

Submitted for

21st EURO Working Group on Transportation Meeting, EWGT 2018, 17-19 September 2018,
Braunschweig, Germany

Exploring the Impact of User Preferences on Shared Autonomous Vehicle Modal Split: A Multi-Agent Simulation Approach

Joseph Kamel^a, Reza Vosooghi^{a,b,*}, Jakob Puchinger^{a,b}, Feirouz Ksontini^a, Göknur Sirin^c

^a*Institut de Recherche Technologique SystemX, Palaiseau 91120, France*

^b*Laboratoire Génie Industriel, CentraleSupélec, Université Paris-Saclay, Gif-sur-Yvette 91190, France*

^c*Direction de la Recherche / Nouvelles Mobilité (DEA-IRM), Technocentre Renault, Guyancourt 78280, France*

Abstract

Shared autonomous vehicles are considered to have a transformative impact on future urban transportation and especially Shared Mobility. In order to assess the transport-related impact of this new mode, various models and simulations are under development. The majority of recently proposed simulations are based on activity-based multi-agent approaches. Thanks to the disaggregated level of data in multi-agent simulation, the traveler decision making mechanisms might be individualized according the attributes. In this paper we try to take in advantage of this granularity in order to explore the impact of user preferences on the modal split of shared autonomous vehicles. To illustrate the proposed methodology the transport system of Paris is simulated by using an activity-based multi-agent tool called MATSim. The traveler preferences toward shared autonomous vehicles use are also summarized based on the literature review.

© 2018 The Authors. Published by Elsevier B.V.

Peer-review under responsibility of the scientific committee of the 21st EURO Working Group on Transportation Meeting.

Keywords: Multi-Agent Simulation; Shared Autonomous Vehicle; User Preferences

1. Introduction

The advent of Autonomous Vehicle (AV) technologies fosters new opportunities for transportation systems by evolving modal shift from private mobility to service-use. It is expected that the next revolution in Shared Mobility will be based on Shared Autonomous Vehicles (SAV). This new service seems to be of high importance for car producers given their recent investments in the technology. However the impact of such systems on traveler behavior, urban form, congestion, and the environment is unknown since no large-scale public SAV service exists today (Stocker & Shaheen, 2018). Nevertheless, considerable efforts are being made nowadays for the planning of future SAV services, and the need for a reliable demand analysis and traffic simulation gains importance.

In order to predict the impact of SAVs on travel behavior and traffic state, new simulation-based approaches have recently been developed where mainly activity-based multi-agent simulations are used (Vosooghi et al., 2017). The latter allow us to model and simulate future SAV systems at a fine level of detail. Actually a multi-agent simulation is developed in which travel demand emerges from the interactions of four types of agents in the transportation system: node, arc, traveler and vehicle (Zheng et al., 2013). Different levels of granularity and interaction complexity between agent types and agents themselves can be supposed in the simulation. While the complex microscopic rules for traffic simulation are already well investigated in this approach, the traveler preferences to use different modes are not

* Corresponding author. Tel.: +33-6-6597-4453.
E-mail address: reza.vosooghi@irt-systemx.fr

adequately reflected. However, multi-agent approaches allow us to consider more details with regards to travelers and to explore how these details affect predictions.

Despite the mentioned advantages of multi-agent approaches, most of the simulations relying on it are based on a homogeneous structure of behavior. This implies that all the travelers have similar preferences when they make mode choice or other decisions. This similarity renders the results less accurate especially considering that user preferences are one of the key components that determine SAV market share and its impact (Stocker & Shaheen, 2018). This paper aims to: first, better understand how user preferences would impact the SAV modal split in a multi-modal transportation system; and second, to gain insight into how to integrate those factors in a multi-agent simulation approach. This study introduces the first large-scale multi-agent activity-based simulation of Paris area developed in MATSim, a multi-agent transport platform. The following sections describe related studies, the methodology of modeling and considering user preferences in a multi-agent simulation, the experiment, key findings, and study conclusions.

2. Literature review

According to the existing literature, large attention is nowadays given to behavioral experimental studies towards using AV services, especially on behavioral characteristics and perceptions, willingness-to-pay and level of awareness and attitudes (Gkartzonikas & Gkritza, 2017). Also there are a limited number of studies specifically addressed to the case of SAV. Some of them are just based on mode choice surveys, but in a few studies more advanced modeling and simulation approach have been conducted. Steck et al. (2018) addressed the impact of autonomous driving on the value of travel time savings (VTTS) and mode choices for commuting trips using stated choice experiments and mixed logit modelling. They compared privately owned autonomous vehicles and SAV to other modes. Their results show that mode-specific parameters such as in-vehicle time and cost play an important role for the mode choice, but the socio-demographic variables (gender, age and income) do not have a significant effect on modal split. The authors also estimated that using SAV will reduce the VTTS up to 10% for commuting trips. Liu et al. (2017) simulated SAV mode requests using agent-based simulation through a stochastic process for four possible fare levels. In their study the out-of-pocket cost and the value of time variables are considered and the modal split of SAVs and conventional vehicles are compared.

In two above-mentioned studies the system-related attributes including travel and waiting time and cost are estimated or considered to have a strong impact on mode choice comparing to the individual-specific characteristics, while the latter have also significant effects according to other studies. Haboucha et al. (2017) developed a model for autonomous vehicle long-term choice decisions and pointed from their panel data that travelers have a strong individual taste-variation for using SAV. What is interesting in their results is that although cost is an important variable in the SAV choice, 25% of travelers would refuse to use SAVs even if they would be completely free. Also their results show that gender, age, educational level and income are the most significant individual-related attributes that have an effect on SAV mode choice. Krueger et al. (2016) selected a stated preferences (SP) mode choice survey to compare SAV with the currently chosen travel modes. In their study the respondents were asked to indicate whether they would switch to SAVs without or with ride-sharing on a recent trip. They further identified the characteristics of potential users by using a mixed logit model. Their results show that SAVs may be more attractive to young travelers and that their acceptance is very sensitive to service attributes. Bansal et al. (2016) analysed individuals' stated frequencies to use SAV under different pricing scenarios and identified the characteristics of potential SAV adopters. In their study gender and age are found to have a significant effect on SAV use.

To the best of the author's knowledge only the above-mentioned studies have specifically dealt with user preferences and SAV mode choice. However, the last three reflect the importance of specific variables on decision mechanisms for choosing SAVs. The surveys investigating those variables and their effects on SAV mode choice from the literature are summarized in Table 1.

Table 1. Socio-demographic variables and their effects on SAV mode choice

Author(s), year	Number of respondents	Variables (positive: +, not significant: *)
Bansal et al. (2016)	347	Gender (male ⁺), age (younger ⁺),
Haboucha et al. (2017)	721	Gender (male ⁺), age (younger ⁺), income (higher ⁺), educational level (higher ⁺)
Krueger et al. (2016)	435	Age (younger ⁺)
Steck et al. (2018)	172	Gender*, age*, income*

According to these studies, the most important attributes impacting SAV use are gender, age, income and educational level. Notwithstanding, the consideration of user preferences according to the mentioned attributes in the simulation remains very limited.

3. Methodology

The work conducted in this research can be described in four major parts: 1) data preparation, 2) multi-agent simulation, 3) adding user preferences, and 4) model set up and calibration.

3.1. Data preparation

As mentioned above in multi-agent approaches, instead of getting a sample representing whole population of a given zone, the behavior of individuals and households is simulated. Thus, such an approach requires to have detailed data. As this microdata is not available for the whole population, a synthetic population is usually generated. This is done by drawing households and individuals from microdata samples on a zonal level. A typical synthetic population is composed of spatially located households which possess certain attributes, such as household size, income, car ownership, etc. Households consist of one or several individuals, who possess personal attributes, such as gender, age, main activity, etc. The preferences of travelers depending on the above-mentioned attributes can be considered in the mode choice decision mechanism in the simulation process.

In this research the synthetic population of around 12 million people for Greater Paris (Île-de-France) is generated from public use microdata (INSEE 2013) and regional transport survey (EGT 2010) relying on a simulation-based synthesis. The population is generated by randomly drawing the adequate number of households (defined by marginal data) from the microdata at a specific spatial zoning called IRIS (5261 zones), see Fig. 1a. The simulation is repeated until some minimum standardized root mean square errors (SRMSE) for other attributes (age ranges, socio-professional categories, etc.) are reached as goodness-of-fit. The simulation-based synthesis has been selected for this study because compared to other popular methods such as iterative proportional fitting (IPF) or combinatorial optimization (CO), this approach is more flexible in terms of adding additional attributes from other sources (Farooq et al., 2013). Also microdata employed in this study consists of an acceptable sample rate (more than 35% of the whole population).

It should be noted that the area for the generation of the synthetic population has been chosen to be bigger than the case study area (i.e. Paris, with a population of around 2,200,000 inhabitants), see Fig. 1b, as this is necessary to capture all trips and activities starting or ending in Paris. Also all modes except cargo and taxi are simulated.

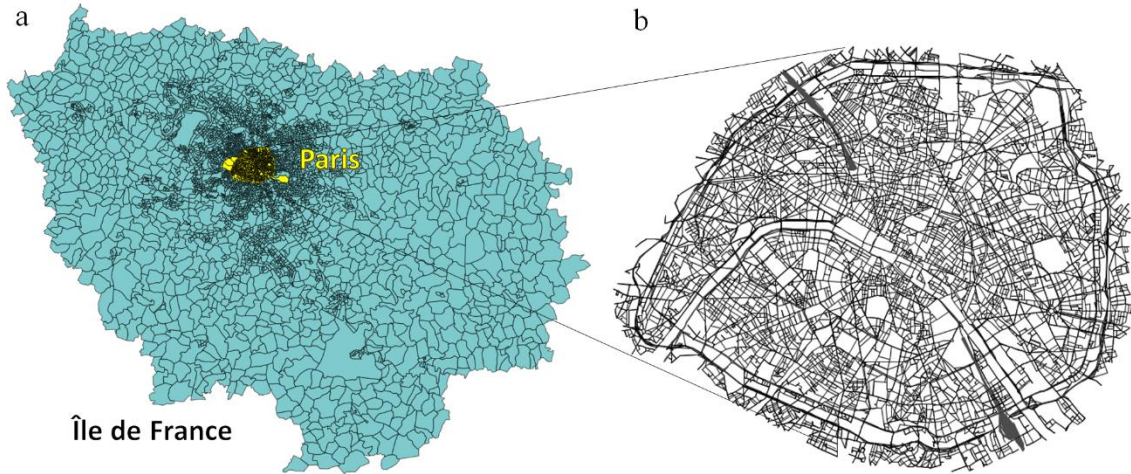


Fig. 1. (a) IRIS spatial zones of Paris and Greater Paris (Île de France); (b) road and transit network of case study area (Paris).

For every individual in a household, activity chains and locations have also been associated to simulate their mobility in a given day, see Fig. 2. Activity chains for each socio-professional category of individuals were derived from a regional transport survey. The locations for home, study and work activities were already available in the bigger aggregated special zoning (commune, 1276 zones for Greater Paris) in the public use microdata. Based on these data, Origin-Destination (OD) matrices have been generated. For other trip purposes the OD matrix was derived from the regional transport survey. To geo-localize each activity inside the zone, facility data has been used. The activities in each zone are randomly distributed and the activity details have been derived from a regional transport survey analysis. In this analysis all the individuals were grouped by socio-professional attributes and for each group the probability of activity start times and durations have been extracted. Activity start times and durations have then been assigned to the synthesized individuals according to the estimated probability. For example, for the activity chain of home-study-home the mean and standard deviation values of start time and activity duration are calculated. Then these values are randomly assigned to each individual (who had the same activity according to their category). To avoid cumulative error at the end of the day, the start time windows for each activity are fixed.

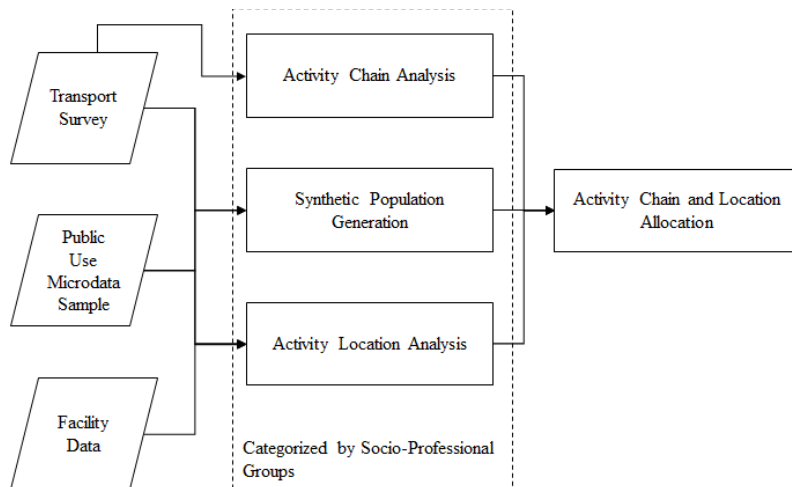


Fig. 2. Data preparation process in order to find synthetic households and individuals with their activity characteristics.

3.2. Multi-agent simulation

Our contribution has been made by using the multi-agent simulation platform MATSim (Horni et al., 2016), which is currently the most widely applied tool for simulating new vehicle sharing systems (Vosooghi et al., 2017). MATSim uses the population with an initial daily plan for each agent as input. The core of MATSim relies on a co-evolutionary algorithm for optimizing agent's plan. In fact, almost all agent-based transport simulations do not involve discrete choice models as they are used in conventional transport demand models, but these simulations are more based on stochastically finding the maximized utility for all agents. In each iteration agents choose or innovate a plan (a set of choices including mode, route and activity end-time) for their day. This plan is then executed with a traffic simulator. The simulation converges once the average performance measures for all agents stabilize. The performance measure in MATSim is presented by a standard scoring function based on the Charypar-Nagel scoring method (Charypar & Nagel, 2005). This scoring function is divided into two parts: activity scoring and travel scoring:

$$S_{plan}^j = \sum_{q=0}^{N-1} S_{act,q}^j + \sum_{q=0}^{N-1} S_{trav,mode(q)}^j \quad (1)$$

For each agent j , with N as the number of activities and trip q which is the trip that follows by activity q , the total score of an executed plan is calculated. The global concept is that the activity score is logarithmically increasing with the activity duration whilst the travel score is usually negative. These two scoring formulas contain various specific or individual-related parameters. The latter are usually set up as been the same for all the agents in the population. However, activity and travel scores can be modified in order to individualize scoring parameters. It is thus possible to integrate preferences to the simulation. Since the orientation of this paper is to study the impact of preferences on the mode choice, only the travel scoring has been modified. Nevertheless activity scoring is also modified in an indirect way by considering the marginal utility of traveling.

3.3. Adding user preferences

Each trip in MATSim is scored based on the mode, travel time and distance. We used in our approach only three major parts of the initial travel scoring function:

$$S_{trav,mode(q)} = C_{mode(q)} + \beta_{trav,mode(q)} \cdot t_{trav,q} + \beta_m \cdot \gamma_{d,mode(q)} \cdot d_{trav,q} \quad (2)$$

For each mode used in trip q the score is calculated from mode-specific constant C_{mode} , marginal disutility of travel $\beta_{trav,mode}$, travel time t_{trav} , marginal utility of money β_m , monetary distance rate $\gamma_{d,mode}$ and $d_{trav,q}$ which is the distance traveled between two activity locations. Instead of having the same coefficient values for all agents, we need to change the scoring according to the agent's attributes. In order to do this, preference factors are introduced. These factors are determined for three attributes: gender, age and income. We assumed that the travel scoring coefficients would vary according to the related factor with specific distributions, see Fig. 3.

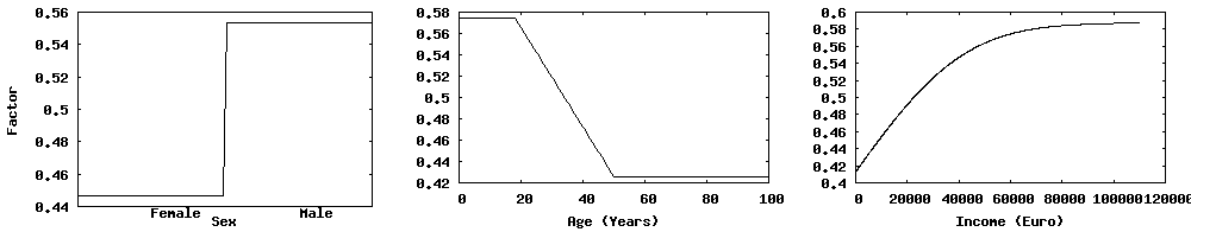


Fig. 3. Distribution graph of preference factors for gender (sex), age range and income.

As there are no data available on SAV user preferences for the Paris area, we use the study of Haboucha et al. (2017) in order to set the variation of each factor. The distribution forms are defined according to our hypothetical assumptions. In general men are more likely to use a SAVs than women, as well as younger persons in comparison to an older ones and wealthy persons to poor ones. For the gender factor κ_{sex} two fix values are assumed. We also assume that the age preference factor κ_{age} for young traveler changes linearly. For income, the preference factor κ_{income} is assumed to grow logarithmically. The preference factor is then multiplied to corresponding coefficients in the SAV scoring function. It should be noted that since the factors are used two times in the calculation, the mean values of the coefficients distribution have been fixed to be 0.5:

$$\begin{aligned}\kappa_{age,sex} &= 2 - (\kappa_{age} + \kappa_{sex}) \\ \kappa_{income} &= 2 - 2 \cdot \kappa_{income}\end{aligned}\tag{3}$$

So the modified travel scoring function for SAV is:

$$S'_{trav,q,sav} = C_{sav} + \kappa_{age,sex} \cdot \beta_{trav,sav} \cdot t_{trav,q} + \kappa_{income} \cdot \beta_m \cdot \gamma_{d,sav} \cdot d_{trav,q}\tag{4}$$

In this function the income factor is multiplied to the monetary part of scoring only, while age and gender factors are multiplied to the second part. The mode-specific constant set during calibration process remains unchanged.

3.4. Model set up and calibration

To set up the travel scoring function, the marginal utility (and disutility) parameters as well as the monetary distance rate have been estimated based on the econometric analysis relying on transport regional surveys. The model is calibrated by mode specific constants which are set up during the simulation according to the overall modal share derived from regional transport statistics. The calibration process is done by trial and error; the goal has been to have a stable modal split that corresponds to reality for the basic scenario. SAVs have then been added to the simulation by means of the hypothetical constant and coefficients. Since the aim of the research is to compare two scenarios, these values are assumed to be hypotheticals, however those values could be set up once survey data for SAVs are available.

4. Experiments and results

Three scenarios have been simulated: 1) the transportation system of Paris area without SAV, 2) SAV added without considering user preferences, and 3) SAV added and user preferences are considered. SAV services are considered to be only car-share and not the ride-share. Four main modes for the basic scenario are considered (PT, private car, walk and bike). In the two last scenarios a fleet of 17000 SAVs with 20 initial distribution points are added to the simulation. The simulation is repeated 900 times and lasted a week (Bullx S6130 - CPU: Intel® Xeon® E7- 88 threads - RAM: 1TBs DDR4). The modal splits of the three scenarios are shown in Table 2.

Table 2. Comparison of modal splits for three scenarios.

Scenario	Modal split (%)				
	PT	Private car	Walk	Bike	SAV
W/O SAV	29.7	21.4	47.1	1.8	-
W/ SAV W/O Preferences	30.1	15.7	45.5	3.4	5.3
W/ SAV W/ Preferences	31.3	15.9	45.6	3.4	3.8

SAV modal shares are estimated to be 5.3% without, and 3.8% with introducing preferences. While the overall modal split of the SAV after introduction of user preferences decreases by 28%, the use of this mode decreases by 38% among the travelers. The modal share of private car decreases up to 26-27% in the two last scenarios. Meanwhile the modal shares of PT increase. That can be because of the users who take SAV for PT access and egress. Also use

of SAV results in the higher use of bike. It can be due to the fact that the previously car users are now more likely to use bike as a well-integrated mode to SAV or PT. In the contrary, walk modal shares decrease in the second and the third scenarios. The longer walking trips could potentially be replaced by SAV in these scenarios. The main difference between two scenarios with SAV is the PT modal split which unusually increases in the third scenario. It seems that this modal split grows because the eliminated SAV sub-trips are replaced by PT. The SAV use fluctuations for each attribute according to their variation before and after the introduction of preferences are also shown in Fig. 4. These changes for older people, females and persons with higher income are more significant.

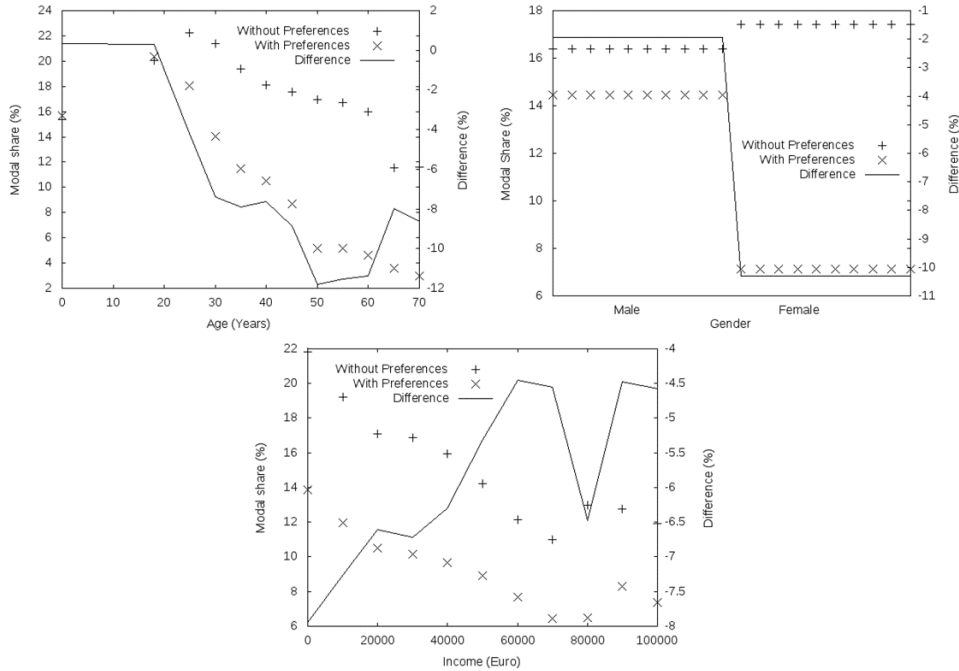


Fig. 4. Modal share (SAV use) changes for each attribute before and after the introduction of preferences.

In order to have a more in-depth analysis of how user preferences (all attributes together) change the SAV modal split, we grouped the users with similar attributes (called profile). The profiling process in the current study is based on one of the attributes that we have generated in the synthetic population. Then the SAV uses by each profile are compared, see Table 3.

Table 3. Comparison between SAV use by profile with and without preferences.

Profile	Population, female / male (% of total)	Average Age (year)	Average Household Income (€)	W/O Preferences (%)	W/ Preferences (%)	Relative Difference (%)
Employed workers	51.2 / 48.8	41.0	35 736	20.1	12.1	-40
Unemployed	52.1 / 47.9	37.9	29 588	15.6	9.8	-37
Retired or pre-retired	60.7 / 39.3	74.5	39 866	11.9	3.3	-72
Students >14 years	54.9 / 45.1	20.3	45 042	19.0	19.2	+1
< 14 years of age	50.6 / 49.4	9.2	45 891	15.9	16.0	+1
Women or men at home	95.6 / 4.4	55.3	51 310	13.3	3.2	-76
Other inactive	48.6 / 51.4	45.6	30 435	16.3	8.6	-47

In Table 3, the second column presents population distribution by gender for each profile. The female rates for retired persons and for women or men at home are higher than other profiles. The fourth column shows the average household income. It should be noted that this value for the unemployed people is significant (even if is less than for

other profiles) because of the unemployment benefits funded by local authority. Concerning SAV use, detailed analysis of the population according to the profiles shows bigger changes, especially for retired persons and for women or men at home. For instance, relative difference for the retired or pre-retired population is about two times bigger than for employees. That is because of the higher average age of retired persons as well as the bigger portion of female population in this profile. The SAV use for students older than 14 years and the persons younger than 14 years remains approximately unchanged. Also the changes for employees and unemployed is estimated to be almost the same. That is because of the similar attributes values. As the values of time for these two profiles are not the same, logically the changes should be more significant. The reason for it is related to the main drawback of today's surveys: the measurements for SAV preferences are related more to the classic socio-demographic attributes, while it seems that profiling could be more useful in this case.

4. Discussion and conclusion

In this article we explored the impact of user preferences on SAV modal share by simulating the multimodal transportation system of Paris area applying a multi-agent simulation. The simulation results reveal that neglecting user preferences in classic multi-agent simulations can significantly change the outputs for future scenarios. Also, a detailed analysis of the population according to the profiles shows that demographic structure plays an important role regarding to the use of SAV. Today, despite the considerable number of studies related to exploring the influence of travelers' preferences on the demand of SAV, the investigation about introduction of these preferences in the simulation remains very limited. We started to bridge this gap, but further explorations are still required.

Future work will involve investigations on how to add further attributes to the mode choice mechanism of agents as well as how to integrate them to the activity-related performance measure of multi-agent simulation.

Acknowledgements

This research work has been carried out in the framework of IRT SystemX, Paris-Saclay, France, and therefore granted with public funds within the scope of the French Program "Investissements d'Avenir".

References

- Bansal, P., Kockelman, K.M., Singh, A., 2016. Assessing public opinions of and interest in new vehicle technologies: An Austin perspective. *Transp. Res. Part C Emerg. Technol.* 67, 1–14.
- Charypar, D., Nagel, K., 2005. Generating complete all-day activity plans with genetic algorithms. *Transportation (Amst)*. 32, 369–397.
- Farooq, B., Bierlaire, M., Hurtubia, R., Flötteröd, G., 2013. Simulation based population synthesis. *Transp. Res. Part B Methodol.* 58, 243–263.
- Gkartzonikas, C., Gkritza, K., 2017. A Literature Review on Surveys for Autonomous Vehicles, in: *Transportation Research Forum Annual Conference Chicago, Illinois*. pp. 1–14.
- Haboucha, C.J., Ishaq, R., Shifan, Y., 2017. User preferences regarding autonomous vehicles. *Transp. Res. Part C Emerg. Technol.* 78, 37–49.
- Horni, A., Nagel, K., Axhausen, K.W. (eds.), 2016. *The Multi-Agent Transport Simulation MATSim*. London : Ubiquity Press.
- Krueger, R., Rashidi, T.H., Rose, J.M., 2016. Preferences for shared autonomous vehicles. *Transp. Res. Part C Emerg. Technol.* 69, 343–355.
- Liu, J., Kockelman, K.M., Boesch, P.M., Ciari, F., 2017. Tracking a system of shared autonomous vehicles across the Austin, Texas network using agent-based simulation. *Transportation (Amst)*. 44, 1261–1278.
- Steck, F., Kolarova, V., Bahamonde-Birke, F., Trommer, S., Lenz, B., 2018. How autonomous driving may affect the value of travel time savings for commuting. *Transp. Res. Rec. J. Transp. Res. Board*.
- Stocker, A., Shaheen, S., 2018. Shared Automated Mobility: Early Exploration and Potential Impacts, in: *Road Vehicle Automation 4*. pp. 125–139.
- Vosooghi, R., Puchinger, J., Jankovic, M., Sirin, G., 2017. A critical analysis of travel demand estimation for new one-way carsharing systems, in: *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, pp. 199–205.
- Zheng, H., Son, Y.-J., Chiu, Y.-C., Head, L., Feng, Y., Xo, H., Kim, S., Hickman, M., 2013. *A Primer for Agent-Based Simulation and Modeling in: Transportation Applications*, Federal Highway Administration.