



HAL
open science

Utility-based agent model for intermodal behaviors: A case study for urban toll in Lille

Azise Oumar Diallo, Guillaume Lozenguez, Arnaud Doniec, René Mandiau

► To cite this version:

Azise Oumar Diallo, Guillaume Lozenguez, Arnaud Doniec, René Mandiau. Utility-based agent model for intermodal behaviors: A case study for urban toll in Lille. *Applied Intelligence*, 2025, 55 (4), pp.282. <10.1007/s10489-024-05869-1>. <hal-04934756>

HAL Id: hal-04934756

<https://hal.science/hal-04934756v1>

Submitted on 7 Feb 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Utility-based agent model for intermodal behaviors: A case study for urban toll in Lille *

Azise Oumar Diallo¹, Guillaume Lozenguez^{2*}, Arnaud Doniec² and René Mandiau³

¹IFP Energies nouvelles (IFPEN), Rond-point de l'échangeur de Solaize, Solaize, 69360, France.

^{2*} CERI Systèmes Numériques, Institut Mines-Télécom (IMT) Nord Europe, Villeneuve d'Ascq, France.

³Univ. Polytechnique Hauts-de-France, CNRS, UMR 8201-LAMIH, 59313 Valenciennes, France.

*Corresponding author(s). E-mail(s):

guillaume.lozenguez@imt-nord-europe.fr;

Contributing authors: azise-oumar.diallo@ifpen.fr;

arnaud.doniec@imt-nord-europe.fr; rene.mandiau@uphf.fr;

Abstract

To reduce the congestion and pollution in urban cities, the political authorities encourage the modal shift from private cars in favor of sustainable trip behaviors such as intermodality (through combinations of private car and public transport). Coercive decisions such as urban tolls are also an increasingly investigated solution. To avoid the cost of toll taxes, agents thus select intermodal transportation modes (private cars and public transport) by parking their vehicles in park-and-ride (*PR*) facilities at the entrance to the area toll. This paper proposes a methodology for an agent-based model (*ABM*), particularly a model called *utility-based agent*, to reproduce intermodal trip behaviors in a city and to assess the impact of an urban toll. In this context, we focus on multinomial logit models, coupled with the agent-and-activity simulation tool *MATSim*, is used to determine the modal choice for each agent. Based on open data (for *European Metropolis of Lille, MEL*), the simulation shows that **20 € (21.75 \$)** of toll tax is sufficient to reduce by **20%** the use of private vehicles.

*This paper is an extension of our work published to IEA/AIE'2022 [1].

Keywords: utility-based agent, multinomial logit, transport, intermodal behavior, urban toll, *MATSim*

1 Introduction

One of the political authorities' concerns nowadays is reducing transportation systems' negative externalities (e.g., traffic congestion, accidents, pollution). These actors are increasingly looking at more sustainable transportation systems (e.g., bike and car sharing, carpooling, and intermodality) to offer economical (for the customer) and ecological solutions. Most of the proposed solutions aim to reduce the use of private cars (vehicles) (*PV*) in favor of soft and collective transportation modes – e.g., cycling, walking and public transport (*PT*) [2]. These solutions consist in applying dissuasive measures such as urban road tolls [3–5], low emission zones (LEZ) (also called clean air zones) [6, 7], and encouraging the use of public transport by adjusting the pricing system [8–10]. Urban road toll consists of making access to an area chargeable and aims to reduce traffic congestion in the city, pollution and other traffic externalities. Moreover, to accelerate this modal shift (from *PV* to sustainable transport modes), transport authorities are encouraged to significantly develop intermodality systems through urban transport infrastructures (e.g., park-and-ride facilities, multimodal¹ exchange hubs) to promote the combination of private car and public transport [12].

Such various measures will help to change the travel behavior of individuals towards more environmentally friendly transportation alternatives (e.g., combination of individual modes and *PT*) leading to a new balance of transport flows. Among the forms of intermodal practices, the combination of private car and public transport remains particularly important [13]. This combination is generally done through park-and-ride (*PR*) facilities located near to main *PT* stations/hubs.

In this context, the design and the assessment of intermodal policies *a priori* are not always straightforward. In fact, cities adopting such measures are often faced with these questions:

- What park-and-ride configurations (e.g., capacity, location, pricing) should be put in place?
- What amount of urban toll tax is applied to achieve a given reduction in private car use and over what perimeter?

To answer these questions, city authorities most often use transportation modeling tools based on aggregated approaches, such as four-step models [14]. In these models, transport demand is estimated in four hierarchical steps: generation, distribution, modal choice and traffic assignment. Therefore, demand

¹By definition, the multimodality refers to the **availability** of several possible transportation modes (alternatives) for a given trip, while the intermodality is characterized as a **combination of these possibilities** during the same trip [11].

is represented in the form of Origin/Destination (O/D) matrices depending on the transportation mode and the time. This approach, although allowing to estimate mobility demand in an aggregate manner, remains less suited to evaluating intermodality behavior and the impact of an urban toll policy. In fact, modeling intermodality requires taking into account the entire daily chain of activities and travel of the individual to know which alternative would be more suitable [13, 15, 16]. Furthermore, the application of urban tolls will generate dynamic and heterogeneous behaviors of individuals depending on their characteristics. Traditional models do not really enable us to fully take into consideration the variability of individual behaviors and their changes (for example, when faced with a new transport alternative or mobility policy). Opposite to these models, more dis-aggregated approaches such as agent-based models (*ABM*) [17, 18] seem to be much more adapted to the complex nature of intermodal transport systems and mobility measures assessment [19–22]. *ABM* describes a system as a set of agents (active entities) interacting with each other; where each agent may be characterized by its own behaviors, following decision-making processes. Thus, the global decisions result from their cognitive processes. The agent perceives and acts on its (local) environment, according to its own knowledge (e.g., public transport schedules, knowledge of congested areas) and its own goals (i.e, activities to perform). Their local decisions are based on utilities evaluating the different transport modes. This naturally leads to consider the approach described by the *utility-based agent* [23]. Hence, there is a natural relation between the individuals (from a transportation system) and the agents (from our model).

In this paper, we present a realistic modeling of intermodal behavior based on *ABM* in order to assess the effectiveness of an urban toll policy in reducing private car usage. Therefore, we propose a comprehensive methodology for designing agents with such behaviors based-on discrete choice models. Within this framework, we assume the availability of data for a given context. We aim to explore novel discrete choice models, particularly based on multinomial logit models, to structure various transportation modes. Specifically, we focus on the combination of *PV* and *PT*, including the park-and-ride facilities. Different intermodal alternatives and the introduction of urban toll are evaluated by different utilities; agents can take their own decisions according to their modal preferences (i.e., based on these utilities). To study these behaviors, the selected discrete mode choice model is subsequently implemented in *MATSim* [24, 25], a multi-agent platform. Finally, the obtained framework is applied at the perimeter of the European Metropolis of Lille (*MEL*²) in France.

The article is organized as follows. Section 2 presents a state of the art on *ABM* approaches for intermodal behaviors. Section 3 describes our methodological approach for modeling utility-based intermodal behaviors of agents. Then, Section 4 evaluates the obtained results based on the real data from the

²French acronym for *Métropole Européenne de Lille*.

MEL and validates our approach. Section 5 discusses our study, and finally, Section 6 concludes the paper and gives some perspectives.

2 Background

According to Brad Jones et al., the intermodality is “the shipment of cargo and the movement of people involving more than one mode of transportation during a single, seamless journey” [11]. Thus, the intermodality requires the transition (change) from one transportation mode (e.g., car) to another (e.g., public transport, bus) during the same trip. The practice of intermodality has many benefits. From a macroscopic viewpoint, it is possible to reduce the pollution (air, noise) and the congestion by the use of private cars [2]. From an individual viewpoint, the intermodality would offer practical alternatives for traveling efficiently by combining, for example, walking, cycling or scootering (over medium distances) and the vehicle (long distance) with public transport [26]. However, from a practical viewpoint, the intermodality can often be restrictive or inconvenient for the user, due to passages and waiting times during the transfer from one transport mode to another one. It is therefore essential to think carefully about the intermodal transport system upstream in order to encourage users to adopt intermodal travel behaviors. To achieve this, practitioners and policy-makers most often use traffic simulation tools [27]. These computer tools offer different levels of modeling depending on the degree of granularity required: macroscopic representing transport demand in the form of aggregated flows (O/D matrices) [28, 29] and microscopic considering the demand in a disaggregated way where each individual is defined (simulated) specifically [20, 21, 30]. All these approaches have their origins in the traditional four-step model [14]:

- *Generation* evaluates the total number of trips in the study area for a given temporal period (transport demand).
- *Distribution* finds a matrix based on Origin/Destination (O/D) for the different areas (areas producing different trips (e.g., home), and attractive areas expressing the purposes for these trips (e.g., work, education, shopping and leisure).
- *Modal choice* defines the choice among the different available transport modes to perform the trip.
- *Route assignment* aims to determine the best routes to move the different flows based on the capacity of the road network.

For microscopic approaches (notably agent-based), the first three steps serve as input to carry out the final simulation. These approaches are distinguished by their ability to model travel behaviors as realistically as possible (e.g., choice of transportation mode, choice of route, driving) and traffic management (e.g., intersection, lane capacity). Therefore, these models are more appropriate than the traditional method for studying the intermodality and evaluating mobility policies.

In *ABM* context, the transport demand is typically derived from synthetic populations. A synthetic population generation [31–37] may be defined as a process consisting in generating a realistic representation of an actual population. For example, the following definition has been proposed by [38]: “The creation of a synthetic population representing individuals and possibly households such that the synthetic population is - in all of its relevant statistical properties - representative for a target population within a given geographical and socio-demographic reference frame”. For agent-based transport modeling, the synthetic population represents a collection of agents characterized by some characteristics such as home location, age, gender, socio-professional categories, and ownership of transport materials (e.g., license, car, and public transport ticket or pass). It also includes the chains of activities and trips performed by the agents.

Although there are many simulation tools enabling multimodality [24, 39, 40], we found very few studies on the evaluation of intermodal travel behaviors of individuals [7, 41–43]. The work carried out in [41] consists in analyzing from the multi-agent perspective [24], the impact of new mobility services aiming to promote the intermodality at the whole population. Their data are based on a survey, focused on a sample (made up of 139 individuals) of this population in Switzerland. The intermodal mode choice was designed through a multinomial logit (*MNL*) applying to each possible intermodal stop. In terms of results, this study revealed that a part of the population could be in favor of using the mobility services offered by the project. Moreover, the proposed model being too generic, does not take in consideration the complexity of the exchanges during the modes combination – e.g., the use of park and ride (*PR*) facilities in the combination *PV+PT*, between vehicle (*PV*) and public transport (*PT*).

The first intermodal model considering the process of choosing the park and ride in the *PV+PT* combination was proposed in [42]. A simple *MNL* model was proposed to take into account this alternative. Based on this work, a study on the evaluation of the driving restriction zone measure was carried out in Paris, France [7]. Through the possibility of combining *PV* and *PT* via *PR* facilities (often located outside the restriction zone), the study showed that this combination was particularly interesting for the agents impacted by this constraint. At the same time, another work introduced by [43], focused on the combination of shared micro-mobility service (bike sharing) with public transport as an access/egress mode in the San Francisco Bay Area. It showed that such a combination can also increase the use of public transport.

The previous studies are often based on cities/regions with their own specific characteristics, and in a very empirical way. Thus, our study aims to propose a data-based methodology for building intermodal behaviors for *ABM* in face of an application of a coercive mobility policy directly on the transport mode choice model. Therefore, it extends the approach on the intermodality in two ways, essentially by i) the proposal of various discrete choice models allowing to capture this alternative (behavior) better, and ii) the evaluation

of the impact of the application of a policy restricting the use of private cars (particularly, the urban toll).

3 Methodology and utility-based intermodal behaviors for *ABM*

This section presents our methodology for *ABM* (3.1). This methodology can be defined by different steps: Analyzing the available data for *ABM* (3.2), modeling intermodal behaviors (3.3), selecting the discrete choice models and defining the utilities of the transport modes (3.4), and finally adapting the data based models with an *ABM* platform. (3.5).

3.1 Methodology principle

The different previous studies did not really introduce a methodology for modeling intermodal behaviors. Thus, we propose an approach as generic as possible to be usable in other regions/countries. Figure 1 summarizes our methodology for performing an agent-based intermodality simulation.

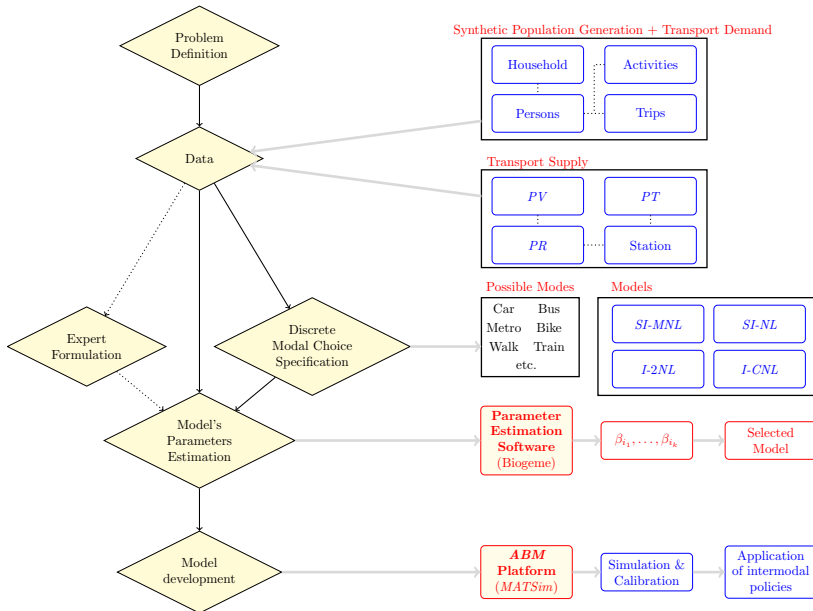


Fig. 1: Methodology for *ABM* transportation simulation

The main modeling steps are thus detailed below. The first step is the problem definition and the different objectives aimed at. The designer can thus study a given problem such as the modification of infrastructures (e.g., positioning of plug-in stations for electric vehicles) or the insertion of a new

transport mode (e.g., the autonomous vehicles in mixed traffic, the intermodality). It is also possible to check their effects for a given locality (e.g., an area of the city, a town, an agglomeration, or a region).

The second step focuses on the obtained data. Thus, we assume that the mobility demand (data on the households, the persons, the activities and the trips) and the transport supply are known for the studied area (for example, *PV* and *PT*). Our approach requires different possible data sources (Cf. 3.2) allowing the generation of synthetic populations.

The third step (the most difficult) focuses on the (possible) consideration of experts formulation, and the specification based on the discrete modal choice. Firstly, the analysis consists in combining the different alternative transportation modes for the description of new intermodal alternatives (Cf. 3.3). Secondly, it aims to select the intermodal alternative based on discrete models (the expert may thus propose and select a particular model among them). In our context, we investigate four discrete models: e.g., *SI-MNL* or *SI-NL*. These models are then processed by a parameter estimation software (e.g., Bio-geme). Parameter values are thus determined for the various utilities required to describe these discrete models (Cf. 3.4).

Finally, the last step consists of the implementation of the discrete models into a *ABM* platform, (*MATSim*, Cf. 3.5). Realistic simulations allow data scientists to reproduce the individual's behaviors. They require an accurate calibration and an analysis of the results for intermodalities policies, according to the target application (our study is focused on the effects of urban toll).

3.2 Available data for *ABM*

As expected, *ABM* approaches require huge amounts of data to allow the generation of a synthetic population. It is usually very difficult to obtain data sources. A pre-processing to exploit consistent data remains often necessary and requires additional works performed by the designer (modeler).

In France, the data sources are anonymized (ensuring the privacy of these data) and are obtained from different databases [44]. In our case study, they come from the population census and household travel (*EMD*, a french acronym for *Enquête Ménage Déplacement*) surveys:

- The population census provides information on individuals' socioeconomic characteristics (e.g., age, socio-professional categories, and household). These data characterize one-third (more exactly, 35%) of the french population. Since it does not concern individuals' travel preferences/habits and mobility behaviors, the census does not provide enriched travel information, but only the main activities (for example, work, leisure and education).
- *EMD* essentially focuses on mobility practices; and it is the unique survey based on a census (around 0.8–2% of the population). The survey considers the detailed reasons (primary and secondary activities), the number of trips, the intermediate trips, and the transportation mode currently used. The data also contain some main users' socio-economic characteristics (e.g., age

and gender), which are often used to establish public mobility policies. It is the only source of data that provides information on intermodal practices.

Traditionally, travel surveys and the population census are the data used for agent-based transport demands [30, 45]. In this goal, we assume that there is a good understanding of the underlying structure of the studied population. Ideally, such information should be collected from real data based on the habits (or the preferences) of individuals. To our knowledge, there is no normalized data for the population or the transport demand. The demand generation should be built by the modeler, based on these available data. Some open-source synthesizers for different case studies are available [46, 47]. Nevertheless, efforts remain necessary to correspond to the studied area.

Regarding the modeling of the transport supply, the main data source is OpenStreetMap (*OSM*)³ [48, 49]. *OSM* data (freely available) provide mapping information (e.g., roads, trails, public transport lines and stations, and buildings). Nonetheless, they give information of uneven quality (the data being more or less detailed for a studied area). To complete the *OSM* data for the creation of transport supply, one can use *BD TOPO* which includes the main axes of the french road infrastructures, as well as secondary roads, paths and surface information (e.g. parking). Finally, public transport (*PT*) data is generally available in the standard *GTFS*⁴ (*General Transit Feed Specification*) format.

In the following, we only use the data sources such as population census survey, *EMD* and *OSM* for the *transport demand* and the *transport supply*; and we show that they are sufficient with regards to our objective. These data sources allow to build intermodal behaviors for *ABM* approaches.

3.3 Modeling the behavior for intermodal alternatives

We propose to describe the main intermodal alternatives. We briefly list these behaviors proposed in our study (3.3.1), and then we focus on the intermodal alternative, between *PV* and *PT* (3.3.2).

3.3.1 Generalities on the intermodal behaviors

Figure 2 gives three simple examples illustrating agents' daily activities chains and the corresponding trips. Agents *A* and *B* are single-modal users: *A* mainly uses *PV*; while *B* is a *PT* user who also walks to perform its activities. Agent *C* has a real intermodal behavior; it is the only one to combine *PV* and *PT* to go shopping.

As above-mentioned, *EMD* is the only data source including intermodal information. The data of this survey are grouped into four files specifically: the *household* file (providing all household characteristics such as motorization, location, and type of housing), *person* file (information on characteristics such

³<https://www.openstreetmap.org/about>

⁴<https://gtfs.org/>

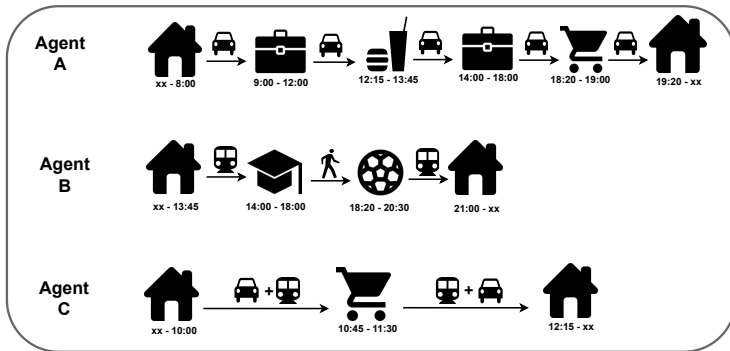


Fig. 2: Examples illustrating agent's trip behaviors according to the daily activities chains

as age, occupation, and driver's license ownership), *trip* and *leg* files (information on the sequence of the transportation mode used, trip motivations, departure time, and distance). It is possible to find the intermodal information based on the two last files.

Note here that we consider, as an intermodal trip, the combination of several mechanized transportation modes (vehicle, public transport, and bicycle), even if some studies also consider the combination of walking (from a certain distance or a walking time) and public transport, as an intermodal practice [7, 50]. Thus, the combination of several public transport (*PT*) modes (e.g., bus and metro during the same trip) is also an intermodal trip [13].

As already underlined previously, we currently focus on intermodal behaviors, combining public transport with other individual transportation modes (*PV* and *Bike*). To rebuild these intermodal trips, we identify all trips by merging these two files *trip* and *leg* (a trip is composed by one or several legs). For example, the agent performs one trip, including two legs (one performed by *PV* and the second by *PT*). Thus, our model considers the following transport modes [44]:

- All public transport modes (bus, subway, tramway, and train) are represented by a single *PT* mode. This aggregation of *PT* modes is needed to enforce this alternative compared with the other transport modes (*PV*, *Walk*, and *Bike*).
- The car mode, as the driver or a passenger, is considered as a private vehicle (*PV*). For example, a driver (or a passenger) taking a private vehicle, and then a public transport will be considered as $PV_driver + PT = PT + PV$ ($PV_passenger + PT = PT + PV$, for a passenger).
- For the intermodal combinations with public transit, we consider all *PT* transfers as unique during the trip. For example, a person taking a private vehicle and then two public transportation modes will be considered as a private vehicle and a public transport ($PV + PT + PT = PT + PV$).

- It is also assumed that we do not consider the intermodal behaviors that are not representative, such as $PV+Bike$ or $PV+Bike+PT$. We also do not consider the combination of PT ($PT+PT$) modes in this study.

Now, we deal with the intermodal behavior combining private vehicles (PV) and public transport (PT). In the following, this alternative combining these two transport modes will be designated by $PV+PT$.

3.3.2 Intermodal alternatives between PV and PT

Private car and public transport combination ($PV+PT$) must be done through a dedicated infrastructure, particularly park-and-ride facilities (PR). To make this possible combination, we propose two new intermodal modes, which are called car_pt and pt_car (respectively for outward and return journeys). Figure 3 illustrates our approach to enable intermodal travel combining a private car (PV) and a public transport (PT). In our illustration, the red line represents the car_pt mode for the outward journey, and the blue line corresponds to pt_car mode for the return.

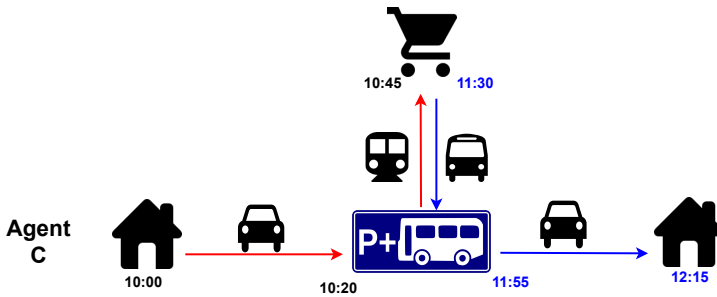


Fig. 3: Illustration of intermodal behavior combining private car and public transport.

The intermodal alternatives car_pt and pt_car are essentially characterized by two modes of transportation, namely private car (PV) and public transport (PT) to describe the two journeys. Actually, we assume that the intermodal trips are made up of two journeys, whose origin is the home. The destination can be any other activity such as going shopping, in the case of the example shown in Figure 3. In addition, we adopt a *tour-based* approach (unlike *trip-based*) to be able to take into account the constraints of modal combination.

The alternative car_pt (shown schematically in red) is used for leaving home with the car as a transportation mode to the parking (first *leg*). The second *leg* is performed by public transport (e.g., bus, metro, tram, train) from PR to the trip's destination, specifically at the nearest station (agent completes its trip by walking). The dual process for the return is similar by using the alternative pt_car . The first *leg* is made by public transport from

the previous destination to the same PR used on the outward journey. Finally, the individuals return home using their car, which remains at PR .

The two transport modes car_pt and pt_car used to represent the intermodal alternatives of private car and public transport ($PV+PT$), have thus the same characteristics.

3.4 Discrete choice models and utilities for considering modal alternatives

Agents' decision-making processes for selecting a transport mode require the specification of models combining transport modes and the definition of their utilities. Firstly, we present a brief description of discrete choice models. Afterwards, we propose discrete choice models based on multinomial logit (MNL) models. After a brief description of the discrete models (3.4.1), we focus on the intermodal alternative based on MNL models (3.4.2) and finally, on the definition of utilities for the transport modes (3.4.3).

3.4.1 Discrete choice model description

The choice of transportation mode is essentially based on discrete choice models [14]. These models take the human dimension into account to describe, explain, and predict choices between two or more alternatives discreet. For a complete description to discrete choice models with simulations, the reader can refer to the work proposed by [51].

Discrete choice models specify the probability of an individual to select an alternative from a given set. This probability is a function of the observed variables related to the alternatives and the individual. In general, discrete choice models are based on the notion of random utility that an individual i has for an alternative (a transport mode, in our case) m [52]. Let $U_{i,m}$ and $U_{i,n}$ be the utilities perceived by i for m and n respectively, i will choose m instead of n if and only if $U_{i,m} > U_{i,n}$. The random utility is a function of explanatory variables observed by the researcher (modeler), composed of the characteristics of the individual (e.g., socio-professional and age) and the attributes of the alternative (e.g., travel time and transport cost) to which we add an error term. Linearly, it is defined as follows:

$$U_{i,m} = \beta \cdot z_{i,m} + \varepsilon_{im} \quad (1)$$

with :

- $U_{i,m}$ defines a function in \mathbb{R} ;
- $z_{i,m}$ is a vector corresponding to the observed explanatory variables;
- β is a vector corresponding to the coefficients (parameter values) of the observed variables;
- $\varepsilon_{i,m}$ is the error term that allows considering the random nature of human behavior, the lack of more precise information in (individual)

decision-making, and the unobservable attributes of the alternatives by the modeler.

3.4.2 Intermodal choice based on *MNL*

To consider new intermodal alternatives in the modal choice, we use multinomial logit models (*MNL*) which are the logit models with more than two alternatives. These *MNL* models are based on one important property, namely the *Independence of Irrelevant Alternatives (IIA)*. These approaches, therefore, assume that the different alternatives are independent (no correlations assumed between alternatives) [51]. This choice is also motivated by the simplicity of their implementation [51, 53].

These models are used to define how agents choose the different transport modes. They need to be structured to provide the best possible description of users' choices. We also study the need (or not) to nest the model and to break up (or not) certain parameters/constants for the modes.

In our study, four discrete models (Figure 4) are proposed for considering modal choices:

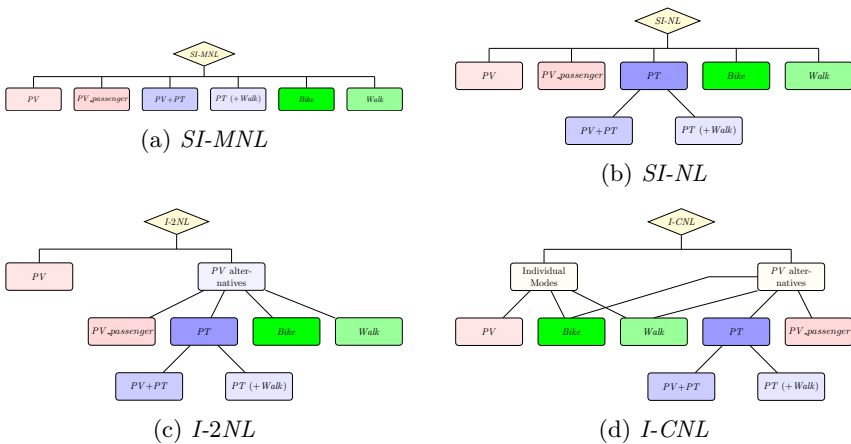


Fig. 4: Our four multinomial logit models (*MNL*) for the transportation modes choices.

Simple Intermodal MNL (SI-MNL)

The *SI-MNL* (Figure 4a) structure was proposed by [42]. The modal alternatives are independent (all the transportation modes are at the same level). This model ensures that *IIA*, is checked; because there is no hierarchy between the different modes. In this case, the probability for Agent i to choose, for example

the private vehicle (*PV*), among the transportation modes j is defined by:

$$P_{i,PV} = \frac{e^{U_{i,PV}}}{\sum_j e^{U_{i,j}}} \quad (2)$$

Simple Intermodal nested logit (SI-NL)

SI-NL (Figure 4b) considers *PT* as the main mode with another mode (*Bike*, *PV*). It is a structure with a nest for considering *PT*. This hierarchical model allows us to take into account changes in *PT* services, which is at the heart of intermodal transport systems. It also assesses the impact of this offer on the main secondary modes.

The nested logit models usually depend on how the nests are built (different nest structures will produce different results). In general, the choice of nest is based on the alternatives that are potentially close or similar to allow correlations between them. The nests are thus assumed to be independent (uncorrelated).

The utility of an alternative (*choice*) in a nest (*Branch*) is defined by: $U(\text{choice}, \text{Branch}) = U(\text{choice}|\text{Branch}) + U(\text{Branch})$, where $U(\text{Branch})$ is a function of variables characteristic of the nest (e.g., price, comfort), and $U(\text{choice}|\text{Branch})$ is a function of certain variables characteristic of the individual (e.g., age) and the modal choice (e.g., travel time). Thus, the probability of choosing an alternative in a nest is determined by the choice (alternative) of this nest (or a branch), and is defined by a conditional probability. The estimation of the choice in a nest requires two models: a model for branch (nest) selection and another one for the choice, given the selection of the branch (nest). For example for our model *SI-NL*, the probability of Agent i choosing the alternative *PV+PT* is defined by:

$$P_{i,PV+PT} = P_{i,(PV+PT)|PT} \times P_{i,PT} \quad (3)$$

with:

$$P_{i,(PV+PT)|PT} = \frac{e^{(U_{i,PV+PT})/\lambda_{PT}}}{v_i}$$

$$P_{i,PT} = \frac{e^{Z_{PT} \cdot \alpha} \cdot v_i^{\lambda_{PT}}}{e^{Z_{PT} \cdot \alpha} \cdot v_i^{\lambda_{PT}} + \sum_{j \in \{PV, PV_passenger, Bike, Walk\}} e^{U_{i,j}}}$$

assuming: $v_i = e^{(U_{i,PV+PT})/\lambda_{PT}} + e^{(U_{i,PT+walk})/\lambda_{PT}}$, the correlation within the nest is equal to $1 - \lambda$ with $\lambda \in]0, 1]$ (no correlation between nesting alternatives, if $\lambda = 1$), Z_{PT} is the vector associated to specific characteristics for the branch (nest) *PT*, and α is a scaling factor assigned to the nest (generally equal to 1).

Although this second model is inspired by [53], we propose the utility functions of the alternatives and the associated calculations (these functions were not initially defined).

Intermodal double nested logit (I-2NL)

I-2NL (Figure 4c) structure is inspired by the recommendations of [54] (the utility functions were not yet defined). The structure consists of 2 nests: a nest for *PT*, and an additional nest focusing on the alternatives to *PV* (designated by *alt_PV*). This second nest considers the relationship between the use of private vehicles with the other modes.

This model looks at the relationship between the use of the private vehicle with other modes, which could be considered as an alternative in this case. The resulting model is more complex than the previous ones. According to [54], the utility of the *alt_PV* nest is better understood by the modeler and is made up of the utilities of modes (*PV_passenger*, *PT*, *Bike*, and *Walk*). To calculate the choice probabilities, we start with the lowest nest. In this way, the utility for the *PT* nest is determined like in the first model. Then, *alt_PV* is estimated similarly, considering the utilities of other transport modes. The probability of an agent *i* choosing *PV+PT* remains unchanged (Cf. Equation 3). Finally, the probability of selecting the branch (nest) *PT* is defined by the choice of *alt_PV*:

$$P_{i,PT} = P_{i,PT|alt_PV} \times P_{i,alt_PV} \quad (4)$$

with:

$$P_{i,PT|alt_PV} = \frac{e^{(U_{i,PT})/\lambda_{alt_PV}}}{v'_i}$$

$$P_{i,alt_PV} = \frac{e^{Z_{alt_PV} \cdot \alpha} \cdot (v'_i)^{\lambda_{alt_PV}}}{e^{Z_{alt_PV} \cdot \alpha} \cdot (v'_i)^{\lambda_{alt_PV}} + e^{U_{i,PV}}}$$

where: $v'_i = \sum_{j \in \{PT, PV_passenger, Bike, Walk\}} e^{(U_{i,j})/\lambda_{alt_PV}}$.

Intermodal cross nested logit (I-CNL)

I-CNL (Figure 4d) with 3 nests, including (i) *PV* alternatives, (ii) individual modes, and (iii) *PT* mode. The aim is to better address (by the designer) the relationships between the different transportation modes (individual or collective modes), as well as the consideration for *PV* alternatives.

The principle for calculating nest probabilities remains similar to other nested models, proceeding step by step to calculate the probability at each hierarchical level. The difficulty here concerns alternatives such as *Bike* and *Walk*, belonging to two nests (individual modes and *PV* alternatives). For the sake of conciseness, we do not detail the probability of choosing these nests.

To conclude, Table 1 summarizes the different proposed models. The two last models (*I-2NL* and *I-CNL*) define a more complex hierarchical structure involving several levels of nesting. Their implementations would require more computing times, without the guarantee of being as much more efficient than the two other models (*SI-MNL* and *SI-NL*).

Table 1: Synthesis of discrete choice models

Model	Taking into account <i>IIA</i>	Number of Nesting	Hierarchical Level
<i>SI-MNL</i>	Yes	0	Simple
<i>SI-NL</i>	Between <i>PT</i> and other modes	1	Medium
<i>I-2NL</i>	All the modes	2	High
<i>I-CNL</i>	All the modes	3	Very high

Following the creation of these models, we can now define the utility functions for each alternative.

3.4.3 Definition of utilities for the transportation modes

The utility functions (called U) of the alternatives should therefore reflect the characteristics of each transportation mode for the individual or the group of individuals. In our context, the utility defines the decisions performed for their travel behaviors (i.e., the choice of transportation modes). The agents select a transportation mode or a combination of modes, maximizing their utility. The modal choice can thus be easily taken into account by the agents. Through an iterative approach, agents can thus learn from their activities (actions) to change their choices.

The utility functions follow the recommendations proposed in [54]. These functions are based on the observed explanatory variables from *EMD*. The utility functions of the private car, public transit, and *PV+PT* are adapted from [42, 44] by adding the processing on toll costs. Here, we focus on the intermodal alternative (the other modes such as *Bike* and *Walk* are not detailed). The utility function (Equation 5) representing the intermodal alternative, is based on *PV* and *PT* functions (Equations 6 and 7, respectively).

$$U_{i,car_pt} = \beta_{ASC,car_pt} + U_{i,PV} - \beta_{inVehicleTime,PV} \times \theta_{parkingSearchPenalty} - \beta_{accessEgressWalkTime} \times \theta_{accessEgressWalkTime} + U_{i,pt} - \beta_{ASC,PV} - \beta_{ASC,PT} \quad (5)$$

$$U_{i,PV} = \beta_{ASC,PV} + \beta_{inVehicleTime,PV} \times x_{inVehicleTime,PV} + \beta_{inVehicleTime,PV} \times \theta_{parkingSearchPenalty} + \beta_{accessEgressWalkTime} \times \theta_{accessEgressWalkTime} + \beta_{cost} \times x_{cost,PV} + \beta_{cost} \times toll_fee \times \begin{cases} 1 & \text{if a part of trip is performed in the toll zone} \\ 0 & \text{else} \end{cases} \quad (6)$$

$$\begin{aligned}
U_{i,PT} = & \beta_{ASC,PT} + \beta_{inVehicleTime,PT} \times x_{inVehicleTime,PT} \\
& + \beta_{accessEgressTime,PT} \times x_{accessEgressTime,PT} \\
& + \beta_{numberTransfers} \times x_{numberTransfers} \\
& + \beta_{transferTime,PT} \times x_{transferTime,PT} + \beta_{cost} \times x_{cost,PT}
\end{aligned} \tag{7}$$

where:

- i corresponds to the individual/agent or group of individuals. Therefore, we could define a utility function according to the specific characteristics of the agent (or group) (e.g., age, gender, and socio-professional category).
- β are parameters (coefficients) of the model to be estimated. They determine an alternative's utility (value) for a given individual during the simulation.
- x are variables corresponding to the known individual's characteristics (e.g., age, gender) and the alternative's attributes (e.g., travel time, cost).
- ASC (*Alternative Specific Constants*) represent constants specific to alternatives, describing the variation of choice not explained exclusively by the known attributes.
- $toll_fee$ is an urban toll tax, and
- θ are calibration parameters to be adjusted manually during the simulation. These parameters make it possible, for example, to consider the constraints of use (search/cost for parking, for example) of the private car, which are not provided in the travel survey data.

These utilities are used in the following to implement agents' decision-making with the two selected multinomial logit models: *SI-MNL* and *SI-NL*.

3.5 Implementation of our intermodal behaviors model in *MATSim*

Combining activity-based and agent-based models effectively reproduces the interactions between the agents during their daily activities [55]. Thus, individuals' decision-making (modal choice) is modeled more explicitly. These approaches make it possible to evaluate mobility policies more realistically. In a previous study, seven most popular *ABM* simulators were compared [19], and it was found that *MATSim* [24] and *SUMO* [39] are particularly useful for examining mobility policies. Specifically, *MATSim* was found to be more suitable in replicating intermodal behaviors through an activity-based approach and the description of utility for the agent's decision-making process. We also use *eqasim*, an adaptation of *MATSim* [25] focusing on the discrete mode choice model.

The work presented in this paper is carried out using the multimodal agent-and-activity-based simulation software *MATSim* (*Multi-Agent Transport Simulation*) [24]. *MATSim* agents are driven by the desire to maximize their *utility*, which are achieved by completing necessary daily activities (such as work, education, leisure, and shopping) while minimizing the unproductive travel time. To reach this compromise, agents choose appropriate travel

modes, routes, and departure times, and draw up a series of travel decisions in each simulation/iteration. Then, they evaluate the resulting daily plan and decide whether to keep it or make marginal changes at the next iteration, using an evolutionary algorithm for plan mutations. Over time, agents optimize their daily mobility plans, taking into account their individual constraints and capabilities, as well as the competitive behavior from other agents.

The process is based on the following steps. Each agent has a daily plan of trips and activities. The initial plans are then represented by an *initial demand* related to the transport supply (e.g., the networks for the road network and *PT*). After an evaluation of plans by a score based on the performed activities and the utility provided by the transport mode. At each iteration, a subset of agents can modify their plans to improve their score by changing either the transport mode, the route or the departure time (re-planning), and thus have the possibility to optimize their daily mobility plans.

In *eqasim*, the decision-making process, and particularly the modal choice is replaced by a discrete choice module. Thus, the discrete choice module chooses the mode(s) according to the approach used (*trip-based*, i.e., a single-mode for each trip or *tour-based*, i.e., the chain of modes for the round trip). This choice is also performed by considering some constraints to avoid generating unrealistic combinations, such as using the car without a driving license. As already underlined by [56], the parameters resulting from utility functions of activities and trips are considered as a real challenge to calculate during the calibration process.

In the *MATSim* current version, the routing for car trips is based on shortest path algorithms such as Dijkstra; whereas public transport routing is based on the dynamic programming connecting stations/stops and lines. Note also that the intermodal routing combining *PT* and *Walk* is already predefined in *MATSim*. For the alternative *PV+PT*, a new specific routing algorithm was defined. The *PV+PT* trip routing algorithm is defined by four steps:

1. Find the *PR* closest to the agent's home.
2. Calculate a route between home and *PR* using shortest path algorithms included in *MATSim*.
3. Perform routing between *PR* and final destination using the public transport routing module of *MATSim*.
4. The return to home follows the previous three steps in the opposite direction.

Thus, this algorithm ensures that the same *PR* is used for both private car and public transport during the trip.

4 Our case study based on *MEL* data

This section provides general information on the study area, the European Metropolis of Lille (France), *MEL* (4.1), a few details on the setup and parameters proposed for the experiments (4.2), and the results obtained during the experiments (4.3).

4.1 Description of the *MEL*

MEL comprises 95 municipalities with 1.1 million inhabitants over an area of 672 km^2 . Figure 5 provides an overview of the spatial distribution of populations⁵ and the car ownership (at least one vehicle) per household⁶ for the municipalities of the *MEL*. The population is gathered around the main metropolis (Lille); while the rest is weakly populated. However, the opposite phenomenon is observed concerning the number of cars per household. The number of vehicles per household is higher in these peripheral areas, whereas around 70% of households in the center own a private vehicle. The public transport network includes two metro lines (43.6 km), two tram lines (22 km), and around 90 bus lines. There are currently 12 *PR* facilities (approximately 5,000 parking spaces).

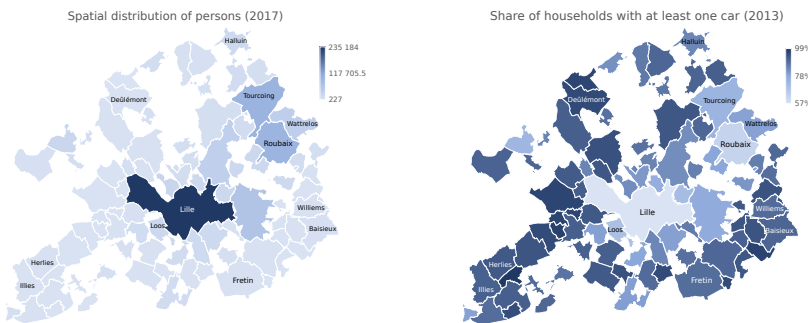


Fig. 5: Spatial distribution of the population (left) and car ownership per household (right) in the municipalities of *MEL*

We note a good heterogeneity within the population, with as many women (52%) as men (48%). The distribution of individuals by age is fairly uniform with a large percentage of young people, 20% are under 14 years old, 43% between 15 and 44 years old, 18% are between 45 and 59 years old, and 19% are over 60 years old. On the working side, 79% of individuals have an activity, including 59% of workers. More than 70% of the professionals work outside their municipality of domicile, leading to a high use of private cars (68%) compared to other transportation modes. It is also noted that almost half of

⁵<https://www.insee.fr/fr/statistiques/3698339#consulter>

⁶<https://www.data.gouv.fr/fr/datasets/taux-de-motorisation-des-menages/>

the people are probably concerned by a parking problem. Its crucial use for the population, involving various problems (noise and air pollution for example), makes our case study particularly interesting, to investigate mobility policies aimed at reducing its usage.

MEL household-travel survey (*EMD*), carried out in 2016, was used in this study. Globally, 9,479 persons belonging to 4,539 households (approximately 0.87% of the actual population) were interviewed about their trips performed the day before. 24,629 were listed in this survey. More than half of all trips (all purposes combined) are performed using a private car (57.9%), walking comes second with nearly 30%, followed by public transport (*PT*) with 6.7% and the bicycle which has less than 2% of modal share. Only 4% of trips are intermodal (a combination of two more mechanized modes of transportation). A public version of this survey is available on the open data site ⁷.

4.2 Experimental setup and parameters

To consider the experiments, it is necessary to calibrate and to determine the setup (4.2.1) and parameters of the discrete choice models (4.2.2).

4.2.1 Experimental setup

The synthetic population generated in this study represents a sample of 20% of the actual population, which corresponds to 225,240 agents with a total of 746,318 trips. This sample is based on recent works about population down-scaling [57] for multi-agent transport simulations. This choice is mainly justified for reasons of computational cost and for obtaining exploitable simulation times. For example, the computation time with such a population for a simulation of 300 iterations on Intel Xeon CPU E5-2650 v4 2.2 GHz and 128 GB of RAM takes about 18 hours. The flow and storage capacities of the road infrastructure are multiplied by 0.18 to consider the down-sampling of traffic flow on the existing road network. The number of iterations is set to 100, and 20% of agents are allowed for re-planning (change of transportation mode).

In the following, we assume a toll area, which leads the individuals to use public transport to enter the city. Figure 6 provides information on the area and the geographical location where the urban toll is applied. The application area (hatched in black line) is thus defined on the perimeter of the city of Lille, avoiding to include the nearby park-and-ride facilities (see the symbol on the figure). The urban toll application configuration is defined as follows:

- *Application time*: We apply the measure for the whole day.
- *Price*: We propose a variation of different amounts between 2 and 20 euros to significantly reduce the modal share of vehicles (10 – 30%).
- *Who*: the toll applies to all agents' trips performed by a private car in the toll area.

⁷<https://opendata.lillemetropole.fr/explore/dataset/enquete-deplacement-2016/information/?location=10,50.65641,3.03338&basemap=jawg.streets>

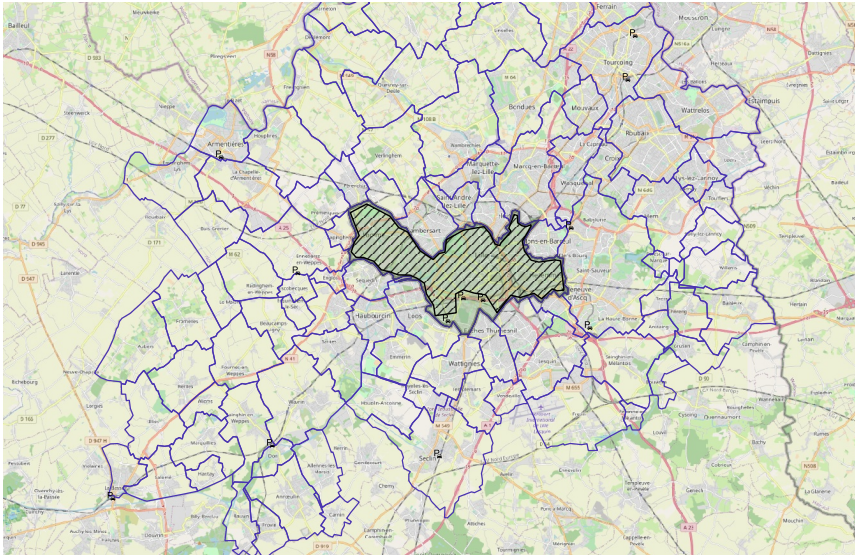


Fig. 6: Urban toll area (hatched in black line) in Lille of the *MEL*

- *intermodality*: We propose to study two typical cases, with or without the *PV+PT* alternative to evaluate urban toll area (through the use of *PRs*).

4.2.2 Parameters estimation

The calibration step is adjusted for the different parameters of our discrete choice models. The obtained results must be closed to the empirical (reference) data (*EMD* 2016), in terms of modal alternatives and distance distribution. In particular, we fit the parameters for *PV+PT* alternatives, as they are not estimated in the discrete models. The calibration is also necessary, because the real data on the use of park-and-ride facilities are lacking. Finally, we set the capacity of the road infrastructure for different values to improve the simulation results. The model is estimated with *PandasBiogeme* [58] (a version of the open source package *Biogeme*, commonly used in transportation modes) for estimating discrete choice models.

Table 2 summarizes the main characteristics of two models *SI-MNL* and *SI-NL* (a prior description with a different sample size for *SI-MNL* model was given in [1, 42, 44]). The total number of observations used for the estimation is 4,771 (a third of the data from the *MEL*) taking into account a good representation for each transportation mode. *SI-NL* as one additional parameter to estimate than the first one (*SI-MNL*), corresponding to the nest of this logit model (to take into account the combinations between *PV* and *PT*).

The convergence is the first indicator for evaluating discrete choice models. A model converges when the values of different parameters do not change or vary slightly before the end of the iterations defined by the selected software. All the parameters of these two models could be estimated, confirming that

the utility functions are well defined. Thus, the two models converge after 78 iterations.

Table 2: Characteristics of the models *SI-MNL* and *SI-NL* (based on *Pandas-Biogeme*) - BFGS algorithm with trust region for simple bound constraints, applied on 4771 samples (Excluded observations: 343)

Statistical Data	<i>SI-MNL</i>	<i>SI-NL</i>
<i>Number of estimated parameters</i>	13	14
<i>Init log likelihood</i>	-6700.209	-6700.209
<i>Final log likelihood</i>	-4138.308	-4138.307
<i>Likelihood ratio test for the init. model</i>	5123.803	5123.804
<i>Rho-square for the init. model</i>	0.382	0.382
<i>Rho-square-bar for the init. model</i>	0.38	0.38
<i>Akaike Information Criterion</i>	8302.615	8304.614
<i>Bayesian Information Criterion</i>	8386.729	8395.198
<i>Final gradient norm</i>	2.2345e - 02	6.7685
<i>Bootstrapping time</i>	0:11:01.052890	1:07:19.546332
<i>Number of threads</i>	4	4
<i>Proportion analytical hessian</i>	0.0%	0.0%
<i>Relative projected gradient</i>	2.247521e - 6	3.161913e - 06
<i>Number of iterations</i>	78	78
<i>Number of function evaluations</i>	221	223
<i>Number of gradient evaluations</i>	72	73
<i>Number of hessian evaluations</i>	0.0	0.0
<i>Cause of termination</i>	2.2e ⁻⁶ <= 6.1e ⁻⁶	4.2e ⁻⁶ <= 6.1e ⁻⁶
<i>Optimization time</i>	0:00:01.795740	0:00:03.016882

The estimated models (*SI-MNL* and *SI-NL*) are statistically significant (and are nearly similar). The decrease of the final value of the *Final log Likelihood* (-4,138.308 for *SI-MNL* and -4,138.307 for *SI-NL*) compared to their initial state (-6,700.209 for *SI-MNL* and *SI-NL*) shows the quality of the obtained values for the parameters. The improvement in the *log Likelihood* values for these two models shows that the explanatory variables introduced are significant. In addition, estimated models are also confirmed by the rho square equals to 0.382. This value shows an acceptable predictive quality for the two models.

Our objective for obtaining significant parameters is crucial to better represent the real modal choice of individuals according to their characteristics and the attributes of transport supply. Table 3 shows the estimated parameters of both models. To evaluate their values, we also consider their sign (negative or positive) and the standard error. The coefficients of the explanatory variables (β and θ) determine the utility value for each alternative. For example, the β_{cost} is intuitively negative, thus reflecting a certain disutility towards transport modes, whose a higher cost for the use of their vehicle.

There are deviations from the error values for some parameters, in particular the constants for the alternatives ($\beta_{ASC,PV}$, $\beta_{ASC,PT}$, $\beta_{ASC,Bike}$, $\beta_{ASC,Walk}$). *SI-NL* thus presents higher standard errors (*Rob. Std err.*) for

Table 3: Determination of estimated parameters for the two models, *SI-MNL* and *SI-NL* (based on *PandasBiogeme*)

	Parameters	<i>SI-MNL</i>		<i>SI-NL</i>	
		Value	Rob. Std err		Rob. Std err
<i>PV</i>	$\beta_{ASC,PV}$	-2.69	0.272	-2.69	2.05
	$\beta_{inVehicleTime,PV}$	0.00133[min^{-1}]	0.00918	0.00134[min^{-1}]	0.0101
<i>PT</i>	$\beta_{ASC,PT}$	-3.83	0.272	-3.84	2.1
	$\beta_{inVehicleTime,PT}$	-0.000842[min^{-1}]	0.00158	-0.000841[min^{-1}]	0.00164
	$\beta_{accessEgressTime,PT}$	0.0228[min^{-1}]	0.0511	0.0227[min^{-1}]	0.0511
	$\beta_{numberTransfers}$	-0.19	2.25	-0.205	2.25
	$\beta_{transferTime,PT}$	-0.907[min^{-1}]	0.412	-0.936[min^{-1}]	0.412
<i>Bike</i>	$\beta_{ASC,Bike}$	-3.07	0.312	-3.07	2.07
	$\beta_{travelTime,Bike}$	-0.108[min^{-1}]	0.0157	-0.108[min^{-1}]	0.0145
	$\beta_{highAge,Bike}$	0.00979	0.0049	0.00979	0.0049
<i>Walk</i>	$\beta_{ASC,Walk}$	0.218	0.274	0.217	2.05
	$\beta_{travelTime,Walk}$	-0.131[min^{-1}]	0.00639	-0.131[min^{-1}]	0.0068
<i>Other</i>	β_{cost}	-0.0142[$euro^{-1}$]	0.0459	-0.0142[$euro^{-1}$]	0.046
	$\beta_{Nest,PT}$	-	-	1	1.61
	$\theta_{parkingSearchPenalty}$	4	-	4	-
	$\theta_{accessEgressWalkTime}$	4	-	4	-
	θ_{ASC,car_pt}	1.25	-	1.25	-

The calibration parameters (the three last rows of this table) are not among those to be estimated by the model.

these parameters. These differences in values are mainly explained by the nature of *SI-NL* model, which tries to take into account the relationships between alternatives, while the *SI-MNL* assumes that there are no relationships. Thus, *SI-NL* model would be less accurate in estimating the values of these parameters compared to the *SI-MNL*. It should be noted at this level that the estimation time (*Bootstrapping time*) of *SI-NL* parameters (more than one hour of calculation) remains important compared to *SI-MNL* (11 minutes). Generally speaking, the standard errors are low, indicating a good capacity to estimate the parameters' values.

The analysis of the signs of the parameters' values indicates a conformity of the models to represent the behavior of individuals in their modal choice. For example, the cost parameter has a negative effect on the choice process, which is rather intuitive. The higher the cost of an alternative, the less interest the individual will have in this alternative. Moreover, the parameters related to the travel time of the alternatives, except for the car, have a negative sign. This situation also expresses some reality, because the more time it takes to travel by a transportation mode, the less likely it is to be used. The private car remains a characteristic with regard to the comfort it would provide in terms of the time spent on the trip. The number of transfers between public transport systems is also seen as a limiting factor for the use of this mode, while access and egress times ($\beta_{accessEgressTime.pt}$) are not perceived as very constraining.

It should be noted here that the calibration parameters ($\theta_{parkingSearchPenalty}$ ⁸, $\theta_{accessEgressWalkTime}$ ⁹, and θ_{car_pt}) are manually estimated to adjust the simulation output. The values thus presented (except for the constants of *car_pt* mode) have been adapted from a previous study [25].

SI-NL with a *PT* nest aims at studying the correlation between the alternatives, *PV* and *PT*. Regarding the value of the nest coefficient ($\beta_{Nest,PT} = 1$), it turns out that these two alternatives are not really correlated. Therefore, *SI-MNL* is well-suited to assess the choice of transportation mode, taking into account the intermodal alternatives. More specifically, *SI-MNL* gives estimated parameters statistically significant, except those related to the duration of time spent for *PV* ($\beta_{inVehicleTime,PV}$) and *PT* ($\beta_{inVehicleTime,PT}$) and the time to reach *PT* ($\beta_{accessEgressTime,PT}$, $\beta_{numberTransfers,PT}$ and $\beta_{transferTime,PT}$, with a p-value greater than 0.5).

To conclude, the two models based on these parameters can thus be used in the simulation to infer the agents' behaviors in terms of modal choice. These two models could be defined for experimental evaluations; but, *SI-NL* requires more computational times than *SI-MNL* for equivalent results. That is why, we consider in the following, the parameters of *SI-MNL* for our experimentations.

4.3 Experimental Results

The experiments aim to evaluate our proposal concerning the previous parameters for the utilities of different transportation modes (including the intermodal alternatives, and in particular *PV+PT*) following *SI-MNL*. We also implement it within *MATSim/eqasim* framework. We show the results concerning an urban toll (4.3.1) and we propose to apply our approach for a larger region (4.3.2).

4.3.1 Impact of an area toll for the agents' behaviors

Each agent estimates the utilities of the different transportation modes and selects its behavior according to its preferences. Once again, it should be remembered that we consider the various information concerning existing infrastructures (including roads, park-and-ride facilities, etc.) and real user behavior. Given the demand from different users characterized by agents and the existing infrastructure, we can assess the benefits of proposing a toll area. In addition, the cost of such a pricing system also needs to be estimated to significantly reduce the use of *PV* (fixed to 20%). Finally, we would also like to study the effects of the intermodal alternative *PV+PT* for this problem.

In this context, we perform different experiments by changing the ad-hoc value of the toll tax, and we observe the agents' global behavior. Thus, the

⁸ $\theta_{parkingSearchPenalty}$ allows to take into account the effects of the search for a parking space for the vehicle.

⁹ $\theta_{accessEgressWalkTime}$ captures the influence of walking time to go to the car and to the final destination.

idea is to check that the aggregation of individual behaviors leads to a consistent overall behavior. Therefore, two situations need to be investigated for the various agents to decide, depending on the estimated utilities. The selection of a transportation mode is performed according to their activities and their estimated utilities: (i) without the consideration the intermodal behavior $PV+PT$ (case 1), (ii) with the consideration of the intermodal alternative $PV+PT$ (case 2). These two cases present the results obtained for the different transportation modes (private *car*, public transport *pt*, bike, walk, other passengers than the driver in a vehicle *car_passenger*, and the intermodal alternative *car_pt*).

Figures 7, 8, 10, and 11 illustrate the global modal shares with/without a toll area for the two cases (without/with the intermodal alternative) in *MEL* and within the main town (Lille) where the toll is applied. We give the results related to reference data (from *EMD* 2016), the simulation without toll area (simulation base case), and the simulations obtained with toll area for different toll taxes.

Thus, we can see the effect of the urban toll on the decrease of private car use. Compared to the reference (base) situation, this reduction is 2%, 5%, 8%, 11%, 14%, 16%, 18%, and 19%, respectively, for taxes of 2.5, 5, 7.5, 10, 12.5, 15, 17.5 and 20 euros without considering the $PV+PT$ alternative on the scale of the entire metropolis (Figure 7). The reduction in car use is slightly greater when considering the intermodal alternative with 5%, 9%, 12%, 15%, 18%, 19%, 20%, and 21% according to the toll tax (Figure 8).

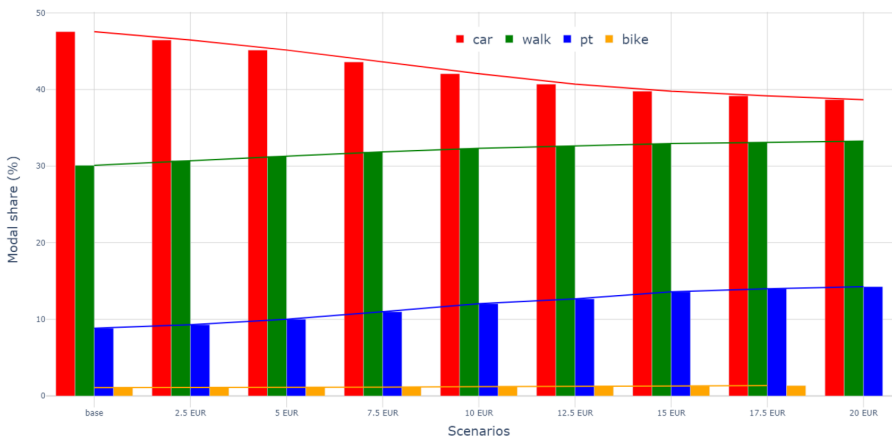


Fig. 7: Modal shares without the intermodal alternative within the *MEL*.

This exciting result shows that combining private car and public transport makes individuals more likely to leave their cars and to switch to the intermodal alternative. Without this alternative, the users prefer to pay the

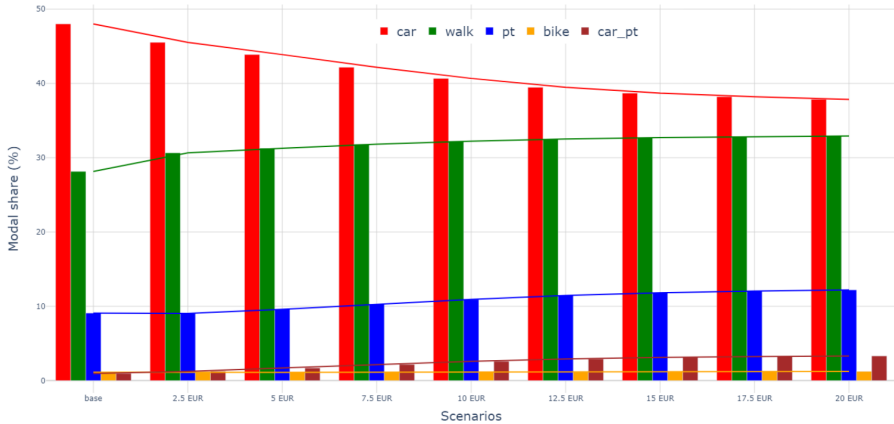


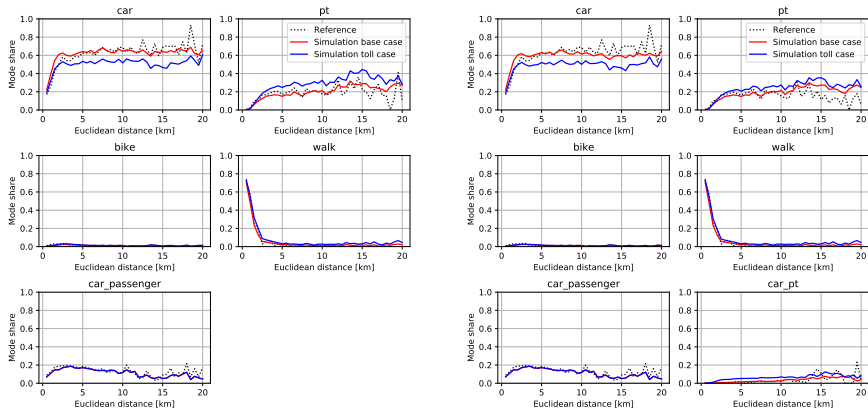
Fig. 8: Modal shares With considering the intermodal alternative $PV+PT$ within the MEL .

toll tax and continue to use their own vehicle. Another interesting observation concerns the reduction in private car use, which stabilizes at a fixed rate regardless of toll increases. In fact, from 15 euros already, the reduction is not really significant.

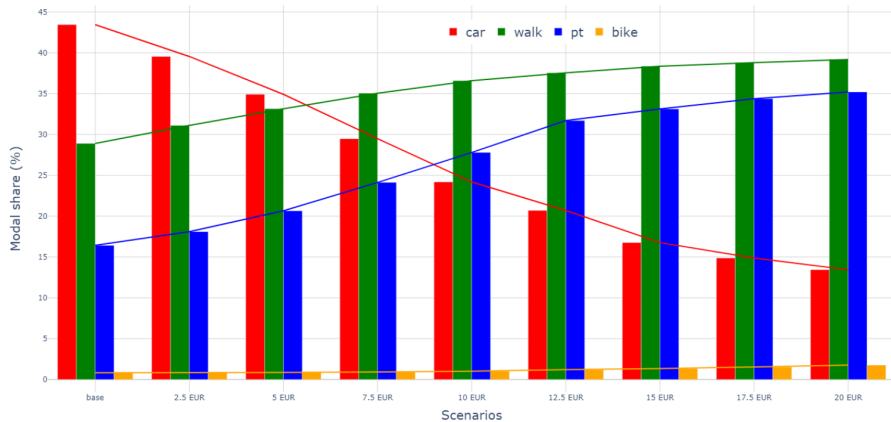
We find that a tax of 20 euros should be applied for reducing significantly (fixed to 20%) the use of PV within the MEL . Considering this scenario and without the $PV+PT$ alternative, the reduction in private car use leads to an increase of 62%, 24%, and 10% for public transport, cycling and walking, respectively (Figure 7). For the $PV+PT$ alternative scenario, the increase in PT , cycling, and walking is 10%, 16%, and 20%, respectively. This reduction results in an exponential increase of 244% in the modal share of the intermodal alternative (Figure 8). Therefore, this trip opportunity is more useful than PT , for the users.

Figure 9 provide additional information on the changing for the transport modes, by considering the travel distance at the MEL -scale simulation with 20 euros for toll taxes. The reference situation (EMD 2016) is in black dotted line, the simulation without road pricing in red color line and the toll application scenario in blue line. Thus, users who abandon their cars for cycling and walking are concerned about short-distance trips (less than 3 km). Meanwhile, the intermodal alternative and public transport are preferred for medium (3–10 km) and long-distance (more than 10 km) trips (Figure 9b). Public transport remains the privileged transport mode for medium and long distances without this alternative (Figure 9a). The modal shift over short distances (less than 3 km) is not globally visible. A more detailed analysis on the scale of the city of Lille (toll area) shows a better understanding of this behavior performed by car within the town.

The various changes due to the application of urban tolls are more striking in the perimeter of Lille (Figures 10, 11 and 12). We note a drop of 69% and



(a) Case 1: Without the intermodal alternative

(b) Case 2: With the intermodal alternative $PV+PT$ **Fig. 9:** Modal shares with/without a toll area (20 euros tax) within the MEL.**Fig. 10:** Modal shares without the intermodal alternative within Lille.

72% for the use of private cars in favor to other alternative, when we applied a tax of 20 euros in the cases 1 and 2, respectively (Figures 10 and 11). Compared to the base situation, the modal share of PT , $Walk$, and $Bike$ increase of 111%, 43%, and 118%, respectively in case without $PV+PT$ alternative (Figures 10). When considering the $PV+PT$ alternative, the modal share of PT , $Walk$, $Bike$, and $PV+PT$ increase of 69%, 40%, 60%, and 514%, respectively (Figures 11). The choice of the alternative to the private car concerning the travel distance, in the case of the application of urban tolls, is more perceptible on a city scale. The choice of walking and cycling for short distances is more outstanding. The intermodal alternative ($PV+PT$) remains the more attractive transport mode

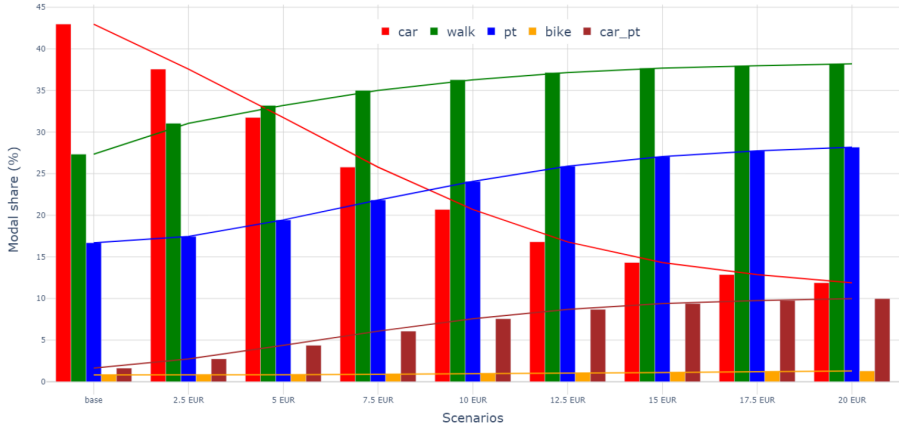
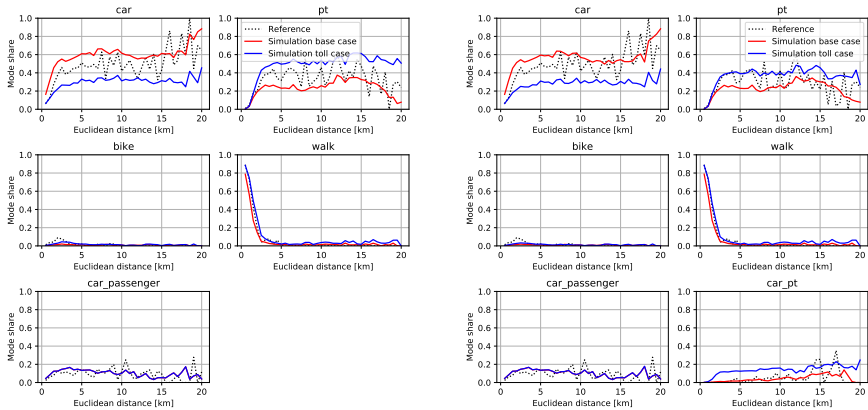


Fig. 11: Modal shares With considering the intermodal alternative $PV+PT$ within Lille.

for the agents (users) in the case of urban toll applications. This combination allows agents moving to Lille to park their cars in the PR facilities, and then to take a public transport to enter the city. Thus, they avoid paying the toll tax.



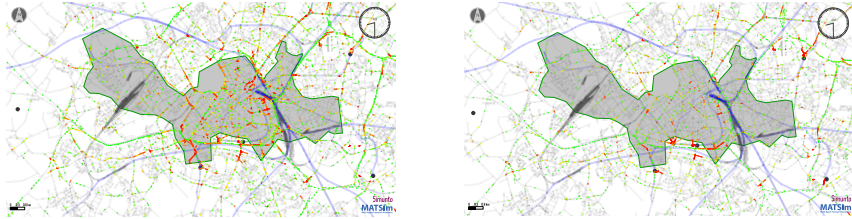
(a) Case 1: Without the intermodal alternative

(b) Case 2: With the intermodal alternative $PV+PT$

Fig. 12: Modal shares with/without a toll area (20 euros tax within Lille).

This phenomenon is visible through the flow of cars at the PR near the city's entrance. Figure 13 gives an overview of this situation from 8:30 am. The vehicles in red color are in a congestion situation. They move very slowly;

those in yellow also circulate slower than when the traffic is fluid (the cars in green). We can see clearly the impact of road pricing on the flow of cars within the city of Lille (Figure 13b). The number of vehicles has dropped considerably compared to the situation without an urban toll (Figure 13a).



(a) Traffic situation without urban toll.

(b) Traffic situation with urban toll.

Fig. 13: Vehicles' flow at *PR* located at the entrance to Lille

4.3.2 Extending the model for a larger region

Based on the estimated parameters of the *MEL*, we extend the framework to a large region (*Hauts-de-France*). In fact, an important part of the daily trips inside the area of the *MEL* involve an origin or a destination outside of this area, i.e. the region. This region includes 5 departments (Aisne, Nord, Oise, Pas-de-Calais, Somme) grouping 3,789 municipalities, with approximately 6 millions inhabitants¹⁰. According to the travel survey from this region, more than 14 million trips are recorded every year, including around 9 million to Lille. The car (be driver or passenger) remains the main transportation mode with 70% of modal shares, followed by the walking (22.5%), the public transport (5%), the bike (2%) and the intermodal combination (1%).

Due to the constraints of computational times, the simulation is based on a restriction of synthetic population, fixed to 10% of the actual population of this region. As for the previous simulations on *MEL*, the urban toll is applied at the entrance to Lille throughout the day, and a tax of 20 euros (hypothetical) is applied for each entry into the toll area.

Figure 14 shows the results of the *PV+PT* alternative and the impact of the congestion charge to Lille. Figure 14a gives an idea of the size of this region compared to the *MEL* (green color) and to the area of application of the urban toll within Lille (yellow color). Figure 14b presents the obtained results for the modal shares. The dotted curve corresponds to the reference situation, and the one in red color represents the modal shares of the calibrated simulation (before the application of the urban toll). Finally, the curve in blue color shows the new shares following the application of the congestion charge. This figure shows that our model can simulate by extension a larger region

¹⁰<https://www.prefectures-regions.gouv.fr/hauts-de-france/Region-et-institutions/Portrait-de-la-region/Chiffres-cles/Chiffres-cles-de-la-region-Hauts-de-France>

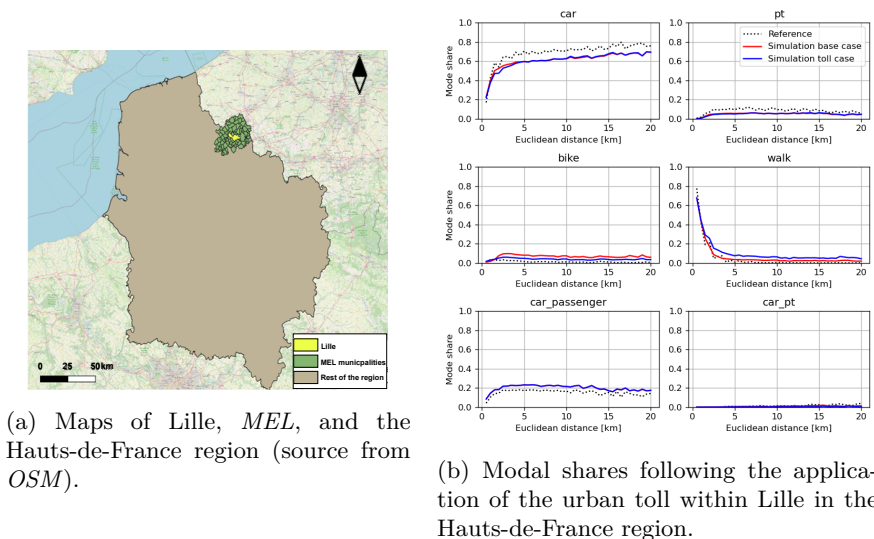


Fig. 14: Simulation results for a larger region, Hauts-de-France

for the different transport modes (including the intermodal alternative). As expected, the urban toll application influences the agents' travel behaviors. As in the case of the *MEL*, this change is not really visible on a larger scale. On the other hand, we note a similar reduction in car use in the toll area.

After selecting an efficient model (*SI-MNL*) for our purpose and determining its parameters, we propose to study the effectiveness of a tax to enter the city center. We show that the tax should reach 20 euros to significantly change the agents' intermodal behaviors (around 20% and 70% in the whole study zone and the toll area, respectively). Finally, we also briefly describe a preliminary work that allows us to extend (with the same parameters) our model for a larger region. These first results should be considered as proof of concept. In fact, our objective is to show that modeling agents with activities and utilities is a relevant proposal for addressing mobility (or intermodality) policies at different scales. A coordinated policy should be applied to reduce the use of private cars significantly. The framework thus proposed can, therefore, support a basis for studying the behaviors of users; for example, if a tax is applied to a toll area, then we could characterize the effects on the categories of individuals or on the types of vehicles (diesel, electric, hybrid). The following section deals with a discussion of our study.

5 Discussion

The simulations based on our model require different consistent data sources. We debate the introduction of other data sources, for example, based on sensors (5.1). Then, the experiments based on an urban toll are also discussed (5.2).

5.1 Other data sources

Our study is essentially based on data sources obtained from *EMD* and population census. Although these data are rich, they have the disadvantage of representing only a small population sample. On the other hand, it is possible to collect data from sensors positioned on the infrastructure (radar) or circulating on the infrastructure (e.g., connected vehicles, smartphones) [44, 59, 60].

Thereby, more and more authorities organizing mobility are deploying sensors to control user access. Wi-Fi data are used, for example, to manage the flow of public transport users [61]. The volume of captured data provides crucial information for sizing the transport offer according to this demand (for example, parking facilities for public transport, by validating tickets and cards or capturing signals).

Personal mobility data can be even more massive, such as FCD (Floating Car Data) or FMD (Floating Mobile Data) [62]. They are collected from cars or phones [63] that send their positions (obtained from a GPS location or GSM triangulation) to third parties (guidance service operators such as Waze or Google Maps, mobile phone operators, car manufacturers) at regular intervals.

In theory, these floating data can provide better knowledge of the traffic flows for different transportation modes (FCD for the vehicles and FMD for the other modes) and better describe travels than the data from standard surveys. Indeed, the information describing users' travel routes is more recent and accurate due to the frequently updated location system.

However, the main limitations for their use are firstly, the lack of contextual information (e.g., used transportation modes, individual characteristics) and secondly, access to raw data from providers. (e.g., mobile phone operators, Google) [64]. These data could also be useful for particular contexts: for instance, the study of parking facilities in an urban context [44], or a better understanding of the individuals' trips within the toll area.

5.2 Study based on a toll area

There are mainly two ways to implement urban tolls: by cordon (it consists of creating a cordon around a given zone and applying the toll to each finite perimeter entry and/or exit) and by area (the toll is used as soon as a car drives in the area concerned; the difference lies in the payment, performed only once to circulate in the area and not at each passage of the cordon). The road toll system studied in this work relates to an application by area on the city of Lille and allows to reduce road traffic inside the town with a toll tax

of 20 euros. More broadly, in the literature, different road pricing systems are depending on the desired objective [65]:

- by distance: this type of toll consists of charging per distance traveled by car. For example, for a 1 euro/km price, a driver who will do 25 km will pay 25 euros.
- by road section: in this case, it is a question of applying a particular charge each time an area (regardless of its length) is crossed (e.g., tunnel and bridge).

A similar method was applied to the city of Zurich (Switzerland) [66]. Unlike our approach based on the toll area, this study focused on using road pricing through a cordon around the city. In addition, the toll is only applied during peak hours (5:30–9:00 and 17:30–19:00) to enter and exit the city. The authors carried out several simulations with different tax amounts of less than 5 euros to obtain a 20% reduction in automobile traffic in the city.

The difference in the tax value may be due to the infrastructures of the city studied, but it also highlights the difficulty to model behaviors for artificially simulated agents. From our point of view, a daily tax of 20 euros seems to be expensive. Data-based model implies a certain inertia in the behaviors. It does not permit huge changes at an individual scale (relocation, equipment purchase, etc.).

Furthermore, our study considers a toll area, which leads individuals to use public transport to enter the city. However, the cordon toll would be more suitable if, for example, one seeks to avoid crossing an area by car to guide them to the roads around the city. The drawback of the toll area compared to the cordon is that the flow of cars crossing the area is not considered. However, the main objective of the toll is to reduce or limit these types of trips in the city by making bypass roads.

Another point not dealt with in this paper is the acceptability of such measures by individuals depending in particular on their income [10, 67]. For example, we have seen the *gilets jaunes* crisis, following the increase in fuel taxes in France (2018). Therefore, it is essential to strike a balance between the application of these measures and the categories of individuals.

6 Conclusion and Perspectives

The main objective of political decisions in the transportation systems is reducing the negative effects such as traffic congestion or pollution. The proposed solutions consist in reducing the use of private car, and in proposing collective modes (public transport). Thus the authorities develop new intermodal modes (combining for example, private vehicles and public transport) and/or urban toll areas. However, these decisions may change the users' behaviors. It is, therefore, necessary to propose adapted approaches to reproduce and/or to evaluate such decisions.

Existing studies provide different models for addressing this need, ranging from macroscopic to microscopic approaches. Macroscopic models are based on aggregated data. They have the advantage of assessing situations where all users exhibit average behavior. However, their dual disadvantage is that it is difficult to design more specific behaviors. Opposite to these traditional models, the agent-based approach (*ABM*) is particularly suitable for describing more precisely these individual activities. Each agent can assess all transportation modes according to its preferences and activities. Estimating the different possibilities thus characterizes the *ABM*, which we refer to utility-based agents. The problem then lies in generating a realistic population of agents for a given application. We also highlighted the fact that only some studies have been performed on intermodality.

In this paper, we proposed a general *ABM* approach of modeling individuals' travel behaviors, in the context of modal choice (including the intermodal alternative) and the application of urban tolls. The agents will be able to estimate their utilities associated with different available modes, according to their preferences and activities, and select the corresponding transportation mode. Among the different intermodal possibilities, we focused on an intermodal alternative the most used, namely the combination between private car and public transport.

We described a methodology consisting of different steps: (i) the objectives to be achieved, (ii) the available data needed to generate a realistic population, (iii) the specification (possibly with the help of experts) for the description of different transport modes, including the intermodal alternative, (iv) the presentation of utilities and selected discrete models, and (v) the choice of a multi-agent platform to support our approach.

The first step was to clearly define the problem and the questions that the designer might ask when modifying an infrastructure or introducing a new mode of locomotion for users (e.g., autonomous vehicles in mixed traffic, positioning of plug-in stations for electric cars) and its effects on users' behavior in a given locality (e.g., an area of the city, a town, an agglomeration, a region).

The second step assumed the existence of different available data sources. Thus, we distinguished between two types of data sources: those that could be used for user demand and, therefore, for generating synthetic populations; and those describing infrastructures. For user demand, we need data on households, users' journeys, and their main activities. We showed that we used data sources from the population census providing information on individuals' socioeconomic characteristics (e.g., age, socio-professional categories, and household) and describing their main activities (for example, work, leisure and education) of users. It was also necessary to obtain data sources (from *EMD*), essentially focusing on mobility practices. Concerning the transport supply, we considered OpenStreetMap (*OSM*) providing sufficient information (for example, roads, public transport lines, and stations). This step involved for maintaining the consistency of the anonymized data (In France, the data are anonymized), detailed in [44].

The third step focused on motivations for describing intermodal alternatives. In this way, we addressed generalities about agents' activities. The data available on users' activities made it necessary to suggest simplifying hypotheses for using transportation modes. For a sequence of agent activities, we have shown whether the transportation mode is modal or intermodal. For the intermodal alternative, we described two new transport modes (*car_pt* and *pt_car*) while defining their usage constraints, intermodal routing through *PR* facilities. Thus, we considered *car_pt* for the trips from the home with a car to the parking and then from this parking to the destination by public transport. The dual process for the return (i.e., from the destination to its home) is considered similar by using the alternative *pt_car*.

The fourth step detailed how to combine the transportation modes and the utilities. We naturally proposed *multinomial logit (MNL)* models. This work described four discrete models (*SI-MNL*, *SI-NL*, *I-2NL* and *I-CNL*) for organizing modal alternatives. The first is a simple *MNL*, while the last three models are based on nested structures. The two last models require more important abstraction levels and higher computational times for determining the different parameters of utilities. For the modal choice decision process, two discrete models (*SI-NL* and *SI-MNL*) are thus kept for the experimentation. Finally, we also gave the utilities for the different modes, including the intermodal alternative.

The fifth and final step presented our choice of a multi-agent platform *MATSim*, often used in transportation applications. This platform has the advantage of describing an activity-based model, as required in our context. The decision process for each agent has thus been modified to consider the utilities associated with the different transport modes. We also adapted this platform to include the discrete choice model and various useful extensions (in particular, the intermodal behaviors to reduce the private vehicle use).

Then, we applied this approach to validate intermodal alternatives and assess the value of a city toll area and its pricing. The experiments are thus carried out on the European Metropolis of Lille (*MEL*) data (1.1 million inhabitants). *MEL* is an agglomeration made up of 95 municipalities, within Lille is the largest with 233,098 inhabitants. We have emphasized that this agglomeration properly verifies our hypotheses: a very high usage of private vehicle by users, despite an extensive infrastructure (a public transport network including two metro lines, two tram lines, and 90 bus lines, and 12 *PR* facilities). Due to the absence of data for the *PRs*, it was necessary to obtain simulation results close to users of park-and-ride facilities (Cf. a previous study previously described in [44]). The number of agents was based on 20% of the actual population (746,318 trips) for two sufficient main reasons: the computational cost and the down-scaling for multi-agent transport simulations. In this study, we assumed an urban toll area to enter the city (Lille), and investigated the intermodal alternatives for significantly reducing the modal share of vehicles. For the calibration, we adjusted the different parameters of our discrete

choice models to the empirical (reference) data (*EMD* 2016). We used *PandasBiogeme*, for estimating these parameters of two discrete models (*SI-NL* and *SI-MNL*). We showed that these two models are statistically significant (and are nearly similar) and could be used for simulations. However, we also underlined that *SI-NL* requires more computational times than *SI-MNL*, for equivalent results. We therefore opted for *SI-MNL* for our experiments. The proposed experiments thus included (or not) the intermodal alternative and the variation of the urban toll tax (from 2.5 to 20 euros). These simulations were very correct in terms of realism, reproducing the shares of individuals practicing intermodality, for persons aged from 18 to 54 years. Our study showed that with 20 euros as toll tax, it was possible to reduce personal car use by 20% and 70% within *MEL* and Lille, respectively. We also briefly described a preliminary work (by assuming that the behaviors for *MEL* remained true for the global region) to extend our model for a larger region, Hauts-de-France (3,789 municipalities, 6 millions inhabitants). Based on the same values for the different variables, we showed (as expected) that the results required more coordinated policies to reduce significantly the use of private vehicles.

Our study was essentially based on data sources obtained from *EMD* and population census. Other data sources could be useful to improve the obtained results. It would be possible to collect other data sources, for example, from sensors positioned on the infrastructure (radar) or circulating on the infrastructure (e.g., connected vehicles, smartphones, sensors from *PR*). As part of future work, we could also plan to extend our simulation framework to other modal combinations such as bike or e-scooter (in particular, the shared micro-mobility services) [68] and public transport, or even the use of taxis or autonomous shuttles [69]. Finally, we would like to investigate other road pricing systems (following the distance or the road section).

Acknowledgments. We would like to thank the anonymous reviewers for many valuable comments improving our paper.

Declarations

- Funding: Not applicable
- Conflict of interest/Competing interests: The authors declare that they have no competing interests.
- Ethics approval: Not applicable
- Consent for publication: Not applicable
- Availability of data and materials: All data generated or analysed during this study are included in this published article.
- Code availability: Not applicable
- Authors' contributions: The authors (A-O. D., G. L., A. D. and R. M.) contributed equally to this work.

References

- [1] Diallo, A.O., Lozenguez, G., Doniec, A., Mandiau, R.: Agent-Based Intermodal Behavior for Urban Toll. In: Fujita, H., Fournier-Viger, P., Ali, M., Wang, Y. (eds.) 35th International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE, July 19-22. Lecture Notes in Computer Science, vol. 13343, pp. 397–408. Springer, Kitakyushu, Japan (2022)
- [2] Dacko, S.G., Spalteholz, C.: Upgrading the city: Enabling intermodal travel behaviour. *Technological Forecasting and Social Change* **89**, 222–235 (2014)
- [3] Turvey, R.: Road and bridge tolls in nineteenth-century London. *The Journal of Transport History* **17**(2), 150–164 (1996)
- [4] Bruxelles, I.-E.: *Le péage urbain aux portes de Bruxelles: la clé d’une ville humaine*. Bruxelles (2007)
- [5] Cervero, R.: Traffic Impacts of Variable Pricing on the San Francisco–Oakland Bay Bridge, California. *Transportation research record* **2278**(1), 145–152 (2012)
- [6] Dias, D., Tchepel, O., Antunes, A.P.: Integrated modelling approach for the evaluation of low emission zones. *Journal of environmental management* **177**, 253–263 (2016)
- [7] Yin, B., Diallo, A.O., Seregina, T., Coulombel, N., Liu, L.: Evaluation of low-traffic neighborhoods and scale effects: the paris case study. *Transportation research record* **2678**(1), 88–101 (2024)
- [8] Cats, O., Reimal, T., Susilo, Y.: Public transport pricing policy: Empirical evidence from a fare-free scheme in Tallinn, Estonia. *Transportation Research Record* **2415**(1), 89–96 (2014)
- [9] Tirachini, A., Hensher, D.A.: Multimodal transport pricing: first best, second best and extensions to non-motorized transport. *Transport Reviews* **32**(2), 181–202 (2012)
- [10] Kilani, M., Proost, S., Van der Loo, S.: Road pricing and public transport pricing reform in paris: complements or substitutes? *Economics of Transportation* **3**(2), 175–187 (2014)
- [11] Jones, W.B., Cassady, C.R., Bowden Jr, R.O.: Developing a standard definition of intermodal transportation. *Transportation Law Journal* **27**, 345 (2000)

- [12] Colard, J., de Lapasse, B., Clément, C.: Mobilités dans les Espaces Périphériques et Peu Denses: Pour Un Territoire Plus accessible? France Stratégie, ??? (2021)
- [13] Oostendorp, R., Gebhardt, L.: Combining means of transport as a users' strategy to optimize traveling in an urban context: empirical results on intermodal travel behavior from a survey in Berlin. *Journal of Transport Geography* **71**, 72–83 (2018)
- [14] McNally, M.G.: The four step model. Technical Report UCI-ITS-AS-WP-07-2, Center for Activity Systems Analysis, Univ. of California, Irvine, CA (2000)
- [15] Janssens, D., Wets, G., Brijs, T., Vanhoof, K., Arentze, T., Timmermans, H.: Integrating bayesian networks and decision trees in a sequential rule-based transportation model. *European Journal Operational Research* **175**(1), 16–34 (2006)
- [16] Joubert, J., de Waal, A.: Activity-based travel demand generation using bayesian networks. *Transportation Research Part C Emerging Technologies* **120** (2020). <https://doi.org/10.1016/j.trc.2020.102804>
- [17] Treuil, J.P., Drogoul, A., Zucker, J.D.: *Modélisation et Simulation à Base D'Agents : Exemples Commentés, Outils Informatiques et Questions Théoriques*. Dunod, Paris (2008)
- [18] Wooldridge, M.: *An Introduction to Multiagent Systems*. John Wiley & Sons, Hoboken, New Jersey (2009)
- [19] Diallo, A.O., Lozenguez, G., Doniec, A., Mandiau, R.: Comparative evaluation of road traffic simulators based on modeler's specifications: An application to intermodal mobility behaviors. In: *Proceedings of the 13th International Conference on Agents and Artificial Intelligence (ICAART)*, vol. 1, pp. 265–272 (2021)
- [20] Bazzan, A.L., Klügl, F.: A review on agent-based technology for traffic and transportation. *The Knowledge Engineering Review* **29**(3), 375–403 (2014)
- [21] Mandiau, R., Champion, A., Auberlet, J., Espié, S., Kolski, C.: Behaviour based on decision matrices for a coordination between agents in a urban traffic simulation. *Appl. Intell.* **28**(2), 121–138 (2008)
- [22] Balbo, F., Mandiau, R., Zargayouna, M.: Extended review of multi-agent solutions to Advanced Public Transportation Systems challenges. *Public Transp* (2024). <https://doi.org/10.1007/s12469-023-00332-9>

- [23] Russell, S., Norvig, P.: Artificial Intelligence: A Modern Approach. Pearson Education, New Jersey, USA (2003)
- [24] Horni, A., Nagel, K., Axhausen, K. (eds.): Multi-Agent Transport Simulation MATSim, p. 618. Ubiquity Press, London (2016). <https://doi.org/10.5334/baw>
- [25] Hörl, S., Balać, M., Axhausen, K.W.: Pairing discrete mode choice models and agent-based transport simulation with matsim. In: TRB Annual Meeting (2019). Transportation Research Board
- [26] Cao, X.J., Mokhtarian, P.L., Handy, S.L.: The relationship between the built environment and nonwork travel: A case study of northern california. *Transportation Research Part A: Policy and Practice* **43**(5), 548–559 (2009)
- [27] Pursula, M.: Simulation of traffic systems-an overview. *Journal of geographic information and decision analysis* **3**(1), 1–8 (1999)
- [28] Lighthill, M.J., Whitham, G.B.: On kinematic waves ii. a theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **229**(1178), 317–345 (1955)
- [29] Richards, P.I.: Shock waves on the highway. *Operations research* **4**(1), 42–51 (1956)
- [30] Balmer, M., Axhausen, K.W., Nagel, K.: Agent-based demand-modeling framework for large-scale microsimulations. *Transp. Research Record* **1985**, 125–134 (2006)
- [31] Beckman, R.J., Baggerly, K.A., McKay, M.D.: Creating synthetic baseline populations. *Transport. Res. A: Policy Pract.* **30**(6), 415–429 (1996)
- [32] Guo, J.Y., Bhat, C.R.: Population synthesis for microsimulating travel behavior. *Transport. Res. Rec.* **2014**(1), 92–101 (2007)
- [33] Cho, S., Bellemans, T., Creemers, L., Knapen, L., Janssens, D., Wets, G.: Synthetic population techniques in activity-based research. In: *Data Science and Simulation in Transportation Research*, pp. 48–70. IGI Global, Pennsylvania, United States (2014). Chap. 3. <https://doi.org/10.4018/978-1-4666-4920-0.ch003>
- [34] Tanton, R.: A review of spatial microsimulation methods. *Int. J. Microsimul* **7**(1), 4–25 (2014)
- [35] Chapuis, K., Taillandier, P.: A brief review of synthetic population generation practices in agent-based social simulation. In: *Social Simulation*

- Conference (SSC'2019) (2019)
- [36] Garrido, S., Borysov, S.S., Pereira, F.C., Rich, J.: Prediction of rare feature combinations in population synthesis: Application of deep generative modelling. *Transportation Research Part C* **120**, 102787 (2020)
- [37] Yaméogo, B.F., Gastineau, P., Hankach, P., Vandanjon, P.-O.: Comparing Methods for Generating a Two-Layered Synthetic Population. *Transportation Research Record* **2675**, 136–147 (2021)
- [38] Rich, J., Flötteröd, G., Garrido, S., Pereira, F.: Review of population synthesis methodologies. In: 8th Symposium of the European Association for Research in Transportation (hEART 2019), Budapest, Hungary (2019)
- [39] Lopez, P.A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y.-P., Hilbrich, R., Lücken, L., Rummel, J., Wagner, P., Wießner, E.: Microscopic Traffic Simulation using SUMO. In: The 21st IEEE International Conference on Intelligent Transportation Systems (ITSC), 4–7 November, pp. 2575–2282. IEEE, Maui, HI, USA (2018). <https://elib.dlr.de/124092/>
- [40] Espié, S., Auberlet, J.M.: Archisim: A behavioral multi-actors traffic simulation model for the study of a traffic system including its aspects. *International Journal of ITS Research* (n1), 7–16 (2007)
- [41] Haslebacher, R.: Intermodal routing in matsim applied to sbb green class. Master's thesis, IVT, ETH Zurich (2018)
- [42] Diallo, A.O., Doniec, A., Lozenguez, G., Mandiau, R.: Agent-based simulation from anonymized data: An application to Lille metropolis. In: Proceedings of the 12th International Conference on Ambient Systems, Networks and Technologies (ANT) (2021)
- [43] Balac, M., Hörl, S.: Simulation of intermodal shared mobility in the San Francisco Bay Area using MATSim. In: 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), pp. 3278–3283. IEEE, Indianapolis, United States (2021)
- [44] Diallo, A.O., Lozenguez, G., Doniec, A., Mandiau, R.: Agent-Based Approach for (Peri-)Urban Inter-Modality Policies: Application to Real Data from the Lille Metropolis. *Sensors* **23**(5), 2540 (2023)
- [45] Kamel, J., Vosooghi, R., Puchinger, J., Ksontini, F., Sirin, G.: Exploring the impact of user preferences on shared autonomous vehicle modal split: A multi-agent simulation approach. *Transportation Research Procedia* **37**, 115–122 (2019)
- [46] Hörl, S., Balac, M.: Introducing the eqasim pipeline: From raw data

- to agent-based transport simulation. *Procedia Computer Science* **184**, 712–719 (2021)
- [47] Ziemke, D., Kaddoura, I., Nagel, K.: The matsim open berlin scenario: A multimodal agent-based transport simulation scenario based on synthetic demand modeling and open data. *Procedia computer science* **151**, 870–877 (2019)
- [48] Lebrun, Y., Adam, E., Mandiau, R., Kolski, C.: A model for managing interactions between tangible and virtual agents on an RFID interactive tabletop: Case study in traffic simulation. *J. Comput. Syst. Sci.* **81**(3), 585–598 (2015)
- [49] Bonhomme, A., Mathieu, P., Picault, S.: Simuler le trafic routier à partir de données réelles. *Revue d’Intelligence Artificielle* **30**(3), 329–352 (2016)
- [50] Olvera, L.D., Guézéré, A., Plat, D., Pochet, P.: Intermodality in a context of poor transport integration: the case of Sub-Saharan African cities. In: *Transport Research Arena (TRA) 5th Conference* (2014)
- [51] Train, K.E.: *Discrete Choice Methods with Simulation*. Cambridge university press, United Kingdom (2009)
- [52] Ben-Akiva, M.E., Lerman, S.R.: *Discrete Choice Analysis: Theory and Application to Travel Demand* vol. 9. MIT Press, USA (1985)
- [53] Krajzewicz, D., Heinrichs, M., Beige, S.: Embedding intermodal mobility behavior in an agent-based demand model. *Procedia computer science* **130**, 865–871 (2018)
- [54] CEREMA: *Modélisation multimodale des déplacements de voyageurs concevoir un modèle de choix modal*. Technical report, Centre d’études et d’expertise sur les risques, l’environnement, la mobilité et l’aménagement (2015)
- [55] Balmer, M.: *Travel demand modeling for multi-agent transport simulations: Algorithms and systems*. PhD thesis, ETH Zurich (2007)
- [56] Kickhofer, B., Hosse, D., Turnera, K., Tirachinic, A.: *Creating an open matsim scenario from open data: The case of santiago de chile*. Working Paper 16-02, TU Berlin, Transport System Planning and Transport Telematics (2016)
- [57] Dor, G.B., Ben-Elia, E., Benenson, I.: *Population downscaling in multi-agent transportation simulations: A review and case study*. *Simulation Modelling Practice and Theory*, 102233 (2020)

- [58] Bierlaire, M.: A short introduction to pandasbiogeme. Technical report, Transport and Mobility Laboratory, ENAC, EPFL (2020)
- [59] Cherchi, E., Bhat, C.: Workshop Synthesis: Data analytics and fusion in a world of multiple sensing and information capture mechanisms. *Transp. Research Procedia* **32**, 416–420 (2018)
- [60] Kamalian, M., Ferreira, P., Jul, E.: A survey on local transport mode detection on the edge of the network. *Applied Intelligence* **52**, 16021–16050 (2022)
- [61] Egu, O., Bonnel, P.: How comparable are origin-destination matrices estimated from automatic fare collection, origin-destination surveys and household travel survey? an empirical investigation in Lyon. *Transp. Res. Part A Policy Pract.* **138**, 267–282 (2020)
- [62] Croce, A.I., Musolino, G., Rindone, C., Vitetta, A.: Estimation of travel demand models with limited information: Floating car data for parameters' calibration. *Sustanaibility* **13**, 8838 (2021). <https://doi.org/10.3390/su13168838>
- [63] Lee, W.K., Sohn, S.Y., Heo, J.: Utilizing mobile phone-based floating population data to measure the spatial accessibility to public transit. *Appl. Geogr.* **92**, 123–130 (2018)
- [64] Li, J., Boonaert, J., Doniec, A., Lozenguez, G.: Multi-models machine learning methods for traffic flow estimation from floating car data. *Transportation Research Part C: Emerging Technologies* **132**, 103389 (2021)
- [65] Nagel, K.: Road Pricing. In: Horni, A., Nagel, K., Axhausen, K. (eds.) *The Multi-Agent Transport Simulation MATSim*, pp. 97–102. Ubiquity Press, London, GB (2016). Chap. 15
- [66] de Freitas, L.M., Schuemperlin, O., Balac, M.: Road pricing: An analysis of equity effects with matsim. In: *Proceedings of the 16th Swiss Transport Research Conference, Ascona, Switzerland*, pp. 18–20 (2016)
- [67] du Plessis, W., Joubert, J.W.: Evaluating the effect of road pricing strategies in matsim, using agent-specific and income-dependent values. In: *Proceedings of the International Conference on Computers and Industrial Engineering, CIE*, pp. 525–532 (2012)
- [68] Diallo, A.O., Lozenguez, G., Doniec, A., Mandiau, R.: Estimation of minority modes of transportation based on machine learning approach. In: *Shakshuki, E.M., Yasar, A. (eds.) The 13th International Conference on*

Ambient Systems, Networks and Technologies (ANT), March 22-25. *Procedia Computer Science*, vol. 201, pp. 265–272. Elsevier, Porto, Portugal (2022)

- [69] Salladarre, F., Le Boennec, R., Bel, M.: Robotaxis or autonomous shuttles? The role of urban representations and travel habits in tomorrow's mode choice in France. *Journal of Transport Economics and Policy (JTEP)* **55**(3), 237–259 (2021)