuel consumption by machine learning from gpsand can bus data," Journal of the Eastern Asia Society for Transportation Studies, vol. 11, pp. 906–921, 2015.

[15] H. Almer, "Machine learning and statistical analysis in fuel consumption ´ prediction for heavy vehicles," 2015.

[16] S. McBride, C. Sandu, A. Alatorre, and A. Victorino, “Estimation of vehicle tire-road contact forces: A comparison between artificial neural network and observed theory approaches,” SAE Technical Paper, pp. 01–0562, 2018.

[17] G. N. Bifulco, F. Galante, L. Pariota, and M. R. Spena, “A linear model for the estimation of fuel consumption and the impact evaluation of advanced driving assistance systems,” Sustainability, vol. 7, no. 10, pp. 14 326–14 343, 2015.

[18] C. M. Atkinson, S. Petreanu, N. N. Clark, R. J. Atkinson, T. I. McDaniel, S. Nandkumar, and P. Famouri, “Numerical simulation of a two-stroke linear engine-alternator combination,” SAE Technical Paper, Tech. Rep., 1999.

[19] C.-M. Vong, P.-K. Wong, and Y.-P. Li, “Prediction of automotive engine power and torque using least squares support vector machines and bayesian inference,” Engineering Applications of Artificial Intelligence, vol. 19, no. 3, pp. 277–287, 2006.

[20] Y. Shi and R. Eberhart, “A modified particle swarm optimizer,” in 1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence (Cat. No