The first agent-based model of greater Jakarta integrated with a mode-choice model

Anugrah Ilahi*, Milos Balaca, Aoyong Li, Kay. W. Axhausen

*ETH Zurich, Stefano-Franscini-Platz 5, 8093 Zurich, Switzerland

Abstract

Agent-based models became popular in recent decades, as more traditional modeling approaches were not suitable to model emerging transportation modes and transport policies. However, creating a detailed scenario, which is the basis of the agent-based model and that represents the supply and the demand for a study area, is not trivial and requires substantial effort. This is even more the case when the study area is large and one that contains a large number of people living and performing activities in it. This is exactly the case for the region of Greater Jakarta, which is the subject of this research, with an area of approximately 8000 km² and 30 million people.

Here we present the synthesis of the Greater Jakarta commuting scenario that can help future researchers in generating large-scale scenarios in regions where the data is scarce or not easily obtainable. First runs, using an agent-based model MATSim, are presented and can be used as a back-bone for future improvements of the Greater Jakarta scenario and investigation of different transport related questions for the area.

© 2019 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Peer-review under responsibility of the Conference Program Chairs.

Keywords: MATSim; Agent-Based Model; Greater Jakarta
1. Introduction

Greater Jakarta is comprised of nine cities including Jakarta (consisting of five cities), Bogor, Depok, Tangerang, Bekasi, and four regencies; Tangerang, South Tangerang, Bogor, and Bekasi. Greater Jakarta, known as Jabodetabek, has a population of approximately 30 million inhabitants. It has a significant role in the national economy in Indonesia producing more than IDR 1,500 trillion (USD 113.51 Billion³). In 2012, Jakarta was the primary contributor to the Jabodetabek GDP with a share of 72.44%. Thus, the economic activities and employment opportunities of Jabodetabek are concentrated in the city of Jakarta. At the national level, Jabodetabek contributes 18.48% of Indonesia’s GDP. As a consequence, around 3.6 million people commute within, into and out of [1]. Therefore, a substantial chunk of the traffic volumes is generated by daily commuters coming to Jakarta for work or education.

In Jakarta and in Indonesia in general, there are different transportation options available, some of them formal, like car, motorcycle, commuter rail, Bus Rapid Transit (BRT), big buses and medium buses, and some informal ones like microbuses (called “angkot”, which is an informal service without a fixed schedule). Simulation of microbuses in an agent-based model has been previously used in the South-African context [2].

To the best knowledge of the authors, this research will be the first to make use of an agent-based model that incorporates all of the mentioned transportation modes for the simulation of daily behavior of the people performing their activities in Greater Jakarta. This paper will also add to the growing literature on modelling large-scale cities especially when the data is scarce or not easily obtainable.

Impacts of carsharing, bikesharing, congestion pricing, automated vehicles or equity effects were hard to investigate on a suitable level using the more traditional modeling techniques. Agent-based modeling frameworks have, on the other hand, shown to be suitable in the recent past to model large-scale cities, and to investigate the impacts of these emerging transportation options and policies. Therefore, we make use of an agent-based simulation framework called MATSim [3] in this research.

The remainder of this paper is structured as follows. The following section describes the MATSim framework. The third section presents in detail scenario development. The fourth section presents first results obtained for the commuting population of the Greater Jakarta. Finally, the last section presents conclusions, limitations and further recommendations.

2. MATSim Framework

This study utilizes a Multi-Agent Transport Simulation (MATSim), which performs a microscopic simulation of daily schedules of synthetic persons performing activities in the study area. The persons in MATSim are called agents, each agent having its own plan that represents its daily schedule of activities connected by trips. The plans are simulated using the mobility simulation (mobsim) for a number of iterations. Before the start of each simulation some of the agents are allowed to change a part of their plan in the re-planning phase. This simulation cycle can be seen in Fig 1.

![Fig 1. The MATSim loop](image)

† USD1 = IDR 14,178 on 01.03.2109

---

Anugrah Ilahi et al. / Procedia Computer Science 151 (2019) 272–278
In this work, however, we use a slightly different iterative approach proposed by Hoerl et al [4, 5] that integrates mode-choice models with the microsimulation in MATSim. In this approach, agents are allowed to change their modes of travel based on a discrete-mode choice model implemented. Since the mode-choice model in the re-planning phase uses estimates on the travel-times, costs, waiting times etc. from the previous iteration, scoring phase is no longer needed. This approach is then utilized in order to investigate mode choices for the commuting population of the Greater Jakarta population.

2.1. Traffic Flow Model

Traffic simulation model in MATSim uses a queue based approach, which has two attributes storage and flow capacity. Storage capacity defines how many cars can be stored at a time on a road link, and flow capacity defines the outflow capacity of a link.

2.2. Population

Commuting population is synthesized using the data gathered by the Japan International Cooperation Agency (JICA) in 2009, which includes 3% of the households in greater Jakarta. To expand the sample and to synthesize complete commuting population, we utilize a Bayesian network approach and Generalized Raking (GR) as shown in [7, 8]. Synthesized population contains approximately 20 million agents, which represent the people who have either work or school activities based on a statistics bureau on a province level (Jabodetabek consists of 3 provinces: Jakarta, West Java and Banten). There are two activity chains in this population: home – work – home, and home – school – home. The locations of home, work, and school activities are based on the addresses provided by the respondents and assigned a coordinate randomly drawn from an area created by drawing 1 km radius circle around the coordinates of the address. Unfortunately, starting times of work activities as well as their duration is not reported in the survey conducted by JICA. Therefore, we distributed the starting times and activity durations based on the behavior of people living in ile-de-France region as the area of the city is similar and the data is publicly available.

2.3. Network

The network of Greater Jakarta used to create the scenario is based on the OpenStreetMap (http://openstreetmap.com). The network of Greater Jakarta was extracted from OSM using Java Osmosis. The network consists of 472,205 nodes and 1,214,769 links. The links are classified by its function and parameter such as speed, lane capacity, and road hierarchy. The road classification is reported in Table 1. The network uses coordinate system EPSG:5330 for Batavia/Jakarta.

### Table 1. Road network classification in the MATSim Greater Jakarta scenario

<table>
<thead>
<tr>
<th>Hierarchy</th>
<th>Highway-type</th>
<th>Lanes</th>
<th>Free speed (m/s)</th>
<th>Free speed Factor</th>
<th>Lane capacity</th>
<th>One way</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Motorway</td>
<td>2</td>
<td>33.33</td>
<td>1.0</td>
<td>2000</td>
<td>True</td>
</tr>
<tr>
<td>2</td>
<td>Trunk</td>
<td>1</td>
<td>22.22</td>
<td>1.0</td>
<td>2000</td>
<td>False</td>
</tr>
<tr>
<td>3</td>
<td>Primary</td>
<td>1</td>
<td>22.22</td>
<td>1.0</td>
<td>1500</td>
<td>False</td>
</tr>
<tr>
<td>4</td>
<td>Secondary</td>
<td>1</td>
<td>8.33</td>
<td>1.0</td>
<td>1000</td>
<td>False</td>
</tr>
<tr>
<td>5</td>
<td>Tertiary</td>
<td>1</td>
<td>6.94</td>
<td>1.0</td>
<td>600</td>
<td>False</td>
</tr>
<tr>
<td>6</td>
<td>Minor</td>
<td>1</td>
<td>11.11</td>
<td>1.0</td>
<td>600</td>
<td>False</td>
</tr>
<tr>
<td>7</td>
<td>Residential</td>
<td>1</td>
<td>4.16</td>
<td>1.0</td>
<td>600</td>
<td>False</td>
</tr>
<tr>
<td>8</td>
<td>Living street</td>
<td>1</td>
<td>2.77</td>
<td>1.0</td>
<td>300</td>
<td>False</td>
</tr>
<tr>
<td>9</td>
<td>Rail</td>
<td>1</td>
<td>44.44</td>
<td>1.0</td>
<td>9999</td>
<td>True</td>
</tr>
</tbody>
</table>
2.4. Public transport network and counting stations

There is no publicly available GTFS (General Transit Feed Specification) data for Greater Jakarta. Therefore, we have manually constructed the public transport schedules using the data from a company called Trafi (https://www.trafi.com/id/jakarta). OSM network and manually constructed GTFS schedules are converted to MATSim format using the pt2matsim extension [9]. BRT lines are categorized as dedicated lanes. In the end, the transit schedule is mapped to the MATSim network using the same extension. Finally, we obtain public transport lines within the network.

There are 1,756 public transport lines in total. There are 325 BRT lines (including other bus companies that operate in BRT lines), 421 Bus lines, 22 commuter rail lines, and 988 microbus lines. There are also 26 counting stations that count number of vehicles in 15min bins.

2.5. Private and public transport vehicles

As the vehicles have different sizes and capacities, we classified private and public transport vehicles as in Table 2.

<table>
<thead>
<tr>
<th>Name</th>
<th>Symbol</th>
<th>Length/Width [m]</th>
<th>Capacity Seats/Standing</th>
<th>Number of lines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private vehicles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car</td>
<td>car</td>
<td>4.3/1.6</td>
<td>7/0</td>
<td>-</td>
</tr>
<tr>
<td>Motorcycle</td>
<td>mc</td>
<td>1.7/1.0</td>
<td>2/0</td>
<td>-</td>
</tr>
<tr>
<td>BRT</td>
<td>pt</td>
<td>18/2.5</td>
<td>50/30</td>
<td>325</td>
</tr>
<tr>
<td>Public transport</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bus</td>
<td>bus</td>
<td>12/2.5</td>
<td>35/15</td>
<td>421</td>
</tr>
<tr>
<td>Commuter rail</td>
<td>rail</td>
<td>240/2.8</td>
<td>1000/2000</td>
<td>22</td>
</tr>
<tr>
<td>Microbus</td>
<td>angkot</td>
<td>4.2/1.6</td>
<td>8/0</td>
<td>988</td>
</tr>
</tbody>
</table>

3. Mode-choice in MATSim

The utility function for different modes are based on the values estimated for the city of Zurich. However, the parameters had to be calibrated in order for the model to fit the behavior of people in Greater Jakarta. As the people in Jakarta have lower income levels and are more cost sensitive than people in Zurich we modified the cost parameter accordingly.

The formulation of the utility functions for modes used in the scenario are presented here:

\[
U_{PT} = \beta_{\text{number of transfers}} \cdot X_{\text{number of transfers}} + \beta_{\text{in vehicle time}} \cdot X_{\text{in vehicle time}} + \beta_{\text{transfer time}} \cdot X_{\text{transfer time}} + \beta_{\text{access time}} \cdot X_{\text{access time}} + \beta_{\text{cost}} \cdot X_{\text{cost}} + \theta_{\text{cost}}
\]

(1)

\[
U_{\text{car}} = ASC_{\text{car}} + \beta_{\text{travel time}} \cdot X_{\text{travel time}} + \beta_{\text{parking time}} \cdot X_{\text{parking time}} + \beta_{\text{distance cost}} \cdot X_{\text{distance cost}} + \beta_{\text{cost}} \cdot X_{\text{cost}} + \theta_{\text{cost}}
\]

(2)

\[
U_{\text{MC}} = ASC_{\text{MC}} + \beta_{\text{travel time}} \cdot X_{\text{travel time}} + \beta_{\text{access time}} \cdot X_{\text{access time}} + \beta_{\text{distance cost}} \cdot X_{\text{distance cost}} + \beta_{\text{cost}} \cdot X_{\text{cost}} + \theta_{\text{cost}}
\]

(3)

\[
U_{\text{walk}} = ASC_{\text{walk}} + \beta_{\text{travel time}} \cdot X_{\text{travel time}} + \beta_{\text{time}}
\]

(4)

Table 3 shows the parameters obtained after calibration, and the result can be seen in Fig 3. We also include the value of per km cost of car and motorcycle, which are 0.05 USD and 0.015 USD respectively, and are based on fuel prices in Jakarta.
Table 3. Calibrated parameters of mode choice model.

<table>
<thead>
<tr>
<th>Mode</th>
<th>Parameters</th>
<th>Estimation</th>
</tr>
</thead>
</table>
| Public Transport | $\beta_{\text{numberOfTransers, PT.}}$  
|               | $\beta_{\text{inVehicleTime, PT [min$^{-1}$]}}$ | -0.170     |
|               | $\beta_{\text{transferTime, PT [min$^{-1}$]}}$ | -0.120     |
|               | $\beta_{\text{accessEgressTime, PT [min$^{-1}$]}}$ | -0.048     |
| Car           | $\alpha_{\text{Car}}$  
|               | $\beta_{\text{travelTime, Car [min$^{-1}$]}}$ | 1.227      |
| Motorcycle    | $\alpha_{\text{MC}}$  
|               | $\beta_{\text{travelTime, MC [min$^{-1}$]}}$ | 1.227      |
| Walk          | $\alpha_{\text{walk}}$  
|               | $\beta_{\text{travelTime, walk [min$^{-1}$]}}$ | -0.141     |
| Other         | $\beta_{\text{Cost}}$                            | -0.030     |
| Calibration   | $\theta_{\text{parkingSearchTime, Car [min]}}$ | 4.000      |
|               | $\theta_{\text{accessEgressTime, Car [min]}}$     | 4.000      |
|               | $\theta_{\text{accessEgressTime,MC [min]}}$       | 2.000      |

4. First Results

4.1. Mode shares

The mode-share results for a simulated sample of 5% are shown in Fig 2. We can observe some differences in the modal share, especially for a car mode. However, looking at motorcycle and car together, the statistics match very well. One of the possible reasons for underestimating the amount of car trips is the unreliability of car and motorcycle ownership data, which is not always available in the JICA dataset.

![Mode shares comparison between JICA study and MATSim](image-url)
4.2. Traffic counts

There are 26 counting locations in our network. The traffic volume results for one of the most important roads in central Jakarta (Thamrin Street) can be seen Fig. 3, which presents the comparison between traffic volume in MATSim and traffic count data, from 1 A.M to 4 P.M. While having lower counts for car makes sense as we only simulate commuter and also have lower share of cars in the simulation, having larger number of motorcycle numbers is worrying. The reasons behind this can be several. The counting data might not be complete, MATSim might overestimate number of motorcycles taking this route or the traffic count data might not be complete. Therefore, further investigations should find suitable explanations and necessary steps should be taken accordingly.

![Fig 3. Count comparison for motorcycle (left) and car (right) on Thamrin Street.](image)

5. Limitations

In this research, we only synthesize the population that performs mandatory activities. The next step, is naturally to expand this with secondary-activities. However, this is currently not possible as the data on travel-behavior in Jakarta does not exist.

The parameters of the mode choice model is based on the Zurich model which is a limiting factor. In the future work, we hope to obtain necessary stated-preference data to be able to estimate a mode-choice model that is more fitting to the population of Jakarta.

We use 5% of the population in our simulations, as it is time costly to simulate a population of 20 million people. Further research, should also focus on using larger samples to make sure the results stay stable.

6. Conclusions

This paper uses an agent-based modelling framework to simulate the commuting population of Greater Jakarta. The methodology presented here also utilizes a novel approach that integrates mode choice with a microsimulation in MATSim. The results show that differences between the MATSim and JICA mode shares is very low. Further research is however needed in order to represent complete travel behavior of people living in Greater Jakarta. Nevertheless, this research provides the backbone on which further research can be built.

Acknowledgements

The authors wish to acknowledge the funding from Lembaga Pembiayaan Dana Pendidikan (LPDP), Indonesia, Grant numbers: 2016042201598.
References


