Postprint

Features extracted from APPES to enable the categorization of heavy-duty vehicle drivers

Iulian Carpatorea   Sławomir Nowaczyk   Thorsteinn Rognvaldsson   Johan Lodin

Abstract—**Improving the performance of systems is a goal pursued in all areas and vehicles are no exception. In places like Europe, where the majority of goods are transported over land, it is imperative for fleet operators to have the best efficiency, which results in efforts to improve all aspects of truck operations. We focus on drivers and their performance with respect to fuel consumption. Some of relevant factors are not accounted for in available naturalistic data, since it is not feasible to measure them. An alternative is to set up experiments to investigate driver performance but these are expensive and the results are not always conclusive. For example, drivers are usually aware of the experiment's parameters and adapt their behavior.**

This paper proposes a method that addresses some of the challenges related to categorizing driver performance with respect to fuel consumption in a naturalistic environment. We use expert knowledge to transform the data and explore the resulting structure in a new space. We also show that the regions found in APPES provide useful information related to fuel consumption. The connection between APPES patterns and fuel consumption can be used to, for example, cluster drivers in groups that correspond to high or low performance.

I. INTRODUCTION

Heavy-duty vehicles are a major component of traffic on motorways, especially in Europe where approximately 70% of the goods are transported by land. Traffic is also a major contributor to pollution, with trucks being responsible for around 20% of CO_2. Finally, fuel cost can account for up to 40% of the operating cost of a truck, cf. Barnes and Langworthy [1]. All those are important reasons to increase efficiency of vehicles in general. In this paper we focus on the effects a driver has over fuel consumption and we aim at providing a comparison method for driver behavior.

One of the ways we express performance of vehicles is the amount of fuel used to travel a given distance, typically liters per 100 kilometers [L/ 100 km]. This metric is used for expressing performance of both the engine and the driver, but it is also commonly accepted to be flawed. Alternative metrics include "key performance indicators" (KPI) usually manually selected, that are used to calculate a score or ranking for driver performance. One of the caveats of KPI metrics is that they do not generalize well.

In real driving scenarios, it is hard to compare performance among various vehicles and drivers as the conditions under which they operate are hard to quantify. On the other hand, controlled experiments tend to be expensive and often fail to capture many of the important factors that affect fuel consumption and driver behavior. In this work we aim to, based on the data coming from driving under real conditions, extract features that are able to cope with such lack of complete knowledge. One example is incorporating the effect of unmeasured factors, e.g. air drag. Equally important, for those features, is to allow for the effect of other factors, e.g. vehicle speed or road gradient, to be isolated and quantified separately.

Existing literature on comparing driver performance is based on predicting what the fuel consumption would be, given different actions. Bifulco et al. [2] developed a method for calculating instantaneous fuel consumption, with some measure of success. Constantinescu et al. [3] derived five categories of aggressiveness for drivers based on a number of parameters related to driving, for example positive acceleration. Typically, aggression is associated with performance and lacks in the same manner as when fuel consumption is used as a performance indicator. Nylund [4], states that the differences between a good and bad driver, with respect to fuel consumption, can be up to 30% for heavy-duty vehicles.

Vehicle speed and acceleration are outcomes of driver decisions and depend on the current driving conditions. They are controlled primarily by the driver, through the use of acceleration and braking pedals, but also by the longitudinal control systems, such as cruise control, see Teetor [5]. Monitoring the use of pedals, specifically the acceleration pedal, is a direct way of quantifying driver intention. However, driver intentions are motivated by the current goal, but also by conditions under which the driver operates, e.g. vehicle and environment characteristics.

Engine speed is one indicator of a vehicle's state that also contains information regarding the efficiency of the engine. Typically, each manufacturer has a recommended use of the vehicle for best performance, including which engine speed is the best, from a performance point of view, for different conditions. This makes engine speed an important parameter to monitor when analyzing driver performance.

We propose a space defined by the accelerator pedal position (APP) and the engine speed (ES), namely APPES. We use a histogram mapping how the driver's intentions are affecting the vehicle's operation. We believe that various useful features can be extracted from this mapping, in particular ones that can express the connection between driver behavior and fuel consumption. The distribution of APPES and its relation with other variables, e.g. fuel consumption, have been investigated in [6]. Authors show a high correlation among certain, clearly visible, areas in APPES and variables of interest. For example, one of the regions, namely "neutral", corresponds to the area where APP is not pressed and ES is low, i.e. the vehicle being in neutral gear. We find that trips where a large percentage of time was spent in the "neutral" region exhibit an increase in fuel consumption.

We further extend the APPES concept by adding new features, namely transitions between regions, and common

sequences of such transitions, to be extracted. Transitions represent how the data travels in APPES space, i.e. which symbol comes next after the current one. Patterns are representative sequences of transitions. We also show that the new features are meaningful with respect to fuel consumption, the main indicator we use for categorizing driver performance.

The paper is organized as follows: Section II presents an overview of the work in the field. A summary of the data we use is presented in section II-A. In section III we describe the methodology, followed by section IV, experiments. We finish with section V.

II. BACKGROUND

Some methods prefer the statistical approach when it comes to driver behavior and performance. With respect to fuel consumption we have, for example, Volvo Trucks I-See [7], which is a system that works together with the cruise control and aims at increased performance by making use of prior knowledge about road topography. This lets it achieve a better speed profile. Such a system makes use of expert knowledge in the field on how it is best to drive in hilly terrain. A similar method has been also developed by Hellström et al. [8].

Another study has been performed by Mensing et al. [9] where they perform an analysis of vehicle trajectory in order to reduce fuel consumption while maintaining the same average speed which leads to, in their study, to a reduction of fuel consumption of up to 16%. Achieving this improvement requires a different speed profile than that of a normal driver which can lead to disruptions in traffic.

There are a few key differences between our work and Guo et al. [10]. We have data collected from vehicles with an automatic gearbox while Guo et al. had manual. This is an important distinction as it enforces certain values for engine speed. We also have many vehicles operating under different conditions. They are also operated by different drivers. In contrast, in Guo et al. there was one vehicle and the whole process resembled more of a controlled experiment.

A. Data

We test and develop our methods on naturalistic data recorded in EuroFOT[1] project, see [11], by Volvo Group Trucks Technology (GTT), and Customer Fuel Follow-up[2] (CuFF) project.

In both projects, the data is recorded at 10 Hz from a variety of sensors, such as vehicle speed, axle weight, ambient air temperature, and many more. Furthermore, the data is enriched with information from off-line databases, such as road gradient. The vehicles are operated by many drivers and are in use throughout Europe, enabling analysis under varied conditions.

---

[1]European Field Operational Test, gathered naturalistic driving data for assessing the impact of use of "Intelligent transportation systems" with respect to safety and fuel efficiency.

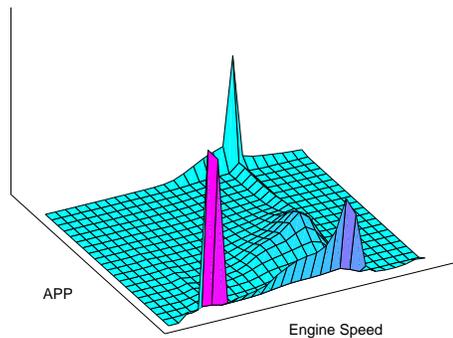[2]GTT project that aims at providing better service to GTT partners.

Fig. 1: APPES

III. METHODOLOGY

APPES is a 2D histogram and it is derived from signals recorded on-board heavy-duty vehicles in normal operation. The two signals, *Accelerator Pedal Position* and *Engine Speed*, are selected based on domain knowledge. Our goal is to have a good representation of time series for qualifying driver actions with respect to fuel consumption.

The first step is to form APPES from the two selected time series. We proceed by symbolizing the data in the new space as described in section III-A. We then further generalize APPES by including additional information not captured by the two primary signals. Finally, we define the concepts of regions and transitions.

A. Data symbolization

An example of APPES can be seen in figure 1. We symbolize the data by assigning different symbols to each of the prominent regions in this space. This assignment can be done in several ways. Some of the options available are manual delimitations for each region, crisp separation such as the one offered by k-means algorithm, or fuzzy regions which can be obtained using, for example, Gaussian Mixture Model (GMM). For robustness we have chosen to use GMM. The mixture model is presented in mathematical form in equation 1.

$$P(\theta) = \sum_{i=1}^{K} \phi_i \mathcal{N}(\mu_i, \Sigma_i) \qquad (1)$$

where ϕ is a vector of weights, μ and Σ are the means and covariance matrices respectively.

We have used an interative approach, we fit models with an increasing number of components to the data from multiple trips, and measure Kullback-Leibler divergence between subsequent GMMs:

$$D_{KL}(P \parallel Q) = \sum_i P(i) log \frac{P(i)}{Q(i)} \qquad (2)$$

where P and Q are probability distributions obtained by using different number of components. The procedure stops when the more complex model is not significantly different from a simpler one. In our case, this procedure leads to selecting six components.

It is worth noting that the time information, i.e. the order in which data points come, is lost when we represent data with APPES, as is typical when transforming time series into histograms.

We assign a symbol to each component in GMM. This is done by computing, for each data point, the probabilities corresponding to each component in GMM and selecting the highest one, for the six symbols associated directly with APPES. In summary, we start with a two time series which we combine into one new time series using APPES. As GMM associate probabilities to each new data point, we use equation 3 to determine the appropriate symbol.

$$symbol(x) = \arg\max_{i=abcdef} GMM(x) \qquad (3)$$

where i is i^{th} component in GMM.

B. Extended APPES

We recognize that the information captured by APPES is incomplete and can be complemented with the use of domain knowledge. For example, a system such as cruise control allows the driver to gain speed without the use of the accelerator pedal. APPES, in the current state, would fail to capture this behavior and could provide misleading estimates of driving performance. However, in our representation, we can easily include the cruise control signal as an additional symbol. Following similar reasoning, we also include the braking pedal as a symbol. We assign the symbols designated for cruise control and braking whenever cruise control is enabled or the brake pedal is pressed, regardless of the values of APP and ES signals.

We have now a symbolic representation of the data, with the symbols "$abcdef$" being associated with regions in APPES, while "gh" represent the cruise control and braking pedal respectively. We can then express each trip as one string consisting of 8 aforementioned symbols.

C. APPES Transitions and Patterns

Transitions represent the information about the order in which symbols occur in the data. The order, in which symbols, are appearing is useful as we can define a sequence of transitions as a pattern, e.g. cruise control followed by braking is different from braking followed by cruise control. Patterns are intuitive, can easily be understood and can be used to define driver behavior and performance. The number of possible patterns grows exponentially, 8^n, where n is the pattern length.

We define interestingness as the frequency of each pattern. This enables the study of a smaller subset of patterns, while still maintaining that the results have a higher applicability since these patterns occur often.
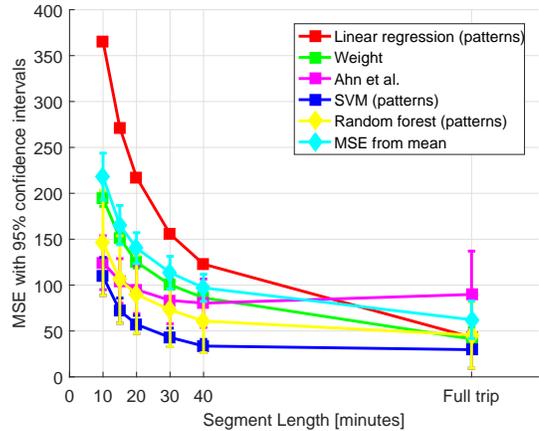

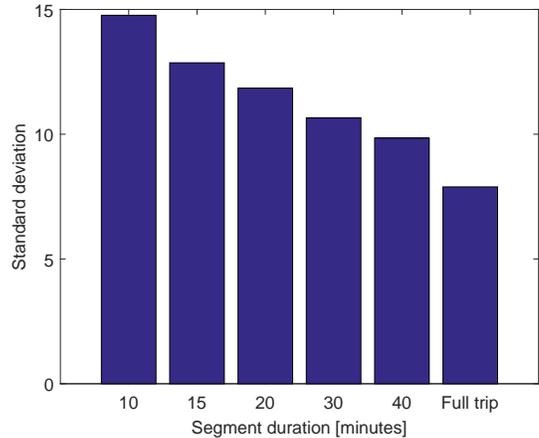Fig. 2: Prediction error for each model depending on segment length.

Fig. 3: Standard deviation of fuel consumption for each segment duration.

IV. EXPERIMENTS

We set up our experiments in order to understand the usefulness for patterns. We lack ground truth regarding driver performance, therefore we focus on relations with fuel consumption. For evaluation we use 800 trips, where each one is at least 30 minutes long and has less than 5% missing data. We select segments of different lengths from these trips, aiming to investigate the short and long term effect of patterns. The shortest selected segment has a duration of 10 minutes while the longest corresponds to a full trip duration, which is on average 3 hours.

We select representative patterns, which means patterns that occur often, in this case at least 500 occurrences, and are diverse, i.e. they are not a subset of longer patterns, within the selected ones. This gives us 81 patterns for the selected trips.

The predictive power of this data, with respect to fuel consumption, is indicative of the information contained within. We argue that APPES patterns capture useful information based on how we construct them, however this assumption needs to be validated. For each segment, we create a binary
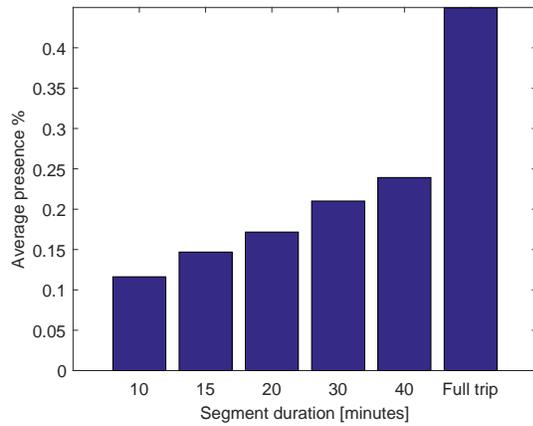
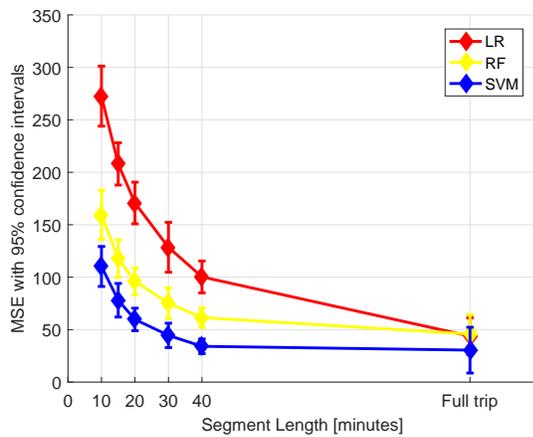Fig. 4: Average number of non-zero elements in the binary vector.

Fig. 6: MSE for data with no outliers.

Fig. 5: MSE for models using segments with at least 1 pattern present.

Fig. 7: Prediction error for SVM and RF for segments with no outliers and at least 1 pattern present

vector for APPES patterns. Each bit in the vector represents the presence or absence of the corresponding pattern. Subsequent instances of same pattern in a segment are ignored.

Fuel consumption is the response variable. We select three regression models: linear (LR), support vector machines (SVM) and random forests (RF). We anticipate that the relation between patterns and fuel consumption is non-linear but still use a linear model for comparison purposes. We have used the implementations provided by Matlab 2016a.

We also choose other methods for comparison. As an initial baseline, we select a model where each prediction is the average of the data, as well as linear regression where the total vehicle weight is the only input. We also use a model based on acceleration and speed inspired by Ahn et al. [12]. Lastly, a model based on FPC, described in [13], which combines an alternative response variable and a linear model with inputs from acceleration and vehicle speed.

The results are presented in figure 2, which shows the relation between MSE for each model for different segment length. For each model we use 10 fold cross validation. The most visible result is a decrease in error as segments become longer. This is an interesting observation and we decided it
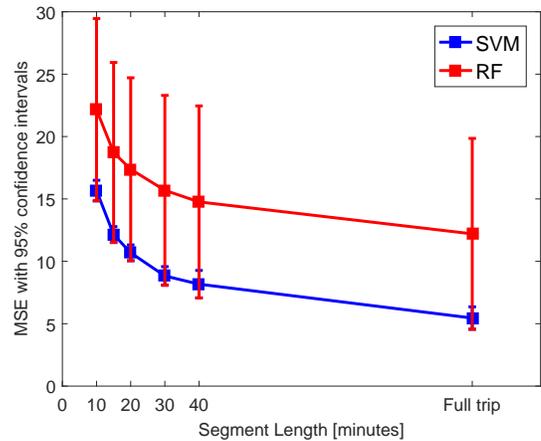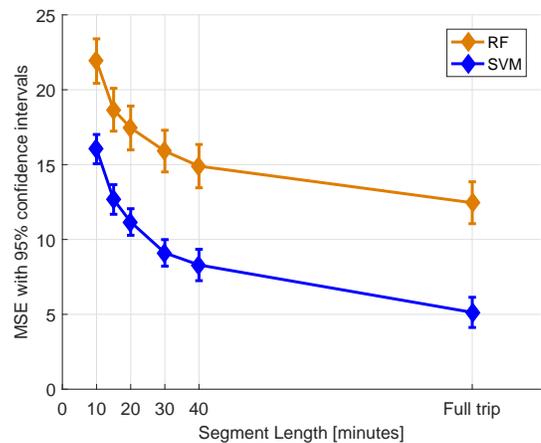
warrants further analysis. We have identified several possible reasons for this.

First, it can partly be explained by the variance in the data. By this we mean that as duration of segments increases, fuel consumption variance decreases. For most models, this leads to a natural decrease in the error. The standard deviation corresponding to each segment duration can be seen in figure 3.

Another reason for this trend is that some segments have the same input feature vector but not the same output. This is true mostly when they have a null binary vector, i.e. none of the representative patterns are present in the given segment, and occurs more often in shorter segments. This is tied to the variance of fuel consumption as each model will try to predict the average over data with the same inputs. Figure 4 shows how many patterns are present, on average, in a segment of a given length.

We want to analyze how segments that have only 0's in the binary vector affect the MSE. We select segments that have at least one pattern present and exclude all other segments from both training and testing. We present the results in figure 5. When comparing the performance of models with and without
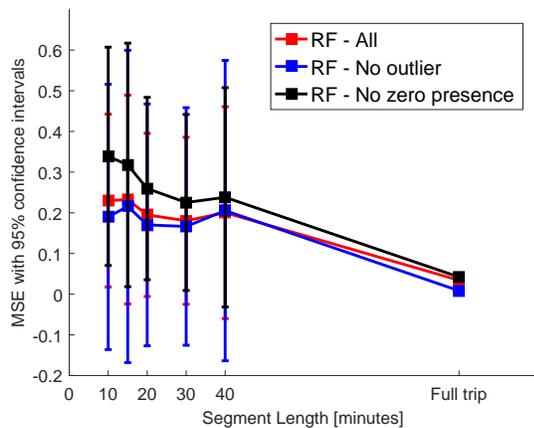
Fig. 8: Prediction error for RF model using response variable modified by FPC.
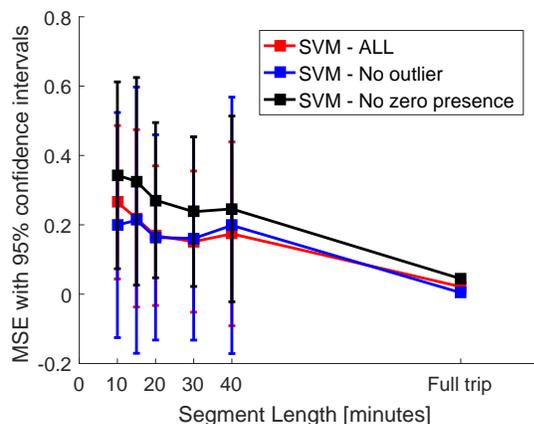

Fig. 9: Prediction error for SVM model using response variable modified by FPC.

segments with 0 presence we expected an improvement for both SVM and RF. However, no significant change is present which would mean that both methods can handle the 0s vectors. LR shows a slight improvement, which is to be expected.

Also we want to see how important are outliers when computing MSE. For a given segment length, we select all segments where the response variable is falling within the 75% percentile. Results are illustrated in figure 6. Both SVM and RF show a significant improvement from which we conclude that outliers contribute the most to the error.

Following the same logic, we use the best subset of our data to do prediction, i.e. no outliers and no segments that do not contain any pattern. The results are presented in figure 7. The major difference observed is the reduction in error uncertainty

For a better understanding of the results, we perform one more analysis. We state that unmeasured factors give a biased result when assessing driver performance and in our other work we argue that we can take into account the effect of said factors. We then want to see how well we can predict a modified response variable based on FPC, see [14].We want to analyze how much of the unmeasured factors we are able
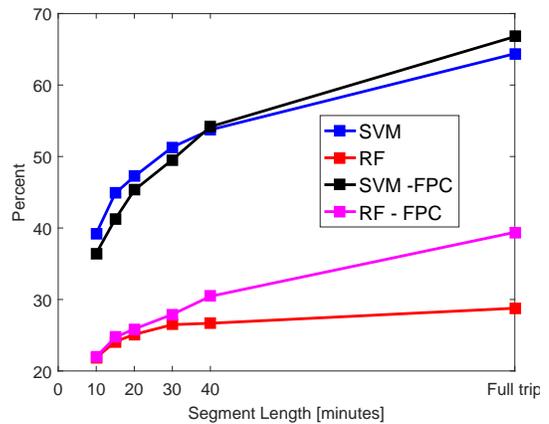

Fig. 10: Percent of predictions with less than 5% error corresponding to each segment length. We compare measured fuel consumption as response variable and modified by FPC response variable.

to capture using the APPES patterns. FPC contains the effect of other factors, such as air drag, facilitating comparison of desired variables, in our case the driver, under similar conditions.

The new output is given by equation 4, where P is the new output, FC is fuel consumption for the segment in question, RGM is fuel used associated with road gradient while FPC is fuel for the trip which segment s_o belongs to.

$$P(s_o) = \frac{FC(s_o) - RGM(s_o)}{FPC} \qquad (4)$$

Figures 8 and 9 show the results. Both, SVM and RF, perform comparably and do not exhibit the same trend as in figure 2. In this case, we do not see the clear improvement correlated with segment length, as we have observed in the other cases.

We end with computing the percent of segments, for each chosen length, that are predicted correctly, i.e. within 5% error. The results are in figure 10. For both response variables, the results are similar, with over 60% of segments being correctly predicted for the full trip when using SVM.

V. CONCLUSIONS AND FUTURE WORK

We have proposed a new space, APPES, and a method on how to extract features from this space. We show that the new features contain relevant information to the task of classifying driver performance. We also conclude that SVM is a good candidate for future predictions using patterns. There is also evidence that the patterns are robust and, together with a powerful classifier like SVM, can handle the incompleteness of data and still perform adequately, while a weak classifier, like LR, will perform poorly. The biggest challenge identified in this work comes in the form of outliers. The selected patterns are unable to capture the high variance in the data, i.e. neither of our patterns can capture spikes in response variable.

Future work includes analysing how well patterns perform when symbols are fuzzy, i.e. instead of a single symbol we have a combination of symbols with associated intensity

for each data point. We also want to analyze what is the contribution of each pattern to driver performance, how it relates to environmental conditions and vehicle characteristics. Understanding how to select the relevant signals depending on the domain, will enable generalization of the concept.

We also want to investigate how we can quantify the importance of each pattern. This is particularly useful for defining interesting patterns, not just based on how often they occur but also based on how important they are.

REFERENCES

[1] G. Barnes and P. Langworthy, "The per-mile costs of operating auto-mobiles and trucks," 2003.

[2] 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.

[3] Z. Constantinescu, C. Marinoiu, and M. Vladoiu, "Driving style analysis using data mining techniques," *International Journal of Computers Communications & Control*, vol. 5, no. 5, pp. 654–663, 2010.

[4] N. Nylund, "Fuel savings for heavy-duty vehicles hdenergy. summary report 2003-2005," Project Report VTT, Tech. Rep., 2006.

[5] R. R. Teetor, "Speed control device for resisting operation of the accelerator," Aug. 22 1950, uS Patent 2,519,859.

[6] I. Carpatorea, S. Nowaczyk, T. Rognvaldsson, and M. Elmer, "Appes maps as tools for quantifying performance of truck drivers," in *Proceedings of the International Conference on Data Mining (DMIN)*. The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), 2014, p. 1.

[7] "I-see," http://www.volvotrucks.com/trucks/uk-market/en-gb/trucks/volvo-fh-series/key-features/Pages/i-see.aspx, accessed: 2016-05-08.

[8] E. Hellström, M. Ivarsson, J. Åslund, and L. Nielsen, "Look-ahead control for heavy trucks to minimize trip time and fuel consumption," *Control Engineering Practice*, vol. 17, no. 2, pp. 245–254, 2009.

[9] F. Mensing, R. Trigui, and E. Bideaux, "Vehicle trajectory optimization for application in eco-driving," in *2011 IEEE Vehicle Power and Propulsion Conference*, Sept 2011, pp. 1–6.

[10] P. Guo, Z. Li, Z. Zhang, J. Chi, S. Lu, Y. Lin, Z. Shi, and J. Shi, "Improve fuel economy of commercial vehicles through the correct driving," in *Fisita 2012 World Automotive Congress*. Springer-Verlag Berlin Heidelberg, 2013.

[11] WWW, "http://www.eurofot-ip.eu/."

[12] K. Ahn, H. Rakha, A. Trani, and M. Van Aerde, "Estimating vehicle fuel consumption and emissions based on instantaneous speed and acceleration levels," *Journal of transportation engineering*, vol. 128, no. 2, pp. 182–190, 2002.

[13] "Ommitted due to double blind review. available upon request."

[14] I. Carpatorea, S. Nowaczyk, T. Rognvaldsson, J. Lodin, and M. Elmer, "Learning of aggregate features for comparing drivers based on naturalistic data," in *IEEE International Conference on Machine Learning and Applications (ICMLA)*, December 2016.