A hypothetical urban layout generation model for exploring land use impacts on travel behavior

Citation for published version (APA):
Lyu, X., Han, Q., & de Vries, B. (2022). A hypothetical urban layout generation model for exploring land use impacts on travel behavior. *Travel Behaviour and Society*, *28*(6), 317-329.
https://doi.org/10.1016/j.tbs.2022.04.009

Document license:
CC BY

DOI:
10.1016/j.tbs.2022.04.009

Document status and date:
Published: 01/07/2022

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 05. Oct. 2023

Travel Behaviour and Society 28 (2022) 317–329

Contents lists available at ScienceDirect

Travel Behaviour and Society

journal homepage: www.elsevier.com/locate/tbs

A hypothetical urban layout generation model for exploring land use impacts on travel behavior

Xiaoming Lyu *, Qi Han, Bauke de Vries

Urban Systems & Real Estate, Department of the Built Environment, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, the Netherlands

ARTICLE INFO

Keywords:
Land use
Travel behavior
Simulation
Urban layout generation

ABSTRACT

In urban planning, land-use policies are commonly applied to reduce automobile travel and encourage active transport because land use is believed to affect travel behavior. However, a debate about such effects continues in literature. Empirical studies differing in context, methodology, and geographic scale come to different results. Although questioned because of self-selection, the accumulated empirical evidence provides a solid foundation for conducting simulation research that helps systematically examine the effects of land use attributes on travel behavior. This paper introduces an urban layout generation (ULG) model specially developed for simulations. The ULG model is not a land use and transport interaction model. It generates hypothetical urban layouts corresponding directly to D-variables, a set of land use measurements widely used in empirical studies. The D-variables are controlled on multiple spatial scales in the ULG model. Besides the spatial aspects of the road network and land use commonly included in virtual city models, the ULG model also simulates the population density. Performance tests showed the ability of the ULG model to generate hypothetical though reasonable urban layouts which meet the requirements of simulation research of land use and travel behavior. Finally, several examples combing the ULG model and a travel behavior simulation model show potential applications of the ULG model.

1. Introduction

Transportation is an essential source of greenhouse gas emissions and energy consumption; therefore, reducing vehicle transport and encouraging active transport (walking and cycling) has become one of the main targets of urban planning. Policymakers believe that changing the attributes of land use, also be termed urban form and built environment in some literature, can affect travel behavior. For example, increasing land use mix can reduce vehicle dependency and encourage active transport and public transport (Ewing and Cervero, 2010; Spears et al., 2014).

A significant amount of land use-travel behavior research was conducted in past decades. Most empirical studies employed the regression method. The regression models use travel-behavior variables as dependent variables, regressed on land use variables and sociodemographic variables as independent variables (Boarnet, 2011). However, a debate continues about the magnitude of regression coefficients of land use, mainly due to self-selection bias and differences in context, methodology, and geographic scale (Ding et al., 2018).

Many empirical studies applied to land use measurements at a local scale, fewer studies applied at a regional scale, and only limited studies compared different geographic scales (Boarnet, 2011; Nasri and Zhang, 2015; Milakis et al., 2015). Studies implicated a stronger correlation between vehicle dependency and land use measurements at the city level than at the neighborhood level because many trips have long-distance in cities (Boarnet, 2011; Ewing and Cervero, 2010).

The self-selection bias implies that people's preferences of travel behavior partly influence their decisions about where to live and work. Therefore, the observed correlation between land use and travel behavior is not fully causal. Practical efforts have been made to control the effects of self-selection (Cao et al., 2009). These methods can reduce but not exclude the impact of self-selection. The self-selection can be avoided by assigning the living and working locations randomly to residents, but it is not realistic (Brownstone, 2008). A good alternative is using panel data that follow households over time, but collecting such data would be costly (Brownstone, 2008); another option is employing joint models of travel behavior, car ownership, and location choice, but the simultaneous estimation of the integrated models is very complex

---

* Corresponding author.

E-mail addresses: x.lyu@tue.nl (X. Lyu), q.han@tue.nl (Q. Han), b.d.vries@tue.nl (B. de Vries).

https://doi.org/10.1016/j.tbs.2022.04.009
Received 30 October 2020; Received in revised form 23 February 2022; Accepted 25 April 2022

(Boarnet, 2011).

Simulation has the potential to apply Brownstone's method of avoiding self-selection. It can assign the living and working locations randomly.

Last century, especially in the early 1990s, many simulation studies were conducted. These simulation studies received strong criticisms (Crane, 2000; S. Handy, 1996). The complaints focused on two main problems. First, these simulations oversimplified the urban form, travelers, and their responses to circumstances. This criticism is not related to the simulation methodology but the simplification of the complexity. It implies those simulations could be more credible if they were not oversimplified. When those simulation models were developed, empirical studies on the relationships between land use and travel behavior were conducted. The measurements of land use were simple, the understanding of travelers' behavior was not deep, and the technology of generating hypothetical urban layout was still immature. With limited empirical and technical support, simulations were crude. However, after decades of development, today, those weaknesses of early simulations have mainly been overcome.

Second, these simulations were criticized for not explaining but assuming particular behavior. The simulation needs input, and specific assumed behavior is an indispensable part of the input. The critical question is whether the assumptions are reasonable. In the early 1990s, with limited empirical evidence, the assumptions were quixotic; nowadays, however, behavior-based assumptions are much closer to reality. The simulation does not explain the assumed behavior, but based on the individual behavior, the simulation intends to explore the group behavior. Even if the individual behavior is apparent, the group behavior may not be understood easily because of the complexity. The emerging group behavior could be explored by simulating all the individuals, which is one of the values of the micro travel-behavior simulation.

This paper aims to introduce an urban layout generation (ULG) model designed especially for simulation studies of the land use impacts on travel behavior using micro travel-behavior simulation models. The ULG model can generate hypothetical urban layouts with the input of D-variables commonly used in empirical studies measuring the built environment. Studies of multi-spatial scales, from the neighborhood level to city level, can be conducted in the ULG model.

2. Background of developing the ULG model

For a better understanding of the design of the ULG model, the purpose of developing the model must be introduced first. The model is developed for simulation research. According to the research framework, requirements for the model are stated. Algorithms for generating urban layouts are collected and selected to meet the requirements.

2.1. The research framework applying the ULG model

The simulation research of land use impacts on travel behavior, employing the ULG model, is based on a regression model applied widely in empirical studies (Boarnet, 2011):

$$Travel-Behavior-Variable = \beta_0 + Land-Use-Variables * \beta_1 + Sociodemographic-Variables * \beta_2 + \varepsilon \quad (1)$$

where travel-behavior variables are regressed on independent land use variables and sociodemographic variables.

Integrating the non-land-use variables (sociodemographic variables) considers self-selection and preference. However, the same sampling of travelers can be applied to the simulation research, making the sociodemographic condition of travelers fixed. The simulation also makes it possible to keep land use attributes fixed among scenarios, changing specific land use variables. As a result, the effects of target land use attributes can be observed directly in the simulation with fixed sociodemographic variables and varying land use variables under different scenarios.

Therefore, generating hypothetical urban layouts and simulating the corresponding travel behavior are the two most essential parts of the simulation research. As shown in Fig. 1, the simulation approach comprises an urban-layout-generation model and a travel-behavior-simulation model. A developed multi-state supernetwork model (MSN model) is employed for the travel-behavior simulation. The ULG model needed for this approach is presented in this paper.

2.2. Requirements for the ULG model

The essential requirement of the ULG model is the ability to generate desired urban layouts. What is a "desired" city? As this simulation system aims to serve the research of land use – travel behavior relationship, the "desired" city here is defined as a city described and restricted by land-use variables widely used in the land use-travel behavior research. These land-use variables are the widely used D-variables: Density, Diversity, Design, Distance-to-transit, and Destination accessibility (Boarnet, 2011; Cervero and Kockelman, 1997). Density is commonly mentioned as population density; Diversity is defined as land use mix; Design means the road pattern; Distance-to-transit relates to public transport. Destination accessibility, however, is not considered in the ULG model because this variable does not describe the whole city but only a part of it. Using the four D-variables, the ULG model can generate three essential layers: population, land use, and road network (including public transport).

For the land use mix research, the generation of land use has specific requirements. First, the generation system should control land use aggregation or dispersion to generate various land use patterns. Second, to allow for research on different spatial scales, the simulation system should control land use at different spatial levels, i.e., district level, neighborhood level, etc.

The ULG model does not simulate the development process of cities; it generates final urban layouts. Consequently, the ULG model does not copy real cities; it generates hypothetical cities for research and development.

2.3. Algorithms of urban-layout-generation models

Many LUTI models (land use and transport integration models) simulate the interaction between land use and travel behavior (Salas-Olmedo et al., 2017; Wang et al., 2015). However, the LUTI model and the ULG model have different aims. As a result, their requirements and algorithms are different. The LUTI model mainly aims to predict land use and transport in the future, and therefore the interactions and evolutionary process between land use and transport are simulated. In contrast, for the ULG model, the effects of transport on land use are not considered. The evolutionary development of the urban layout is also excluded. The most striking difference between the LUTI and ULG models is that the LUTI model always inputs actual urban layouts instead of generating hypothetical cities. Therefore, the urban land use simulation algorithms of the LUTI model are not suitable for the ULG model.

Procedural modeling was applied to urban-layout simulation nearly two decades ago. Various algorithms have been developed to simulate the urban physical environment (Aliaga et al., 2008; Beneš et al., 2014; Thomas Lechner et al., 2003; Parish and Müller, 2001; Weber et al., 2009). However, the entertainment industry initially motivated the development and application of procedural modeling, which generated plausible but not solid hypothetic cities. Therefore, these procedural models may have problems applying to academic urban studies.

Regarding simulation of population density distribution, some procedural models do not take this into account (Aliaga et al., 2008; Beneš

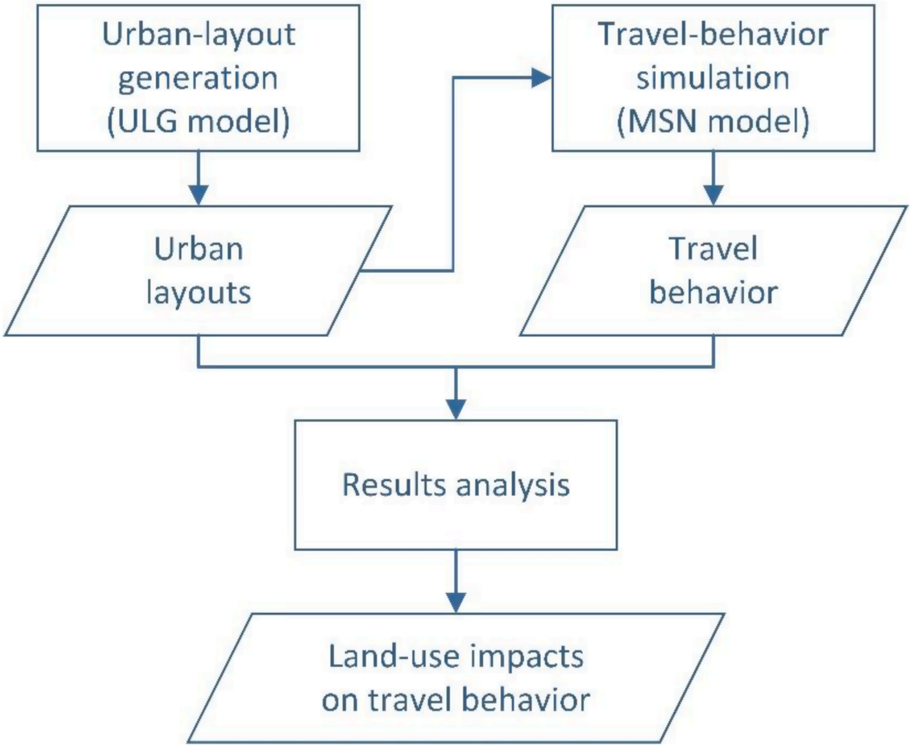


Fig. 1. The simulation research framework.

et al., 2014; Weber et al., 2009), while others do not simulate but input population density maps of actual cities. One or several fixed population density maps cannot meet the requirements of the simulation study in which multiple urban layouts would be generated with different population densities.

There are plenty of algorithms generating road networks. These algorithms can be divided into two types: static and dynamic. The static algorithm generates a fixed road network in one stroke; thus, the result is easier to control. The L-system generates a road network by controlling the direction, angle deviation, and segment length of a road (Parish and Müller, 2001), the example-based algorithm expands a city from an urban image as the input (Aliaga et al., 2008), the tensor-fields algorithm generates the road network by generating a tensor field (Chen et al., 2008). However, there are weaknesses for the static methods applied directly in academic simulation research. In the L-system, for example, highways connect population centers, giving the population density a significant influence on road network forms. It is not very reasonable because the main destinations are not always population centers. In the example-based algorithm, the road network of an entire city is determined by a small piece of urban image, which may contain bias, and the city's road network can hardly be controlled. In the tensor-fields algorithm, the tensor field is often too abstract. The dynamic algorithm generates a growing road network from one or more start points over a growth period (Beneš et al., 2014; Thomas Lechner et al., 2003; Weber et al., 2009). However, the simulation system presented in this paper does not aim to simulate the developing process of the road network, as mentioned above.

Realistic land use allocation, however, could be generated in the procedural urban models (Thomas Lechner et al., 2004; Tom Lechner et al., 2007; Weber et al., 2009). These models usually aim to maximize land value by designing developers' behavior. Hypothetical urban layouts in our simulation framework could be generated by such behavior-based algorithms, including land use development; however, two significant problems will arise. First, the inclusion of land use development in hypothetical cities is needless. More importantly, such an algorithm has difficulty controlling the final urban layout. A slight adjustment of the assumption underlying the behavior-based algorithm could significantly differ in the urban layout.

Therefore, directly using the existing urban-layout-generation algorithms does not fit our approach to simulation of the land use impacts on travel behavior as presented in Fig. 1. This has motivated us to develop the ULG model, as presented in the following sections.

3. The urban layout generation model

3.1. Introduction

The generated urban layout comprises the three layers mentioned in section 2.2: population density, land use, and road network. The road network includes four road types: Highway, Arterial, Distributor, and Local (Lay, 2009).

The geographic scale is an essential aspect of the land use-travel behavior research. Automobile trips are more associated with the regional scale, while non-motorized trips are more heavily influenced by the neighborhood scale (S. L. Handy et al., 2002). Therefore, spatial units at different geographic scales are necessary. In the system, a city is divided into districts and neighborhoods. A district is a spatial unit surrounded by arterials and highways, corresponding to an urban region such as the Central Business District, an administrative district, or a postcode district. A neighborhood is a spatial unit surrounded by distributors (sometimes arterials). The neighborhood is supposed not to exceed walking distance.

Various land classifications are employed in empirical studies of land use – travel behavior. Frank and Pivo (1994) considered seven land uses: single-family, multifamily, retail and services, office, entertainment, institutional, and industrial; Hong et al. (2014) chose *residential* instead of the single-family and multifamily, and used *others* to include entertainment and other uses. Some studies even used a rough classification, including only residential, retail, service, and others (Ding et al., 2018; Nasri and Zhang, 2015). Considering the most common daily activity types in a city, five types of land use are generated in our system: *Residence*, *Industry*, *Commerce*, *Office*, and *Green&Open*. We use *Green&Open*

but not *Others* to emphasize the entertainment activity like Frank and Pivo (1994) did.

If the input variables are the same as the D-variables mentioned in section 2, the target urban layouts are generated accurately and easily. However, the D-variables cannot provide enough information describing an urban layout. More information such as the urban area and population is needed. Some variables represent a simple aspect of the urban layout, such as the urban area and the density, thus they can be inputted directly and be easily allocated to smaller spatial units. Some variables like the Design and Distance-to-transit can be translated into a set of variables, such as the segment length of the road network and the distance between bus stops, and then be well controlled. Diversity, however, describes a complex phenomenon. Thus, its value (the entropy) is hard to allocate to smaller spatial units directly. Therefore, in the system, we choose other more straightforward but indirect variables as input to control the Diversity. The values of the indirect variables need to be tested or calibrated to generate the target result. The primary input is shown in Table 1. We used the V*i* to number the variables.

Fig. 2 shows the flowchart of the ULG model. We used the D*i* to number the documents input to or generated from the processes. With the *Urban-Area*, the built area of the hypothetic city is defined, on which the three layers of the urban layout are based.

3.2. Population

Unlike most existing urban layout simulation models that directly take population density as an input, the ULG model generates it. From the literature, urban population density functions (Li and Ryohei, 2006) are employed in the ULG model to generate the population. In the ULG model, three classic population density models (see Table 2) are currently applied, and more functions could be added when necessary. During the generation process, one model is selected by the user, and then the values of its parameters are entered.

Employing these mathematical models have one problem. It is not hard to infer from the functions of the models that the generated result is a series of concentric circles, which is an abstraction of the real world. One way to make such outcomes more realistic is to design an algorithm generating a reasonable polycentric concentration of population. Another approach is just leaving the results as an original population density map to participate in the land use allocation. After interacting with allocated land use, residents would move to locations of residential land, with a realistic final population density map generated. The interaction algorithm is introduced in the land use section.

Table 1
The primary input to the ULG model.

No.	Variable	Description	Work module
V1	*Urban-Area*	Area of hypothetic city	Generating urban area Land use allocation
V2	*Population*	Population of hypothetic city	Population density
V3	*If-Highway*	Whether to build Highway	Road network
V4	*Length-i*	Segment length of road type *i*. For *i* = Highway, Arterial, Distributor	Road network
V5	*Min-D* *Min-N*	Minimal size of districts and neighborhoods.	Generating districts and neighborhoods
V6	*Land-use-Percent-i*	Land use percentage. For land use *i* = Residence, Industry, Commerce, Office, Green&Open	Land use allocation to districts
V7	*Allocation-Rounds-D*	Allocation rounds of districts	Land use allocation to districts
V8	*Allocation-Rounds-N*	Allocation rounds of neighborhoods	Land use allocation to neighborhoods
V9	*Dispersion-D-i*	Dispersion of land use in districts. For land use *i* = Industry, Commerce, Office, Green & Open	Land use allocation to districts
V10	*Dispersion-N-i*	Dispersion of land use in neighborhoods. For land use *i* = Residence, Industry, Commerce, Office, Green&Open	Land use allocation to neighborhoods

3.3. Road network

Different patterns are built to generate a road network at different scales. Since the road network optimization is not the focus in the research of land use – travel behavior interaction, at this stage, we applied only a simple network, the grid network. The commonly used parameters of street expansion in procedural urban modeling are direction, angle deviation, and segment length (Parish and Müller, 2001; Weber et al., 2009). Roads in the grid network grow forward and vertically from the city center, with fixed direction and no angle deviation, so the only control parameter is the segment length. We use *Length-H, Length-A, Length-D* to control the Highway, Arterial, and Distributor segment length, respectively (see Fig. 3). The *If-Highway* holds whether to build highways.

After building the road network, the city is divided into districts and neighborhoods. Some spatial units, especially near the fringe, could be tiny. To prevent the fragmentary districts and neighborhoods and to adjust the size, two variables named *Min-D* and *Min-N* are employed to control the minimal area of districts and neighborhoods. Any district smaller than the *Min-D* would be merged into one of its neighboring districts. Users can reserve the central district because some old cities have a central historic district, or some modern cities have a CBD.

After the land use allocation, the roads within neighborhoods are built, interacting with the land use. If Green dominates a neighborhood, no roads would be built. If Industry dominates a neighborhood, a low density of Distributor would be built. If Residence, Commerce, and Office mainly occupy a neighborhood, a high density of Local would be built. Otherwise, a normal density of local would be built.

A public traffic net is designed to follow the road network pattern. The bus lines are built vertically or horizontally along the Arterials and Distributors. Such a pattern could make travelers arrive at their destinations with at most one transfer.

3.4. Land use

3.4.1. Algorithm

As mentioned in the requirements of the ULG model, the model is not intended to simulate the development process of cities but only to generate the desired urban layouts. The underlying logic of generating a hypothetical urban layout is like planning a new town. Therefore a land use planning support system – the What-If? model (Klosterman, 2001; Pettit et al., 2015) – is employed for land use allocation.

The *What-If?* system consists of three sub-processes: to analyze each spatial unit's land use suitability, predict the amount of land use, and allocate the land use to the spatial units with the highest suitability.

The suitability analysis process of the *What-If?* system has four steps. The first step is to identify suitability factors, such as distance to the city center, population density, and interactions among land use types considered in our system. Different land use may have various factors. For example, highways may be more critical for Industry but less important for Green. The second step is to specify the suitability factors weights that indicate the relative importance for land use. The third step is to determine the suitability factor ratings that indicate the relative suitability for specified land use. The last step is to determine permissible land use conversions, which is needless in our system in which the land uses are allocated on "blank" areas. The suitability score is calculated by multiplying the factor weights by the corresponding factor rating and then summing these values. The resulting suitability scores indicate the relative suitability of each spatial unit for each land use.

Instead of predicting land use demands, the amount of each land use is calculated in our system by multiplying the variable *Urban-Area* by the *Land-use-Percent-i* (see Table 1). In the land use allocation process, the land use is first allocated to the spatial unit with the highest score of

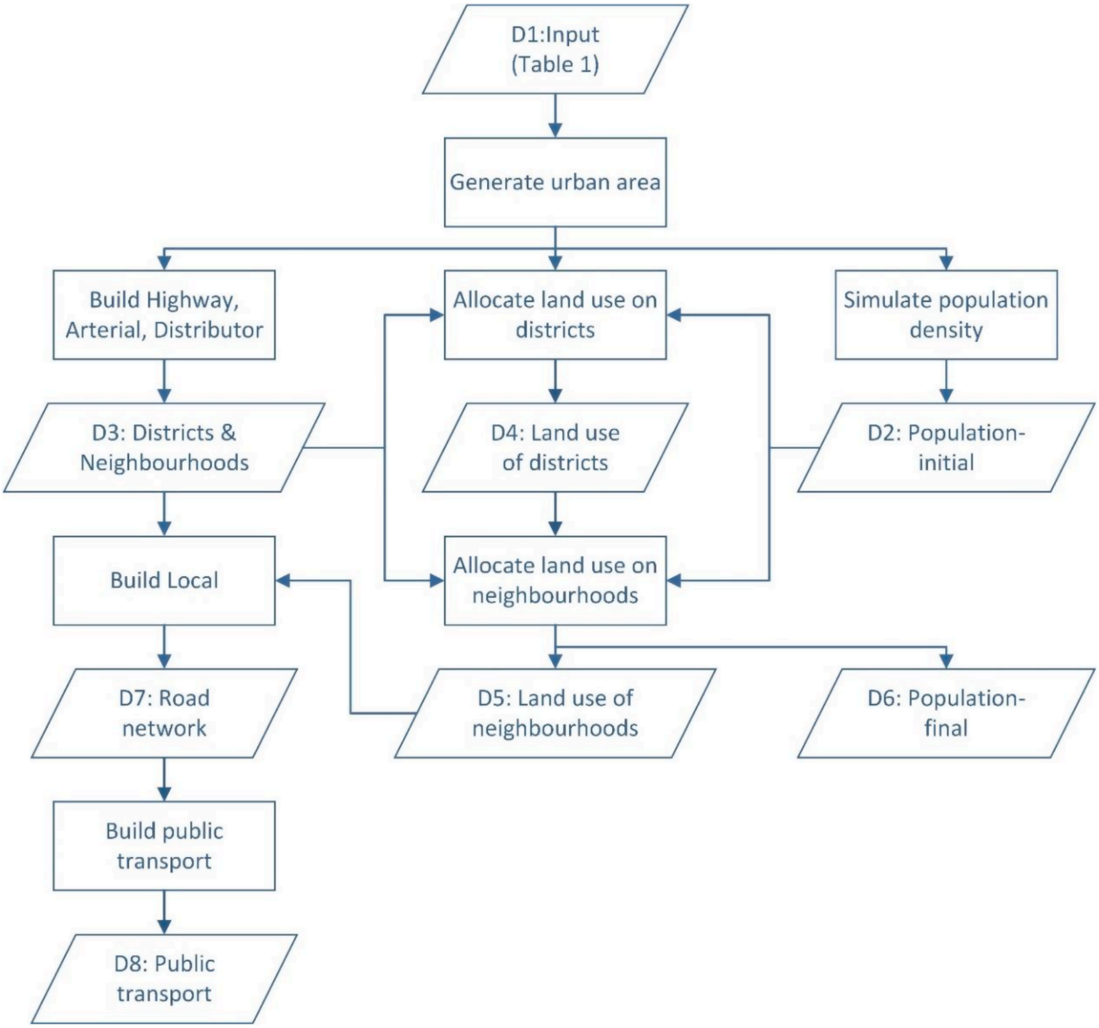


Fig. 2. The flowchart of the ULG model.

Table 2
Population density models.

Model	Function	Population distribution
Negative exponential model	$D(x) = D_0 e^{-\gamma x}$	Concentrated
Normal distribution model	$D(x) = D_0 e^{\gamma x - \beta x^2}$	Diffused
Quadratic model	$D(x) = D_0 + \gamma x - \beta x^2$	Diffused

suitability and then to the one with the second-highest score, and so on. If some units have the same score, one is randomly selected to allocate. The land use is first allocated at the district level and then at the neighborhood level.

The *What-If?* system does not consider the interaction among land use types. For example, there are two pieces of blank area A and B; A adjoins B; A is suitable for Industry, while B is suitable for Residence. In the *What-If?* model A would be allocated to Industry, and B would be allocated to Residence. But the allocation of A to Industry makes B not suitable for Residence anymore. To represent such interaction, in the ULG model, the suitability analysis and allocation deal with only one type of land use at one time. More specifically, a type of land use is allocated after calculating its suitability, then calculates suitability for another type of land use and allocates that land use. Thus, the latter land uses will interact with the former land uses.

However, allocation priority, which is the order to allocate land use, will affect the result significantly. For example, the piece of blank area A is suitable for both land use *i* and j; if *i* is allocated first, it would occupy area A and make *j* another place; conversely, *j* would occupy area A, and *i* is allocated on other sites. To reduce the effects caused by allocation order, we introduced two variables, *Allocation-Rounds-D* and *Allocation-Rounds-N*, to control the number of land use allocation rounds at the district and neighborhood levels, respectively. For example, there are 10 km^2 land *i* and 15 km^2 land *j* for allocation, and area A is 5 km^2. If the *Allocation-Rounds* equals 5, there would be five rounds in the allocation. 2 km^2 of land i and 3 km^2 of land j are allocated within each round. There would be a chance that both land *i* and *j* are allocated within area A. The larger value of the *Allocation-Rounds*, the smaller effects that are caused by the allocation order.

To control the compactness or dispersion of land use, the variables of *Dispersion-D-i* and *Dispersion-N-i* are introduced to hold at the district level and neighborhood level, respectively. In each round of land use allocation, the land use is allocated simultaneously at the number of *Dispersion* spatial units with the highest suitability. Thus, a smaller Dispersion value leads to compact land patterns, while a larger Dispersion value prefers dispersed land patterns.

3.4.2. Land use allocation at the district level

In the district-level allocation, the amounts of each land use are first calculated according to the composition of land use (the *Land-use-Percent-i*), see Fig. 4. After that, the allocation of Industry, Commerce, Office, and Green&Open follows the algorithm introduced in the previous section, which calculates the amounts of land uses in the initial

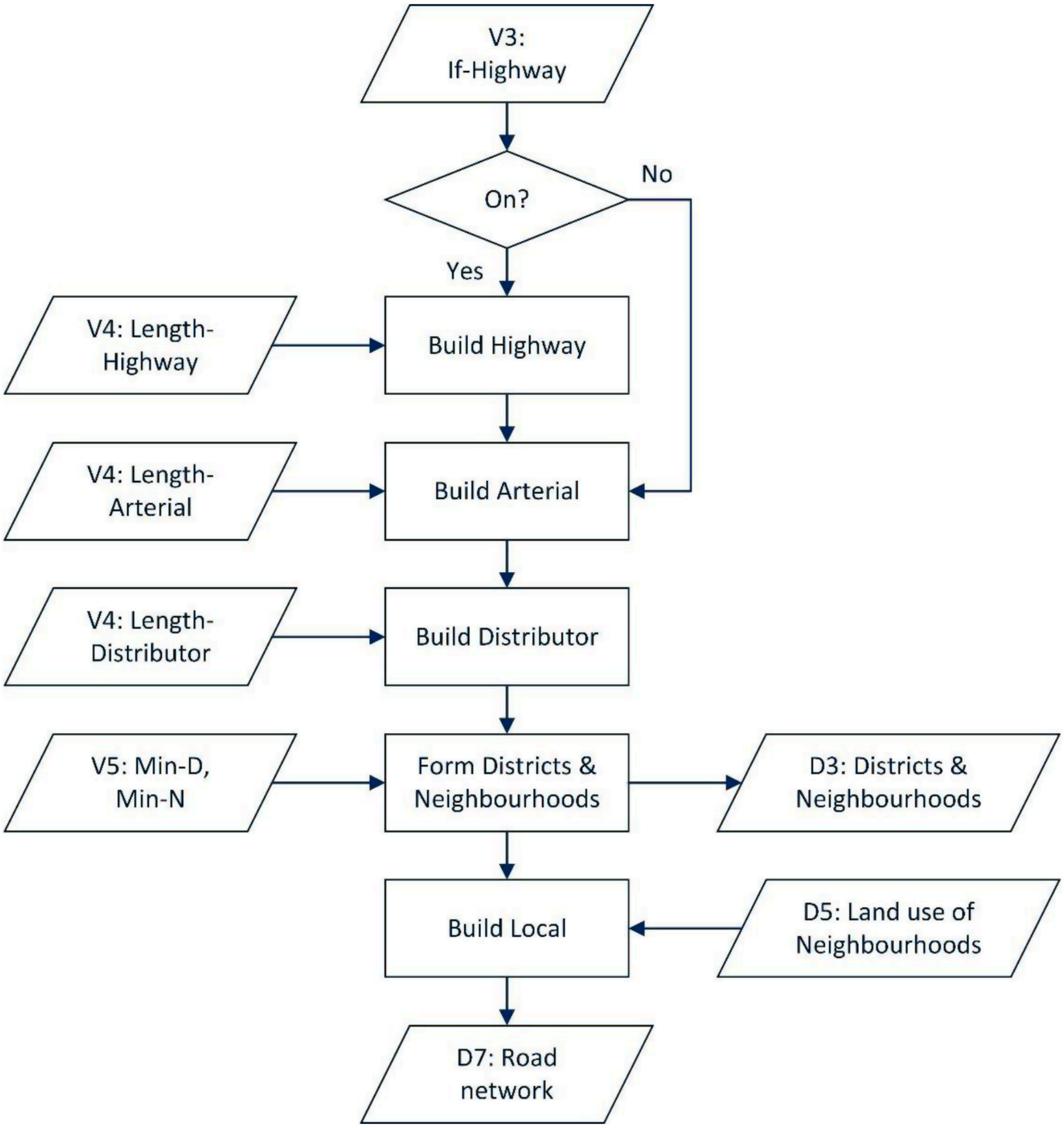


Fig. 3. The flow of road network generation.

allocation round, then repeats the following procedures in all next rounds: calculating the scores of the suitability of all districts, then allocating the amount of land use to districts with the highest scores of suitability.

The allocation of Residence, however, follows a different method. The distribution of Residence should be in accordance with the distribution of population, at least at the district level should be mainly close to the original population density map generated in section 3.2; otherwise, the urban spatial structure that the user chose to simulate in section 3.2 would not be realized. Therefore, the allocation of Residence is not based on the suitability but the number of residents of each district.

As shown in Fig. 4, after calculating the total amount of Residence, the numbers of residents of all districts are calculated according to the population density map, the *Population-initial*. Then Residence is allocated to each district to provide just enough space for their residents to live in.

After the allocation, each district has a record of the land use composition. The figures are the input to the neighborhood-level allocation.

3.4.3. *Land use allocation at the neighborhood level*

Land use allocation to neighborhoods within districts is processed one by one. Within a district, the amounts of all land uses are recorded. The allocation follows the algorithm introduced in section 3.4.1, see Fig. 5, except for the last step of immigration.

During the allocation process, it might happen that the available blank land of a neighborhood is not enough for its residents because other types of land use have been allocated in the neighborhood. Thus, the surplus residents need to ‘immigrate’ to other neighborhoods with blank areas. To be close to the initial population density map, the surplus residents move to neighborhoods with similar distances to the city center with the neighborhood that they move from. Simultaneously, the population density map is adjusted.

After the allocation, each neighborhood has a record of the amounts of land use types in the neighborhood.

4. Performance test

4.1. *Urban layout*

Considering the further application of the ULG model in the

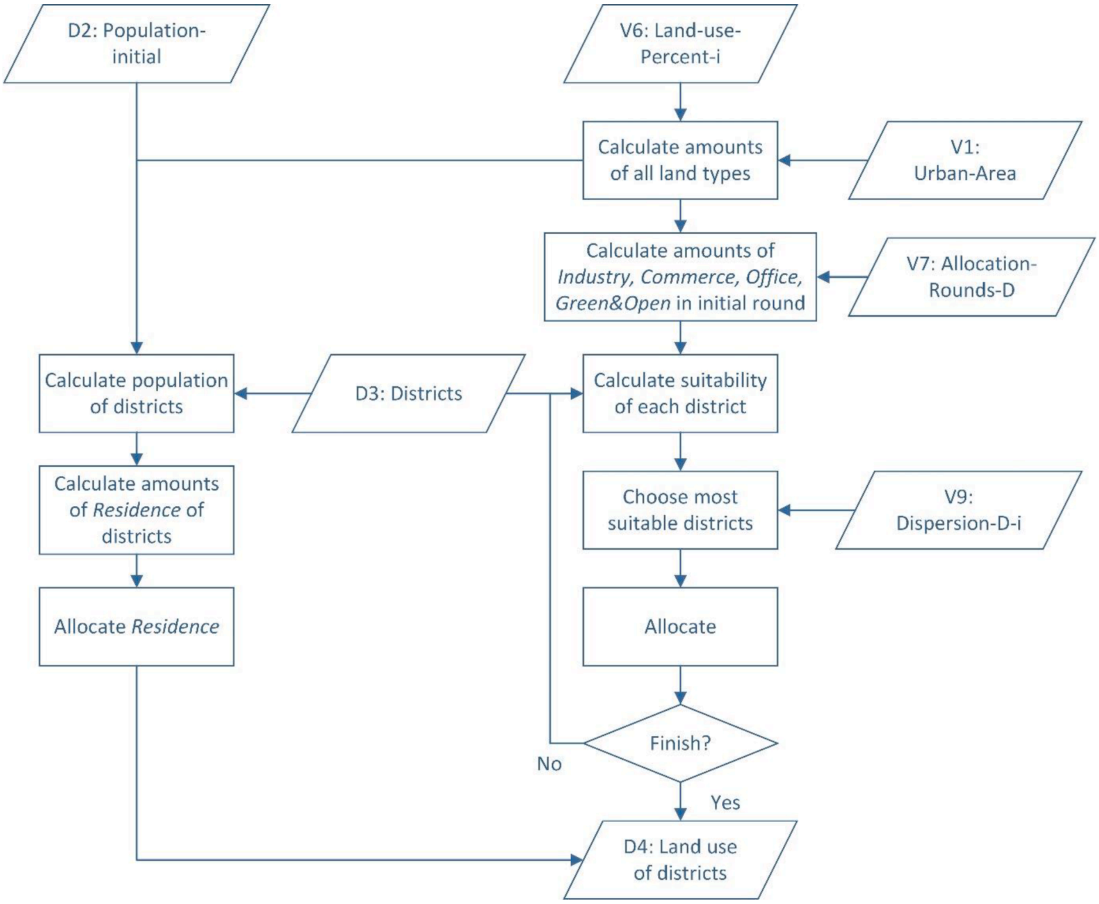


Fig. 4. Land use allocation at the district level.

simulation research mentioned in section 2.1, a 90-square-kilometers mono-centric city with 200,000 residents was generated in the tests. A possible urban layout is shown in Fig. 6:

4.2. The *Allocation-Rounds* variables

The *Allocation-Rounds* variables are designed to control the effects caused by the allocation order of land use. It is expected that larger *Allocation-Rounds* variables would bring smaller effects of the allocation order. Therefore, we generated two sets of urban layouts with different allocation orders and compared the layout difference between the two sets to see how the allocation-order effect would change when the values of *Allocation-Rounds* variables were adjusted.

As the city center has more attraction for Commerce and Office, the allocation-order effect would make earlier allocated land use occupy areas closer to the city center. Similarly, the urban fringe has more attraction for Industry and Green. Thus, the allocation-order effect would make earlier allocated land use occupy areas closer to the urban edge. Therefore, comparing the *Average-distance-to-city-center* of a land use type in different allocation orders can reflect the allocation-order effect.

In this case, at the district level, the first set of allocation order A was Industry, Commerce, Office, and Green; the other set of allocation order B was inversely Green, Office, Commerce, and Industry. Increasing the value of *Allocation-Rounds-D* from 1 to 7, all else equal, eight urban layouts were generated. The *Average-distance-to-city-center* values are shown in Fig. 7. When the *Allocation-Rounds-D* was set 1, the *Average-distance-to-city-center* of Commerce in order A and B differed significantly, but other land uses were close. However, when the *Allocation-Rounds-D* was set larger than 3, the *Average-distance-to-city-center* of all land uses for A and B were very close.

At the neighborhood level, a district-level land use map was first generated as the foundation of further neighborhood-level land use allocation. The allocation order A was Industry, Residence, Commerce, Office, and Green, while order B was inversely Green, Office, Commerce, Residence, and Industry. Increasing the value of *Allocation-Rounds-N* from 1 to 9, all else equal, ten urban layouts were generated. The *Average-distance-to-city-center* values are shown in Fig. 8. When *Allocation-Rounds-N* was set 7 or 9, the *Average-distance-to-city-center* values in order A and B were very close.

4.3. Application

4.3.1. Travel behavior simulation

This part introduces the simulation of travel behavior, including the MSN model as an example of a micro travel behavior simulation model and an algorithm for allocating individuals in a hypothetic city.

The MSN model is employed to simulate the corresponding travel behavior of the hypothetic urban layouts. The MSN model is based on the utility theory, optimizing travelers' choices according to their individual-level least generalized disutility.

The input into the MSN model consists of the environment and the travelers. The environmental input is mainly the land use, road networks, and public transport. The travelers' information includes the sociodemographic profiles, travel preferences, and activity programs of travelers. The model simulates the activity locations, travel mode choices, and travelers' route choices.

A more detailed explanation of the MSN model is given by Liao, Arentze, and Timmermans (2010, 2013). An application of the MSN model can be found in Liao et al. (2017).

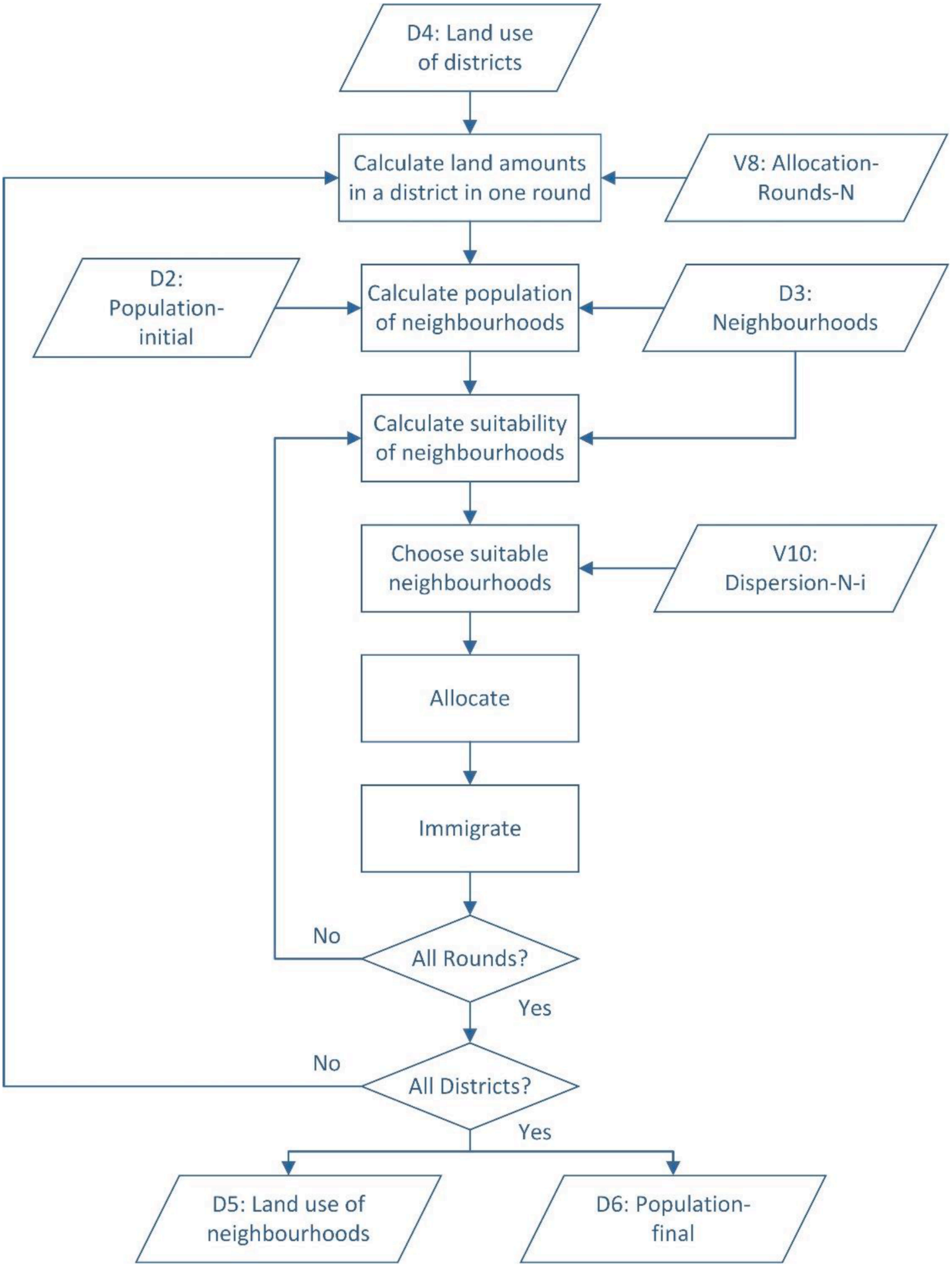


Fig. 5. Land use allocation at the neighborhood level.

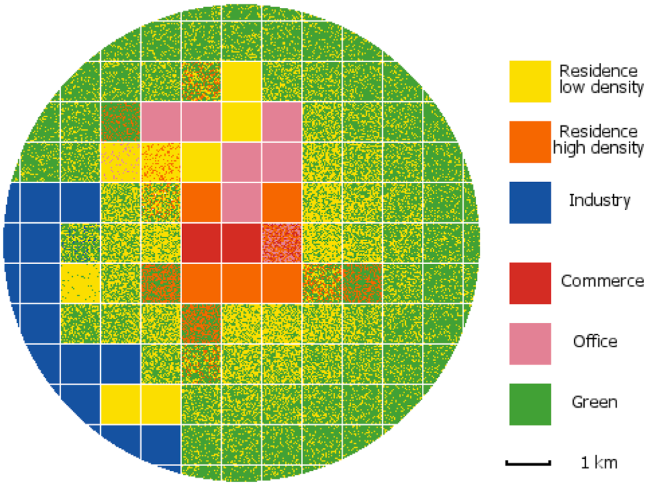
In the simulation system presented in this paper, the activities take place in a hypothetical city. Therefore, individuals' home addresses and work locations cannot be extracted from travel surveys but should be allocated by the simulation approach.

The allocation algorithm is designed based on Monte Carlo methods. The proportion of residents in a neighborhood is the probability of allocating an individual to this neighborhood. The sampled travelers are then randomly assigned to neighborhoods based on the probabilities. The allocation of work locations is similar. According to the amount of Industry and Office land uses, the employment opportunities of each neighborhood are estimated. The proportion of employment opportunity is the probability of allocating a workplace to the neighborhood.

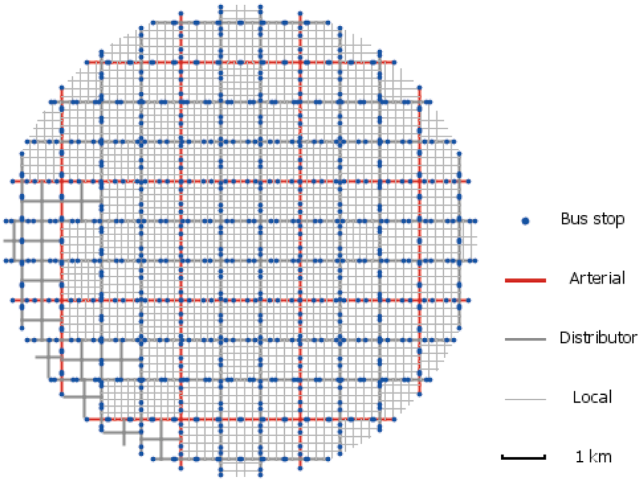
We employed the Dutch national travel survey data for travel behavior simulation. Specifically, the data of a Dutch city Eindhoven was extracted from the surveys. The travel preference information was extracted from an empirical study in the Netherlands (Arentze and Molin, 2013).

4.3.2. *Stability test*

The travel behavior simulation has some randomness, making the



(a) Land use



(b) Transport infrastructure

Fig. 6. An example of urban layout generated by the ULG model.

results fluctuate. The magnitude of the fluctuation, which illustrates the stability of the simulation, affects the credibility of the results. Thus, the acceptable magnitude range should be determined before conducting the simulation. To prove the credibility of the results, the travel behavior magnitude of fluctuation will be analyzed.

Two aspects may affect the stability of the simulation of travel behavior. One is the stability of the MSN model. Tests showed that the MSN model always generates the same result with the same input series. The other is the locations of home and work. They are randomly allocated based on the Monte Carlo method so that different runs can fluctuate.

We used the standard deviation to measure the variation of results. Ten results were simulated based on the same urban layout and the same sample of travelers, and then the standard deviation (SD) of every travel mode was calculated. Fig. 9 shows the SD of travel mode percentage with the size of travelers' sample increased from 4,000 to 30,000. When the sample size was not very large, the standard deviation of results decreased with an increase in the size. When the sample size was larger than 20,000, the SD fluctuated within a small range of less than 0.10%. Therefore, to get stable results, the sample size of simulated travelers should be no less than 20,000. We simulated 20,000 travelers in all the other cases in this paper.

4.3.3. *Land use mix*

This section provides application examples of the whole simulation approach. The application focused on land use, with the dispersion variables adjusted and other variables fixed. The fixed variables and input of the ULG and MSN models were based on the city of Eindhoven. The corresponding urban layouts had five districts, and each district had around 20 neighborhoods. According to the number of the districts, assigning 1, 3, and 5 to the *Dispersion-D-i* would generate the low, medium, and high levels of district land use mix, respectively. Similarly, assigning 1, 11, and 21 to the *Dispersion-N-i* would generate the low, medium, and high levels of neighborhood land use mix, respectively.

We employed entropy to measure the land use mix. The entropy values of districts and neighborhoods were firstly calculated. Then, to reflect the land use mix of the whole city, we calculated the area-weighted average entropy at both the district level and the neighborhood level, with the area of a particular district or neighborhood as the weight. A larger value of the entropy indicates a better-mixed land use.

One of the main aims of the simulation approach is to explore the impacts of land use mix on travel behavior. It is expected that the *Dispersion-D-i* controls district-level land use mix, the *Dispersion-N-i* controls neighborhood-level land use mix, and the land use mix affects travel behavior.

To test the *Dispersion-D-i* variables, their values increased from 1, 3, to 5, with *Dispersion-N-i* fixed at 1 in Fig. 10 and fixed at 21 in Fig. 11. The results showed that an increase of *Dispersion-D-i* would increase the entropy of both district and neighborhood. With increased entropy, the usage of cars and bikes dropped but walking increased. Compared with Fig. 10, the travel mode in Fig. 11 showed a more significant extent of variation, which implied the magnitude of the land use effects might change with land use mix degree.

To test the *Dispersion-N-i* variables, their values increased from 1, 11 to 21, with *Dispersion-D-i* fixed at 1 in Fig. 12 and set at 5 in Fig. 13. The results showed that an increase of *Dispersion-N-i* would increase the entropy on the neighborhood level. With increased neighborhood entropy, the usage of cars and bikes dropped but walking increased. The different magnitude of the land use effects was observed again. Compared with Fig. 12, the travel mode in Fig. 13 showed a more significant extent of variation.

Another main aim of the simulation approach is to explore the impacts of land use types on travel behavior. This application took the Commerce land use as an example. In the tests, only the *Dispersion-D-Commerce* and *Dispersion-N-Commerce* changed.

Fig. 14 showed that when the *Dispersion-D-Commerce* increased from 1 to 5, the entropy increased as well, and the percentage of car and bike trips dropped while walking trips increased. It suggested the allocation of commerce land use could affect travel behavior. The magnitude of district-level Commerce's effects changed, which is larger when the *Dispersion-D-Commerce* increased from 1 to 3.

Fig. 15 showed that when the *Dispersion-N-Commerce* increased from 1 to 21, the entropy of the neighborhood grew, and the percentage of car and bike trips dropped while walking trips increased. It could also be observed that the magnitude of neighborhood-level Commerce's effects is larger when the *Dispersion-N-Commerce* was risen from 1 to 11 and is smaller when increased from 11 to 21.

Although the simple examples above aim to show the application of the ULG model and the simulation approach, the results of the examples provide helpful information. First, the results support the assumption that an increase in land use mix brings a decrease in car usage and an increase in active transport. Second, the elasticity of travel mode with respect to land use mix is not fixed. The marginal effect of the land use mix showed a decreasing trend. Third, the land use mix at district and neighborhood levels influences travel behavior. Forth, different land use types may have a different magnitude of impact.

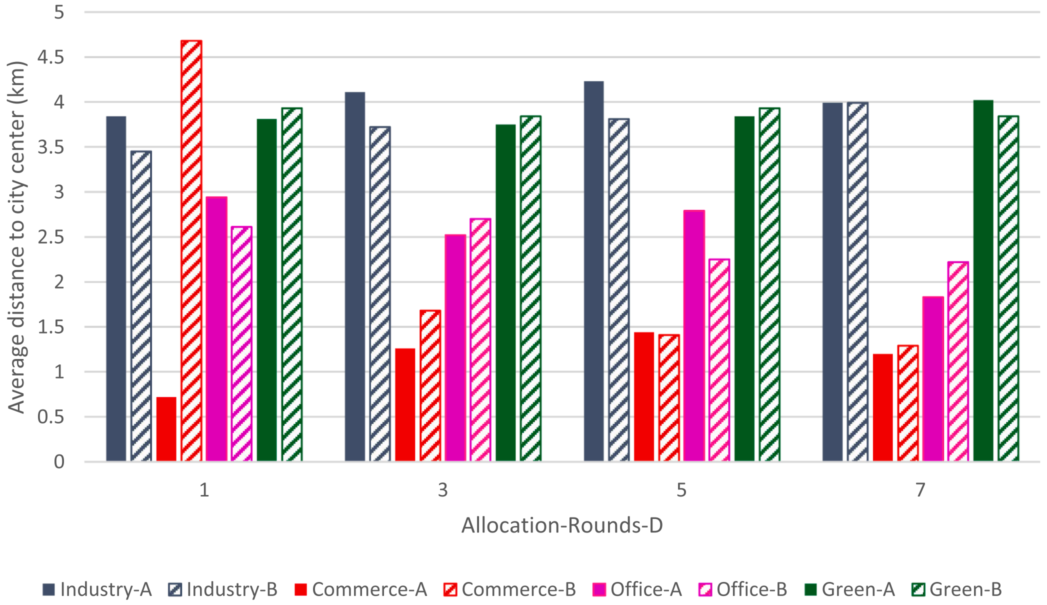


Fig. 7. The Average-distance-to-city-center of land uses under different Allocation-Rounds-D.

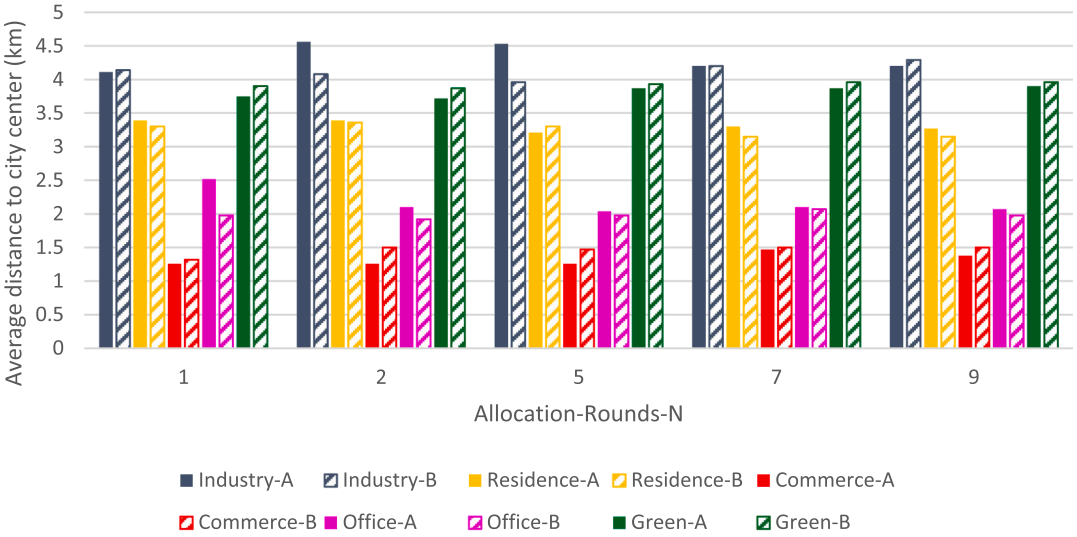


Fig. 8. The Average-distance-to-city-center of land uses under different Allocation-Rounds-N.

5. Discussion and conclusions

This paper proposes the Urban Layout Generation model that is newly developed to explore land use impacts on travel behavior via a simulation framework. The ULG model generates hypothetical urban layouts described by land use measurements, specifically the D-variables widely used in the land use – travel behavior studies. The urban layouts are then inputted into a travel behavior simulation model to simulate the corresponding travel behavior. Thus, the land use measurements and the travel behavior are correlated in the simulation.

The function of the ULG model in the simulation framework does not aim to copy an actual city and predict its future development, which distinguishes the ULG model from the LUTI models. The ULG model is closer to the procedural urban models whose aim is to generate virtual cities. The main difference between the two models is that the procedural models simulate the whole evolutionary history of the virtual city. In contrast, the ULG model only aims at the final urban layout. Therefore, the ULG model is more suitable for planning a new city or testing new ideas of land use planning.

The aim of the ULG model is generally achieved by proving that the ULG model can well control most of the D-variables. The *Density,* which refers to population density, is a new component in the ULG model because most previous research takes it as input or does not consider it. The ULG model controls the *Density* by selecting the population density function with appropriate parameter values and then interacting with neighborhood residential land use during the land allocation process. Accordingly, the ULG model can control the *Density* in districts but with some randomness in neighborhoods. The *Design,* which refers to the road network, is controlled by selecting the road pattern with appropriate parameter values of the road network. The *Distance-to-transit,* which

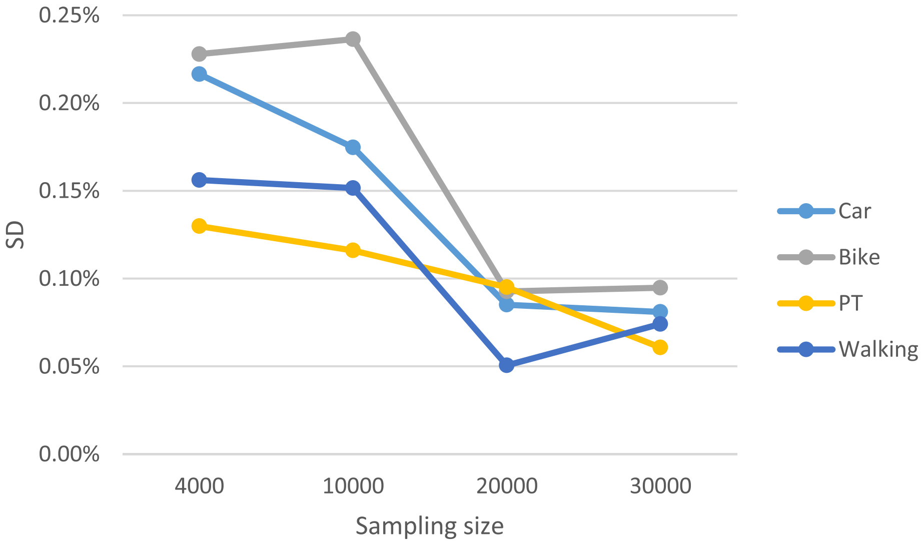


Fig. 9. The standard deviation of simulated results with different sampling sizes.

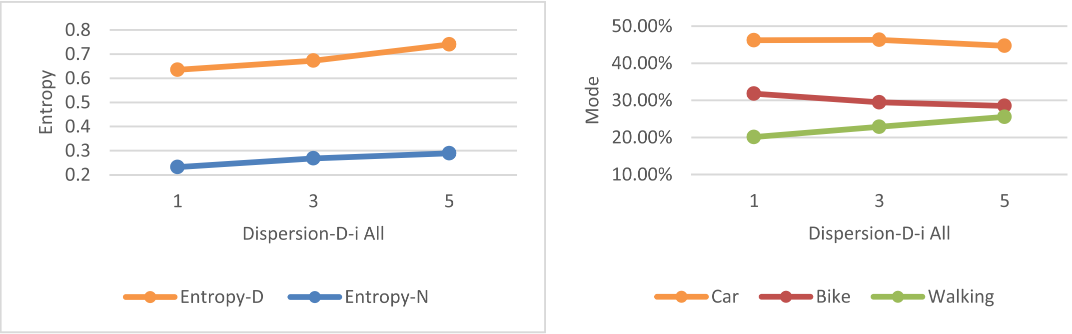


Fig. 10. Effects of Dispersion-D-i (Dispersion-N-i = 1, left: Entropy, right: Travel Mode).

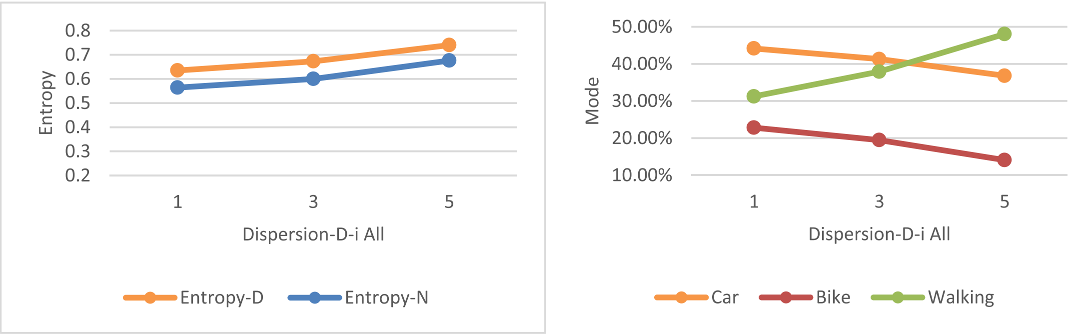


Fig. 11. Effects of Dispersion-D-i (Dispersion-N-i = 21, left: Entropy, right: Travel Mode).

refers to public transport, is controlled by entering the average distance between bus stops.

The *Diversity*, which refers to land use mix, however, is hard to control. This is because entropy (Diversity) value can be transformed into various land use compositions. Therefore, *Diversity* is indirectly controlled in the ULG model by holding the dispersion of each land use type. The ULG model can generate a city with an increased or decreased entropy value but cannot generate a specific value directly. To generate a definite value of entropy, several runs are needed. *Diversity* is also affected by the allocation order of land uses. However, this effect can be almost eliminated with proper values of the Allocation-Rounds variables.

Besides generating desired urban layouts corresponding to the D-variables, the ULG model can control the D-variables both on district-level and neighborhood-level, enabling to execute studies at different spatial scales.

Exploring how the D-variables affect travel behavior is the core application of the ULG model – keeping all D-variables fixed except those whose impact will be analyzed. Section 4.4 showed simple examples focusing on the land use mix. Another important potential

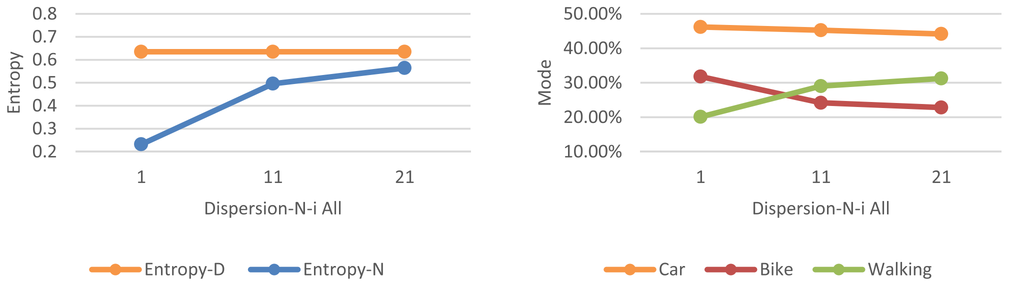


Fig. 12. Effects of Dispersion-N-i (Dispersion-D-i = 1, left: Entropy, right: Travel Mode).

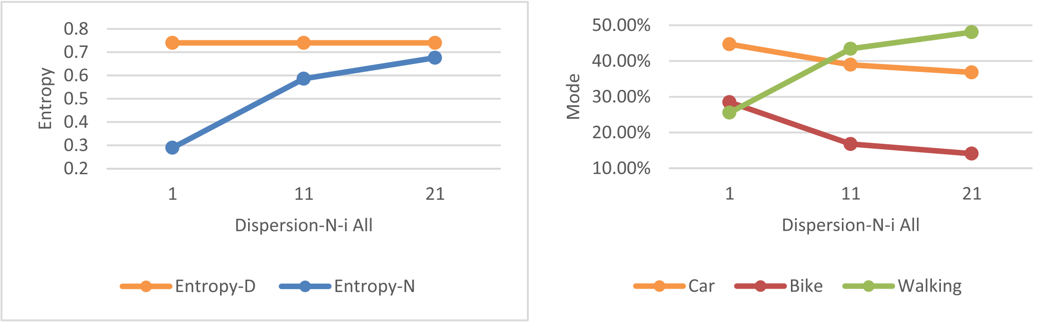


Fig. 13. Effects of Dispersion-N-i (Dispersion-D-i = 5, left: Entropy, right: Travel Mode).

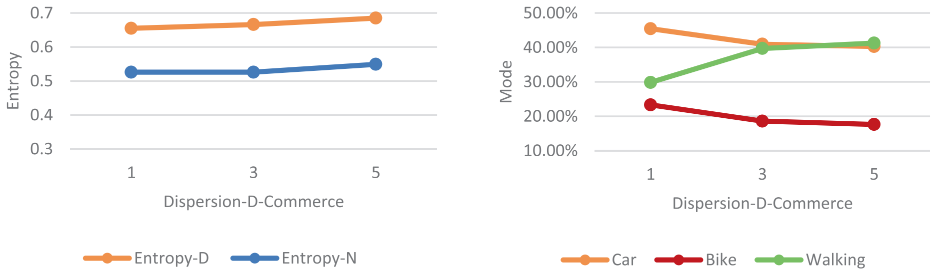


Fig. 14. Effects of Dispersion-D-Commerce (Dispersion-D-i = 3, Dispersion-N-i = 11, left: Entropy, right: Travel Mode).

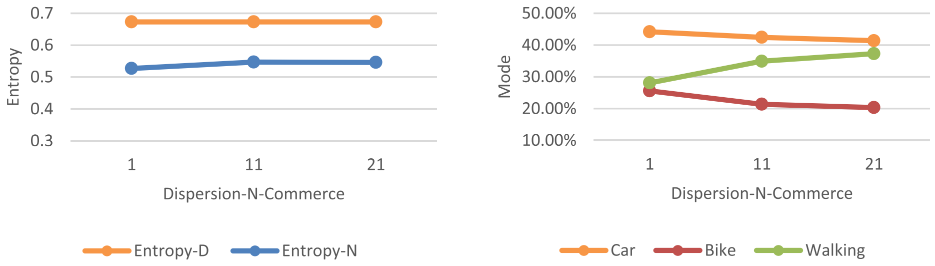


Fig. 15. Effects of Dispersion-N-Commerce (Dispersion-D-i = 3, Dispersion-N-i = 11, left: Entropy, right: Travel Mode).

application of the ULG model is land use optimization. A simple application on land use mix already showed nonlinearity and complexity. Great complexity can be expected when integrating land use, population density, road network, and public transport system. The ULG model and the simulation framework can contribute to exploring typical patterns of urban forms that aim at achieving objectives such as decreasing vehicle dependency or encouraging active transport.

Applying the ULG model can provide credible insights for testing different land use-travel behavior scenarios in real-world cities. To test land use-travel behavior scenarios, this paper presents a novel ULG model for generating different urban layouts. The credibility of a model should be examined by verification, calibration, and validation (Wilensky and Rand, 2015). Verification tests the correctness of the model itself, specifically the design and algorithm of the model. In this paper, the framework, requirements, and algorithms of the ULG model are verified. Additionally, tests and applications are presented to demonstrate the model's performance.

Before being applied to simulation studies, the ULG model needs to be calibrated and validated in its application context. We intend to present this in future publications.

In this paper, the ULG model was applied to a medium-sized city. The used population density functions, road network, and bus system are suitable for a medium city. To generate reasonable metropolitan layouts, more population density functions (urban spatial structure), road network patterns, and transport infrastructures such as railway, subway, tram, and airport will be added in the future.

CRediT authorship contribution statement

Xiaoming Lyu: Conceptualization, Methodology, Software, Writing – original draft. **Qi Han:** Supervision, Writing – review & editing. **Bauke de Vries:** Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The author Xiaoming Lyu was sponsored by China Scholarship Council [grant number 201306620006]. The authors would like to thank Feixiong Liao's support for the MSN model.

References

Aliaga, D.G., Vanegas, C.A., Beneš, B., 2008. Interactive Example-Based Urban Layout Synthesis. ACM Trans. Graph., 27(5), 160:1–160:10. 10.1145/1409060.1409113.

Arentze, T.A., Molin, E.J.E., 2013. Travelers' preferences in multimodal networks: Design and results of a comprehensive series of choice experiments. Transp. Res. Part A: Policy Pract. 58, 15–28. https://doi.org/10.1016/j.tra.2013.10.005.

Beneš, J., Wilkie, A., Křivánek, J., 2014. Procedural Modelling of Urban Road Networks. Comput. Graphics Forum 33 (6), 132–142. https://doi.org/10.1111/cgf.12283.

Boarnet, M.G., 2011. A Broader Context for Land Use and Travel Behavior, and a Research Agenda. J. Am. Plann. Assoc. 77 (3), 197–213. https://doi.org/10.1080/01944363.2011.593483.

Brownstone, D., 2008. Key Relationships Between the Built Environment and VMT.

Cao, X., Mokhtarian, P.L., Handy, S.L., 2009. Examining the impacts of residential self-selection on travel behaviour: A focus on empirical findings. Transp. Rev. 29 (3), 359–395.

Cervero, R., Kockelman, K., 1997. Travel Demand and The 3Ds: Density, Diversity, and Design. Transp. Res. Part D: Transp. Environ. 2 (3), 199–219. https://doi.org/10.1016/S1361-9209(97)00009-6.

Chen, G., Esch, G., Wonka, P., Müller, P., Zhang, E., 2008. Interactive procedural street modeling. ACM Trans. Graph., 27(3), 103:1-10. 10.1145/1360612.1360702.

Crane, R., 2000. The Influence of Urban Form on Travel: An Interpretive Review. J. Plann. Literature 15 (1), 3–23. https://doi.org/10.1177/08854120022092890.

Ding, C., Wang, Y., Tang, T., Mishra, S., Liu, C., 2018. Joint analysis of the spatial impacts of built environment on car ownership and travel mode choice. Transp. Res. Part D: Transp. Environ. 60, 28–40. https://doi.org/10.1016/j.trd.2016.08.004.

Ewing, R., Cervero, R., 2010. Travel and the Built Environment: A Meta-Analysis. J. Am. Plann. Assoc. 76 (3), 265–294. https://doi.org/10.1080/01944361003766766.

Frank, L.D., Pivo, G., 1994. Impacts of Mixed Use and Density on Utilization of Three Modes of Travel: Single-Occupant Vehicle, Transit, and Walking. Transp. Res. Rec. 1466, 44–52. https://doi.org/10.1073/pnas.1100480108.

Handy, S., 1996. Methodologies for exploring the link between urban form and travel behavior. Methodol. Explor. Link Urban Form Travel Behav. 1 (2), 151–165.

Handy, S.L., Boarnet, M.G., Ewing, R., Killingsworth, R.E., 2002. How the built environment affects physical activity: Views from urban planning. Am. J. Prev. Med. 23 (2), 64–73. https://doi.org/10.1016/S0749-3797(02)00475-0.

Hong, J., Shen, Q., Zhang, L., 2014. How do built-environment factors affect travel behavior? A spatial analysis at different geographic scales. Transportation 41 (3), 419–440. https://doi.org/10.1007/s11116-013-9462-9.

Klosterman, R.E., 2001. The What If ? Planning Support System. Planning Support Systems: Integrating Geographic Information Systems, Models, and Visualization Tools, 263–284.

Lay, M.G., 2009. Handbook of Road Technology, 4th ed. Spon Press, Abingdon.

Lechner, Thomas, Watson, B., Ren, P., Wilensky, U., Tisue, S., Felsen, M., 2004. Procedural Modeling of Land Use in Cities.

Lechner, Thomas, Watson, B., Wilensky, U., Felsen, M., 2003. Procedural city modeling. In: 1st Midwestern Graphics Conference. St. Louis, MO, USA.

Lechner, T., Watson, B., Wilenski, U., Tisue, S., Felsen, M., Moddrell, A., Brozefsky, C., 2007. Procedural modeling of urban land use. North Carolina State University, Dept. of Computer Science.

Li, J., Ryohei, N., 2006. An Econometric Study on the Urban Population Density Functions: A Survey. Urban Plann. Int. 21 (1), 40–47.

Liao, F., Arentze, T., Molin, E., Bothe, W., Timmermans, H., 2017. Effects of land-use transport scenarios on travel patterns: a multi-state supernetwork application. Transportation 44 (1), 1–25. https://doi.org/10.1007/s11116-015-9616-z.

Liao, F., Arentze, T., Timmermans, H., 2010. Supernetwork Approach for Multimodal and Multiactivity Travel Planning. Transp. Res. Rec.: J. Transp. Res. Board 2175, 38–46. https://doi.org/10.3141/2175-05.

Liao, F., Arentze, T., Timmermans, H., 2013. Incorporating space-time constraints and activity-travel time profiles in a multi-state supernetwork approach to individual activity-travel scheduling. Transp. Res. Part B: Methodol. 55, 41–58. https://doi.org/10.1016/j.trb.2013.05.002.

Milakis, D., Cervero, R., Van Wee, B., 2015. Stay local or go regional? Urban form effects on vehicle use at different spatial scales: A theoretical concept and its application to the San Francisco Bay Area. J. Transp. Land Use 8 (2), 59–86. https://doi.org/10.5198/jtlu.2015.557.

Nasri, A., Zhang, L., 2015. Assessing the Impact of Metropolitan-Level, County-Level, and Local-Level Built Environment on Travel Behavior: Evidence from 19 U.S. Urban Areas. J. Urban Plann. Dev. 141 (3), 04014031. https://doi.org/10.1061/(ASCE)UP.1943-5444.0000226.

Parish, Y.I.H., Müller, P., 2001. In: Procedural Modeling of Cities. ACM Press, Los Angeles, CA, USA, pp. 301–308. https://doi.org/10.1145/383259.383292.

Pettit, C.J., Klosterman, R.E., Delaney, P., Whitehead, A.L., Kujala, H., Bromage, A., Nino-Ruiz, M., 2015. The Online What if? Planning Support System: A Land Suitability Application in Western Australia. Appl. Spatial Anal. Policy 8 (2), 93–112. https://doi.org/10.1007/s12061-015-9133-7.

Salas-Olmedo, M.H., Wang, Y., Alonso, A., 2017. Assessing accessibility with local coefficients for the LUTI model MARS. Comput. Environ. Urban Syst. 64, 194–203. https://doi.org/10.1016/j.compenvurbsys.2017.02.004.

Spears, S., Boarnet, M. G., Handy, S., Rodier, C., 2014. Policy Brief on the Impacts of Land-Use Mix on Passenger Vehicle Use and Greenhouse Gas Emissions.

Wang, Y., Monzon, A., Ciommo, F.D., 2015. Assessing the accessibility impact of transport policy by a land-use and transport interaction model - The case of Madrid. Comput. Environ. Urban Syst. 49, 126–135. https://doi.org/10.1016/j.compenvurbsys.2014.03.005.

Weber, B., Müller, P., Wonka, P., Gross, M., 2009. Interactive Geometric Simulation of 4D Cities. Comput. Graphics Forum 28 (2), 481–492. https://doi.org/10.1111/j.1467-8659.2009.01387.x.

Wilensky, U., Rand, W., 2015. An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo. MIT Press.