A Machine Learning Model for Average Fuel Consumption in Heavy Vehicles

Avala Jaswanth[a], Dr.Deevena Raju[a]

Dept. of Computer Science and Engineering, Faculty of Science and Technology (IcfaiTech, ICFAI Foundation for Higher Education Hyderabad)

Abstract-I look into ways to forecast fuel use in large vehicles using machine learning. I look at data from several sources that describe the characteristics of the road, the car, the driver, and the weather, and I identify a regression to a fuel consumption expressed in litres per distance. The thesis was written for Scania and uses information that Scania has access to. I assess the most effective machine learning techniques, how the frequency of data collection impacts the prediction, and which features have the most bearing on fuel usage. The lower collecting frequency of 10 minutes is, in my opinion, superior than the greater frequency of 1 minute. The evaluated models perform similarly, and the factors that affect fuel consumption most are the weight of the vehicle, the speed at which it is travelling, and the slope of the road.

Keywords: Machine learning, regression, fuel consumption

1. INTRODUCTION

In this work, techniques for estimating fuel consumption in heavy trucks using machine learning (ML) and statistical analysis are evaluated. The goal is to forecast fuel usage in litres per mile using past data detailing driving circumstances. A huge number of attributes that describe a fuel consumption situation must be analysed, and ML techniques must be used to find a regression line connecting these attributes to fuel consumption. Environmental factors, vehicle design, driver behaviour, and weather conditions may all be considered. How to make such predictions for aircraft [1], [2], engines [3], passenger cars [4], and heavy vehicles [5], [6], [7] has been studied. The earlier research offers recommendations for the ML techniques that predict fuel consumption most accurately as well as the characteristics that have the biggest impact on road vehicles' fuel consumption. How to anticipate fuel usage for Scania's heavy vehicles using the data sources accessible to Scania is the specific issue examined in this study.

The study is a component of the COMPANION project, an initiative by the haulage firm Transportes Cerezuela in Spain, Scania CV AB, Volkswagen Group Research, KTH, Oldenburger Institut für Informatik (OFFIS), IDIADA Automotive Technology, and Science & Technology in the Netherlands [8]. By investigating how to build a fuel model and predict fuel consumption using platooning as a factor, this study contributes to the COMPANION project.

The evaluation of several ML-based techniques for regression from a set of descriptive attributes to a fuel consumption in litres per km is one objective of the study. Examples of characteristics taken into account in the study include the vehicle's weight, engine power, speed, slope, and speed limit, as well as weather information like wind speed and direction. The Fleet Management (FM) system of Scania, a GPS navigation system, and weather monitoring are some of the sources from which the information is gathered information on the configuration of the vehicle and SMHI data. On the basis of this data, a variety of ML techniques are trained to identify a regression to fuel usage.

The frequency at which active vehicles transmit the data from Scania's FM system can vary from vehicle to vehicle. The messages include details about the vehicle's location, the reading on the odometer at the moment, the amount of fuel it is using, and other descriptive information. The primary objective of the study is to determine if the current standard sampling rate of 10 minutes is adequate and how the frequency of data collecting influences prediction.

2. BACKGROUND

The study is explained and key ideas are outlined in the parts that follow. Section 2.1 introduces the COMPANION project, of which this study is a part, and discusses the platooning concept. The FM system, which is the study's most significant data source and a requirement for performing statistical fuel consumption prediction, is described in Section 2.2. The field's prior research is presented in later sections, which further highlight the study's importance. The following queries will be addressed by this study:

- How does the frequency of data collection influence the accuracy of prediction? Can a sampling rate of 10 minutes be used, or is a greater frequency needed?
- Which characteristics in the data are most important for fuel prediction?
- Which of the assessed ML techniques is most appropriate for this issue?

The COMPANION project and platooning

This research is a component of the COMPANION project, a study of the design, management, and use of vehicle platoons or road trains [8]. It has been demonstrated that platooning while driving lowers air resistance and lowers fuel consumption [6], [9], [10]. The COMPANION project's objective is to create a real-time coordination system that dynamically forms, maintains, and disbands platoons in accordance with a decision-making mechanism that takes into account both historical and current data about the condition of the infrastructure (traffic, weather, etc.) [8]. By investigating how to build a fuel model and predict fuel consumption using platooning as a factor, this study contributes to the COMPANION project.

Scania CV AB, Volkswagen Group Research, KTH, Oldenburger Institut für Informatik (OFFIS), IDIADA Automotive Technology, Science & Technology in the Netherlands, and the Spanish haulage firm Transportes Cerezuela are all participating in the COMPANION project [8].

FLEET MANAGEMENT

The management of a business's transportation fleet is known as fleet management FM. In the Scania example, the customer has enrolled in the Scania Fleet Management programme in order to track and manage all Scania trucks. The FM system for Scania has a car. Vehicles communicate signals with their locations, current fuel levels, and other attributes at predetermined intervals as part of a tracking component. The frequency can be set to any unit of time, but a frequency of 10 minutes is the most typical. For some automobiles, the frequency could reach one minute. The FM system also has a driver behaviour analysis component that monitors data on the frequency of hard breaks, the amount of time the car is idle, and other factors. More information about the FM data is provided in Section 4.1. A truck-mounted onboard computer records the data for each vehicle, which is then transmitted via a telecommunications link to a backend system. The information is kept in a database system and made accessible to Scania employees as well as to consumers outside of Scania.

3. METHODOLOGY

A data mining phase and a training phase can be used to separate the process of gathering the data set, developing the models, and assessing the outcomes. The phase of data mining involves gathering, analysing, and combining data from many sources. ML models are fitted to the pooled data during the training phase. The highest performing model will be identified and its most influential attributes will be assessed after the training is complete.

In the topic of artificial intelligence known as machine learning, systems that can learn from examples in various ways are built. The concept can be explained using Tom M. Mitchell's definition [11]. If a computer program's performance at tasks in a class of tasks T, as assessed by a performance measure P, increases with experience E, then it is said to learn from experience E with respect to that class of tasks T and performance measure P. Mr. Tom Mitchell

The statistics about fuel use and other information gathered from various data sources constitute the experience E in the context of this study. Estimating fuel usage is task T, and performance measurement P has an impact on how big the inaccuracy is in the prediction

Performance evaluations of regression techniques

Several metrics can be used on the trained model to judge how well an ML regression method captures the underlying relationship. Below is a description of the metrics I plan to employ.

BIAS AND VARIANCE

The bias of a model is the discrepancy between the estimated value and the true value of the parameter being evaluated in statistical ML applications like regression. As a result, bias serves as a gauge of how well the model can estimate things. Underfitting is associated with high bias [12]. Variance relates to how well the model holds up to fresh training instances. It can be defined as the variation in estimates across various model realisations. We have a high variance model, for instance, when training a model on one set yields a considerably different outcome from training the model on another set although both sets of training data describe the same underlying relation [12]. Error = Bias + Variance is a formula that can be used to express a model's overall error. The issue of concurrently reducing these two characteristics to produce a low error is known as the bias-variance tradeoff [12]. A model should ideally be chosen so that it can both properly represent the regularities in its training data and generalise well to new data. These two characteristics can be used to assess the model's generalizability.

MEAN SQUARED ERROR

The average of the squares of the prediction errors is what is known as a model's mean squared error, or MSE. The discrepancy between the estimate and the true value is what is meant by the term "error" in this situation. The MSE is expressed by (3.1) [13] and takes into account both the estimator's bias and variance.

Bi [1]as(Estimate, TrueValue)2 + V arianceEstimate = MSE(V variance).

In light of an estimator's variance and level of bias, the MSE evaluates its quality. The MSE's square root is known as the root mean squared error, or RMSE. The findings of using the RMSE as a metric are identical to those of using the MSE, but the RMSE can be seen as a more accurate depiction of the error.

Linear Regressive

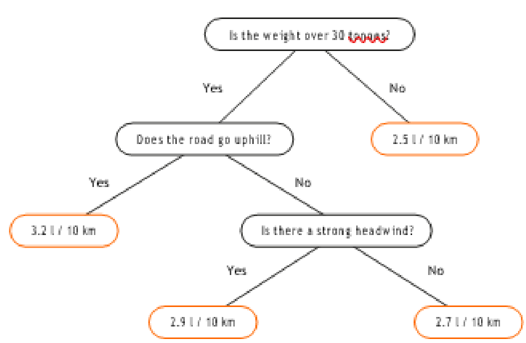
The fitting of a linear function of one or more inputs to an output is known as linear regression. The fitting of a straight line with input x and output y on the form y = w1x + w0, where w0 and w1 are real-valued coefficients to be learned [14], is what is done in the univariate case. The most typical approach for determining the weights in a linear regression issue is to minimise the squared loss function. Finding the weight vector w in accordance with (3.3) is then the issue. This method of selecting the weights ensures that we will discover a singular global minimum [14]. In comparison to the univariate example, the multivariate case of linear regression is not significantly more complex. The goal of multivariate linear regression is to determine the hyperplane that best fits the outputs y according to some loss function, most frequently the squared loss function. Each example xj in this process is an n-element vector. The collection of functions of the kind specified by (3.4) [14] provides the hypothesis space. (3.5) [14] provides the weights vector that minimises the loss function.

Cook's radius

Cook's distance is a popular metric for assessing the impact of a single data point on a linear regression model. When using least squares regression, Cook's distance, or Cook's D, is used to calculate the impact of a particular data point. The definition in mathematics is given by (3.6), where p is the number of fitted parameters in the model, MSE is the mean squared error of the model, and Yj is the prediction for observation j from the full regression model, Yj(i) is the prediction for observation j from a regression model trained on data where observation i has been omitted [15].

REGRESSION TREES

Using a single value, or decision, as the output, decision trees are a straightforward supervised learning technique. The final model accepts a vector of attributes as inputs. In a decision tree, the decisions are represented by the leaf nodes, and the conjunctions of qualities that influence those decisions are represented by the branches [14].



An example of a simple regression tree.

ARTIFICIAL NEURAL NETWORKS

The original idea for ANNs was to connect a large number of basic neurons into a network that could handle complex problems. A node in the neural network is referred to as a neuron in an ANN. It "fires" roughly speaking when a linear combination of its inputs exceeds a certain threshold [14]. Each of the nodes' inputs has a corresponding weight, and there may be one or more of them. In the node, the inputs are multiplied by the respective weights, and the sum is sent to an activation function that provides a binary answer indicating whether or not the sum surpassed the threshold [14]. The construction of an ANN node is shown in Figure 3.2.

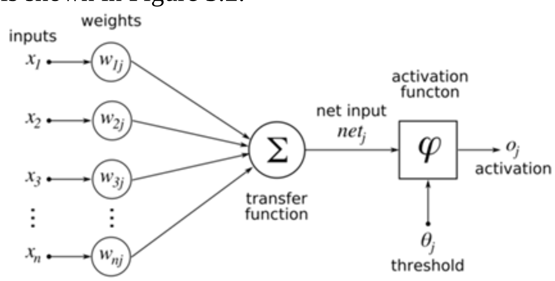
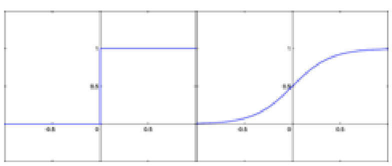
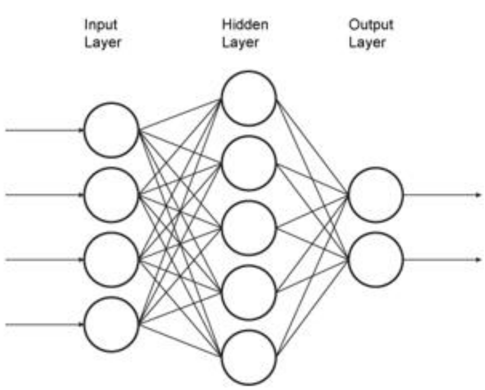


Illustration of an artificial neuron.



Activation functions commonly used in ANN nodes.

To form a network, the nodes of an ANN are arranged in layers and connected by directed links where each link has an associated weight. The layers in between the input and output layers are referred to as hidden layers [14]. Figure 3.4 illustrates how an ANN may be constructed.



An example of an ANN with four inputs and a hidden layer

4. DATA COLLECTION AND PROCESSING

There are four sources of the data that will be used for training. The FM data that Scania collects and stores in their own computer system is the primary and possibly most significant data source. In addition to map data and information on road features from Scania's parent company Volkswagen, which can be accessible through SMHI's web interface [16], vehicle configuration data from an internal Scania system, and weather data are all available to supplement the FM data.

FLEET MANAGEMENT DATA

Variable name	*Description*
Heading	Vehicle heading
Latitude	Latitude of the vehicles position
Longitude	Longitude of the vehicles position
Speed	Vehicle speed
Time position	The time the message was recorded
Vehicle ID	Vehicle identifier
Odometer	Accumulated distance in total
Total fuel	Accumulated fuel in

Table: Variables of interest from the FM data.

The FM database also includes data on driving habits, such as how often brakes are applied over a given amount of time or how much time is spent in ratios that aren't appropriate for the pace. The database only contains data on driver behaviour that has been aggregated and has a temporal resolution of about 1 hour. I calculated averages of the driver behaviour data over time to get around this issue. Table 4.2 provides more information on the relevant factors.

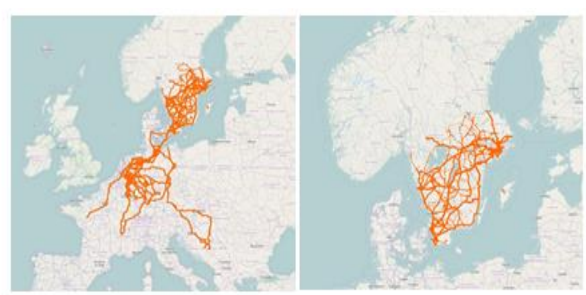
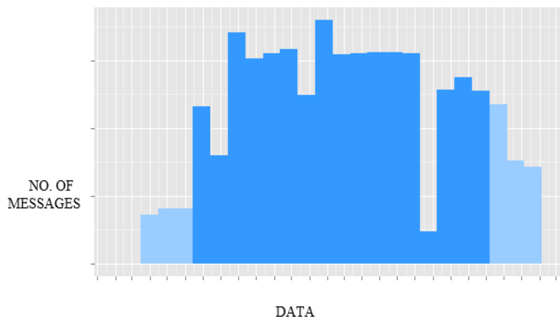


Figure 4.1. Maps overlayed with positions reported from the Transport Laboratory's vehicles

VEHICLE DATA

There is no descriptive data regarding vehicle configuration in the FM database. Instead, the distinct

PIS system must provide this data. A C# wrapper for a web-based SOAP XML API is used to access PIS. Access to vehicle-specific information, such as the type of engine, gearbox, cab, tyres, etc., is made available by the service.



The amount of messages delivered from the Transport Laboratory's fleet of vehicles each month since they began collecting FM data in 2013 is depicted in a histogram in Figure 4.2. The estimated date range used in the end is shown in the highlighted region. The histogram dips that appear over the summer and holiday seasons of 2013 and 2014 can be attributed to the drivers taking vacations. From the Transport Laboratory, about 11 million legitimate messages had been received overall.

ROAD DATA

The system DigitalReality 3.0, which is offered by Volkswagen, the parent company of Sca- nia, is where the road data is obtained. Volkswagen is creating the DigitalReality system as a part of the COMPANION project. The desktop version of the GPS routing system consists of a frontend GUI, often known as a workbench, that enables high level interactions with the underlying map data. The workbench provides access to methods for routing between two or more waypoints as well as the ability to visualise map data and routes. The system also comes with a low level Java API that offers ways for utilising the Java programming language to query the map database.

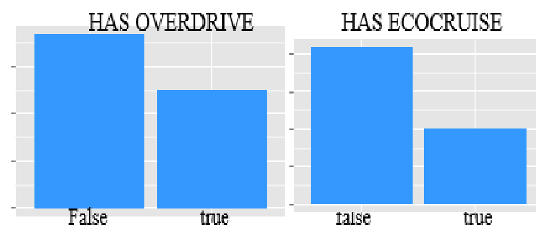
Weather information

The SMHI public web based interface can be used to view the weather data for this study [16]. Since the data made available via the API has better temporal and spatial resolution than the data sets used in Lindberg's study [17], this is an upgrade over that work. According to SMHI [16], historical observations are only valid if they are more than three months old. For observations that are older than three months, SMHI's corrective and quality control process cannot be guaranteed to be complete.

Table 4.3. Variables of interest in the PIS system.

Variable name---- *Description*

Product class-Whether the vehicle is a truck or a bus
Technical total weight-The weight of the vehicle
GTW technical-The maximum allowed gross trailer weightEngine stroke volume-The volume of the engine
Horsepower-The power of the engineRear axle gear ratio-The rear axle gear ratio
Emission level-The emission level, one of 3 classes
Overdrive-Whether the vehicle has overdrive or not
Ecocruise-Whether the vehicle has ecocruise or not



Control system descriptions are shown in Figure 4.5. The left diagram displays how many of the cars have an overdrive transmission, which, if present, might lower fuel consumption. The right image demonstrates how many of the cars had ecocruise systems installed, which could further cut down on fuel use.

DATA CONSOLIDATION:

A schematic overview of the data consolidation process is shown in Figure 4.8. According to the data selection section above, the data is retrieved locally and filtered to only include pertinent data points. Then, using matching criteria, the data points are matched with one another. Pre-processing operations are performed on the consolidated data set that results to compute features and get the data ready for training. The following sections provide a description of this procedure's specifics.

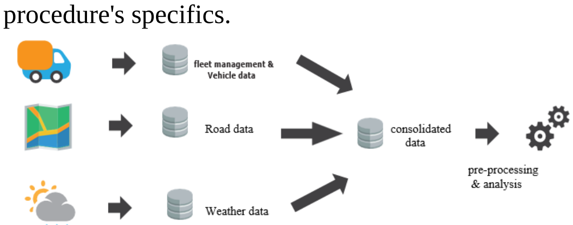
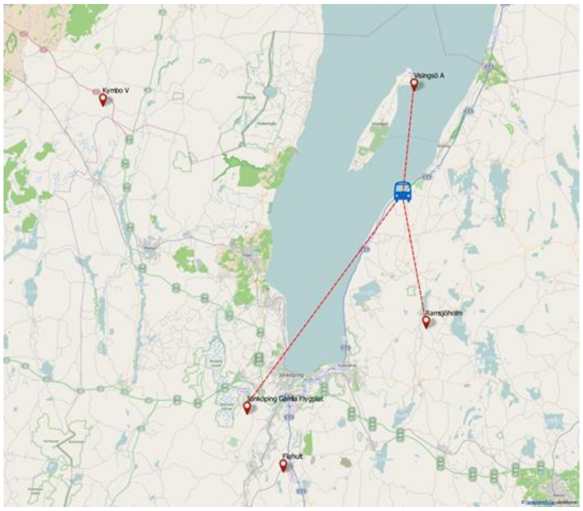


Figure 4.8. Schematic view of the data consolidation process.



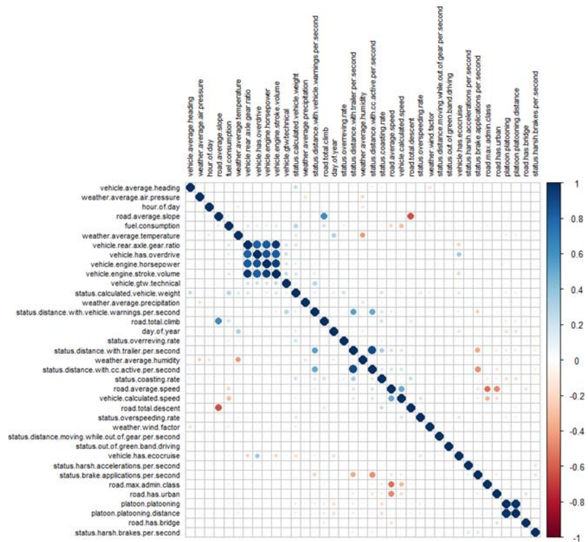
Example of how weather measurements are matched to a position message in Figure 4.9. The closest stations are fetched and examined for the location immediately south of Gränna in order of their proximity to the spot. In this instance, Visingsö is the nearest station and provides data on 5 of the 7 characteristics. The second closest station at Ramsjöholm does not monitor the other 2 characteristics, however the third closest station in Jönköping can provide them. The search is finished when all seven parameters have been located.

6. RESULTS

Data division and normalisation

The data is separated into training, validation, and test data sets before the various models are built. In iterative training, the training set is used to train many models, and the validation test is used to validate and assess the models in order to choose the optimum meta parameters. The final model is evaluated on the test data in order to assess its performance after the meta parameters have been set and a finalised model has been chosen. Data is divided by partitioning it according to date information. The information gathered between 2013-06-01 and 2014-05-31 was used for training. The remaining data that was gathered between 2014-06-01 and 2014-10-31 is the data that was used for testing and validation. In order to prevent the validation and test data sets from including observations from the same dates, the two sets are additionally divided based on date partitioning. The use of date-based partitioning assures that there are no significant dependencies or correlations between the data in the various groups. There is a chance that observations from the same vehicle and the same time period would show in both the training and test sets if a random sample strategy had been used, resulting in a significant correlation between the data sets.

DATA ANALYSIS



The correlation map of the features in the 1-minute data set is shown in Figure 6.1. Strong positive correlation is represented by dark blue circles, whereas strong negative correlation is represented by dark red circles. Depending on the correlation's absolute value, the circles expand. There is no link between the characteristics, as indicated by empty boxes. The diagram demonstrates the significant correlation between the engine parameters as well as the correlation between the three parameters that describe a road's slope profile. There are other correlations between the speed limit and the type of road.

LINEAR REGRESSION

The median value and variance of the two data sets, with sampling rates of 10 and 1 minutes, respectively, were compared. Figure 5.2 presents the findings. Compared to the 10-minute data set, the 1-minute data set has a larger variation and heavier tails.

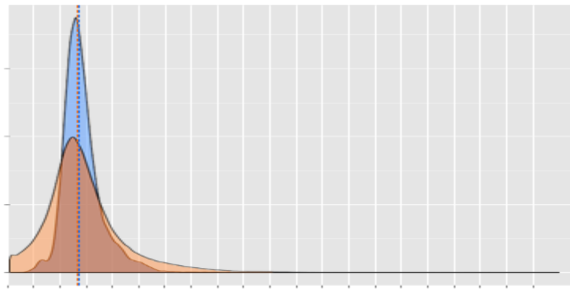
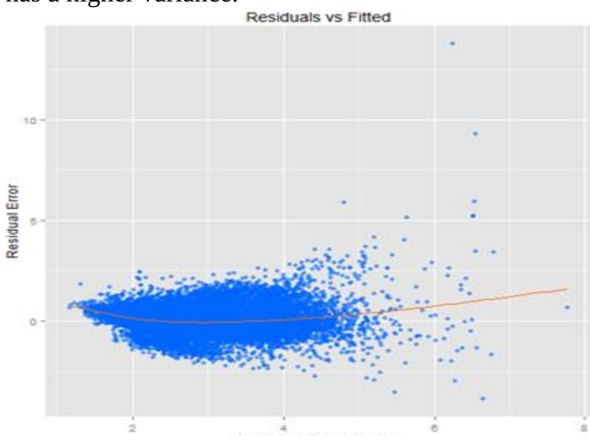
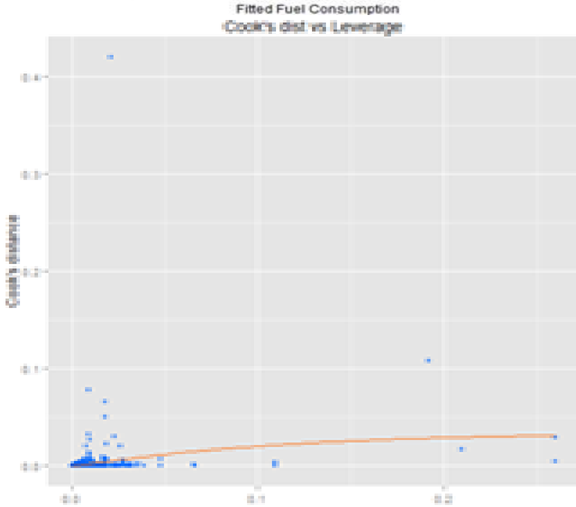
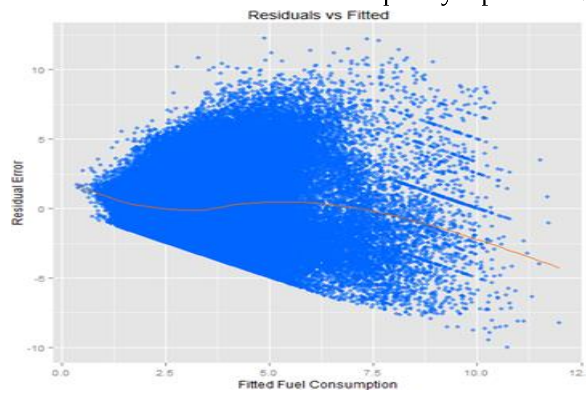


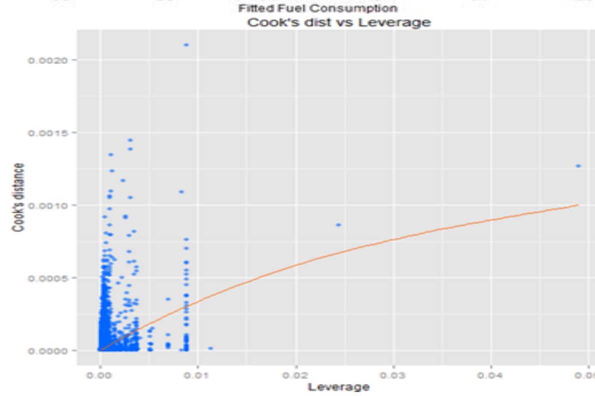
Figure 6.2 shows a Gaussian estimate of the fuel consumption's probability density function (PDF) for several data sets. Plots for the 1-minute and 10-minute data sets are orange and blue, respectively. The median values of the various distributions are shown by the dashed lines. The diagram demonstrates that while the median values for both distributions are quite similar, the data set with the finer sampling rate has a higher variance.





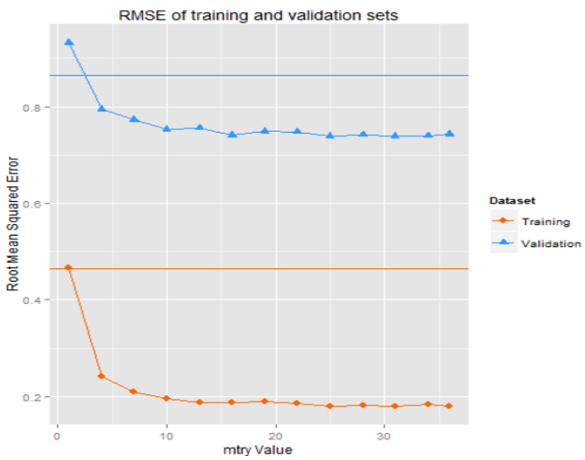
Plots illustrating how well the linear regression model fits the 10-minute data set are shown in Figure 6.3. The smoothed means of the blue dots are shown by the orange lines. The left graphic demonstrates how the residual errors change based on the fuel consumption value that was applied. The average residual is close to zero and the greatest absolute residual is about 2-2.5 when the values are close to 3 l/10 km. In contrast, the variance is higher when the forecast is higher and lower when the prediction is higher. For the extreme values, the mean error also increases. This suggests that the distribution is heteroscedastic and nonlinear, and that a linear model cannot adequately represent it.





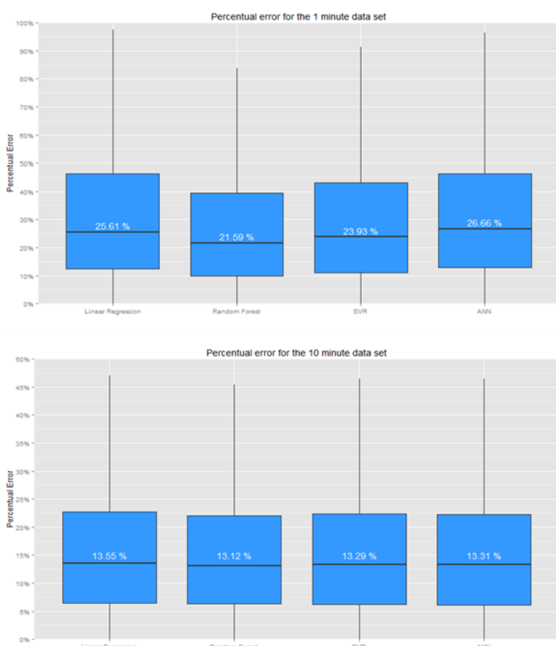
Residuals versus fitted and Cook's distance vs leverage plots for the 1-minute data set following the removal of outliers are shown in Figure 6.4. The smoothed means of the blue dots are shown by the orange lines. The 0 number for fuel usage is indicated by the sharp bottom limit in the left image. Similar characteristics are displayed in both figures as in the diagrams for the 10-minute data set. Due to the constricted y-axis, the right plot seems different yet depicts the same conditions. Compared to the 10-minute data set, the left image's heteroscedasticity is much more obvious. This suggests that employing a finer sampling rate results in a larger range in fuel use. It also shows that a linear model cannot adequately represent either of the data sets.

SUPPORT VECTOR REGRESSION



VARIABLE IMPORTANCE

The random forest fitted to the 10-minute data set was trained using subsets of the features to evaluate the contribution of the various features in the data sets. Driver behaviour, platooning, road, vehicle, and weather features were used to categorise the features. Figure 6.5 shows how removing each of the feature groups from the data set alters the prediction error. It is obvious that deleting the elements that affect driver behaviour minimises the error. This suggests that the driving behaviour features are noisy and of low quality.



Comparison of error distributions between various models that were fitted to the data is shown in Figure 6.6. The 10-minute data set is shown in the bottom image, while the 1-minute data set is shown in the top image. The illustrations display boxplots of the percent-based mistake. The horizontal lines, annotated numbers, and boxplots' four quartiles all represent the distributions' median values. The illustrations demonstrate that in both situations, the random forest fit yields the best test set results.

7. CONCLUSIONS

According to the training data, a sampling rate of 10 minutes is not only enough for prediction but also superior to the lower sample rate of 1 minute. The results show that prediction on the 10-minute data set produces less inaccuracy, despite the 1-minute data set having orders of magnitude more data than the 10-minute data set. This is explained by the fact that the 1-minute data set has a higher variation as well as a higher impact of noise on the data due to its higher sampling rate.

But only when assessing the model's capacity to forecast fuel consumption in litres per kilometre is this accurate. Another issue, which may be more intriguing for real-world use, is how well the models do in foretelling the total fuel used over a longer path. An experiment would need to be conducted to determine which sample rate provides the best outcomes for projecting total fuel use. Such an experiment might be set up in the manner described below:

1. Gather fresh information about specific trucks driving particular routes from the various data sources, such as a route between Helsingborg and Södertälje.
2. Determine the truck's overall fuel consumption for each route.
3. Calculate the fuel consumption of each route using the predicted model.
4. Compare the total consumption that was anticipated with the actual consumption.

Models should be fitted to both the 1- and 10-minute data sets for this experiment, and the results should be compared using an error measurement in litres. In this method, actual fuel consumption would be considered when evaluating the models fitted to the two data sets rather than how accurately they predicted a momentary value of fuel consumption in litres per kilometre. The fact that the models trained on the 1-minute data set only used a subsample of the complete data set whereas the models trained on the 10-minute data set utilised the entire data set is another pertinent issue when comparing the various sampling rates. This sub sampling was carried out to reduce the amount of time and memory needed to train the models, but it might have produced a model that suited the data less well than if the whole data set had been used. Additionally, the parameter search for the SVR model was done in two steps, first looking for the value and then using a grid search to locate the values for C. Due to the dependence between the three variables, it's possible that by segmenting the search in this way, I missed the ideal parameter value. Finding the best parameter values would have been easier with a full grid search in three dimensions. However, it was found that a three-dimensional parameter search would take too much time.

	RMSE of my models [l/10 km]	RMSE of Svärd's models [l/10 km]
Linear regression	1.65	1.43
Random forest	1.29	1.05
SVR	1.41	0.99
ANN	1.55	1.13

RESULTING CONCLUSIONS

The study's findings show that:

• road slope, vehicle speed, and vehicle weight are the most important parameters for predicting fuel consumption;

• the random forest, SVR, and ANN models produce results that are comparable; and

• a sampling rate of 10 minutes is preferable to a sampling rate of 1 minute for predicting fuel consumption measured in litres per distance.

It is obvious that the lower sampling rate of 10 minutes results in a smaller prediction error when comparing the models fitted to the 10-minute data set to the models fitted to the 1-minute data set.

The Friedman test and two-way ANOVA analysis show that the reduction in error is statistically significant at the 0.05 level. This can be explained by the 10-minute data set's reduced variance and the smaller impact of noise on the data. The relationship between road slope, vehicle speed, and vehicle weight is demonstrated to be one of the most significant parameters for prediction. The research of Svärd [7]and Lindberg [17] supports these findings. No statistically significant differences between the random forest, SVR, and ANN models could be identified, so I must draw the conclusion that the models are equivalent when used to forecast fuel consumption for large trucks in litres per km.

REFERENCES

[1] G. D. Schilling, Modeling aircraft fuel consumption with a neural network. Master's thesis, Virginia: Polytechnic Institute and State University , 1997.

[2] Antonio Trani, F. Wing-Ho, Glen Schilling, Hojong Baik, and Anand Seshadri AIAA 4th Aviation Technology, Integration and Operations (ATIO) Forum, 2004.

[3] N. K. T. a. S. Baysec., Prediction of torque and specific fuel consumption of a gasoline engine by using artificial neural networks, Applied energy, 87:349-355, 2010.

[4] L. F. Y. Z. a. H. L. Haikun Wang, "Modelling of the fuel consumption for passenger cars regarding driving characteristics.," *Transportation Research Part D: Transport and Environment,* vol. 13, no. 7, pp. 479-482, 2008.

[5] A. Viswanathan., Data driven analysis of usage and driving parameters that affect fuel consumption of heavy vehicles, Stockholm: Master's thesis, Linköping University, Statistics,The Institute of Technology, 2013.

[6] K. Hansson, Data-driven analysis of the fuel saving potential of road vehicle platooning., Master's thesis, KTH, School of Computer Science and Communi- cations (CSC), 2013.

[7] C. Svärd., Predictive modelling of fuel consumption using machine learning techniques., Technical report, Scania CV AB, 2014.

[8] Companion - cooperative dynamic formation of platoons for safe and energy- optimized goods transportation, 2015. [Online]. Available: URL: http://www.companion-; 2015-02-13. [Accessed 13 02 2015].

[9] A. G. a. K. H. J. Asad Al Alam, "An experimen- tal study on the fuel reduction potential of heavy duty vehicle platooning," *In Intelligent Transportation Systems (ITSC)*, vol. 13th International IEEE Conference on, pp. 306-311, 2010.

[10] A. D. J. D. K. B. a. A. N. Michael P Lammert, Effect of platooning on fuel consumption of class 8 vehicles over a range of speed, following distances, and mass., Technical report,SAE Technical paper , 2014.

[11] T. S. ., Heavy truck modeling for fuel consumption simulations and measurements, Master's thesis, Linköping University, 2001.

[12] P. R. a. H. S. Christopher D. Manning, Intro- duction to Information Retrieval, New York, NY, USA: Cambridge University Press, 2008.

[13] W. M. a. R. L. S. Dennis Wackerly, Mathematical Statistics with Applications, Cengage Learning, Boston, MA, USA, 7th edition, 2007.

[14] S. R. a. P. Norvig., Artificial Intelligence: A Modern Approach., Prentice Hall Press, Upper Saddle River, NJ, USA, 3rd edition, 2009.

[15] W. M. a. T. Sincich., A Second Course in Statistics: Regres- sion Analysis., Prentice Hall, 2003.

[16] F. Rosenblatt., Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms., Washington D.C., USA: Spartan Books, 1962.

[17] J. Lindberg., Fuel consumption prediction for heavy vehicles using machine learning on log data., Master's thesis, KTH, School of Computer Science and Communications (CSC), 2014.