Validation of a simulation system exploring land-use impacts on travel behavior

Xiaoming Lyu, Qi Han, Bauke de Vries

Abstract

This paper aims to validate a simulation system for exploring the land-use impacts on travel behaviour. The system simulates urban layout and travel behaviour. The key parameters and input for simulating urban layout are calibrated by Eindhoven, and a set of hypothetic cities are then generated to validate the travel behaviour, with the measurement of entropy employed measuring the land use pattern. The results show that the simulation system can generate a variety of land use patterns, and the different travel behaviour of the land use patterns are captured. The results also show that when the land use patterns are similar to a real city pattern, their travel behaviour is similar as well. The system is proved to be a helpful simulation tool to explore the impacts of land use on travel behaviour.

1. Introduction

To relieve the traffic congestion in cities and to reduce the energy consumption and transport emissions, people try to reduce the vehicle travel and encourage active travel (walking and cycling) and public transport. Besides traffic policies and management, the land use policies are believed helpful to affect travel behaviour. For example, mixing land use could reduce vehicle use and encourage active travel (Ewing and Cervero, 2010; Spears et al., 2014).

* Corresponding author. Tel.: +31-40-2472754.
E-mail address: x.lyu@tue.nl
The causal effects of land use (also named built environment or urban form) on travel behaviour received more concern (Ding et al., 2016). In many related studies, the travel-behaviour variables are regressed on the independent land-use variables and sociodemographic variables (Boarnet, 2011).

\[
Travel - Behavior - Variable = \beta_0 + Land - Use - Variables \cdot \beta_1 + Sociodemographic - Variables \cdot \beta_2 + \varepsilon
\]  

(1)

Despite a large amount of existed studies, there is little consensus due to the different empirical contexts, methodologies, geographic scale, and residential self-selection (Ding et al., 2016).

Most studies have measured the land-use variables at the neighbourhood level (Boarnet, 2011). Only a few research has compared different geographic scales (Hong et al., 2014; Milakis et al., 2015; Nasri and Zhang, 2015). Results suggest that measures of land use variables at a regional or metropolitan level usually have a larger magnitude of correlation with travel behaviour than those at the neighbourhood level (Boarnet, 2011; Ewing and Cervero, 2010). This is because of a good portion of long trips in metropolitan areas (Boarnet, 2011). Thus, the metropolitan-level scale is more important.

The residential self-selection means that people choose their residential locations based on their travel preferences, which implies that the effect of land use on travel is not wholly causal. The ideal method for solving the self-selection problem is to randomly assign persons to different locations and then observe differences in their travel behaviour, which is rarely realistic (Brownstone, 2008). One of the best alternatives is to collect panel data following households over time, but the data are very expensive (Brownstone, 2008); another is to integrate models of location choice, car ownership, and travel behaviour, but the nonlinearity in the joint models makes simultaneous estimation complex (Boarnet, 2011). Besides the residential self-selection, the self-selection of employment locations is also important.

Simulation method has the potential to solve the self-selection problem perfectly that is hard for empirical studies. In a simulation, the residents can be randomly assigned to locations of residential land use, and their workplaces may also be allocated randomly to employment locations. This is the ideal way mentioned by Brownstone. Besides, it is convenient for a simulation system to test the possibly non-linear impacts of the land-use variables on travel behaviour by keeping other variables fixed, which is almost impossible for empirical research.

The objective of this paper is to validate a simulation system designed to explore the impacts of land use on travel behaviour. After a brief description of the system, it is calibrated by Eindhoven and then the results are validated. The system is proved as a useful tool for exploring the impacts of land use on travel behaviour.

2. Simulation system

2.1. Overview of the simulation framework

The basic idea of the simulation research is to simulate hypothetic urban layouts and their travel behaviour, and then to explore the impacts of land use on travel behaviour via analysing the simulated results, see Figure 1. The simulation system consists of a sub-system of urban-layout simulation and another sub-system of travel-behaviour simulation. A multi-state supernetwork model (MSN model) developed by Liao et al. (Liao et al., 2010, 2013) was employed to simulate the travel behaviour; an urban-layout simulation module was programmed on the NetLogo platform. In this section, we briefly describe the simulation system, especially the urban-layout module, because the system has many new advancements since its early prototype (Lyu et al., 2017).

The hypothetic urban layout consists of three layers: population density, land use, and road network. The hierarchy of road network includes Highway, Arterial, Distributor, and Local, which divide a city into districts and neighbourhoods. A district is defined as a spatial unit surrounded by arterials and highways, and a neighbourhood is surrounded by distributors. Five types of land use are considered: Residence, Commerce, Office, Industry, Green&Open.
Table 1 shows the input and key parameters of the urban-layout generator. With the *Urban Area*, the area of the simulated city, a clear urban built area could be generated firstly (Figure 2).

Table 1. The input and key parameters of the urban-layout generator

<table>
<thead>
<tr>
<th>Input variable</th>
<th>Function</th>
<th>Input variable</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban Area</td>
<td>Global</td>
<td>Land use Amount</td>
<td>Land use allocation</td>
</tr>
<tr>
<td>Population</td>
<td>Population</td>
<td>Iteration</td>
<td>Land use allocation</td>
</tr>
<tr>
<td>Distance-Arterial</td>
<td>Road network</td>
<td>Dispersion-D-i</td>
<td>Land use i at districts</td>
</tr>
<tr>
<td>Distance-Distributor</td>
<td>Road network</td>
<td>Dispersion-N-i</td>
<td>Land use i at neighbourhoods</td>
</tr>
</tbody>
</table>

Fig. 1. The framework of the simulation system.

Fig. 2. The flowchart of urban-layout simulation.
2.2. Simulation of population

To simulate population distribution, users first input the Population of the city, and then select the Density-Model from a population density model library and input its parameters to generate an ideal population map. The map would then interact with the land-use allocation module to generate a realistic population density map.

2.3. Simulation of road network

The Highway was not obligatory to build. The construction of Arterial and Distributor was controlled by parameters of road span, the Distance-Arterial and Distance-Distributor, as well as a road pattern selected from a road-pattern library. The Arterial and Distributor formed a skeleton of the road network and divided the city into Districts and Neighbourhoods. After the land use allocation on neighbourhoods, the Local was built within each neighbourhood, interacting with land use. The density of Local was higher in neighbourhoods with a large percent of the Residence, Office, and Commerce, and lower with the Industry and Green&Open.

Based on the road network layer, a public transport system was built. The bus stations were allocated along with arterials and distributors, controlled by a parameter of station span. Another parameter controlled the time span to form the timetable of buses.

2.4. Simulation of land use

The land use was firstly allocated to districts and then to neighbourhoods. The algorithm of allocation was based on the What-If? land-use planning support system (Klosterman, 2001), consisting of two steps: to calculate land use suitability of each spatial unit, and to allocate land use to areas with the highest suitability. The suitability was calculated by the following formula:

\[ S_{ki} = \sum_j w_{kj} r_{ij} \]  

where \( S_{ki} \) is the suitability of land unit \( i \) for land use \( k \), \( w_{kj} \) is the weight of factor \( j \) for land use \( k \), and \( r_{ij} \) is the rating of land unit \( i \) on factor \( j \).

The parameter Iteration (see Table 1) was used to control the degree of interaction among land uses in the allocating process. The parameters Dispersion-D-i and Dispersion-N-i controlled the compactness of every land use, at the district level or neighbourhood level.

The allocation of the Residence was an exception. The Residence was directly allocated at the district level according to the population of each district, not calculating the suitability. At the neighbourhood level, however, the allocation was similar with other land uses. If the available land of a neighbourhood was not enough for the residents, the surplus people were then moved to other neighbourhoods with high residence-suitability, adjusting the population density map simultaneously.

2.5. Simulation of travel behaviour

The MSN model was employed to simulate the travel behaviour. The MSN model integrates the networks of transport, land uses as locations of facilities and services, and individuals’ activity programs, which could model the travel mode, route, and activity location. The suitability of the MSN model for evaluating land use-transport scenarios was demonstrated by an application in Rotterdam (Liao et al., 2017).

The activity program, an input of the MSN model, was extracted from real travel-survey data. In hypothetic cities, the home and work locations in the activity program needed to be allocated in accordance with the land use condition. The home locations were allocated based on population density, while the work locations were based on employment density estimated by the land use of Office and Industry, see Figure 3.
3. Data

The city of Eindhoven was selected to calibrate and validate the simulation system. Eindhoven is a Dutch city, with an area of 89 km² and a population of 223,220 in the year 2012. The population density of Eindhoven came from the *Wijk-en Buurtkaart 2012* (District and Neighbourhood Map), the land use was extracted from the *Bestand Bodemgebruik 2012* (Soil usage file), the road network was drawn from the *Nationaal Wegen Bestand 2011* (National Roads File), and the public transport data came from 9292 (a platform providing travel information on public transport in the Netherlands).

The activity programs were extracted from the *OViN 2011 – 2013* (Dutch national travel-diary surveys). The travellers of Eindhoven were selected, and the inter-city trips were excluded. The activities or trip-motivations of OViN were recoded, to be in accordance with the land use map. The adjustment was shown in Table 2.

<table>
<thead>
<tr>
<th>OViN code</th>
<th>Explanation OViN</th>
<th>Recode to land use</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>To work</td>
<td>Industry or Office</td>
</tr>
<tr>
<td>2</td>
<td>Business visit</td>
<td>Industry or Office</td>
</tr>
<tr>
<td>3</td>
<td>Transport profession</td>
<td>Industry or Office</td>
</tr>
<tr>
<td>4</td>
<td>Pick up / bring people</td>
<td>Industry or Office</td>
</tr>
<tr>
<td>5</td>
<td>Pick up / bring goods</td>
<td>Industry or Office</td>
</tr>
<tr>
<td>6</td>
<td>Education</td>
<td>Office</td>
</tr>
<tr>
<td>7</td>
<td>Shopping</td>
<td>Commerce</td>
</tr>
<tr>
<td>8</td>
<td>Visit / stay</td>
<td>Residence</td>
</tr>
<tr>
<td>9</td>
<td>Tours / hiking</td>
<td>Green &amp; Open</td>
</tr>
<tr>
<td>10</td>
<td>Sports / hobby</td>
<td>Green &amp; Open</td>
</tr>
<tr>
<td>11</td>
<td>Other leisure</td>
<td>Green &amp; Open</td>
</tr>
<tr>
<td>12</td>
<td>Personal services</td>
<td>Commerce</td>
</tr>
<tr>
<td>13</td>
<td>Others</td>
<td>Random</td>
</tr>
</tbody>
</table>
4. Calibration and validation

Based on the regression model of formula (1), it is reasonable to assume that the same group of travellers (same values of sociodemographic variables) would have the same pattern of travel behaviour in different cities with a same land-use pattern (same values of land-use variables). Therefore if a hypothetic city can capture key land-use variables of a real city, its travel behaviour would be reasonably similar to the real travel behaviour, and then the result of land-use impacts on travel behaviour, which is concluded from simulated results, would be credible.

The strategy of validation was to demonstrate that the simulation system could generate hypothetic cities which captured key land-use variables of real cities, and the simulated travel behaviour was also close to the real situation. Eindhoven was selected as a benchmark, with the urban layout for calibration and the travel behaviour for validation.

4.1. Calibration

With the population density of Eindhoven, a curve estimation was processed with SPSS, see Table 3. The Quadratic model was chosen to simulate population.

<table>
<thead>
<tr>
<th>Model</th>
<th>Curve Estimation</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clark model</td>
<td>$D(x) = 25053.9e^{-0.84x}$</td>
<td>0.645</td>
</tr>
<tr>
<td>Newling model</td>
<td>$D(x) = 5065.8e^{-0.056x-0.033x^2}$</td>
<td>0.911</td>
</tr>
<tr>
<td>Quadratic model</td>
<td>$D(x) = 5238.2 - 721.8x + 9.2x^2$</td>
<td>0.922</td>
</tr>
</tbody>
</table>

The length of arterials in Eindhoven was the criteria to value the Distance-Arterial, while both of the length of distributors and the number of neighbourhoods were considered to set the value of Distance-Distributor because the neighbourhoods were formed by the distributors. The values of Land use Amount were calculated by Eindhoven’s land use map. The key input values were shown in table 4, and the calibrated results were shown in table 5.

<table>
<thead>
<tr>
<th>Input variable</th>
<th>Input way</th>
<th>Input variable</th>
<th>Input way</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area</td>
<td>89 km²</td>
<td>Residential Amount</td>
<td>35.23%</td>
</tr>
<tr>
<td>Population</td>
<td>223220</td>
<td>Industrial Amount</td>
<td>10.27%</td>
</tr>
<tr>
<td>Density-Model</td>
<td>$D(x) = 5238.2 - 721.8x + 9.2x^2$</td>
<td>Commercial Amount</td>
<td>2.04%</td>
</tr>
<tr>
<td>Distance-Arterial</td>
<td>2.7 km</td>
<td>Office Amount</td>
<td>5.86%</td>
</tr>
<tr>
<td>Distance-Distributor</td>
<td>0.9 km</td>
<td>Green&amp;Open Amount</td>
<td>46.60%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measures</th>
<th>Eindhoven</th>
<th>Hypothetic city</th>
</tr>
</thead>
<tbody>
<tr>
<td>Districts</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Neighbourhoods</td>
<td>116</td>
<td>113</td>
</tr>
<tr>
<td>Length-Arterial (km)</td>
<td>68.4</td>
<td>69.1</td>
</tr>
<tr>
<td>Length-Distributor (km)</td>
<td>153.8</td>
<td>154.5</td>
</tr>
</tbody>
</table>

4.2. Validation

The key variables of population and road network had been calibrated, with land use variables of Dispersion-D-i and Dispersion-N-i to be tested in this section (see Table 6, the suffixes of -R, -I, -C, -O, and -G represent the land use of residence, industry, commerce, office, and green). The two groups of variables control the compactness of land uses at the district level and the neighbourhood level, together controlling the land use pattern of hypothetic cities.
The testing strategy was to generate different land use patterns and then to see how the land use pattern affect travel behaviour. We used entropy to measure the land use pattern. The entropy is a widely used variable measuring land use mix, as well as an important variable in the regression land use – travel behaviour model. The Entropy-D and Entropy-N (see Table 6), the weighted average entropy at the district level and neighbourhood level, were employed to measure the degree of land use mix.

To test the effect of land use mix over all, cities of lowest land use mix, mid-range mix, and highest mix were generated. In H1, H2, and H3 the Dispersion-D-i and Dispersion-N-i increased simultaneously, resulting in an obvious increase of Entropy-D and Entropy-N. With the increase of land use mix, the car-use dropped obviously from 59.36% to 45.86%, while the active travel (biking and walking) increased steadily, although the cycling dropped and the walking increased.

To test the change in land use pattern at different levels, the H4, H5, and H6 were generated. From H4 to H2 and then to H5, the Dispersion-D-i increased, with Dispersion-N-i stable. Both the entropies increased. So the Dispersion-D-i affected the district-level and the neighbourhood-level land use mix. From H6 to H5 and then to H3, the Dispersion-D-i kept stable, with Dispersion-N-i increased. This time the Entropy-D stayed, while the Entropy-N increased. So the Dispersion-N-i only affect the neighbourhood-level land use mix. Moreover, the same trend appeared that when land use mix increased, the car-use decreased and active transport increased.

To test the impact of residence and employment, the H7, H8, H9, and H10 were generated. In H1, H7, and H8, only the Dispersion-N-R increased. It was interesting that the impact of residence was nonlinear: with an increasing mix of residence, the car-use dropped first and then increased, while the active transport increased first and then dropped. However, from H9 to H2 and then to H10, the dispersion of the industry and office increased (Dispersion-D-I, Dispersion-N-I, Dispersion-D-O and Dispersion-N-O), making the car-use dropped and the active transport increased.

### Table 6. Urban layout and travel behaviour of Eindhoven and hypothetic cities.

<table>
<thead>
<tr>
<th>Control variables</th>
<th>Eindhoven</th>
<th>H1</th>
<th>H2</th>
<th>H3</th>
<th>H4</th>
<th>H5</th>
<th>H6</th>
<th>H7</th>
<th>H8</th>
<th>H9</th>
<th>H10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dispersion-D-I</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Dispersion-D-C</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Dispersion-D-O</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Dispersion-D-G</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Dispersion-N-R</td>
<td>1</td>
<td>6</td>
<td>11</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>6</td>
<td>11</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Dispersion-N-I</td>
<td>1</td>
<td>6</td>
<td>11</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Dispersion-N-C</td>
<td>1</td>
<td>6</td>
<td>11</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Dispersion-N-O</td>
<td>1</td>
<td>6</td>
<td>11</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Dispersion-N-G</td>
<td>1</td>
<td>6</td>
<td>11</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Urban layout</td>
<td>Eindhoven</td>
<td>H1</td>
<td>H2</td>
<td>H3</td>
<td>H4</td>
<td>H5</td>
<td>H6</td>
<td>H7</td>
<td>H8</td>
<td>H9</td>
<td>H10</td>
</tr>
<tr>
<td>Entropy-D</td>
<td>0.632</td>
<td>0.406</td>
<td>0.491</td>
<td>0.512</td>
<td>0.424</td>
<td>0.512</td>
<td>0.512</td>
<td>0.449</td>
<td>0.438</td>
<td>0.473</td>
<td>0.502</td>
</tr>
<tr>
<td>Entropy-N</td>
<td>0.357</td>
<td>0.152</td>
<td>0.325</td>
<td>0.392</td>
<td>0.225</td>
<td>0.377</td>
<td>0.368</td>
<td>0.184</td>
<td>0.213</td>
<td>0.245</td>
<td>0.359</td>
</tr>
<tr>
<td>Travel behaviour</td>
<td>Eindhoven</td>
<td>H1</td>
<td>H2</td>
<td>H3</td>
<td>H4</td>
<td>H5</td>
<td>H6</td>
<td>H7</td>
<td>H8</td>
<td>H9</td>
<td>H10</td>
</tr>
<tr>
<td>Mode-car</td>
<td>46.30%</td>
<td>59.36%</td>
<td>49.47%</td>
<td>45.86%</td>
<td>57.30%</td>
<td>47.81%</td>
<td>48.04%</td>
<td>56.09%</td>
<td>57.87%</td>
<td>49.13%</td>
<td>49.61%</td>
</tr>
<tr>
<td>Mode-bike</td>
<td>32.04%</td>
<td>20.58%</td>
<td>19.52%</td>
<td>15.43%</td>
<td>20.11%</td>
<td>16.68%</td>
<td>17.94%</td>
<td>21.42%</td>
<td>19.37%</td>
<td>18.43%</td>
<td>20.77%</td>
</tr>
<tr>
<td>Mode-PT</td>
<td>0.75%</td>
<td>1.23%</td>
<td>0.74%</td>
<td>0.69%</td>
<td>1.16%</td>
<td>0.64%</td>
<td>0.61%</td>
<td>0.96%</td>
<td>1.05%</td>
<td>0.65%</td>
<td>0.71%</td>
</tr>
<tr>
<td>Mode-walk</td>
<td>20.90%</td>
<td>18.83%</td>
<td>30.27%</td>
<td>38.01%</td>
<td>21.43%</td>
<td>34.87%</td>
<td>33.42%</td>
<td>21.53%</td>
<td>21.72%</td>
<td>31.79%</td>
<td>28.91%</td>
</tr>
<tr>
<td>Bike &amp; Walk</td>
<td>52.94%</td>
<td>39.41%</td>
<td>49.79%</td>
<td>53.44%</td>
<td>41.54%</td>
<td>51.55%</td>
<td>51.36%</td>
<td>42.95%</td>
<td>41.09%</td>
<td>50.22%</td>
<td>49.68%</td>
</tr>
</tbody>
</table>
Compared Eindhoven with hypothetic cities, although no cities had exactly same entropies with Eindhoven, it was obvious that when the entropies were close to those of Eindhoven, the travel behaviour was also close to Eindhoven’s case. The cycling and walking, when separately, were not very precise to Eindhoven, but when together as the active transport, the results became much better.

5. Conclusion

The objective of this paper is to validate the simulation system. The process of calibration and validation provided an inspection of how the system works. The results showed that the system could generate a variety of land use patterns, and when the land use patterns of hypothetic cities were similar to those of real cities, the travel behaviour would be similar as well.

From the series of urban layouts, the impacts of land use pattern on travel behaviour could be observed. The impacts were somehow nonlinear and complex. The system presented in this paper is a useful tool that can help to improve our understanding of how does the land use affect the travel behaviour.

It is recognized that the following work for further research should be conducted. First, only a few variables were initially tested in this study, a detailed sensitivity analysis needs to be applied. Second, more measures of land use pattern besides land use mix, for example the concentration, clustering, and proximity (Galster et al., 2001), should be introduced to the study. Third, this paper tested the system via only one city. To get more robust results, the system should be tested in more cases.

References


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


