Passenger Car Unit Values of Connected Autonomous Vehicles in Urban Road Networks

Christina Gasparinatou
National Technical University of Athens
Department of Transportation Planning and Engineering
5 Iroon Polytechniou St., GR-15773, Athens, Greece
Email: xrigasparinatou@gmail.com

Maria G. Oikonomou
National Technical University of Athens
Department of Transportation Planning and Engineering
5 Iroon Polytechniou St., GR-15773, Athens, Greece
Email: moikonomou@mail.ntua.gr

Athina Tympakanaki
Aimsun SLU
Ronda Universitat 22B, 08007 Barcelona, Spain
Email: athina.tympakianaki@aimsun.com

Jordi Casas
Aimsun SLU
Ronda Universitat 22B, 08007 Barcelona, Spain
Email: jordi.casas@aimsun.com

Tamara Djukic
Aimsun SLU
Ronda Universitat 22B, 08007 Barcelona, Spain
Email: tamara.djukic@aimsun.com

Eleni I. Vlahogianni
National Technical University of Athens
Department of Transportation Planning and Engineering
5 Iroon Polytechniou St., GR-15773, Athens, Greece
Email: elenivl@central.ntua.gr

George Yannis
National Technical University of Athens
Department of Transportation Planning and Engineering
5 Iroon Polytechniou St., GR-15773, Athens, Greece
Email: geyannis@central.ntua.gr

Word count: 6,411 words + 2 tables (250 words per table) = 6,911 words

Submitted [July 29, 2021]

ABSTRACT
Connected autonomous vehicles (CAVs) are expected to gradually penetrate urban traffic and significantly affect traffic operations in a microscopic and macroscopic level. In this work, we aim to estimate the Passenger Car Unit (PCU) value of a CAV under different market penetration rate scenarios and further quantify its relationship between road geometry and control (road type, control type etc.) using microscopic simulation. The PCU value is estimated as the capacity change observed in the (network and link) Macroscopic Fundamental Diagram (MFD), when different mixtures of vehicle technologies may exist on the road network. For the purpose of this work, eleven future mobility scenarios are executed in the Aimsun Next mobility modeling software and the resulting PCU values are estimated. Classical statistical and machine learning regression models are further developed to identify the factors that may affect the estimated PCU values. Findings show that, in a network level, there exists a polynomial relationship between CAVs' PCU and their penetration rate in the traffic mix. In a link level, the CAV PCU value is found to be highly affected by the observed lane flow, the section length, the control type, the road type, the number of lanes, the number of public transport lines of the road segment, as well as the market penetration rate of CAVs. The paper ends with a discussion on the implications of the results for the macroscopic modeling and the testing of CAV related management policies.

Keywords: Connected Autonomous Vehicles (CAVs), Passenger Car Unit (PCU), Macroscopic Fundamental Diagram (MFD), urban traffic, microscopic simulation, random forest

INTRODUCTION

According to navigation and mapping company TomTom (*1*), drivers in Athens spend 37% more time on the road due to increased traffic congestion. Regarding road safety, the National Highway Traffic Safety Administration – NHTSA (*2*) estimated that 94% of accidents are caused by human error. In addition, 90% of the harmful air pollutants come from 25% of cars (*3*). Connected Autonomous Vehicles (CAVs), given automation that is said to eliminate human error and connectivity increasing user/vehicle perception about other users and the system conditions, can be considered a promising solution towards increasing road safety, reducing traffic congestion and environmental pollution (*4, 5*).

But along with the promise of these new vehicle technologies comes the perils. Although, from the early stages of research on CAVs, researchers have suggested that autonomous vehicles are expected to improve traffic flow by increasing network capacity (*6–8*) and hence the penetration of such technologies is, nowadays, expected to be gradual. Until their full dominance, legacy and CAVs will coexist on the roads creating inhomogenous traffic conditions. This may cause an explosive mixture of interactions which could lead to critical roadway traffic and safety conditions. Especially in the early stages of implementation, CAVs impacts on traffic conditions and capacity should be taken into serious consideration. There exist studies with evidence that – at an early stage – there will be detrimental effects to network capacity and overall traffic performance (*9, 10*).

In the new traffic landscape, estimating the Passenger Car Unit (PCU) for CAVs is of outmost importance for meaningful capacity analysis, signal design, and traffic management. Various research has been conducted concerning the estimation of PCU of conventional vehicles and different methods have been developed. Some of the methods are based on speed modeling (*11–14*), on headway (*14–17*), on space occupancy (*18–20*) as well as on time occupancy (*17–21*). The research landscape regarding PCU values in urban road networks is fragmented; PCU have been considered to take a static value or values in relation to traffic characteristics, road geometrics and other factors (*22*).

Literature on autonomous or/and connected vehicles and PCU values have been mainly focused on highways (*23–27*) and signalized intersections (*28–33*). The research on possible PCU values for CAVs in urban road networks where a greater variability of geometry and control conditions may arise is still at its infancy.

We extend past research by proposing a method to estimate the PCU factor of CAVs in urban road networks in relation to the market penetration rate and other geometry and control factors. The approach is based on quantifying the effect of CAVs' penetration rate to changes in capacity at a network – and link – level as depicted in the Macroscopic Fundamental Diagram (MFD). The analyses are based on simulated experiments of various future mobility scenarios. The proposed methodology and outcomes enable the detailed estimation of CAV impacts on traffic in a variety of traffic conditions and road infrastructure.

The paper is organized as follows: the next section describes the methodological approach, in which the PCU factor estimation approach and MFD method are presented. Following, we present the simulation framework and its parameterization and discuss the network-level impacts, provided from the MFD approach, as well as link-level impacts. Next, we present the results from the development and training of the machine learning model to identify the factors that may affect the link-level PCU values. The paper ends with a summary of specific research outcomes, as well as some concluding remarks.

METHODOLOGICAL APPROACH

Passenger Car Unit (PCU) measures the impact of a transport mode (passenger cars, heavy vehicles, buses, etc.), as a function of vehicle dimensions and operating capabilities, on the traffic flow efficiency compared to a standard unit of passenger car. Hence, a PCU factor of 1 is used as the unit for conventional cars. In the case of AVs, as field data are not yet available, the estimation of the PCU values can only be based on simulation output data (*34, 35*). Depending on the percentage of AVs in the traffic demand, different capacity levels are expected to occur, hence, the PCU values will vary accordingly. One approach to derive the equivalence in PCU factors is to calculate the proportion of capacity reached by vehicles of different types (e.g., buses, trucks, AVs) with respect to Conventional Vehicles (CV) (reference vehicles) as follows (**Equation 1**):

$$PCU_{AV} = PCU_{CV} \frac{capacity_{CV}}{capacity_{AV}} \qquad (1)$$

where PCU_{AV} is the PCU factor for autonomous vehicles, PCU_{CV} is the PCU factor for conventional vehicles and is assumed to be equal to 1, $capacity_{CV}$ is the derived capacity for conventional vehicles and $capacity_{AV}$ is the derived capacity for autonomous vehicles. In this study the same PCU value for conventional vehicles is considered for all the vehicle types (e.g., private cars, trucks, buses). Nevertheless, the PCU estimation can be refined to consider different PCU values for conventional vehicles depending on the vehicle type.

Capacity change is derived from the network's Macroscopic Fundamental Diagram (MFD). Experiments and simulations performed by Daganzo and Geroliminis (*36*) showed that in many cases the average volume and the average density of a road network are related with a reproducible curve, known as the Macroscopic Fundamental Diagram (MFD). The theory behind this diagram allows a conversion of prediction models from a microscopic level in macroscopic control and monitoring, which is of utmost significance, since microscopic data are rarely used for large networks (*37*). The MFD is a model that depicts the road's network traffic state and basically represents the relationship between traffic demand and traffic supply in the road network. Through this method, it is possible to estimate the traffic conditions of the network and thus, determine its level of service. In addition, the MFD theory is appropriate for predicting traffic-related variables, as in a research conducted by Gao et al. (*38*), in which the predicted speed was estimeted using the MFD model.

The concept of the urban road network MFD is utilised in order to derive the network capacities for different simulation scenarios with various AV penetration rates. One approach to represent the network's MFD considers the total number of vehicles passing through the network (i.e. that have reached their destination) versus the total number of vehicles inside the network (which equals the density divided by the sum of all link lengths of the road network), within a specified interval. Alternatively, the total network production (veh × distance travelled per unit time) and accumulation (veh) can be calculated from loop detectors on links (*39, 40*). The first approach for obtaining the network's MFD is preferred in homogeneously congested networks, while the second can be more accurate for the estimation of the average network speed and density in less homogeneously congested networks. For this analysis, the first approach is adopted as the values for the macroscopic variables to define the MFD can be derived directly from the simulation.

The above methodology may have far reaching managerial implications in relation to the rate of penetration of CAVs in road networks. Macroscopic route choice and assignment models apply Volume Delay Functions (VDFs) to model the travel time or the cost on a section or connection as a function of different parameters such as the road segment traffic volume, capacity, length, etc. VDFs further use the concept of PCU to convert the capacity and volumes into passenger car equivalents for each vehicle type (*41–43*). With the proposed approach, the derived functional relationship between the PCU factors and AV penetration rates is used as input to the VDFs in the macroscopic simulation models in order to forecast the implications of AVs on the network performance under various future mobility scenarios (**Figure 1**).

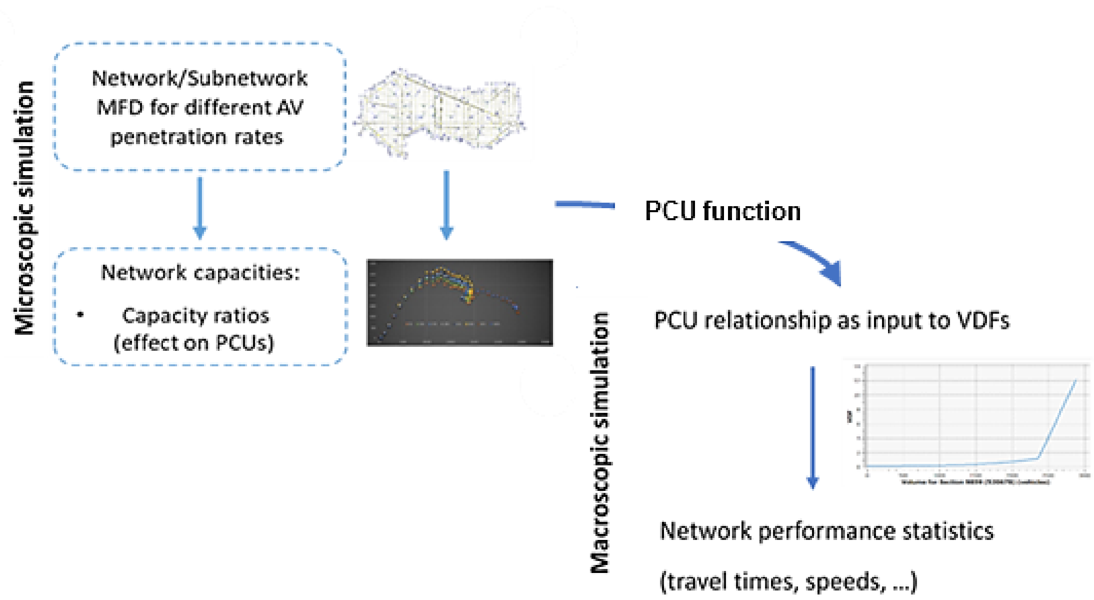

Figure 1 Transition from microscopic to macroscopic analysis using PCU function

SIMULATING AUTONOMOUS VEHICLES IN URBAN ROAD NETWORKS

Microscopic simulation method provides information related to individual vehicles by modeling traffic flows at a high level of detail (*44*). The simulation inputs concern data from various sources such as the network geometry, traffic volume and modal split. In addition, data exported from the microsimulation can provide an initial, descriptive estimation of several impacts. Each vehicle is tracked as it interacts with surrounding traffic as well as with the environment. In microscopic simulation, interactions between vehicles at intersections are represented and thus, every vehicle in the network is recorded (*45*). This results to the basic principle of microsimulation, which depicts traffic conditions as realistically as possible. Moreover, microscopic simulation is widely used to evaluate new traffic control and management technologies as well as performing analysis of existing traffic operations (*46*). Furthermore, Giuffrè et al., (*34*) and Kang et al., (*35*) used a microscopic simulation model to estimate the PCE value of heavy vehicles in roundabouts and in freeways, respectively.

For the above reasons, the microscopic simulation method could examine impacts on traffic and provide insights into the impacts of microscopic flow characteristics under several traffic simulation scenarios and evaluate the influence of different CAV penetration rates on a microscopic level. The microscopic simulation analysis explores the implications of CAVs on the network efficiency. Therefore, in the present research, the microscopic simulation method was selected to extract data of different market penetration rates of CAVs in order to be further analyzed and then the PCU values of CAVs to be estimated in relation to traffic, control and road network conditions.

Study Network

The study network that has been simulated in the Aimsun Next mobility modeling software is the city center of Athens as shown in **Figure 2** (left). The network created in Aimsun is presented in **Figure 2** (right) and consists of 1,137 nodes and 2,580 road segments. The network data for each road segment concerned geometric as well as functional characteristics, namely, length, width, number of lanes, directions, free flow speed and capacity. The respective characteristics of nodes that were included in the model network were the allowed movements, number of lanes per movement, priority, traffic light control plans, free speed flow and capacity. More specifically, the road segments of the network are 1,424 secondary streets, 1,033 signalized streets and on/off ramps and 123 arterials. In addition, in 909 road segments of the network there are traffic lights, in 133 yield signs, in 354 stop signs and in the rest 1,184 road segments there are no signals. The total length of road sections is 348 km and the network size reaches approximately 20 km^2.

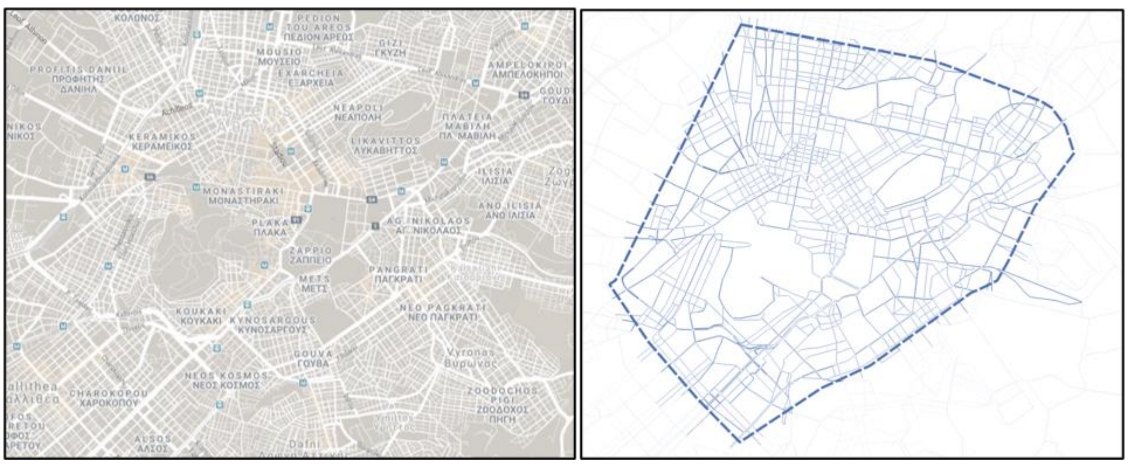

Figure 2 The city of Athens network in a conventional map (left) and in the Aimsun software (right)

Moreover, the microscopic model was calibrated using data that were collected for year 2019 from 107 detectors, which are recording traffic volume in main roads in Athens network. Extensive tests of the traffic model were performed so that the results of the model correspond to real traffic conditions. Since the R^2 coefficient value was greater than 0.90 (specifically 0.98), the model represents a good validation. Additional data of field measurements were also considered. The OD matrices that were extracted consisted of 290×292 centroids and a total number of 82,270 passenger car trips and 3,110 truck trips for the morning peak hour. Furthermore, the Athens model included public transport, namely 95 bus and 14 trolley lines as well as their 1,030 public transport stations, the service frequencies and waiting times at stops.

Modeling Autonomous Vehicles

For modeling connected and autonomous vehicles (CAVs) within the present research, we implemented an aggressive CAV profile, which relates to short clearance in car-following, short anticipation distance for lane selection, short clearance in gap acceptance in lane changing, limited overtaking, no cooperation and small gaps. Different driving profiles were investigated as in other studies (*47–49*). The behavior of CAVs was modeled by using the Gipps car following model (*50*). This model is able to mimic the behavior of real traffic; the parameters involved correspond to obvious driver and vehicle characteristics and affect the behavior of the simulated flow in logical consistent ways. More specifically, the model predicted the response of the following vehicle based on the assumption that drivers set limits to their desired braking and acceleration rates. The alterations in the parameters of the Gipps car-following model to simulate directly CAVs are shown in **Table 1**. In addition, in modeling CAVs, it was necessary to take into account their lane-changing behavior as it is considered to be different than human driven vehicle behavior. For this reason, the Gipps lane changing model was applied in the present research (*51*), as well. This model analyzes the decisions that drivers have to make before changing lanes and ensures that the simulated drivers behave logically in situations that are similar to real traffic conditions. The vehicle parameters of the car-following and lane-changing behavior that were used in microsimulation are presented in **Table 1**.

For the present study, eleven (11) microscopic simulation scenarios of vehicle substitution with CAVs were executed. In each scenario, we gradually increased the penetration rate of CAVs (every 10%) along with the equivalent decrease in the conventional vehicles. The base scenario consisted solely of conventional vehicles and heavy vehicles. Each scenario included 2-hour simulation (before and during morning peak) and was executed multiple times (10 replications with random seeds generating stochastic results) and the aggregated output was extracted. Finally, the simulation step was 2 minutes.

TABLE 1 Parameters of the Car Following and Lane Changing Model per Vehicle Type

Model	Factor		Human-Driven vehicle	Aggressive CAV	Explanation on factors of Aggressive CAV
	Reaction Time	in car following	0.8 sec	0.4 sec	Vehicle connectivity
		at STOP	1.2 sec	0.4 sec	
		at traffic light	1.6 sec	0.4 sec	
Car Following Model	Sensitivity	Mean	100%	50%	Implication of shorter headways compared to CVs
		Min	100%	10%	
		Max	100%	90%	
Lane Changing Model	Overtake Speed Threshold		90%	85%	Implication of caution on overtaking manouvers compared to CVs
	Cooperate in Creating a Gap		YES	NO	Smaller gaps compared to CVs
	Distance Zone	Min	0.80	1.00	Logner distance at which lane change to diverge from a motorway compared to CVs
		Max	1.20	1.25	
	Safety Margin		1.00	0.75 - 1.25	Longer clearance compared to CVs

Network-Level Impacts

Figure 3 depicts the resulting network-level MFD for each of the market penetration rate scenarios as a result of the average of 10 replications. This MFD is considered to be relatively well-defined during free-flow and peak conditions. It can also be observed that, the increase of CAV market penetration rate leads to increased network throughput and therefore, increased capacity. The capacity increase can be attributed to the lower reaction times that CAVs can achieve due to their enhanced capabilities compared to conventional vehicles. More specifically, the average increase in capacity is 8% (140 veh/ 2 min) when CAVs reach 100% market penetration rate, compared to the capacity of the base-case scenario (0% CAVs).

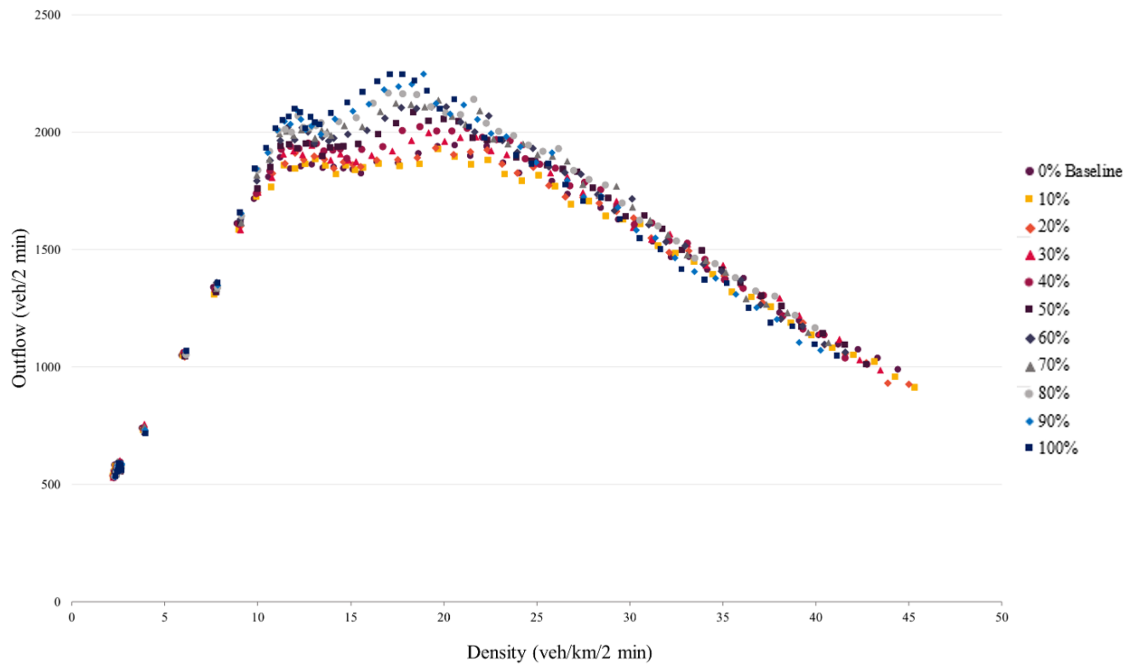

Figure 3 MFD for different CAV penetration rate

In addition, the network capacity is increased by an average rate of 13% (200 veh/ 2 min) compared to the base-case scenario during peak hour conditions, whereas the highest increase in capacity appeared to be 18% (347 veh/ 2 min). Furthermore, the change in the penetration rate of CAVs from 20% to 30% appears to have the highest increase in capacity. In this case, the network capacity is increased by 5% as an average value. Finally, the MFD illustrates that, when more CAVs existing the network the density values are decreased improving network traffic conditions.

Moreover, through a regression analysis in network-level, the scatter plot of the estimated PCU value for CAVs (based on **Equation 1**) with the market penetration rate of CAVs in the examined network was extracted (**Figure 4**). The polynomial model results show that, for low market penetration rates the PCU value of CAV is greater than 1, indicating that the CAVs will have a negative impact in the traffic flow. Reduction in PCU (below 1) is observed when CAV market penetration rate exceeds 40%. The estimated functional relationship and the R^2 coefficient value (R^2=0.93) of the model is also presented in **Figure 4**. The high R^2 value indicates a good fit. In addition, the t-statistics and p-value demonstrate that the estimated model is statistically significant at a 95% confidence level.

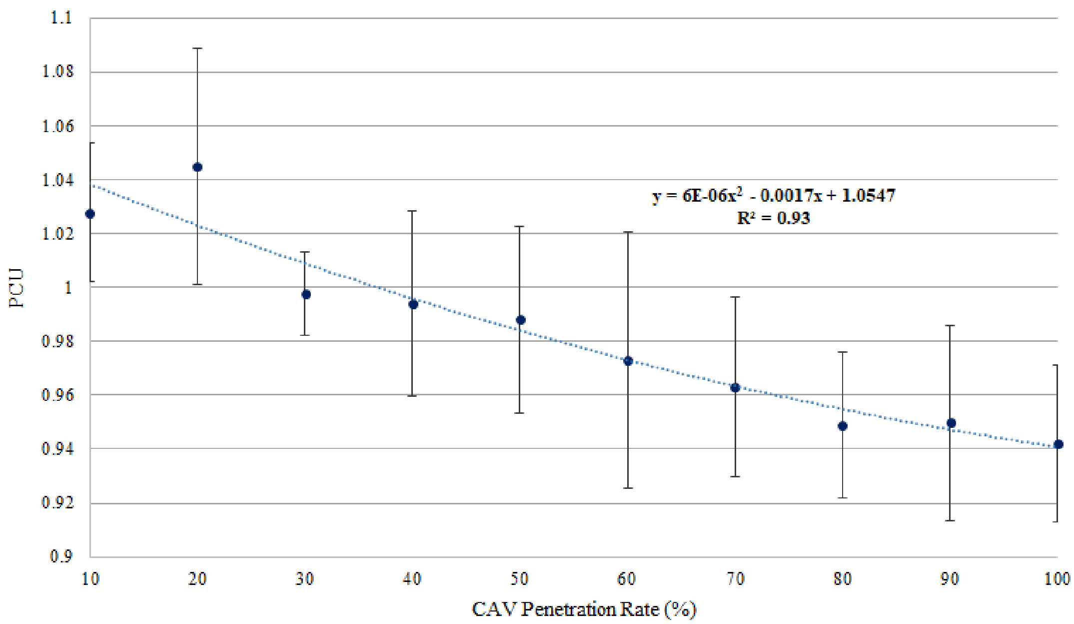

Figure 4 Estimated PCU factors versus CAV market penetration rate

Link-Level Impacts

The final output concerns link-level data, extracted from the microscopic simulation, and consists of the PCU estimation required variables, according to **Equation 1**, as well as additional traffic data, namely lane flow that were also recorded every 2 minutes (simulation step). These data were enriched with the number of lanes, the public transport lines and the length of each road segment. In addition, the corresponding road type and traffic control type of the road segment were also included. Therefore, the final dataset consists of the the variables presented in **Table 2**.

TABLE 2 Variables in Dataframe

Variable	Type	Description
PCU	Continuous	Passenger Car Unit factor for CAVs
Lane flow	Continuous	Road segment traffic flow per lane (veh/h/lane)
Length	Continuous	Length of road segment (m)
Control type	Categorical	0: No signal, 1: Yield sign, 2: Stop sign, 3: Traffic light
Road type	Integer	0: Secondary Street, 1: Signalized street - On/off ramp, 2: Arterial
Penetration rate	Integer	Market penetration rate of CAVs (0-100%)
Number of lanes	Integer	Number of lanes of road segment
Number of public transport lines	Integer	Number of public transport lines occurring in road segment

Figure 5 illustrates that road type and traffic control type significantly affect the PCU factors of CAVs. It can also be observed that arterials indicate lower PCU factor range of CAVs comparing to secondary and signalized streets. This is justifiable, since higher speeds are observed in arterias and therefore CAVs operate more efficiently. Respectively, non-controlled

road segments present lower PCU factors. Therefore, higher PCU factors of CAVs are presented when there is a stop sign or traffic light compared to when there is a yield sign or no control. Finally, **Figure 6** displays the distribution of PCU values that follows the typical pattern for all normal distibutions with a range of 0.55-1.4 and an average value of 0.95.

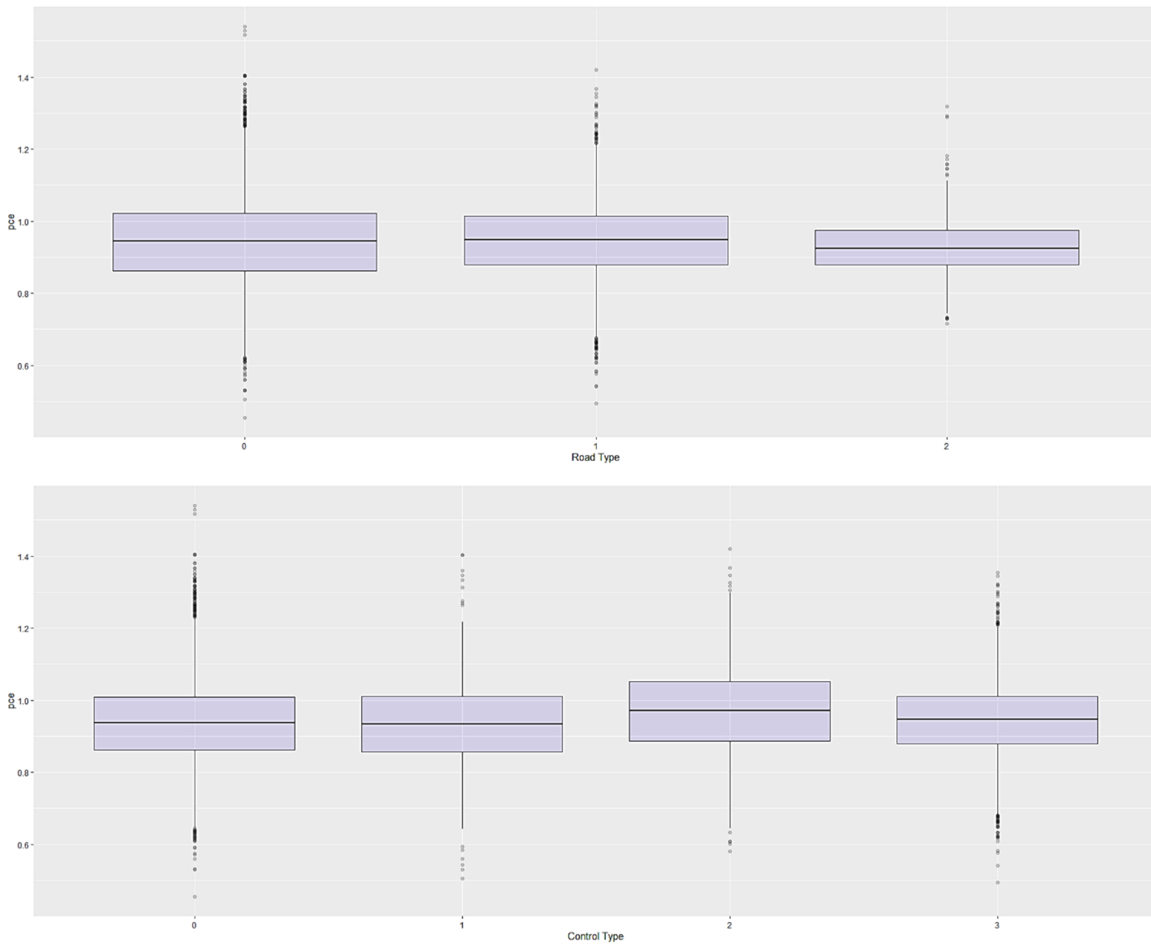

Figure 5 Estimated PCU factors in relation to road type and traffic control type

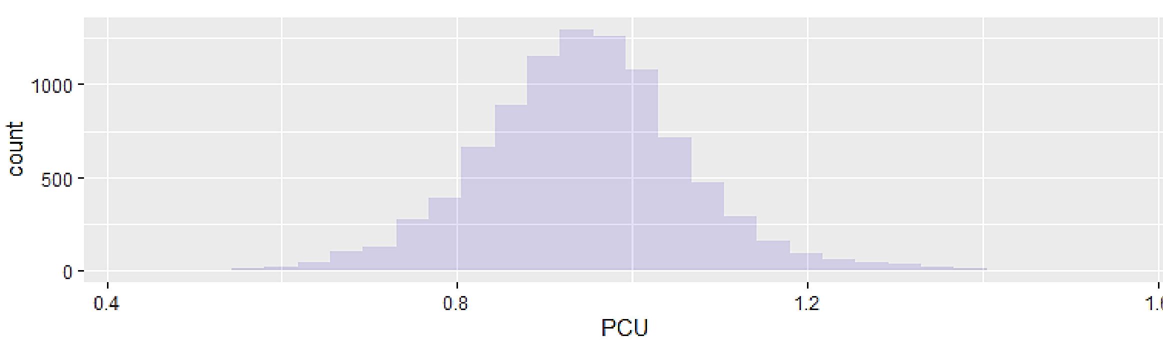

Figure 6 PCU distribution

FACTORS AFFECTING PCU VALUES USING MACHINE LEARNING

In order to estimate the factors converting autonomous vehicles to PCU, a Random Forest (RF) regression model was developed. The RF is a decision tree-based algorithm that combines the prediction of several simpler models, i.e., Decision Trees (DT), in order to improve robustness and generalization. In RF, a predefined number of decision trees is developed using a random sample drawn from the training set, following certain limitations about the size of the trees and the number of leaves. The randomness introduced in the algorithm is considered to decrease the variance and chances of overfitting, two issues that are common with simple decision trees. Another advantage of RF over DT is that it is flexible to the data on which it is trained, resulting to less significant effects when training data are differentiated (*52, 53*). The pseudocode of RF for regression is the following:

1. For b=1 to B:
    a. Draw a bootstrap sample Z^* of size N from the training data.
    b. Grow a random-forest tree T_b to the bootstrapped data, by recursively repeating the following steps for each terminal node of the tree, until the minimum node size n_{min} is reached.
        i. Select m variables at random from the p variables.
        ii. Pick the best variable/split-point among the m.
        iii. Split the node into two daughter nodes.
2. Output the ensemble of trees $\{Tb\}_1^B$.

The prediction at a new point x is given as follows (**Equation 2**):

$$\hat{f}_{rf}^{B}(x) = \frac{1}{B} \sum_{b=1}^{B} T_b(x) \qquad (2)$$

In contrast to simple decision trees, which provide simple classification rules, which can also be applied manually, in each of their nodes, Random Forests are not so easy to interpret, due to their complexity. Interpretability can be distinguished in the following categories (*54*): (1) Model-specific interpretation tools that are limited to specific model classes and deal with weights and structural properties (e.g. the practices followed in linear models), and (2) Model-agnostic tools that can be used on any Machine Learning (ML) model and are applied after the model has been trained (post hoc) and aim to analyze the relationship between feature input and output pairs without considering weights or structural information specific to the model. Moreover, interpretation can be seen as global, when the entire behavior of a model is attempted to be explained or local when the focus is on why a prediction for a certain instance is given. Some well-known methods for ML interpretability include the Partial dependence plots, the permutation importance and conditional permutation importance, the Individual conditional expectation, the Local surrogate models, the Accumulated Local Effects (ALE) and the SHAP (*54*).

In the specific study, we use the Mean Decrease Accuracy method of computing the feature importance on permuted out-of-bag (OOB) samples based on mean decrease in the accuracy, which is similar to permutation-based methods. Given a base value of MSE in out-of-bag-CV, we permute the values of each variable at a time and compute the % relative change in MSE (%IncMSE); the most important variables will have the higher values of %IncMSE.

To complement the above approach, researchers used the node imputity index (IncNodePurity), which relates to the decrease in the loss function after a certain split using a specific variable; the larger the decrease of impurity after a certain split, the more informative the corresponding input variable (*55*). The two indices are shown to be highly correlated and share the same bias (*56*).

For dataframe analysed above, a random forest model was developed with dependent variable the PCU factor of CAVs and independent variables the simulated lane flow, the road type and traffic control type, the length, the number of lanes and the number of public transport lines of each road segment, as well as the penetration rate of CAVs (**Table 2**). The model's parameterization is as follows:

- Base calssifier: Decision Tree,
- Number of tree estimators: 100,
- Maximum depth limitation: No (inf),
- Minimum number of samples to split an internal node: 2,
- Minimum number of samples for each leaf node: 1

Training results show that the RF model can explain the 45.9% of the variance in the dataset, with a mean squared residuals of 0.0227. The % explained variance is a measure that describes how well the model can make predictions of a test set that explains the target variance of the training set. Therefore, it is considered that the RF model presents a good fit. In addition, the Mean Absolute Percentage Error (MAPE) is investigated and it seems that the RF model can predict the PCU factors of CAVs with an error of 8.35%.

The variable's importance as unveiled from the RF models is very helpful to define which variables are significant. **Figure 7** demonstrates the importance rankings of the RF variables based on %IncMSE and IncNode Purity. The variables with the higher values of both measures are the lane flow and length of the road segment and followed by the market penetration rate of CAVs and the number of public transport lines variables. The variables presenting lower values are the number of lanes, control type and road type of road segment variables. Additionally, the variable with higher %IncMSE rate is the road segment length, whereas the ones with higher IncNode Purity are the number of transit lines occuring in road segment and its lanes' flow.

Finally, it can be concluded that the lane flow and length of road segment variables are considered as the most important variables. It should be noted that the specific variable importance should be interpreted as a relative ranking of predictors (*57, 58*). Although, in RF models, the magnitude of the effect and the sign of each variable are not easily identified, specific xAI techniques exist that can be implemented in order to extract the sign of the influence of each predictor to the PCU value (*59*).

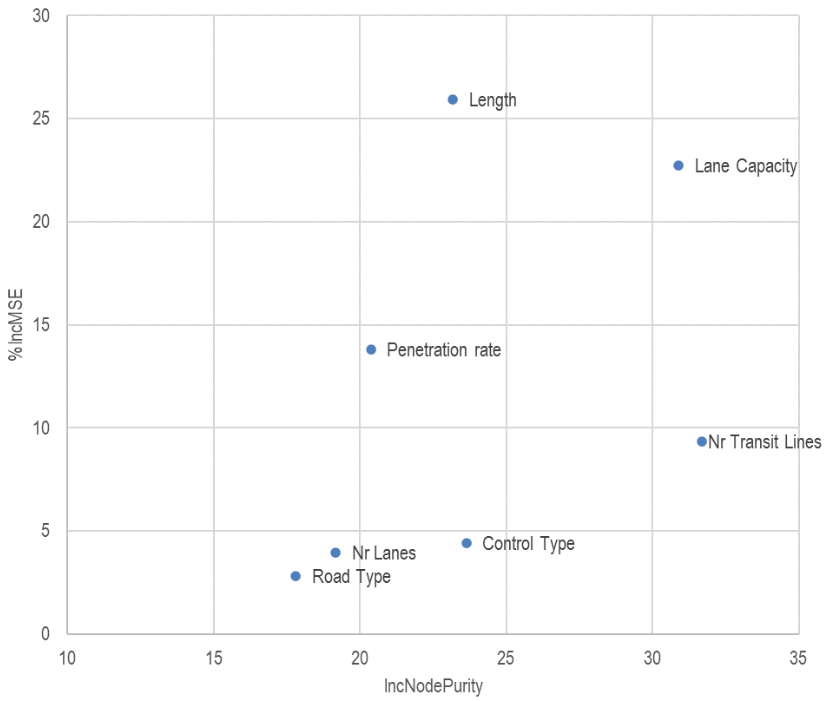

Figure 7 %IncMSE in relation to IncNode Purity

CONCLUSIONS

The recent technological innovations lead to a new mobility paradigm, which will be mostly marked by autonomous vehicles. The integration of Connected and Autonomous Vehicles (CAVs) to existing traffic will take place in the near future with great effects on the urban environment. For transport researchers, as well as stakeholders, it is prudent to anticipate the impacts that automation will introduce. Within this framework, the present research proposed an approach to estimate the impact of CAVs on traffic were compared to a typical Passenger Car Unit (PCU) as the ratio of the capacities of conventional vehicles to the corresponding capacities of the CAVs on link level of an urban network based on a network-level Macroscopic Fundamental Diagram (MFD).

To that end, a set of simulation-based analyses were first conducted on the Athens (Greece) microscopic simulation tesbed, using Aimsun Next modeling software, for the identification and measurement of the impacts of CAVs on the network. Eleven (11) market penetration rate scenarios (0-100%) were established with a simulation duration of two hours and a simulation time step of extracting data of two minutes. Furthermore, the advent of automation is modelled within the network by the examination of an aggressive CAV profile. Based on the results of the microscopic simulation, the MFD was used in order the network capacities to be derived. In particular, the capacities were then translated into PCU factors for each of the investigated CAV penetration rates and afterwards a polynomial relationship between PCU values of CAVs and their penetration rate was indentifed. A deeper look at the link level impacts of CAVs on capacity showed that PCU factors of CAVs could be linked to the geometry and control characteristics of each road segments. Finally, a Machine Learning (ML) model was trained using the link-level results obtained by the microscopic simulation experiments, in order to further investigate these interrelationships.

Findings demonstrated evidence that, automation and connectivity will work towards beneficially improving traffic conditions in cities and especially network capacities. In network level, the average increase in capacity, when CAVs reach 100% market penetration rate, is identified to be 8%, compared to the capacity of no automation scenario. In addition, when more CAVs existing the network the density values are expected to be decreased by improving network traffic conditions. Moreover, another important outcome of the performed network level analysis is that, there exists a significant relationship between PCU factors of CAVs and their market penetration rate in the traffic mix. More specifically, for low market penetration rates the PCU value of CAV is greater than 1, indicating that the CAVs will have a negative impact in the traffic flow, while a reduction in PCU (below 1) is observed when CAV market penetration rate exceeds 40%. In a link level approach, the PCU value of CAVs is found to be highly affected by the observed lane flow, length, control type, road type, number of lanes, and number of public transport lines of the road segment, as well as the market penetration rate of CAVs. The most significant variables are the lane flow and length of the road segment followed by the market penetration rate of CAVs and the number of public transport lines variables. Finally, in arterials, the PCU values of CAVs are lowaer compared to those of the secondary and signalized streets, probably due to the higher observed speeds. Non-controlled road segments present lower PCU values.

The above network and link level impacts of CAVs to traffic capacity may have far-reaching implications for the deployment of citywide Connected and Automated Traffic (CAT) strategies. The extracted PCU values are extremely useful for evaluating and forecasting the long-term impacts on travel demand (e.g. mode and destination choice), as well as wider impacts such as changes in land use (area attractiveness, employment, parking spaces, etc.). The approach can be used by city authorities, operators and other stakeholders to adjust the macroscopic volume delay functions and test the impact of future mobility strategies in a strategic level. The estimated values and functional relationships should be further evaluated with respect to the ability of the approach to scale up to different networks. The further investigation of scalability and transferability would provide robust PCU functional relationships that can be directly applied to macroscopic models of other urban networks in order the corresponding CAVs impacts to be analyse, without the need of a microscopic simulation analysis. Nevertheless, the critical role of connectivity and its impact in relation to simple automation needs is to be further investigated, taking into account the particularities of different areas (i.e., vehicle types, number of signalised intersections, dedicated lanes).

ACKNOWLEDGMENTS

The present research was carried out within the research project “LEVITATE - Societal Level Impacts of Connected and Automated Vehicles”, which has received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement No 824361.

AUTHOR CONTRIBUTION STATEMENT

The authors confirm contribution to the paper as follows: study conception and design: C. Gasparinatou, M. G. Oikonomou, A. Tympakianaki, J. Casas, T. Djukic, E. I. Vlahogianni, G. Yannis; data collection: C. Gasparinatou, M. G. Oikonomou; analysis and interpretation of results: C. Gasparinatou, M. G. Oikonomou, A. Tympakianaki, J. Casas, T. Djukic, E. I. Vlahogianni, G. Yannis; draft manuscript preparation: C. Gasparinatou, M. G. Oikonomou. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. International company TomTom, 2020, https://www.tomtom.com/.

2. National Highway Traffic Safety Administration (NHTSA). Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey, 2015.

3. Richard, M. G. 25% of Cars Are Causing 90% of the Air Pollution That We Breath, New Study Finds, 2021.

4. Forrest, A., and M. Konca. Autonomous Cars & Society, May 1, 2007.

5. Althoff M., O. Stursberg, and M. Buss. Model-Based Probabilistic Collision Detection in Autonomous Driving. *IEEE Transactions on Intelligent Transportation Systems*, 2009, pp. 299-310.

6. Shladover, S., D. Su, and X. Y. Lu. Impacts of Cooperative Adaptive Cruise Control on Freeway Traffic Flow. *Transportation Research Record: Journal of the Transportation Research Board*, 2324, 2012, pp.63–70.

7. Litman, T. Autonomous Vehicle Implementation Predictions. *Victoria Transport Policy Institute 28*, 2014.

8. Friedrich, B. The Effect of Autonomous Vehicles on Traffic. In Autonomous Driving. Springer, Berlin, Heidelberg., 2016, pp. 317-334.

9. Calvert, S. C., W. J. Schakel, and J. W. C. Van Lint. Will Automated Vehicles Negatively Impact Traffic Flow?. *Journal of Advanced Transportation*, 2017.

10. Overtoom, I., G. Correia, Y. Huang, and A. Verbraeck. Assessing the Impacts of Shared Autonomous Vehicles on Congestion and Curb Use: A Traffic Simulation Study in The Hague, Netherlands. *International Journal of Transportation Science and Technology*, S2046043020300265, 2020 –. doi:10.1016/j.ijtst.2020.03.009.

11. Aerde, V. M., and S. Yagar. Capacity, Speed, and Platooning Vehicle Equivalents for Two Lane Rural Highways, 1984.

12. Chandra, S., V. Kumar, and P. K. Sikdar. Dynamic PCU and Estimation of Capacity of Urban Roads. *Indian Roads Congress*, Vol. 23, Issue 4, 1995.

13. Rahman, M., and F. Nakamura. Measuring Passenger Car Equivalents for Non-Motorized Vehicle (Rickshaws) at Mid-Block Sections. *J. Eastern Asia Soc. Transp.*, 2005, Stud. 6: 119–126. https://doi.org/10 .11175/easts.6.119.

14. Patil, S. V. and P. R. Adavi. Development of Passenger Car Units: Case Study-NAL Stop, Pune. *Int. J. Sci. Eng*, 2015, pp. 89-95.

15. Ahmed, U. Passenger Car Equivalent Factors for Level Freeway Segments Operating

Under Moderate and Congested Conditions, 2010.

16. Ahmad, A., and R. Rastogi. Regression Model for Entry Capacity of a Roundabout Under Mixed Traffic Condition-An Indian Case Study, 2017.

17. Mohan, M., and S. Chandra. Three Methods of PCU Estimation at Un-Signalized Intersections, 2018.

18. Chandra, S., and U. Kumar. Effect of Lane Width on Capacity Under Mixed Traffic Condition in India, 2003.

19. Paul, P. K., and P. K. Sarkar. Determination Of Dynamic PCUs Of Different Types of Passenger Vehicles on Urban Roads: A Case Study, Delhi Urban Area, 2013.

20. Swamy, C. A. R., C. R. Munigety, and M. V. L. R. Anjaneyulu. Passenger Car Unit Based on Influence Area, 2016.

21. Sonu, M., A. Dhamaniya, S. Arkatkar, and G. Joshi. Time Occupancy as Measure of PCU at Four Legged Roundabouts, 2016.

22. Sharma, M., and S. Biswas. Estimation of Passenger Car Unit on urban roads: A literature review. *International Journal of Transportation Science and Technology, 2020*.

23. Kanaris, A., P. Ioannou, and F. S. Ho. Spacing and Capacity Evaluations for Different AHS Concepts: Automated Highway Systems. Edited by P. Ioannou. Berlin: Springer, 1997, pp. 125–171.

24. Vander Werf, J., S. E. Shladover, M. Miller, and N. Kourjanskaia. Effects of Adaptive Cruise Control Systems on Highway Traffic Flow Capacity. *Transp. Res. Rec.,* 1800 (1), 2002, pp. 78–84. https://doi.org/10.3141 /1800-10.

25. Van Arem, B., C. J. G. van Driel, and R. Visser. The Impact of Cooperative Adaptive Cruise Control on Traffic-Flow Characteristics. IEEE *Trans. Intell. Transp. Syst.*, 7 (4), 2006, pp.429–436. https://doi.org/10 .1109/TITS.2006.884615.

26. Kesting, A., M. Treiber, M. Schonhof, and D. Helbing. Adaptive Cruise Control Design For Active Congestion Avoidance. *Transp. Res. Part C: Emerging Technologies 16* (6), 2008, pp. 668–683. https://doi.org/10 .1016/j.trc.2007.12.004.

27. Hurtado-Beltran, A., and L. R. Rilett. Impact of CAV Truck Platooning on HCM-6 Capacity and Passenger Car Equivalent Values. *Journal of Transportation Engineering, Part A: Systems*, Vol. 147, 2021.

28. Dresner, K., and P. Ston. A Multiagent Approach to Autonomous Intersection Management. *J. Artif. Intell. Res.,* 31, March 2008, pp. 591–656. https://doi.org/10.1613/jair.2502.

29. Li, Z., M. Chitturi, D. Zheng, A. Bill, and D. Noyce. Modeling Reservation-Based Autonomous Intersection Control in VISSIM. *Transp. Res. Rec.*, 2381 (1), 2013, pp. 81–90. https://doi.org/10.3141/2381-10.

30. Li, Z., L. Elefteriadou, and S. Ranka. Signal Control Optimization for Automated Vehicles at Signalized Intersections. *Transp. Res. Part C: Emerging Technologies 49*, Dec 2014, pp. 1–18. https://doi.org/10.1016/j.trc .2014.10.001.

31. Le Vine, S., A. Zolfaghari, and J. Polak. Autonomous cars: The Tension Between Occupant-Experience and Intersection Capacity. *Transp. Res. Part C: Emerging Technologies*, 52, Mar 2015, pp. 1–14. https://doi.org/10.1016/j.trc.2015.01.002.

32. Mohan, M., and S. Chandra. Queue Clearance Rate Method for Estimating Passenger Car Equivalents at Signalized Intersections. *Journal of Traffic and Transportation Engineering (English Edition),* Vol. 4, 2017, pp. 487-495.

33. Şentürk Berktaş Ecem, Tanyel Serhan. Effect of Autonomous Vehicles on Performance of Signalized Intersections. *Journal of Transportation Engineering, Part A: Systems*, 146(2), 2020, 04019061–. doi:10.1061/JTEPBS.0000297

34. Giuffrè, O., A. Granà, R. Mauro, A. B. Silva, and S. Chiappone. Developing Passenger Car Equivalents for Freeways by Microsimulation. *Transportation Research Procedia,* Vol. 10, 2015 pp. 93 – 102.

35. Kang, N., and H. Nakamura. An Analysis of Heavy Vehicle Impact on Roundabout Entry Capacity in Japan. *International Symposium on Enhancing Highway Performance,* 2016, pp. 308–318.

36. Daganzo, C. F., and N. Geroliminis. An Analytical Approximation for the Macroscopic Fundamental Diagram of Urban Traffic, 2008, pp. 771-781.

37. Geroliminis, N., and C.F. Daganzo. Macroscopic Modeling of Traffic in Cities. 86th Annual Meeting of the Transportation Research Board, Washington, D.C., 2007.

38. Gao, Y., J. Zhao, Z. Qin, Y. Feng, Z. Yang, and B. Jia. Traffic Speed Forecast in Adjacent Region between Highway and Urban Expressway: Based on MFD and GRU Model. *Journal of Advanced Transportation,* 2020.

39. Kouvelas, A., J. P. Perrin, S. Fokri, and N. Geroliminis. Exploring the Impact of Autonomous Vehicles in Urban Networks and Potential New Capabilities for Perimeter Control. *5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS),* 2017, pp. 19-24.

40. Lioris, J., R. Pedarsani, F.Y. Tascikaraoglu, P. Varaiya. Platoons of Connected Vehicles Can Double Throughput in Urban Roads. *Transportation Research Part C: Emerging Technologies,* Apr. 1, 2017, pp. 292-305.

41. Friedrich, M., J. Sonnleitner, and E. Richter. Integrating Automated Vehicles into Macroscopic Travel Demand Models. *Transportation Research Procedia 41,* 2019, pp. 360–375.

42. Sonnleitner, J., and M. Friedrich. AV-Ready Macroscopic Modelling Tool. *CoExist Project,* Deliverable 2.7, Version 5, 2020a.

43. Sonnleitner, J., and M. Friedrich. Guide for the Simulation of AVs With a Macroscopic Modelling Tool, *CoExist Project*, Deliverable 2.8, Version 5, 2020b.

44. Ehlert, P. A. M., and L. J. M. Rothkrantz. A Reactive Driving Agent for Microscopic Traffic Simulation. *Proceedings of the 15th European Simulation Multiconference (ESM2001),* 2001, pp. 943–949.

45. Taplin, J. Simulation Models of Traffic Flow, 2008.

46. Owen, L. E., Y. Zhang, L. Rao, and G. McHale. Traffic Flow Simulation Using CORSIM. *Winter Simulation Conference Proceedings,* Vol. 2, 2000, pp. 1143–1147. https://doi.org/10.1109/WSC.2000.899077

47. Sukennik, P. Micro-Simulation Guide for Automated Vehicles. *CoExist Project,* Deliverable D2.5, 2018.

48. Mesionis, G., J. Roca, and M. Brackstone. Modelling CAVs on the UK's Motorway Network – Outputs from the Humandrive Project. *Proceedings of the 17th Annual Transport Practitioners' Meeting,* Oxford, 2019.

49. Casas, J., A. Lenorzer, and J. Perarnau. AV Behavioral Parameters in AIMSUN Next. *LEVITATE project*, 2020.

50. Gipps, P. G. A behavioural Car-Following Model for Computer Simulation. *Transportation Research Part B,* 15(2), 1981, pp. 105–111. https://doi.org/10.1016/0191-2615(81)90037-0

51. Gipps, P. G. A model for the Structure of Lane-Changing Decisions. *Transportation Research Part B,* 20(5), 1986, pp. 403–414. https://doi.org/10.1016/0191-2615(86)90012-3

52. Parsa, A. B., R. Shabanpour, A. K. Mohammadian, A. Joshua, and T. Stephens. A Data-Driven Approach to Characterize the Impact of Connected and Autonomous Vehicles on Traffic Flow. *The International Journal of Transportation Research. Transportation Letters*, 2020.

53. Donges, N. A Complete Guide to the Random Forest Algorithm, 2021.

54. Molnar, C. Interpretable Machine Learning. *In Book*. 2021.

https://christophm.github.io/interpretable-ml-book

55. Dewi, C., and R. C. Chen. Random Forest and Support Vector Machine on Features Selection for Regression Analysis. *International Journal of Innovative Computing, Information and Control,* Volume 15, Number 6, December 2019.

56. Strobl, C., A. L. Boulesteix, A. Zeileis, and T. Hothorn. Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution, *BMC Bioinformatics,* Vol. 8, Article number: 25, 2007.

57. Strobl, C., Malley, J., and Tutz, G. An Introduction to Recursive Partitioning: Rational, Application, and Characteristics of Classification and Regression Trees, Bagging, and Random Forests. *Psychological Methods,* 14(4), 2009a, pp. 323- 348.

58. Strobl, C., J. Malley, and G. Tutz. An Introduction to Recursive Partitioning: Rational, Application, and Characteristics of Classification and Regression Trees, Bagging, and Random Forests. Vol. 14, No. 4, 2009b, pp. 323-348. Retrieved from: http://dx.doi.org/10.1037/a0016973.supp

59. Laña, I., J. J. Sanchez-Medina, E. I. Vlahogianni, and J. Del Ser. From Data to Actions in Intelligent Transportation Systems: A Prescription of Functional Requirements for Model Actionability. Sensors, 21(4), 1121, 2021