COMPUTER-AIDED CIVIL AND INFRASTRUCTURE ENGINEERING

Computer-Aided Civil and Infrastructure Engineering **33** (2018) 209–219

Modeling Relationship between Truck Fuel Consumption and Driving Behavior Using Data from Internet of Vehicles

Zhigang Xu & Tao Wei

School of Information Engineering, Chang'an University, Xi'an, Shaanxi, P.R. China

Said Easa

Department of Civil Engineering, Ryerson University, Toronto, ON, Canada

Xiangmo Zhao

School of Information Engineering, Chang'an University, Xi'an, Shaanxi, P.R. China

&

Xiaobo Qu*

Department of Architecture and Civil Engineering, Chalmers University of Technology, Gothenburg, Sweden

Abstract: *In this research, by taking advantage of dynamic fuel consumption–speed data from Internet of Vehicles, we develop two novel computational approaches to more accurately estimate truck fuel consumption. The first approach is on the basis of a novel index, named energy consumption index, which is to explicitly reflect the dynamic relationship between truck fuel consumption and truck drivers' driving behaviors obtained from Internet of Vehicles. The second approach is based on a Generalized Regression Neural Network model to implicitly establish the same relationship. We further compare the two proposed models with three well-recognized existing models: vehicle specific power (VSP) model, Virginia Tech microscopic (VT-Micro) model, and Comprehensive Modal Emission Model (CMEM). According to our validations at both microscopic and macroscopic levels, the two proposed models have stronger performed in predicting fuel consumption in new routes. The models can be used to design more energy-efficient driving behaviors in the soon-to-come era of connected and automated vehicles.*

*To whom correspondence should be addressed. Email: *drxiaoboqu@gmail.com*.

DOI: 10.1111/mice.12344

1 INTRODUCTION

The air pollution has been becoming more and more frequent and severe in China, and is increasingly considered as a major factor to affect a city's livability and residents' well-being. Indeed, the greenhouse gas (GHG) emissions have attracted much attention by government agencies, academia, as well as the general public (Kolmanovsky et al., 2011; Liao et al., 2012; Slavin et al., 2013; Yin et al., 2015; Martani et al., 2016; Jung et al., 2016; Levin et al., 2016; Zockaie et al., 2016; Taheri et al., 2016; Chen and Guan, 2017). Almost 30% of GHG emissions are determined by the amount of fuels consumed by the transportation sector, which has been a major contributor to energy consumption and

GHG emissions worldwide (U.S. Environmental Protection Agency, 2015). As a fast-growing developing country, China faces many severe environmental challenges (MEPPRC, 2013). It should be pointed out that fuel consumption and GHG emissions generated by trucks are highly over-represented: less than 5% of trucks contribute to over 30% of fuel consumptions and GHG emissions (MEPPRC, 2013). This, in combination with the projected shortage of crude oil, imposes great degree of urgency to develop a reliable approach that is able to accurately predict truck fuel consumption.

Numerous researchers have contributed their efforts in estimating and predicting fuel consumption of trucks in an attempt to identify the best driving behaviors that are able to not only reduce fuel consumption but also decrease GHG emissions. There are basically two types of models. The first type of models is based on physical principle. The representative models for this type are vehicle specific power (VSP) model, Comprehensive Modal Emission Model (CMEM), and EMIssions from Traffic (EMIT) model. Jiménez (1998) proposed the concept of VSP to estimate fuel consumption as well as vehicle emissions. The model essentially summarizes the loads resulting from aerodynamic drag, acceleration, rolling resistance, and hill climbing. This model has been further applied and validated by Jiménez et al. (1999), Zhai et al. (2008), and Duarte et al. (2015). CMEM was initially developed in the late 1990s with sponsorship from the National Cooperative Highway Research Program (NCHRP) and the U.S. Environmental Protection Agency (EPA) to fulfill the need for microscopic emissions modeling (An et al., 1997; Barth et al., 1996, 1997, 1999, 2000; Boriboonsomsin et al., 2012). The model has been applied by numerous researchers and practitioners (e.g., Nesamani et al., 2017; Qi et al., 2004). Cappiello et al. (2002) proposed an EMIT model to estimate the instantaneous emissions and fuel consumption for light-duty vehicles based on instantaneous speed, acceleration and vehicle specific parameters. The performance of this model is comparable to CMEM. Another category of models refers to data driven models. Rakha et al. (2004) developed the Virginia Tech microscopic (VT-Micro) model for estimating hot stabilized light duty vehicles and truck fuel consumption and emissions. The model was developed based on experimentation with numerous polynomial combinations of speed and acceleration levels. In particular, linear, quadratic, cubic, and quartic terms of both speed and acceleration were tested in a trial and error manner against experimental data. Chen et al. (2015) used Next Generation SIMulation data (NGSIM) and CMEM to analyze fuel consumption/emissions aggregated at both lane and link levels. Some researchers also carried out research with respect to the impact of driving behavior on fuel consumption and emissions (Ericsson, 2001; Giannelli et al., 2005; Wu et al., 2015; Walnum and Simonsen, 2015).

The above models primarily focus on driving behaviors (i.e., speed and acceleration). It is known that truck fuel consumption is determined by not only driving behaviors, but also road geometry, environment, and vehicle conditions. Even for two vehicle journeys with exactly the same trajectories, fuel consumption might be significantly different. The reason that these factors are not incorporated in this family of fuel consumption models is that they are not intended to be used by vehicle manufacturers. In contrast, they are used to regulate or optimize the driving behaviors as guidance for eco-driving practices by transport authorities and individual drivers. If a truck driver is required to transport goods from one origin to a destination along with a predetermined route, vehicle conditions (i.e., truck condition), environment (i.e., weather), and road geometry are all fixed. The only thing the driver can adjust is real-time speed.

However, two issues limit the applicability of the above-mentioned models in the practice of eco-driving. First, second-by-second speed-fuel data are needed for CMEM, VT-Micro, and VSP models, which are very difficult to obtain in the current techniques of Internet of Vehicles. Second, the models require too many coefficients that need to be calibrated. For example, there are 32 calibration parameters in the VT-Micro model. In order to address these limitations, we develop two models to predict truck fuel consumption by taking advantage of the dynamic fuel consumption–speed data obtained from the preinstalled device of Internet of Vehicles by Shaanxi Automotive Group. The first model is a semiphysical model based on a new concept of energy consumption index (ECI), which is to quantify the impact of truck drivers' driving behavior on fuel consumption. The second model is a data-driven model based on a Generalized Regression Neural Network (GRNN) to explore the same relationship. The main objective of this research is to establish an implicit and an explicit relationship between fuel consumption and truck drivers' driving behaviors, which can be used as a guide for eco-driving practices.

The two proposed models are compared with three well-recognized existing fuel consumption estimation models: VSP, VT-Micro, and CMEM. We further validate our models at both microscopic level by using the data from vehicle experiments at test beds and macroscopic (aggregated) level by using the data from Internet of Vehicles. The results show that our models have stronger applicability in predicting fuel consumption for new routes. It should be highlighted that the proposed energy consumption index, which has a very simple closed form representation, is naturally suitable for examining the driving behaviors that are more energy efficient and environment friendly. As a result, it can be applied by vehicle manufacturers to design the energy efficient vehicle motion controller or trajectory planners in the soon-to-come era of connected and automated vehicles.

The main contribution of this research is threefold. First, based on the law of energy conservation, we develop an explicit index that is able to reasonably well estimate fuel consumption based on speed data collected from Internet of Vehicles. According to the data from vehicle experiments, this model is comparable or superior to VSP, VT-Micro, and CMEM. According to the data from Internet of Vehicles, this model is superior to VSP and VT-Micro. Second, a GRNN-based model is developed to predict fuel consumption. In this model, we use the past trajectory–fuel relationship to train our model and establish an implicit relationship (without a closed form) between dynamic speeds and fuel consumption for new routes. Third, our model has a comparable performance under high-speed conditions, and a much better performance under low-speed conditions.

The remainder of this article is organized as follows. Section 2 describes the operational data from Internet of Vehicles and presents the data preprocessing procedure. Section 3 introduces three well-recognized models for fuel consumption estimation. Section 4 presents the first model based on physical principle. The second data-driven model based on GRNN is presented in Section 5. Their performances are compared with the two existing models in Section 6, and it is followed by the conclusions in Section 7.

2 DATA DESCRIPTION AND PREPROCESSING

2.1 Data description

Truck operational data are obtained from the system of Internet of Vehicles by Shaanxi Automotive group. All trucks manufactured by the group have a preinstalled device to collect 18 parameters including time, latitude, longitude, torque, speed, engine revolution speed, accumulative fuel consumption, instantaneous fuel consumption, mileage, braking state, gear, air conditioner state, pedal state, temperature of coolant, air pressure, oil pressure, voltage, and temperature of engine oil. When the engine is in operation, the real-time operational data are sent to the vehicle monitoring center through 3G or 4G networks. To the best of our knowledge, the data set is the first of this kind that is obtained from the preinstalled Internet of Vehicles devices. The time resolution of our data is 2–5 seconds. This data set makes possible to examine the relationship between truck fuel consumption with the microscopic driving behaviors as large-scale vehicle trajectories are obtainable.

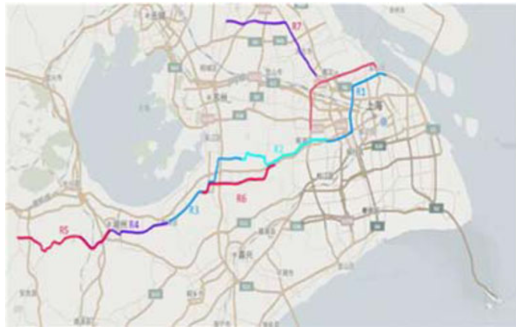


Fig. 1. Geographical layout of the seven routes.

In this research, we use five routes with respect to three trucks to calibrate our models, and two routes with respect to three trucks for validations. The details of these routes are presented in Figure 1 and Table 1. Each route is 50 km long. We aggregate the data into 2 km intervals and there are 25 segments for each route. The total sample size for training/calibration is thus 125 (Routes 1–5) and that for testing/validation is 50 (Routes 6–7). The selected routes are spatially (intercity and intra-city) and hierarchically (expressway, arterial, and freeway) representative.

Table 1
Coordinates of the seven routes

#	*Origin (longitude, latitude)*	*Destination (longitude, latitude)*	*Road type*
R1	121.537857E, 31.379787N	121.23089E, 31.158955N	Expressway
R2	121.22187E, 31.158112N	120.83069E, 31.081255N	Freeway
R3	120.788925E, 31.08342	120.415237E, 30.881762N	Freeway
R4	120.415237E, 30.881762N	119.955177E, 30.828056N	Expressway
R5	119.645241E, 30.82712N	120.092384E, 30.837137N	Arterial
R6	120.553825E, 30.962564N	120.955605E, 31.080454N	Freeway
R7	121.180038E, 31.373402N	120.685844E, 31.571182N	Arterial

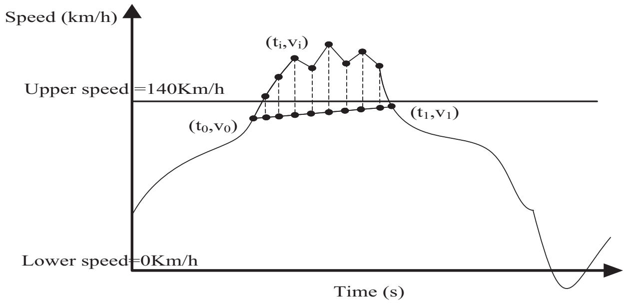


Fig. 2. Data cleansing Rule 1.

Data collection and aggregation play essential roles in the credibility of traffic model development (Qu et al., 2017). Although the data collection system is very advanced, some noises still exist due to a variety of reasons such as device temporary malfunction and/or earth magnetic field jamming. In this regard, a data preprocessing procedure is proposed next.

2.2 Data cleansing

The basic principle of the data cleansing is very straightforward. First, both the speed of sample i (v_i) and instantaneous fuel consumption of sample i (ε_i) have a reasonable range, and the transition between adjacent data should be relatively smooth. As such, we can define our data cleansing Rules 1 and 2 based on simple smoothing principles as described below.

Data Cleansing Rule 1

As shown in Figure 2, we use interpolation to clear both single and continuous polluted Speed data as below.

$$\text{If } v_i < v^{lower}, \text{ then } v_i{}' = v^{lower}$$

$$\text{Else if } v_i > v^{upper}, \text{ then}$$

$$v_i{}' = v_0 + \frac{v_1 - v_0}{t_1 - t_0} \cdot (t_i - t_0)$$

where $v_i{}'$ is the updated speed of samples i, and v^{lower} and v^{upper} are the lower and upper bounds of the speed, respectively; otherwise, $v_i{}' = v_i$.

In this research, the lower and upper bounds for speeds are set as 0 and 140 km/hour. The upper bound is determined by the design manual of trucks.

Data Cleansing Rule 2

$$\text{If } \varepsilon_i < \varepsilon^{lower} \text{ or } \varepsilon_i > \varepsilon^{upper}, \text{ then } \varepsilon_i{}' = (\varepsilon_{i-1}{}' + \varepsilon_{i+1}{}')/2$$

where $\varepsilon_{i-1}{}'$, $\varepsilon_i{}'$, and $\varepsilon_{i+1}{}'$ are updated instantaneous fuel consumption of samples i–1, i, and i+1, respectively, and ε^{lower} and ε^{upper}, are the lower and upper bounds of the instantaneous fuel consumption of samples i–1, i, and i+1, respectively; otherwise, $\varepsilon_i{}' = \varepsilon_i$.

The lower and upper bounds of fuel consumption rates are calibrated based on 1-year operational data. The 5th and 95th percentile-based values are used as the lower and upper bounds, which are 0.173 L/2 km and 0.936 L/2 km, respectively. Although the excluded data might be factual, they are associated with extreme conditions (e.g., bad weather, traffic congestions caused by traffic crashes) that are not suited to be used in eco-driving practice.

Further, the speed and instantaneous fuel consumption are the first order derivatives of the mileage (m_i) and the cumulative fuel consumption $\widehat{\varepsilon}_i$, respectively. Based on this principle, we can define the data cleansing Rules 3 and 4 as below.

Data Cleansing Rule 3

$$\text{If } \left|1 - \frac{m_{i-1}{}' + v_i{}'t}{m_i}\right| > \xi, \text{ then } m_i{}' = m_{i-1}{}' + v_i{}'t$$

where ξ is the tolerance level, t is the interval between two readings, $m_{i-1}{}'$ and $m_i{}'$ are the updated mileages of samples i–1 and i, respectively.

Data Cleansing Rule 4

$$\text{If } \left|1 - \frac{\widehat{\varepsilon}_{i-1}{}' + \varepsilon_i{}'t}{\widehat{\varepsilon}_i}\right| > \xi, \text{then } \widehat{\varepsilon}_i{}' = \widehat{\varepsilon}_{i-1}' + \varepsilon_i{}'t$$

where t is the interval between two readings, $\widehat{\varepsilon}_{i-1}{}'$ and $\widehat{\varepsilon}_i{}'$ are the accumulative fuel consumption as of readings i–1 and i, respectively.

In this research, we use 5% as our tolerance levels. We further check the quality of our cleansed data by comparing the trajectory generated from updated speeds and the recorded trajectory. Figure 3 shows a representative sample with one unrealistic speed value. Red crossed curve and blue dashed curve refer to the trajectories generated from updates speeds and original speeds, respectively. The solid curve is the actual recorded trajectory. As can be seen in Figure 3, the red one is more consistent with the actual one, which minimizes the impact of the unrealistic speed data.

Based on the preceding four data cleansing rules, our data have been cleaned and the updated data are now ready for further analysis.

3 THREE PROMINENT EXISTING MODELS

3.1 VSP model

The VSP model was first developed by Jiménez (1998) in his PhD thesis. It represents the engine load against

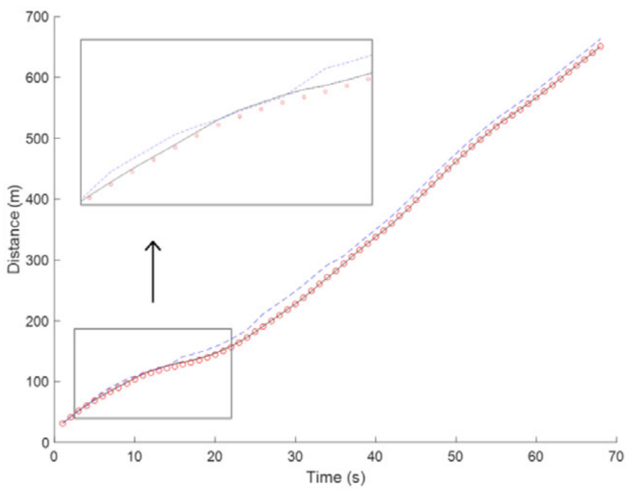


Fig. 3. Actual trajectory versus predicted trajectories.

aerodynamic drag, acceleration, rolling resistance, and the kinetic and potential energy of the vehicle, mathematically,

$$VSP = \frac{power}{mass} = \frac{\frac{d}{dt}\left(E_{kinetic} + E_{potential}\right) + F_{rolling} \cdot v + F_{aerodynamic} \cdot v}{m} \tag{1}$$

where $E_{kinetic}$, $E_{potential}$, v, $F_{rolling}$, and $F_{aerodynamic}$ are kinetic energy, potential energy of the vehicle, speed of the vehicle, rolling resistance, and aerodynamic drag, respectively, and m is vehicle mass. The equation can be further transformed as

$$VSP = v \cdot (a \cdot (1 + \varepsilon_i) + g \cdot grade + g \cdot C_R) + \frac{1}{2}\rho_a \frac{C_D \cdot A}{m}(v + v_w)^2 \cdot v \tag{2}$$

where v_w is headwind into the vehicle (m/s), a is instantaneous vehicle acceleration, ε_i is "mass factor" which is the equivalent translation mass of the rotating components (wheels, gears, shafts, etc.) of the powertrain (which is assumed to be 0.1), the suffix i indicates that ε_i is gear-dependent, *grade* is vertical rise/slope length, g is acceleration of gravity (9.8 m/s^2), C_R is coefficient of rolling resistance (dimensionless), C_D is drag coefficient (dimensionless), A is frontal area of the vehicle, and ρ_a is ambient air density (1.207 kg/m^3 at 20 °C = 68 °F).

In this research, we use our fuel consumption–speed data from Internet of Vehicles to recalibrate the VSP model and its calibration results are presented as

$$VSP = v[a + 9.807grade(\%) + 0.244543)] + 0.000196171(v + v_w)^2 \cdot v \tag{3}$$

This calibrated model is used in our performance analyses in Section 6.

3.2 VT-Micro model

Rakha et al. (2004) established the fuel consumption based on speed and acceleration under thermal steady-state microscopic statistical model of emission. In this model, the vehicle is divided into several categories. According to the test data of the laboratory, the best fitting for each kind of vehicle is determined by the product combination of different acceleration and speed, mathematically,

$$\ln(MOE_e) = \begin{cases} \sum_{i=0}^{3}\sum_{j=0}^{3}\left(L_{i,j}^{e} \cdot s^i \cdot a^j\right), a \geq 0 \\ \sum_{i=0}^{3}\sum_{j=0}^{3}\left(M_{i,j}^{e} \cdot s^i \cdot a^j\right), a < 0 \end{cases} \tag{4}$$

where MOE_e is instantaneous fuel consumption or emission rate (*l/s* in the case of fuel consumption or *mg/s* in the case of vehicle emissions), $L_{i,j}^e$ is model regression coefficient for MOE "*e*" at speed power "*i*" and acceleration power "*j*" for positive accelerations, $M_{i,j}^e$ is model regression coefficient for *MOE* "*e*" at speed power "*i*" and acceleration power "*j*" for negative accelerations, *s* is instantaneous speed, and *a* is instantaneous acceleration.

3.3 CMEM model

CMEM is microscopic in the sense that it predicts second-by-second fuel consumption based on different modal operations from vehicles. In the model, the fuel consumption is broken down into components that are associated with physical phenomena with respect to different vehicle operational status (Barth et al., 2000). The required input for CMEM includes second-by-second speed data. However, as mentioned in Section 3, the time resolution of our data from Internet of Vehicles is 2–5 seconds. In this regard, we cannot use our data to compare the performance of CMEM and the proposed models. In order to compare our proposed models with CMEM, we use the data from two experiments of vehicles at test beds by Cappiello et al. (2002).

4 MODEL I—AN ENERGY CONSUMPTION INDEX-BASED MODEL

4.1 Model development

To estimate fuel consumption, we first categorize the journey into three states: cruise, acceleration, and deceleration based on a sample aggregation principle: if the speed difference is less than 5 km/hour for two consecutive records, we consider this interval as a cruise journey; otherwise, the states are considered as deceleration or acceleration. Figure 4 shows an example of state categorization.

As fuel consumption is used to generate energy for vehicles to cruise, acceleration, or deceleration, we propose the specific equations to estimate fuel consumptions under these three different operational states based on the law of conservation of energy. In other words, we assume fuel consumption is proportional to the cruise speed at the cruise condition, and to the difference of the squared speeds at the acceleration and deceleration conditions, mathematically,

$$ECI = k_1\overline{v_i} + k_2\sum_{j=1}^{J}\left({v_j}^2 - {v_{j+1}}^2\right) + k_3\sum_{m=1}^{M}\left({v_{m+1}}^2 - {v_m}^2\right) + k_0 \qquad (5)$$

where ECI is the proposed energy consumption index, k_0, k_1, k_2, and k_3 are calibration coefficients, $\overline{v_i}$ denotes average cruise speed under cruise state i; v_j denotes speed at Record j under deceleration states; v_m denotes speed at Record m under acceleration states; and I, J, and M represent total number of records with respect to the three different states.

Due to the differences in energy consumption mechanism, we use different coefficients for cruise, deceleration, and acceleration states. Having had the speeds–fuel consumption data, we can calibrate the

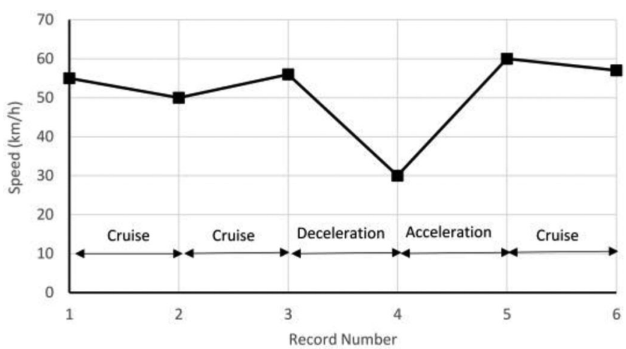


Fig. 4. Categorization of three states from six records.

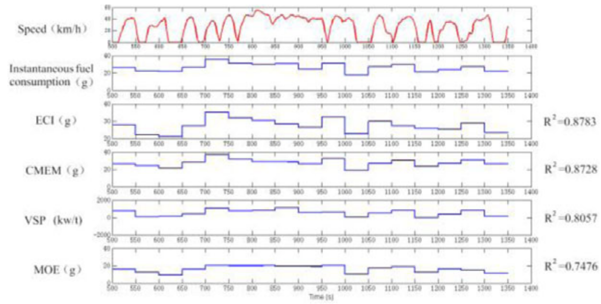


Fig. 5. Comparison based on Experiment 1.

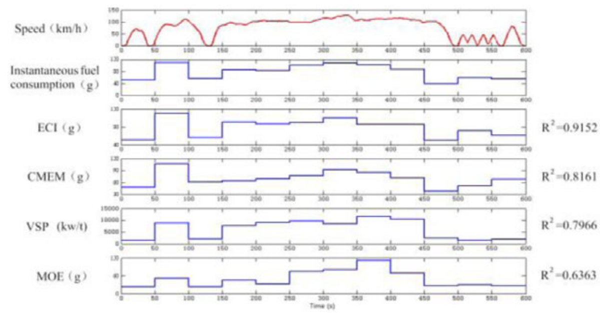


Fig. 6. Comparison based on Experiment 2.

ECI coefficients by simple linear regression models. To be more specific, we use fuel consumption as dependent variable and speed functions (i.e., $\overline{v_i}$ for cruise states, $({v_j}^2 - {v_{j+1}}^2)$ for deceleration states, $({v_{m+1}}^2 - {v_m}^2)$ for acceleration states, and 1 for idle states) as independent variables. In this way, the coefficients can be easily calibrated. The calibrated ECI function by using data from Routes 1–5 is given as

$$ECI = 9.50 \times 10^{-6}\overline{v_i} + 3.94 \times 10^{-5}\sum_{j=1}^{J}\left({v_j}^2 - {v_{j+1}}^2\right) + 9.690 \times 10^{-4}\sum_{m=1}^{M}\left({v_{m+1}}^2 - {v_m}^2\right) + 0.432 \qquad (6)$$

4.2 Microscopic validation using vehicle experiments at test bed

Two vehicle experiments at testbeds (25 minutes each) were conducted by Cappiello et al. (2002). The instantaneous speeds and fuel consumption were recorded. This data set, due to its high resolution, is suitable to be used in comparing performance of ECI, CMEM, VSP, and VT-Micro at microscopic level. Figures 5 and 6 report the results for the comparative analyses.

As can be seen in the comparative analysis, all these four models can generally predict the fuel consumptions and the predicted fuel curves are very consistent with the recorded fuel curves. ECI and CMEM are superior to VSP and MOE based on data from both experiments.

5 MODEL II—A GENERALIZED REGRESSION NEURAL NETWORK-BASED MODEL

5.1 Model development

The proposed Model I is essentially a semiphysical model by taking into account the energy conservation in order to quantify the relationship between speeds and fuel consumption. In this section, we propose a data driven model based on neural networks. Neural networks have been widely used to establish an implicit relationship between inputs and outputs (e.g., Adeli and Hung, 1994; Adeli, 2001; Ahmadlou et al., 2010; Wang and Adeli, 2015; Hirschauer et al., 2015; Calviño et al., 2016). In this research, we employ a novel GRNN. In GRNN, we first group different driving behaviors and establish nonlinear prediction function for each group of driving behaviors (trajectories). Intuitively, GRNN will have very strong performance in predicting fuel consumption (Rooki, 2016; Ni and Li, 2016). This model can be considered as an advanced lookup table when it predicts the fuel consumptions for new driving behaviors. GRNN is composed of input layer, hidden layer, summation layer, and output layer. With the ability of local approximation and less artificial adjustment parameter, GRNN can learn faster and avoid potential erroneous results from subjective parameters. The structure of GRNN is shown in Figure 7.

The GRNN equation Equation (7) is as below:

$$\mathrm{E}(\mathrm{Y}|X) = \frac{\sum_{i=1}^{n} Y_i \exp(-(X - X_i)^T (X - X_i)/2\sigma^2)}{\sum_{i=1}^{n} \exp(-(X - X_i)^T (X - X_i)/2\sigma^2)} \quad (7)$$

$\mathrm{E}(\mathrm{Y}|X)$ can be seen as a weighted average of all the output of the training samples, Y_i, where each Y_i is weighted exponentially according to the Euclidian distance between X(the input of the testing sample) and X_i (the ith input of the training samples). In this theory

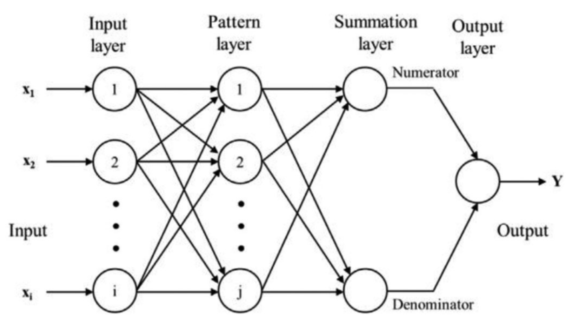


Fig. 7. GRNN model structure diagram.

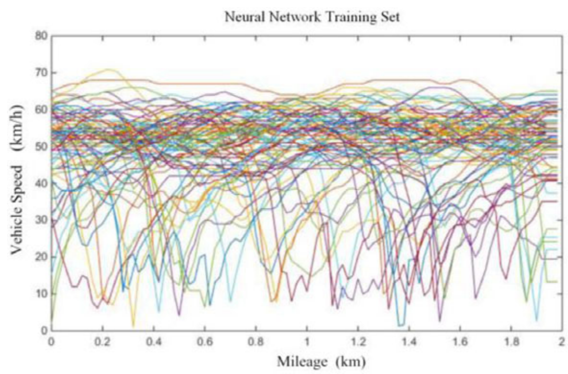


Fig. 8. The speed–mileage curves (input vector X).

σ is the smoothing factor, and optimum-smoothing factor can be determined after many runs according to the mean squared error (MSE) of the estimated values. This process is referred to as the training of the network. We again use data from Routes 1–5 to train the GRNN. The input vector X is described in the Appendix, including mean speed, standard deviation of speeds, peak speed, bottom speed, acceleration share, deceleration share, cruise share, acceleration interval, time duration, is an array that characterizes the feature from speed–mileage curves, which is shown in Figure 8. The output Y is the corresponding accumulative fuel consumption for each 2 km interval. With the speed–fuel consumption data from Routes 1–5, we establish an implicit relationship between Input X and Output Y in our GRNN.

5.2 Identification of eco-driving behaviors

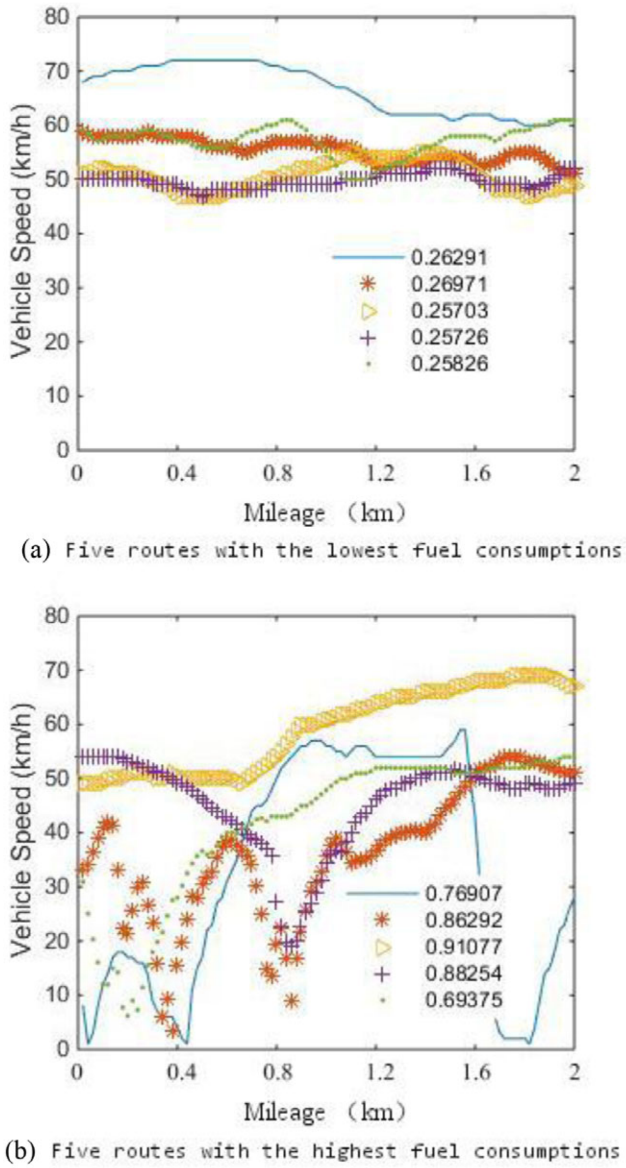
We further analyze 10 segments of 2 km, five of which are associated with the lowest fuel consumptions and the other five are with the highest (Figure 9). It is found that the fuel consumption is positively related to speed fluctuation and acceleration share.

As can be seen in Figure 9, all five segments with the lowest fuel consumptions are either cruise states with an average speed of 50–70 km/hour, or generally at decelerations. Significant oscillations or accelerations are observed for the five segments with the highest fuel consumptions. This provides an empirical evidence for our ECI model proposed in Section 4.

6 PERFORMANCE IN PREDICTING FUEL CONSUMPTIONS FOR NEW ROUTES

6.1 Qualitative performance analyses

We use the speed–fuel consumption data of Routes 1–5 to train/calibrate the models, and use the data of Routes 6 and 7 to test/compare these models. As



(a) Five routes with the lowest fuel consumptions

(b) Five routes with the highest fuel consumptions

Fig. 9. Ten selective routes of 2 km.

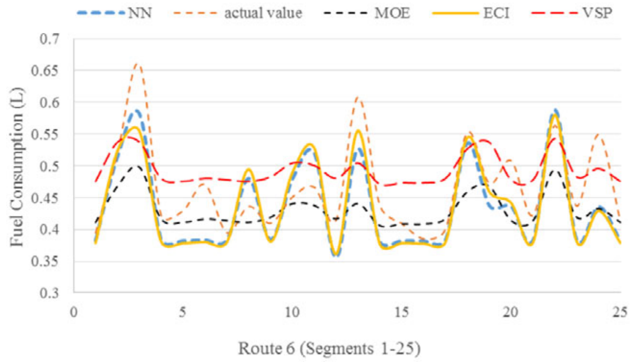


Fig. 10. Performance comparison of Route 6.

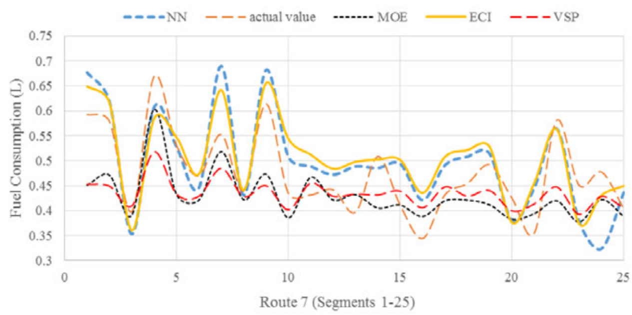


Fig. 11. Performance comparison of Route 7.

our time resolution is 2–5 seconds, we do not have second-by-second data. Consequently CMEM is not compared in this section and it is only validated based on the data from two experiments in test beds.

Routes 6 and 7 represent two crowded corridors leading to the central business district (CBD) of Shanghai. As a result, there are oscillations in fuel consumption along with different segments. However, as can be seen in Figures 10 and 11, the two proposed models can reasonably well reestablish the pattern of actual fuel consumptions.

6.2 Quantitative performance indicators

In order to check the applicability of the two proposed models in comparison with VSP and VT-Micro models, we use relative error (RE), Root Mean Squared Error (RMSE), R^2, and adjusted R^2, which have been widely used (Qu et al., 2014; Qu et al. 2015; Jin et al., 2015; Kuang et al., 2014; Ma et al., 2017; Yao et al., 2017). The RE is calculated as the absolute error divided by the actual value,

$$RE = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|f_i - \hat{f}_i\right|}{f_i} \tag{8}$$

where $\hat{f}_i$ is fuel consumption for observation i predicted by various models, f_i is the actual fuel consumption, and n is number of observations in the range. The RMSE represents the sample standard deviation of the differences between actual data and predicted values, which can be calculated by the root of the mean square error. Mathematically,

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(f_i - \hat{f}_i)^2} \tag{9}$$

Table 2
Performances in predicting Route 6

Model	*RE*	*RMSE*	R^2	*Adj.* R^2
GRNN	0.089	0.051	0.700	N.A.
ECI	0.093	0.053	0.663	0.593
VT-Micro	0.100	0.064	0.626	0.502
VSP	0.125	0.063	0.558	0.496

Table 3
Performances in predicting Route 7

Model	*RE*	*RMSE*	R^2	*Adj.* R^2
GRNN	0.097	0.054	0.561	N.A.
ECI	0.099	0.054	0.571	0.485
VT-Micro	0.170	0.113	0.550	0.400
VSP	0.163	0.105	0.541	0.476

The *RE*, *RMSE*, R^2, and adjusted R^2 of the four models (ECI, GRNN, VT-Micro, VSP) in predicting Routes 6 and 7 are presented in Tables 2 and 3. As can be seen in the two tables, all the four models calibrated/trained based on data from Routes 1–5 have reasonably well performed in predicting fuel consumptions in Routes 6 and 7. GRNN has the best performance in terms of *RE*, *RMSE*, and R^2.

6.3 Speed specific error test

We further carry out the speed-specific error tests. According to our results, all four approaches have a reasonably good prediction when the speed is higher than 20 km/hour. Under low-speed conditions, the relative errors are generally greater. This is mainly because some neglected factors (e.g., fuel consumptions from electrical, transmission, transitional, and operational systems and from pavement conditions) constitute a relatively larger proportion of the overall fuel consumption, while they are relatively marginal under mid to high-speed conditions. It should be pointed out that the proposed GRNN approach significantly outperforms the other three models under low-speed conditions.

7 CONCLUSIONS

It has been widely recognized that the fuel consumption and GHG emissions generated by trucks are highly overrepresented, which imposes a great degree of urgency to develop a reliable approach that is able to accurately and dynamically predict truck fuel consumption. In this research, by taking advantage of the data from Internet of Vehicles, we develop one semiphysical model and one data driven model to more accurately estimate truck fuel consumptions. The first model is on the basis of a new index named ECI and the second model is based on a GRNN.

To measure the performance of the proposed model, we further compared the two proposed models with three well-recognized existing models (VSP, VT-Micro, and CMEM). According to the comparative analysis, the proposed models have a comparable or stronger performance in predicting fuel consumptions. As the ECI is directly generated from truck drivers' driving behaviors, it can be used to design the more energy efficient driving behavior in the soon-to-come era of connected and automated vehicles.

ACKNOWLEDGMENTS

This research is jointly supported by the National Natural Science Foundation of China (51278058), the Fundamental Application Research Program of China Ministry of Transport (no. S2013JC9397), the 111 Project (no. B14043), the Funds for Key Scientific and Technological Innovation Team of the Shaanxi Province (no. 2017KCT-29), and the Joint Laboratory of Internet of Vehicles sponsored by Ministry of Education and China Mobile. The authors are grateful for the data support by Shaanxi Automotive Group. The authors appreciate six anonymous reviewers for their thorough and most helpful comments, and for the journal's rigorous review process.

REFERENCES

Adeli, H. (2001), Neural networks in civil engineering: 1989–2000, *Computer-Aided Civil and Infrastructure Engineering*, **16**(2), 126–42.

Adeli, H. & Hung, S. L. (1994), *Machine Learning: Neural Networks, Genetic Algorithms, and Fuzzy Systems*, John Wiley & Sons, Inc., Hoboken, NJ.

Ahmadlou, M., Adeli, H. & Adeli, A. (2010), Fractality and a wavelet-chaos-neural network methodology for EEG-based diagnosis of autistic spectrum disorder, *Journal of Clinical Neurophysiology*, **27**(5), 328–33.

An, F., Barth, M., Norbeck, J. & Ross, M. (1997), Development of comprehensive modal emissions model: operating under hot-stabilized conditions, *Transportation Research Record: Journal of the Transportation Research Board*, **1587**, 52–62.

Barth, M., An, F., Norbeck, J. & Ross, M. (1996), Modal emissions modelling: a physical approach, *Transportation Research Record: Journal of the Transportation Research Board*, **1520**, 81–88.

Barth, M., An, F., Younglove, T., Scora, G., Levine, C., Ross, M. & Wenzel, T. (2000), The development of a comprehensive modal emissions model, Final report to the National Cooperative Highway Research Program. Available at http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_w122.pdf, accessed 8 August 2017.

Barth, M., Scora, G. & Younglove, T. (1999), Estimating emissions and fuel consumption for different levels of freeway congestion, *Transportation Research Record: Journal of the Transportation Research Board*, **1664**, 47–57.

Barth, M., Younglove, T., Wenzel, T., Scora, G., An, F., Ross, M. & Norbeck, J. (1997), Analysis of modal emissions from diverse in-use vehicle fleet, *Transportation Research Record: Journal of the Transportation Research Board*, **1587**, 73–84.

Boriboonsomsin, K., Barth, M., Zhu, W. & Vu, X. (2012), Eco-routing navigation system based on multisource historical and real-time traffic information, *Intelligent Transportation Systems, IEEE Transactions on*, **13**(4), 1694–704.

Calviño, A., Grande, Z., Sánchez-Cambronero, S., Gallego, I., Rivas, A. & Menéndez, J. M. (2016), A Markovian–Bayesian network for risk analysis of high speed and conventional railway lines integrating human errors, *Computer-Aided Civil and Infrastructure Engineering*, **31**(3), 193–218.

Cappiello, A., Chabini, I., Nam, E. & Lue, A. (2002), A statistical model of vehicle emissions and fuel consumption, *Proceedings of the IEEE International Conference on Intelligent Transportation Systems*, **2002**, 801–09.

Chen, P. & Guan, W. (2017), Mapping to cells: a simple method to extract traffic dynamics from probe vehicle data, *Computer-Aided Civil and Infrastructure Engineering*, **32**(3), 252–67.

Chen, Z., Yang, C. & Chen, A. (2015), Estimating fuel consumption and emissions based on reconstructed vehicle trajectories, *Journal of Advanced Transportation*, **48**(6), 627–41.

Duarte, G., Gonçalves, G., Baptista, P. & Farias, T. (2015), Establishing bonds between vehicle certification data and real-world vehicle fuel consumption—a vehicle specific power approach, *Energy Conversion and Management*, **92**, 251–65.

Environmental Protection Agency (2015), *Greenhouse GAS Inventory Report*, EPA, Washington, D.C., U.S., 1990–2014.

Ericsson, E. (2001), Independent driving pattern factors and their influence on fuel-use and exhaust emission factors, *Transportation Research Part D*, **6**(5), 325–45.

Giannelli, R., Nam, E., Helmer, K., Younglove, T., Scora, G. & Barth, M. (2005), *Heavy-Duty Diesel Vehicle Fuel Consumption Modeling Based on Road Load and Power Train Parameters*, SAE Technical Paper 2005-01-3549, https://doi.org/10.4271/2005-01-3549.

Hirschauer, T. J., Adeli, H. & Buford, J. A. (2015), Computer-aided diagnosis of Parkinson's disease using enhanced probabilistic neural network, *Journal of Medical Systems*, **39**(11), 1–12.

Jiménez, J. L. (1998), Understanding and quantifying motor vehicle emissions with vehicle specific power and TILDAS remote-sensing. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Jiménez, J. L., McLintock, P. M., McRae, G. J., Nelson, D.D. & Zahniser, M. S. (1999). Vehicle specific power: a useful parameter for remote sensing and emissions studies (PDF), *Ninth CRC On-Road Vehicle Emissions Workshop*, San Diego, CA, April 1999.

Jin, S., Qu, X., Zhou, D., Xu, C., Ma, D. & Wang, D. (2015), Estimating capacity and bicycle equivalent units for electric bicycles on shared bicycle paths, *Transportation Research Part A*, **77**, 225–48.

Jung, J., Jayakrishnan, R. & Park, J. Y. (2016), Dynamic shared-taxi dispatch algorithm with hybrid-simulated annealing, *Computer-Aided Civil and Infrastructure Engineering*, **31**(4), 275–91.

Kolmanovsky, V., McDonough, K. & Gusikhin, O. (2011), Estimation of fuel flow for telematics-enabled adaptive fuel and time efficient vehicle routing, *11th International Conference on ITS Telecommunications (ITST) 2011*, St. Petersburg, Russia, pp. 139–44.

Kuang, Y., Qu, X. & Wang, S. (2014), Propagation and dissipation of crash risk on saturated freeways: a case study in Queensland, Australia, *Transportmetrica B: Transport Dynamics*, **2**(3), 203–14.

Levin, M. W., Boyles, S. D., Duthie, J. & Pool, C. M. (2016), Demand profiling for dynamic traffic assignment by integrating departure time choice and trip distribution, *Computer-Aided Civil and Infrastructure Engineering*, **31**(2), 86–99.

Liao, T. Y., Ho, W. M. & Hu, T. Y. (2012), Fuel consumption and emission models development and application for advanced traffic management strategies, *15th International IEEE Conference on Intelligent Transportation Systems (ITSC) 2012*, Anchorage, AK, pp. 846–51, 2012, ISSN 2153-0009.

Ma, Z., Xie, J., Qi, X., Xu, Y. & Sun, J. (2017), Two-dimensional simulation of turning behavior in potential conflict area of mixed flow intersections, *Computer-Aided Civil and Infrastructure Engineering*, **32**(5), 412–28.

Martani, C., Jin, Y., Soga, K. & Scholtes, S. (2016), Design with uncertainty: the role of future options for infrastructure integration, *Computer-Aided Civil and Infrastructure Engineering*, **31**(10), 733–48.

Ministry Environmental Protection of People's Republic of China (2013), Environment Statistics Annual Report, Beijing, China.

Nesamani, K. S., Saphores, J.-D., McNally, M. G. & Jayakrishnan, R. (2017), Estimating impacts of emission specific characteristics on vehicle operation for quantifying air pollutant emissions and energy use, *Journal of Traffic and Transportation Engineering* (English Edition), **4**(3), 215–29.

Ni, Y. & Li, M., (2016), Wind pressure data reconstruction using neural network techniques: a comparison between BPNN and GRNN, *Measurement*, **88**, 468–76.

Qi, Y., Teng, H. & Yu, L. (2004), Microscale emission models incorporating acceleration and deceleration, *Journal of Transportation Engineering*, **130**(3), 348–59.

Qu, X., Wang, S. & Zhang, J. (2015), On the fundamental diagram for freeway traffic: a novel calibration approach for single-regime models, *Transportation Research Part B*, **73**, 91–102.

Qu, X., Yang, Y., Liu, Z., Weng, J. & Jin, S. (2014), Potential crash risk of expressways on-ramps and off-ramps: a case study in Beijing, China, *Safety Science*, **70**, 58–62.

Qu, X., Zhang, J. & Wang, S. (2017), On the stochastic fundamental diagram for freeway traffic: model development, analytical properties, validation, and extensive applications, *Transportation Research Part B*, **104**, 256–71.

Rakha, H., Ahn, K. & Trani, A. (2004), Development of VT-micro model for estimating hot stabilized light duty vehicle and truck emissions, *Transportation Research Part D*, **9**(1), 49–74.

Rooki, R. (2016), Application of general regression neural network (GRNN) for indirect measuring pressure loss of Herschel–Bulkley drilling fluids in oil drilling, *Measurement*, **85**, 184–91.

Slavin, D., Abou-Nasr, M. A., Filev, D. P. & Kolmanovsky, I. V. (2013), Empirical modeling of vehicle fuel economy based on historical data, *The 2013 International Joint Conference on Neural Networks (IJCNN)*, Dallas, TX, pp. 1–6, 2013, ISSN 2161–4407.

Taheri, E., Kolmanovsky, I. & Atkins, E. (2016), Enhanced smoothing technique for indirect optimization of minimum-fuel low thrust trajectories, *Journal of Guidance, Control, and Dynamics*, **39**(11), 2500–11.

Walnum, H. & Simonsen, M. (2015), Does driving behavior matter?An analysis of fuel consumption data from heavy-duty trucks, *Transportation Research Part D*, **36**, 107–20.

Wang, N. & Adeli, H. (2015), Self-constructing wavelet neural network algorithm for nonlinear control of large structures, *Engineering Applications of Artificial Intelligence*, **41**, 249–58.

Wu, X., Freese, D., Cabrera, A. & Kitch, W. A. (2015), Electric vehicles' energy consumption measurement and estimation, *Transportation Research Part D*, **34**(34), 52–67.

Yao, B., Chen, C., Cao, Q., Jin, L., Zhang, M., Zhu, H. & Yu, B. (2017), Short-term traffic speed prediction for an urban corridor, *Computer-Aided Civil and Infrastructure Engineering*, **32**(2), 154–69.

Yin, X., Li, Z., Shah, S. L., Zhang, L. & Wang, C. (2015), Fuel efficiency modeling and prediction for automotive vehicles: a data-driven approach, *2015 IEEE International Conference on Systems Man and Cybernetics (SMC)*, Kowloon, China, pp. 2527–32s.

Zhai, H., Frey, H., Rouphail, C. & Nagui, M. (2008), A Vehicle-Specific Power approach to speed and facility-specific emissions estimates for diesel transit buses, *Environmental Science and Technology*, **42**(21), 7985–91.

Zockaie, A., Aashtiani, H. Z., Ghamami, M. & Nie, Y. M. (2016), Solving detour-based fuel stations location problems, *Computer-Aided Civil and Infrastructure Engineering*, **31**(2), 132–44.

APPENDIX: THE DESCRIPTION OF THE INPUT X OF GRNN

Input	*Formula*	*Description*
Mean speed	$\overline{v} = \frac{2}{t}$	t is the travel time for 2 km.
Standard deviation of speeds	$std = \sqrt{\frac{\sum_{i=1}^{n}(v_i - \overline{v})^2}{n}}$	v_i is the recorded truck speeds, n is the total number of all v_i.
Peak speed	$v_{peak} = \max\{v_1, v_2, v_3, \ldots, v_n\}_{2km}$	v_{peak} is the maximum recorded speed in an interval.
Bottom speed	$v_{bottom} = \min\{v_1, v_2, v_3, \ldots, v_n\}_{2km}$	v_{bottom} is the minimum recorded speed in an interval.
Acceleration share	$Accs = \frac{\sum_{i=1}^{p}\int_{t_i^-}^{t_i^+} vdt}{\int_0^t vdt}$	t_i^- is the start time of the ith acceleration segment in an interval. t_i^+ is the end time of the ith acceleration segment in an interval.
Deceleration share	$Decs = \frac{\sum_{j=1}^{q}\int_{t_j^-}^{t_j^+} vdt}{\int_0^t vdt}$	t_j^- is the start time of the jth deceleration segment in an interval. t_j^+ is the end time of the jth deceleration segment.
Cruise share	$Crus = \frac{\sum_{k=1}^{r}\int_{t_k^-}^{t_k^+} vdt}{\int_0^t vdt}$	t_k^- is the start time of the kth cruise segment in an interval. t_k^+ is the end time of the kth cruise segment in an interval.
Acceleration interval	num	The number of the acceleration segments of the speed curve in an interval.
Acceleration time duration	t_a	Total acceleration time duration.
Deceleration time duration	t_d	Total deceleration time duration.
Cruise time duration	t_c	Total cruise time duration.