TASHA-MATSim Integration and its Application in Emission Modelling

By

JIANG YANG HAO

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science
Department of Civil Engineering
University of Toronto

© Copyright by Jiang Yang Hao 2009
TASHA-MATSim Integration and its Application in Emission Modelling

Jiang Yang Hao
Master of Applied Science
Department of Civil Engineering
University of Toronto
2009

Abstract

Microsimulation is becoming more popular in transportation research. The purpose of this research is to explore the potential of microsimulation by integrating an existing activity-based travel demand model with an agent-based traffic simulation model. Differences in model precisions from the two models are resolved through a series of data conversions, and the models are able to form an iterative process similar to previous modelling frameworks. The resulting model is then used for emission modelling where the traditional average-speed model is improved by exploiting agent-based traffic simulation results. Results from emission modelling have demonstrated the advantages of the microsimulation approach over conventional methodologies that rely heavily on temporal or spatial aggregation.
Acknowledgements

This dissertation could not have been written without all the people who kindly provided support and encouragement during the course of this research. First and foremost I offer my sincerest gratitude to my supervisor, Professor Eric Miller, who has supported me throughout the year with his patience and knowledge.

I would like to thank the MATSim development team led by Professor Kai Nagel and Professor Kay Axhausen for the intensive but wonderful two weeks of MATSim training in Berlin and Zurich, and more importantly, the on-going support throughout the year.

Special thanks also go to Marcel Rieser and Dr. Michael Balmer for all their timely replied emails packed with valuable comments and suggestions; Dr. Marianne Hatzopoulou, for her guidance on emission modelling methodology and ArcGIS; and James Vaughan, our wonderful programmer, for the programming support throughout the year, especially the crash courses in C++, C#, and JAVA. Finally, I would like to thank my mother for providing logistical support throughout the year.
# Table of Contents

Chapter 1. Introduction....................................................................................................................1

Chapter 2. Literature Review .........................................................................................................2

2.1 Travel Demand Models ............................................................................................................2

2.2 Traffic Flow Models ................................................................................................................3

2.3 Emission Models .....................................................................................................................5

Chapter 3. Model Background ......................................................................................................7

3.1 TASHA ..................................................................................................................................7

3.1.1 TASHA Scheduler ............................................................................................................9

3.1.2 TASHA Mode Choice Model ..........................................................................................9

3.1.3 TASHA-EMME/2 Interactions .........................................................................................9

3.1.4 TASHA Validation ...........................................................................................................9

3.2 MATSim-T .............................................................................................................................10

3.2.1 Overview of MATSim Environment ................................................................................10

3.2.2 Scoring Function ............................................................................................................10

3.2.3 Genetic Algorithm ..........................................................................................................12

3.2.4 Mode Choice Model ......................................................................................................12

3.2.5 Queue-based Traffic Simulation ....................................................................................13

3.2.6 Visualization ..................................................................................................................13

3.3 Car Allocation Model ...........................................................................................................13

3.4 Mobile 6.2C and the Emission Calculation Program .............................................................14

Chapter 4. TASHA – MATSim Integration .......................................................................................17

4.1 Architecture ..........................................................................................................................17

4.2 TASHA to MATSim Data Conversion ..................................................................................20

4.2.1 Network Attributes ........................................................................................................22

4.2.2 Schedules ........................................................................................................................24

4.3 MATSim Simulation Configurations .....................................................................................26

4.4 MATSim to TASHA Data Conversion ..................................................................................27

4.4.1 Event File .......................................................................................................................27
4.4.2 Link Statistics ........................................................................................................... 27
4.4.3 Constructing the Travel Time Matrix ................................................................... 28

Chapter 5. TASHA-MATSim Modelling Results ................................................................. 29
5.1 Score Convergence .................................................................................................... 29
5.2 Distributions .............................................................................................................. 30
  5.2.1 Trip Departures and Arrivals .............................................................................. 30
  5.2.2 Trip Length and Duration .................................................................................. 31
  5.2.3 Mode Split ............................................................................................................ 33
  5.2.4 Travel Time Matrix ............................................................................................. 34
5.3 Visualization ............................................................................................................... 35
5.4 Effects of the Time Allocation Mutator (TAM) ......................................................... 36

Chapter 6. Modelling Auto Emissions .............................................................................. 40
  6.1.1 Data Preparation for the Car Allocation Model .................................................. 40
  6.2 MATSim Simulation .................................................................................................. 41
  6.3 Emission Calculations using MATSim Outputs ....................................................... 42
    6.3.1 Emission Factors ............................................................................................... 42
    6.3.2 Calculation Procedures ..................................................................................... 46

Chapter 7. Emission Modelling Results and Analysis ...................................................... 48
  7.1 Total Emissions ........................................................................................................ 48
  7.2 Emissions by Hour ................................................................................................... 50
  7.3 Emissions by Person and Household ....................................................................... 52
  7.4 Link Emissions ....................................................................................................... 57

Chapter 8. Future Work ..................................................................................................... 60
  8.1 Improving TASHA-MATSim Interactions ............................................................... 60
  8.2 Improving the Emission Model .............................................................................. 61

Chapter 9. Conclusions .................................................................................................... 62

References ....................................................................................................................... 63

Appendices ....................................................................................................................... 65
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 2-1</td>
<td>Fully Agent-based Approach of Dynamic Traffic Assignment (Balmer, 2007)</td>
<td>4</td>
</tr>
<tr>
<td>Figure 3-1</td>
<td>TASHA Model Flowchart (Roorda et al., 2007)</td>
<td>8</td>
</tr>
<tr>
<td>Figure 3-2</td>
<td>Descriptions of Mobile6.2 Emission Types</td>
<td>15</td>
</tr>
<tr>
<td>Figure 3-3</td>
<td>Flowchart of the Emission Calculation Program</td>
<td>16</td>
</tr>
<tr>
<td>Figure 4-1</td>
<td>TASHA-MATSim Configuration I</td>
<td>17</td>
</tr>
<tr>
<td>Figure 4-2</td>
<td>TASHA-MATSim Configuration II</td>
<td>18</td>
</tr>
<tr>
<td>Figure 4-3</td>
<td>TASHA-MATSim Configuration III</td>
<td>19</td>
</tr>
<tr>
<td>Figure 4-4</td>
<td>TASHA to MATSim Data Conversions</td>
<td>21</td>
</tr>
<tr>
<td>Figure 4-5</td>
<td>Assignment of Links to Activity Locations (Rieser et al. 2007)</td>
<td>22</td>
</tr>
<tr>
<td>Figure 4-6</td>
<td>MATSim Configuration File Layout</td>
<td>26</td>
</tr>
<tr>
<td>Figure 5-1</td>
<td>MATSim Simulation Score Statistics</td>
<td>29</td>
</tr>
<tr>
<td>Figure 5-2</td>
<td>MATSim Simulation Trip Departures and Arrivals (hour 0 = 4am)</td>
<td>31</td>
</tr>
<tr>
<td>Figure 5-3</td>
<td>Average Trip Length by Hour</td>
<td>32</td>
</tr>
<tr>
<td>Figure 5-4</td>
<td>Average Trip Duration by Hour</td>
<td>32</td>
</tr>
<tr>
<td>Figure 5-5</td>
<td>Mode Split by Hour of Day</td>
<td>33</td>
</tr>
<tr>
<td>Figure 5-6</td>
<td>Mode Share Percentages for the Whole Day and Peak Periods</td>
<td>33</td>
</tr>
<tr>
<td>Figure 5-7</td>
<td>Travel Time Distribution</td>
<td>34</td>
</tr>
<tr>
<td>Figure 5-8</td>
<td>Simulation Visualization (8:35am)</td>
<td>35</td>
</tr>
<tr>
<td>Figure 5-9</td>
<td>Score Statistics for MATSim Simulation with Time Allocation Mutator</td>
<td>36</td>
</tr>
<tr>
<td>Figure 5-10</td>
<td>Distribution of Trip Departures and Arrivals at Iteration 100 with TAM Enabled</td>
<td>37</td>
</tr>
<tr>
<td>Figure 5-11</td>
<td>Effect of the Time Allocation Mutator on Zone-to-Zone Travel Time</td>
<td>38</td>
</tr>
<tr>
<td>Figure 5-12</td>
<td>Trip Length Comparison</td>
<td>39</td>
</tr>
<tr>
<td>Figure 5-13</td>
<td>Trip Duration Comparison</td>
<td>39</td>
</tr>
<tr>
<td>Figure 6-1</td>
<td>Emission Factor File Format</td>
<td>42</td>
</tr>
<tr>
<td>Figure 6-2</td>
<td>HC Exhaust Emission at 7AM</td>
<td>43</td>
</tr>
<tr>
<td>Figure 6-3</td>
<td>HC Exhaust Emission Factor by Hour and Speed</td>
<td>43</td>
</tr>
<tr>
<td>Figure 6-4</td>
<td>3-D Plot of CO Exhaust Emission Factor</td>
<td>44</td>
</tr>
<tr>
<td>Figure 6-5</td>
<td>HC Start Emission Factors</td>
<td>45</td>
</tr>
<tr>
<td>Figure 6-6</td>
<td>HC Soak Emission Factors</td>
<td>45</td>
</tr>
<tr>
<td>Figure 6-7</td>
<td>Emission Calculation Procedures</td>
<td>46</td>
</tr>
<tr>
<td>Figure 7-1</td>
<td>Emission by Type for each Pollutant (HC, CO, NOx in tons, CO2 in kilotons)</td>
<td>49</td>
</tr>
<tr>
<td>Figure 7-2</td>
<td>Exhaust Emission Comparisons between Different Modelling Methods</td>
<td>49</td>
</tr>
<tr>
<td>Figure 7-3</td>
<td>Total HC Emission by Hour</td>
<td>50</td>
</tr>
<tr>
<td>Figure 7-4</td>
<td>Different Types of HC Emission by Hour</td>
<td>51</td>
</tr>
<tr>
<td>Figure 7-5</td>
<td>Percentage Share of Different Types of HC Emission by Hour</td>
<td>51</td>
</tr>
<tr>
<td>Figure 7-6</td>
<td>HC Exhaust Emission per Person by Age</td>
<td>52</td>
</tr>
<tr>
<td>Figure 7-7</td>
<td>Percentage of Trips by Hour for Different Age Groups</td>
<td>52</td>
</tr>
<tr>
<td>Figure 7-8</td>
<td>HC Start Emission per Person by Age</td>
<td>53</td>
</tr>
</tbody>
</table>
Figure 7-9. HC Emission per Person by Gender .......................................................... 53
Figure 7-10. HC Emission per Person by Occupation Type ....................................... 54
Figure 7-11. Household HC Emission Histogram ....................................................... 55
Figure 7-12. HC Emission per Household by Number of Vehicles ............................. 55
Figure 7-13. Weighted Average HC Emission per Household by Household Location ... 56
Figure 7-14. CO Exhaust Emission ............................................................................ 57
Figure 7-15. CO Idling Emission ............................................................................... 58
Figure 7-16. Household and Link Emission Overlay .................................................. 59

List of Appendices

Figure A-1. MATSim-T .............................................................................................. 66
Figure A-2. Complete Flowchart .............................................................................. 67
Figure A-3. TASHA-EMME/2 Flowchart ................................................................. 68
Figure A-4. TASHA-MATSim Flowchart ................................................................. 69
Figure A-5. Emission Modelling Flowchart .............................................................. 70
Figure A-6. MATSim Simulation Configuration File .................................................. 71
Figure A-7. MATSim Simulation Snapshots ............................................................. 73
Figure A-8. Hourly HC Exhaust Emission ............................................................... 79
Figure A-9. Hourly HC Idling Emission .................................................................... 85
Figure A-10. Hourly HC Start Emission .................................................................. 89
Figure A-11. Hourly HC Soak Emission .................................................................. 95
Chapter 1. Introduction

Population growth and urbanization have brought great changes to cities in the past decade. There are more reasons for people to travel and more methods for them to do so. As a result, there is increasing interest and necessity to survey, understand, and predict how people make the decision to travel; how travelling affects the transportation network; and how a utilized but often stressed transportation network influences the growth of cities and brings changes to the environment.

Modelling the transportation environment and its impact on society requires a great variety of models. There are models that focus on specific aspects such as demographic forecasting, car ownership, or trip mode choice. There are also macroscopic frameworks that link multiple sub-models in order to capture the interactions between sub-systems. Despite the great variety in scale and scope, current operational models are mostly static in nature and rely heavily on spatial and temporal aggregations in order to satisfy data limitations and computational constraints. Interactions between models may be hindered by different aggregation techniques, and the resolution of the overall system is often determined by the weakest link. Moreover, while such conventional methods provide acceptable system wide forecasts, they lack the sophistication required for policy analysis. Forecasting people’s sensitivity to new policies requires modelling of the psychological and behavioural processes that lead to people’s decisions.

Responding to the growing demand for more efficient, realistic, and policy-sensitive transportation models, microsimulation is becoming a more popular approach to transportation and land use modelling practices. Microsimulation is a methodology that explicitly models the dynamic behaviour of each individual agent within a modelling environment. The disaggregate nature of this technique makes it capable of providing a much finer resolution of the system state.

This research has two main objectives. The first one is to develop an agent-based travel demand modelling framework by integrating existing agent-based software. The second objective is to use this newly integrated model to obtain a finer resolution for vehicle emission modelling, thus demonstrating the advantages of the microsimulation approach over conventional methodologies that rely heavily on temporal or spatial aggregation.
Chapter 2. Literature Review

This literature review focuses on the role of microsimulation in three subject areas: travel demand modelling, traffic flow modelling, and emission modelling. Studying the modelling philosophy and methodologies in these areas will provide better insights into the models used in this research: TASHA, MATSim, and the new emission model. Since this research involves the integration of multiple existing modelling components, the review will also help identify strengths and weaknesses of each model and understanding the imbalance in model precisions, so that better trade-off decisions can be made during data conversions.

2.1 Travel Demand Models

A review of the recent literature on travel demand modelling clearly indicates that the microsimulation approach is maturing and is being adopted for an increasing number of real-world applications. (Roorda et al., 2007) For modelling daily schedules, activity-based models start with the modelling of four basic dimensions of activities: activity choice; duration; location; and sequencing. (Doherty and Axhausen, 1998) Models developed over the years differ greatly in how these activity dimensions are represented and predicted.

Timmermans (2001) provides summary of activity scheduling models. The models are classified into simultaneous models and sequential models. Simultaneous models work within a utility-maximizing framework and rely on observed activity-travel patterns to forecast new schedules. The main criticism for this approach is that people are unlikely to be aware of all possible patterns available, and thus optimization achieved through utility-maximization does not realistically represent people’s behaviour. (Doherty and Axhausen, 1998) Earlier implementations used multinomial logit models to predict the choice of activity pattern. This was later replaced by nested logit models, which separated different aspects of activity-travel patterns into multiple levels. The use of different nests introduced some flexibility into the rather rigorous utility maximizing approach of schedule pattern selection, however, the computational cost for estimating such models increased substantially.

Sequential models, on the other hand, model activity choices explicitly. Another classification dimension was suggested by Doherty and Axhausen (1998), which distinguish models by whether they adopt an econometric or rule-based approach. Using a sequence of heuristics, the formation of activity patterns is the combined results of a series of state-actions. The computer algorithm is essentially based on a set of
If...Then...Else rules. One such model is Albatross, a fully operational computational process model that uses a decision tree to represent a pre-defined order of decisions. (Arentze and Timmermans, 2000)

Recent models emphasise the concept of project and schedule in modelling travel demand. Axhausen (1998) defines a project as a coordinated set of activities tied together by a common goal or outcome. Projects directly reflect people’s goals while schedules do not. Travel is not a project, but only a means used by people to perform activities in order to reach certain goals. Therefore, this project-driven approach truly models the behavioural and psychological process that will eventually lead to a schedule pattern, not the pattern itself.

Advancement in travel demand forecasting techniques has placed heavy stress on data requirements. Agent-based modelling requires disaggregate data with resolution down to the agent-level. Traditional zonal travel data such as screen line counts etc can no longer accommodate the sophistication of behavioural models. Also, previous data collection techniques focus primarily on the outcome rather than the process that leads to the outcome. Observed activity patterns are in fact the result of unobserved activity scheduling and rescheduling process. New data collection methods have been experimented with in an attempt to gain more insight into the mental process that creates schedules. The CHASE survey, for example, traces household member’s scheduling decisions during an entire week. The results demonstrate that people under different circumstances will use different strategies to create and modify their schedules. This type of survey will help travel demand models to go beyond replicating observed outcomes and truly model people’s behaviour. (Doherty, 2000)

### 2.2 Traffic Flow Models

Current research on the regional scale still relies on traditional aggregated models such as EMME/2, which aggregate data both spatially and temporally, and uses a static deterministic user-equilibrium approach for traffic assignment. These models are designed for the traditional four-step process of transportation modelling. The detailed methodology of EMME/2 was reviewed by Gao (2009) and its performance was compared to that of MATSim, a dynamic agent-based simulation package.

Compared to its applications in travel demand models, microsimulation has been much more widely accepted and used in traffic flow models in recent years. Transportation network simulation such as TRANSIMS and PARAMICS all use microsimulation to model vehicle dynamics on the road. Most of these models are best suited for short-range forecasting. Long-range, regional-scale forecast raises computational issues, and the level of precision necessary for such analysis has been debated over the
years. In addition, the behaviour of these models is often difficult to predict due to their stochastic nature. (Miller, 2003)

Without any form of aggregation, dynamic agent-based traffic simulation models are well suited for integration with agent-based travel demand models. Miller (2003) described the information exchange between these two processes and pointed out that the typical dynamic route assignment procedures break the traditional sequential process in travel demand generation by simultaneously model route choice and departure time choice. This will naturally cause traffic simulation models to “intrude” into the micro-scheduling component of activity-based travel demand models. Balmer (2007) summarized previous work in this area and suggested the possibility of forming a fully agent-based modelling framework. (Figure 2-1) Each individual is modeled as an autonomous entity throughout the iterative process.

![Fully Agent-based Approach of Dynamic Traffic Assignment (Balmer, 2007)](image-url)
2.3 Emission Models

The latest report on air quality trends produced by the Environmental Protection Agency (EPA, 2008) indicates that highway vehicles alone accounts for 52% of carbon monoxide (CO) emissions, 33% of nitrogen oxide (NOₓ) emissions, and 23% of volatile organic compound (VOC) emissions. (EPA, 2008) Increase in car usage induced by urbanization and population growth has led to increasing attention being placed on vehicle emissions. More methodologies for measuring and predicting vehicle emissions have been developed in recent years in response to the increasing number of emission regulations. Estimating emission levels for the current fleet can be achieved through various inspection and surveying techniques, which are almost always more reliable than estimates from emission models. (Hatzopoulou, 2008) Nevertheless, emission models are useful, especially for forecasting and comparing emission levels under different policy scenarios, and determining whether the proposed project conforms to regulatory requirements.

The fundamental building blocks of emission modelling are the Base Emission Rates (BER) established through vehicle testing. The BERs are then multiplied by correlation factors to produce emission factors that are sensitive to variables such as temperature, fuel type, humidity, etc. In the United States, emission factors are prepared by either the MOBILE series developed by the EPA or the Motor Vehicle Emission Inventory model series developed for California by the California Air Resources Board (CARB), with the latest models being MOBILE6 and EMFAC2000, respectively. (Niemeier, 2003)

Once the base emission factors are generated, the next task in emission modelling is to combine these factors with vehicle activities. There are mainly two different approaches for interpreting vehicle activity information, one of them uses average vehicle speed on the road, and the other uses a microscopic approach that explicitly models dynamics of each vehicle. Choice between these two approaches depends on the scope of research. For large, region-wide analysis, the aggregated average speed approach provides sufficient resolution without becoming computationally prohibitive. In contrast, the microscopic approach is more suited for small-scale analysis that compares emissions produced by different types of vehicles or the effect of different driving behaviour. The model series developed by EPA and CARB both use the average speed approach for generating emission rates for different speed categories.

Hatzopoulou (2008) provides a review of a number of operational microscopic emission models. These models fully utilize the microsimulation results from dynamic traffic flow models by combining them
with second by second, or even instantaneous emission rates. They are capable of capturing effects of differences in vehicle operation, vehicle technology, and driver behaviour.

Due to their reliance on travel demand forecasts and close relation to policy development and evaluations, emission models are often applied within an integrated modelling framework that has a much broader scope. There is on-going research and development of the ILUTE model, a fully agent-based, integrated microsimulation model for land use, transportation, and environmental analysis. (Miller, 2005)

Developing macro-models of this scale faces many challenges. Such models are “computationally intensive, data hungry, make extreme demands on our theoretical understanding of urban spatial process and our methodological capabilities for capturing that understanding within operational computer code, and are difficult to estimate and validate.” (Miller et al., 2004) When data from multiple sources need to be pooled together for analysis, differences in aggregation method, precision, surveying location and time may greatly reduce the amount of useful information extracted, and may lead to biased modelling results. Therefore, a large, consistent database such as the Transportation Tomorrow Survey is very valuable for research of this magnitude.
Chapter 3. Model Background

This section provides a brief description of the models and programs used in this research: TASHA, MATSim, Car Allocation Model, and the Emission Calculation Program. EMME/2 has also been used, but it is only for the purpose of comparing simulation results; thus it is not part of the integration process. MATSim consists of a large variety of models. The description provided below will only focus on specific sub-models within MATSim that are either used in this research or have a counterpart in the TASHA framework.

3.1 TASHA

The Travel/Activity Scheduler for Household Agents (TASHA) is a microsimulation model for forecasting travel demand. It is designed to improve upon conventional four-stage models. The model operates based on three hypotheses: (Miller and Roorda, 2001)

“First, a fundamental assumption is that scheduling is an event-driven, sequential process, in which individual episodes are provisionally scheduled as they arise out of personal and household projects.”

“Given this, it follows that, in general, activity-travel is not an optimizing procedure”

“A third fundamental assumption is that travel mode choice (and the associated allocation of household vehicles for individual person travel, as required) is inherent in the activity scheduling process.”

The primary components of TASHA are the scheduler and mode choice model. An additional trip processing module, named Merge_tt, links the two primary components. Its purpose is to build trip chains based on the feasibility of individual trips. Figure 3-1 shows the conceptual design of TASHA as described in detail by Miller and Roorda (2001). Although the program is agent based, it still follows a sequential decision making process similar to that of a conventional four-stage model. The program generates activities for each person based on observed joint probability distributions; predicts activity locations based on a series of entropy models; schedules activities following a set of rules; and finally assigns mode through a random utility tour-based mode choice model.
Figure 3-1. TASHA Model Flowchart (Roorda et al., 2007)
3.1.1 TASHA Scheduler

TASHA scheduler adopts the project based approach described in the literature review. Activity episodes that can help an agent complete a certain project in mind are placed into the agent’s schedule. This bottom-up approach is different from other schedulers that use a top-down approach which begins with the selection of activity patterns. It is considered to be a better representation of people’s dynamic scheduling behaviour. (Miller and Roorda, 2001) Input for the schedule relies mainly on activity distributions and population information from the 1996 Transportation Tomorrow Survey. Other survey years have also been used, as well as a synthetic population generated based on the survey.

3.1.2 TASHA Mode Choice Model

The mode choice component of TASHA has evolved since the operational prototype. The current model includes more recent work that adds GO Rail related modes. This new hybrid model was developed for a project funded by the City of Toronto to research the effect of Transit City, a proposed expansion plan of the existing transit network in Toronto. Eight modes are considered in this version: Auto-drive, Transit, Walk, Ride, Passenger, GO – Drive Access, GO – Drive Egress, and GO – Non-drive Access/Egress.

3.1.3 TASHA-EMME/2 Interactions

This research uses EMME/2 only for comparison purposes, as both the travel demand module and emission modelling module being used here were previously dependant on network information generated through EMME/2. Figure A-3 in the Appendix shows the interactions between these two programs. TASHA outputs are converted into OD flows, and sent into EMME/2 for static traffic assignment. In return, EMME/2 produces inter-zonal travel times as input for TASHA scheduler. For non-drive modes, EMME/2 also updates various level of service information. For activities that become infeasible due to travel time constraints, TASHA scheduler will draw a new activity from trip generation distributions. The process iterates three times.

3.1.4 TASHA Validation

TASHA has been validated using 2001 Transportation Tomorrow Survey data. (Roorda et al., 2007) The validation exercise uses TASHA with parameters calibrated by 1996 TTS data to forecast 2001 activities. The results demonstrate that TASHA is capable of accurately predicting distributions for activity frequencies, activity start times, activity durations, as well as trip distance. The accuracy is highest for work trips due to the relative abundance of data.
3.2 MATSim-T

3.2.1 Overview of MATSim Environment

The Multi-Agent Transportation Simulation Toolbox (MATSim-T, also referred to as MATSim) is currently being developed jointly at TU Berlin, ETH Zürich, and CNRS Lyon. It consists of a great variety of microsimulation tools for modelling travel demand and traffic flow. MATSim is a fast agent-based simulation designed to handle large networks and millions of agents. Large scale scenarios have been tested in Zurich, Berlin, and other cities. (Balmer, 2007) Figure A-1 in the Appendix presents a complete flow chart of the various MATSim components and how they interact with each other. To manage the wide range of sub-models and algorithms, MATSim adopts a modular approach with standardized data format, so that new modelling and analysis modules can be integrated into the framework in a plug-and-play style.

MATSim was originally planned to be an alternative to TRANSIMS outside the U.S. until TRANSIMS became open source. (Balmer, 2007) Both generate individual activity plans for feeding into dynamic traffic assignment, but MATSim has various advantages, such as the use of hierarchical XML formats, multiple plans per agent, and considerably faster simulation speed.

The Iterative Demand Optimization Process – Evolutionary Algorithm (MATSim-EA) within the MATSim framework is similar to the TASHA-EMME/2 framework, but in a pure microsimulation environment. Starting with an initial travel demand, the simulation is an iterative process between travel activity rescheduling and traffic assignment. The traffic assignment component in MATSim uses event-driven queue-based microsimulation to model vehicles on the network.

3.2.2 Scoring Function

Schedules can be modified through three strategies. The first one is a Dynamic Dijkstra Router that changes the travelling route. The second one is called Time Allocation Mutator (TAM), which shifts activity start times by a specified amount. The last strategy is to reschedule the whole activity using a household scheduler module named planomat. Comparing the scheduling and rescheduling process in MATSim to those in TASHA, the two programs differ in both the way schedules are evaluated, and how new schedules are generated.

Unlike TASHA which uses a set of rules to assess feasibility of the schedules, MATSim evaluates each schedule based on a utility based scoring function. A summary of the scoring method is provided below.
Detailed explanation on this concept and methodology can be found in the original paper by Charypar and Nagel (2005).

\[
U_{\text{plan}} = \sum_{i=1}^{n} U_{\text{act}}(\text{type}_i, \text{start}_i, \text{dur}_i) + \sum_{i=2}^{n} U_{\text{trav}}(\text{loc}_{i-1}, \text{loc}_i) \quad - (1)
\]

Schedules are referred to as plans in MATSim. As shown in Equation 1, the score of an activity plan \( U_{\text{plan}} \) is the sum of utilities of all activities and disutility of travelling between activity locations. The utility of performing an activity is given by the sum of five terms: utility of performing the activity \( (U_{\text{dur},i}) \), disutility of early arrival \( (U_{\text{wait},i}) \), later arrival \( (U_{\text{late},ar,i}) \), and early departure \( (U_{\text{early},dp,i}) \), and another disutility term for performing an activity with very short duration \( (U_{\text{short},dur,i}) \). (Equation 2)

\[
U_{\text{act},i} = U_{\text{dur},i} + U_{\text{wait},i} + U_{\text{late},ar,i} + U_{\text{early},dp,i} + U_{\text{short},dur,i} \quad - (2)
\]

Gains and losses in the utilities identified in Equation 2 are related to the time spent on the corresponding portion of the activity. A series of parameters reflects the marginal utilities. While the latter four utility terms are modelled to be linear with respect to time, the utility gain for performing the activity is logarithmically related to its duration. This reflects the diminishing marginal utility gains as the duration increases.

From the utility expressions, it can be observed that the emphasis is on the activities themselves rather than the travelling between activities. The disutility for travel is linearly related to the overall travel time. Delays and other complications in travel are indirectly represented as opportunity loss for not performing activities. In contrast, TASHA focuses more on the travelling, and accepts activities as long as they are feasible. This is because TASHA considers multiple modes of travel, and thus a more complicated procedure is required to analyze travel time for different modes. Travel time in transit, for example, is broken apart into walk, wait, and in vehicle travel times. TASHA values different components of travel time using different utility parameters. To be even more realistic, transfers between transit lines, or mode switching during a trip chain should also be considered, so that the model can consider the complexity of the trip.

The scores can strongly affect simulation results since they are used in many simulation modules whenever an assessment of trip or schedule performance is required. For example, the rerouting algorithm uses the scoring function to determine the best route between two activity locations; the rescheduling process selects the next schedule based on scores of existing schedules. The scoring
methods currently being used in MATSim may be unrealistic for non-auto modes, but should be sufficient for schedules only involving auto travel.

### 3.2.3 Genetic Algorithm

A schedule needs to be modified or replaced when it is infeasible as viewed by TASHA, or from the MATSim’s perspective, has low performance/score. MATSim comes with a rescheduling module called *planomat*. Comparing to the scheduler in TASHA, *planomat* uses a very different concept for generating new schedules. TASHA scheduler replaces infeasible activities within a schedule by drawing a new one from activity frequency distributions, while *planomat* uses a Genetic Algorithm to create a new schedule based on existing ones. After multiple iterations, schedules from TASHA still represent the initial activity frequency distributions. MATSim schedules may drift towards a set of more optimal schedules that “make best use of time”, but the definition of optimal is dependent on the specifications of the scoring functions.

To use the Genetic Algorithm, MATSim keeps not only one schedule for each person, but a list of possible schedules originally generated from the agent database. At the beginning of a new iteration, the agent may choose to pick an existing schedule from the list based on scores, or try a new schedule generated by *planomat*. The new schedule is generated by choosing two existing schedules as parents, randomly selecting activities from them, and mutating start times and durations of the activities. The list gets updated again at the end of the iteration, at which point the schedule with the lowest score is removed. The list not only reflects what the person did during every iteration, but also what he/she wants to do, and his/her previous experience with some of the schedules. To a certain extent, this approach gives MATSim the ability to model people’s memory. The developers in Berlin have plans to expand the capability of *planomat* to incorporate more aspects of travel demand such as location choice, mode choice, and development of activity patterns. (Meister, Balmer 2006)

### 3.2.4 Mode Choice Model

MATSim is currently being used to simulate only auto traffic. Multiple mode choice modules exist in the MATSim environment. A very simple approach for modelling mode choice is to let the simulation duplicate existing plans and assume the duplicates all use non-drive mode for travelling. This new mode has adjusted parameters for trip score calculations. A more sophisticated model has recently been developed by Francesco et al. (2008) to produce more realistic and behavioural sensitive initial mode choice information. Another model is being developed for the optimization stage.
3.2.5 Queue-based Traffic Simulation

MATSim uses stochastic, queue based agent traffic simulation with a set of very simple rules. Vehicles enter and exit links on a first-in first-out basis. Vehicles cannot exit unless they have spent enough time on the link for them to travel the full length of the link. Vehicles cannot enter a link whose storage capacity has been reached, and cannot leave a link that reaches flow capacity. (Balmer, 2007) The result is a very fast simulation that produces outputs meaningful and useful for most transport planning purposes. No lane changes, signal schedules, or turn restrictions are modelled. However, it may be possible to indirectly model some of these scenarios by modifying input network or other simulation parameters.

3.2.6 Visualization

MATSim comes with a variety of modules for visualizing simulation results. These visualizers can be used to view a snapshot of the system at certain simulation time, or view the system dynamically during simulation. The latter option of course will have much higher requirement on the processing and graphical capability of the computer hardware. The most basic visualizer is Netvis, which can be found in the MATSim tutorial, as well as in the MATSim utility folder. Vehicles are displayed as square blocks on the network with no indication of vehicle speed. However, the visualizer does have options to view network link and node IDs. Netvis was used at the early stage of this research to visually check the network converted from EMME/2.

Another more advanced visualizer is OTFVis. This module is capable of either displaying snapshots of the system, or dynamically displaying the network during the simulation. The degree of congestion is reflected by the color of each car, ranging from green for free-flow speed to red for severe congestion. Since displaying large number of vehicles would significantly slow down simulation, this most efficient alternative is to create a movie using multiple snapshots taken throughout the simulation.

3.3 Car Allocation Model

The Car Allocation Model was developed by Hatzopoulou et al.(2007) using the Python programming language. The purpose of the model is to add the missing link between activity schedules and vehicle emission estimations: vehicle usage information. The model assigns a vehicle to each auto trip within each person’s activity chain, taking into account the availability of vehicles, as well as scheduling conflicts between different people in the same household. Due to lack of data on vehicle types and people’s preferences when choosing a vehicle, the model simply assumes that all vehicles are identical.
and assigns them on a first come first serve basis. People’s preferences have been simplified to a basic rule that makes people prefer the first car assigned to them.

An earlier plan was to use observed survey trips for emission modelling in order to avoid being bottlenecked by the progress made in MATSim-TASHA integration. However, TTS data lack certain information required by the Car Allocation Model to detect conflicts between people’s schedules. First of all, the model requires trip chain information. TTS records are all trip based and therefore do not have sufficient information for detecting conflicting requests for car usage. Secondly, the model requires information on both the start time and end time of each trip. Neither TTS nor TASHA provides the end time of each trip, only the start time. Trip duration was previously obtained by running EMME/2 using travel demand generated by TASHA. The EMME/2 estimation on travel time was temporally aggregated by hour and spatially aggregated by zone. An agent-based traffic assignment simulation such as MATSim is expected to provide a finer resolution on trip durations.

In addition to allocating vehicles, the model also computes vehicle soak and start emissions. These types of emissions are produced while the vehicles are stopped at, or just leaving an activity location. Therefore, schedules from TASHA provide sufficient data to calculate these emissions. Calculating exhaust emissions, on the other hand, requires information on vehicle dynamics on roads. For this research, only the car allocation portion of the program is enabled, since all emission calculations are based on MATSim outputs.

3.4 Mobile 6.2C and the Emission Calculation Program

Different types of emissions can be calculated separately. As mention above, in the TASHA-EMME/2 framework, start and soak emissions can be calculated directly from TASHA output. For this project, all types of emissions were calculated using MATSim output so that the calculations would be based on a consistent time scale. In addition, MATSim output allows exhaust emission calculations to become more sensitive to the effect of congestion by taking into account the time when vehicles are idling on congested roads.

A new program was written to perform emission calculations based on the events output by MATSim. To comply with future development plans of TASHA and ILUTE, the program was written in C# programming language. Detailed description on the calculation steps is presented in Section 6.3.2. Despite reading schedules from a different source, the program still uses the same scope and concept for calculating emissions as in the TASHA-EMME/2 framework. The program uses emission factors from
<table>
<thead>
<tr>
<th>Emission</th>
<th>Description</th>
<th>Addressed in</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exhaust running</td>
<td>Refer to emissions which exit a vehicle’s tailpipe while the vehicle is operating in a warmed-up condition. Exclude emissions occurring during a vehicle start and warm-up</td>
<td>USEPA, 2004a, USEPA, 2003a, USEPA, 2002a, USEPA, 2001k, USEPA, 2001m, USEPA, 2001x</td>
</tr>
<tr>
<td>Start</td>
<td>Refer to engine start emissions. Two types of starts exist: cold and hot starts</td>
<td>USEPA, 2004a, USEPA, 2003a, USEPA, 2002c, USEPA, 2001w, USEPA, 2001x</td>
</tr>
<tr>
<td>Hot soak emissions</td>
<td>Refer to hydrocarbon losses from fuel vapours in the intake manifold and fuel system, driven off the vehicle by the heat of the engine immediately after shut down. Usually due to small leaks in the evaporative emission control system (joints, lines, valves) and permeation of the fuel hoses and tank. Fuel tank temperature is usually close to ambient but can increase in fuel injected vehicles due to fuel returning from the hot engine compartment. Typically, tank temperatures in fuel injected vehicles can exceed ambient temperatures by 5 to 15°F. Hot soak emissions are not a direct function of ambient temperature. If the vehicle is restarted, the full hot soak effect is interrupted, resulting in fewer hot soak emissions.</td>
<td>USEPA, 2004a, USEPA, 2003a, USEPA, 2001f, USEPA, 2001h, USEPA, 2001l</td>
</tr>
<tr>
<td>Diurnal emissions</td>
<td>Refer to hydrocarbon losses from fuel vapors driven off the vehicle from the increasing temperature of the fuel in the tank and other locations on the vehicle while the engine is shut down and during times of day when the ambient temperature is rising. If the vehicle is restarted, the full diurnal effect is interrupted, resulting in fewer diurnal emissions.</td>
<td>USEPA, 2004a, USEPA, 2003a, USEPA, 2001c, USEPA, 2001d, USEPA, 2001e, USEPA, 2001g, USEPA, 2001n</td>
</tr>
<tr>
<td>Resting losses</td>
<td>Small but continuous seepage and minor leakage of gasoline vapor through faulty connections, permeable hoses, and other components of the fuel system</td>
<td>USEPA, 2004a, USEPA, 2003a, USEPA, 2001c, USEPA, 2001g</td>
</tr>
<tr>
<td>Running losses</td>
<td>Evaporative emissions which have escaped from a vehicle while the engine is operating. May appear from the evaporative canister, the fuel inlet, the top of the gas tank, and other spots where the integrity of the evaporative system has broken down or the purge system has become inoperative. May be an artifact of a particular evaporative system design or the result of poor maintenance. Because of greater heating of the fuel and evaporative system on longer trips, running loss emissions are not constant throughout a trip. The rate is assumed to continually increase as a function of trip length until it reaches a plateau at a trip length of about 50 to 60 minutes.</td>
<td>USEPA, 2004a, USEPA, 2003a, USEPA, 2001i, USEPA, 2001m</td>
</tr>
<tr>
<td>Crankcase emissions</td>
<td>Evaporative blow-by emissions, resulting primarily from defective positive crankcase ventilation (PCV) systems</td>
<td>USEPA, 2004a, USEPA, 2003a</td>
</tr>
<tr>
<td>Refuelling emissions</td>
<td>Vapours that escape into the atmosphere when incoming liquid fuel displaces vapours in the vehicle fuel tank</td>
<td>USEPA, 2004a, USEPA, 2003a</td>
</tr>
</tbody>
</table>

Figure 3-2. Descriptions of Mobile6.2 Emission Types
the same source as before, which is the Canadian version of the Mobile6.2 model developed by the United States Environmental Protection Agency (USEPA). Compared to the original, the Canadian version (Mobile6.2C) has various modifications made by Environment Canada to reflect the GTA fleet characteristics and emission control technology. HC, CO, NO$_x$ and CO$_2$ emissions are estimated for light-duty private vehicles. Figure 3-2 shows the USEPA definition for different types of emissions. Depending on the type of emission, the emission factors may be further classified by road type, speed, or time of the day. The details are presented in Section 6.3.1.

Figure 3-3 shows the program flowchart. Processes above the dashed line are called sequentially by the Program class. In order to reduce the amount of post-processing, the program accumulates emissions for the network, links/roads on the network, and households as soon as emissions are calculated. The program is also flexible enough that it is capable of outputting each calculation result with associated link or person attributes. This is useful for debugging and detailed analysis with a small sample dataset. The output for all agents in TASHA-MATSim would be very large in size, close to the size of the event file.

![Flowchart of the Emission Calculation Program](image-url)
Chapter 4. TASHA – MATSim Integration

This chapter begins with a description of the different configurations considered for TASHA-MATSim integration. Various issues encountered during the data conversion process are then described in detail. Most of the custom programs developed for dealing with these issues can be found in a special folder named Toronto playground within the MATSim environment.

4.1 Architecture

TASHA and MATSim can be integrated using different configurations. At the early stage of this research, discussions with the MATSim developers yielded three distinct configurations.

1)

![Diagram of TASHA-MATSim Configuration I]

This configuration is the one being used for this research. Among all possible configurations, this is considered to be the base, or first trial. It is very similar to the TASHA – EMME/2 configuration. TASHA generates personal tour information, but instead of building an OD trip matrix, that information is converted into a format that can be read by MATSim. MATSim simulates the tours and attempts different routes to minimize and stabilize individual travel time. Once the simulation is complete, inter-zonal travel time can be extracted from the simulation results and feedback into TASHA. Thus, the iterative procedures between TASHA and EMME/2 can be implemented in a similar fashion for TASHA and MATSim. MATSim’s performance relative to that of EMME/2 can be assessed by comparing results of the two configurations. Since the current version of MATSim can only handle auto trips, TASHA still
has to rely on EMME/2 for non-drive related data. For example, transit time and level of service data from EMME/2 are still being used. As a result, this configuration requires the least amount of data transformation. MATSim does not require any information on person or household attributes. When feeding back travel time, despite the disaggregate trip travel times available from MATSim, TASHA scheduler still needs to aggregate travel times both temporally and spatially. This greatly limits the advantages of using MATSim for travel time calculation. Although this configuration may be too similar to TASHA-EMME/2, it is a logical first step to test the feasibility of integrating TASHA with MATSim.

Figure 4-2. TASHA-MATSim Configuration II

Travel demand modelling in MATSim is a two phase process. The first phase generates an initial travel demand, and during the second phase the demand is optimized through iterative travel simulations. (Francesco et al., 2008) Following this concept, a simpler configuration is to use TASHA to generate only the initial schedule. Comparing to existing planning algorithms in MATSim, the TASHA scheduler is more sophisticated and has been proven to be able to reproduce travel patterns observed from travel surveys. Therefore, the TASHA scheduler should provide a good starting point for the optimization process in MATSim. The schedule will be later modified by the rerouting and replanning modules within MATSim. Unlike the TASHA scheduler, replanning modules within MATSim are designed to handle disaggregate travel times. Therefore, this configuration will fully utilize the advantage of having travel time information for individual trips. There is ongoing research by the MATSim developers to improve various
planning modules in MATSim. These planning modules would require information on personal and household attributes. The plan file, which is currently storing activity schedules, will likely be expanded to handle the additional information.

3)

Figure 4-3. TASHA-MATSim Configuration III

The developers from Zurich suggested a long term plan to implement TASHA within the MATSim Environment. TASHA will serve as a replanning module that modifies people’s plans during each MATSim iteration. Three main challenges have been identified for implementing this configuration. The first one is the same challenge faced in the previous configuration, which is to make MATSim capable of managing person and household attributes, as well as other data used by TASHA scheduler. The second challenge is to modify the scheduling algorithm in order to consider travel time in disaggregate form. This is a major change to the existing algorithm, and is likely to be implemented in the next version of TASHA. Finally, a new mode choice model needs to be implemented in MATSim. Moving from the simple mode choice model as described in Section 3.2.4 to one similar to that of TASHA requires coding not only for the model itself, but also a whole range of supporting utilities such as a transit network handler, or a new routing algorithm.

Comparing to the previous configuration, this configuration performs rescheduling on a different level. During each MATSim iteration, the program chooses to change the route or reschedule the activity. Depending on the parameter set in the configuration file, rescheduling can happen much more
frequently than before, and thus significantly raises computational cost. In addition, MATSim passively optimizes schedules through a trial and error process that keeps the best five schedules for each person. Allowing TASHA to make changes to the schedule at each MATSim iteration may quicken the process of optimization since people can change both route and timing. However, the goal of TASHA is to create a feasible schedule that reflects observed trip distributions. This goal may conflict with the optimization process within MATSim. In reality, people are not global optimizers and will not always generate or choose the best possible schedule. Before integrating TASHA into MATSim, more research is required in order to properly control schedule optimization.

4.2 TASHA to MATSim Data Conversion

MATSim requires input files to be in XML format. A format conversion program was written in C# to convert network files from EMME/2 into XML format. The program was written at the very early stage of this research, and thus does not belong to the MATSim environment and does not utilize the network handlers in MATSim. The development team in Berlin helped create a similar converter for activity schedules, which has already been integrated into the MATSim development environment. (Converter.java) Because this schedule converter was written while the mode choice sub-model in TASHA was still in development, significant modifications and enhancements have since been made to the converter to adapt to the new mode choice output. The current version of the converter is named ModdedConverter.java in the MATSim Toronto playground. Modifications include checking for duplicate records; shifting TASHA time forward by four hours so that MATSim simulation starts at hour 0; and improving the random coordinate generation process such that people in the same household start their trip chains at the same household coordinates.

Figure 4-4 graphically shows the numerous conversions from the EMME/2 network to MATSim network, and from TASHA schedules to MATSim plans. Preparation of the network file is relatively simple since most node and link variables required by MATSim are readily available in EMME/2. For converting the schedules, the main challenge is that information from TASHA is trip based, whereas MATSim requires an event based schedule. This means that each schedule in TASHA needs to be broken into data associated with the start and end of each activity. In addition, the coordinates of each activity need to be generated based on the coordinates of activity centroids. A built-in algorithm in MATSim was used for this task. The following paragraphs describe numerous conversions made during data conversion, each focusing on a particular attribute of the network or the schedule.
Figure 4-4. TASHA to MATSim Data Conversions
4.2.1 Network Attributes

Centroids

Both EMME/2 and TASHA locate households and activity locations at zonal activity centroids. In contrast, locations in MATSim are represented by Cartesian coordinates. The coordinates are generated by drawing a random distance that is up to a certain percentage of the distance between a zonal centroid and the centroid of the next closest zone. The percentage is suggested to be 70%, which has been used in previous MATSim implementations and proved to provide an acceptable coverage area. The distribution is not uniform. The density increases the closer it is to the activity centroid. An implication of this centroid-to-coordinates transformation is that centroids and nodes become indifferent from a modelling perspective once loaded into MATSim. The same happens to centroid connectors and actual links.

Figure 4-5. Assignment of Links to Activity Locations (Rieser et al. 2007)

(a) Given a region consisting of zones, each zone having a centroid, and a network with nodes and links.
(b) Circles are defined around zones’ centroids.
(c) Activity locations are randomly chosen within the circle for each zone. 
(d) The activity locations are assigned to the nearest link

The process of randomizing locations helps disperse the population so that they can start and end their activities on different links while still being within proximity of the zonal activity centroid. This method, however, has its limitations, and may not be any more realistic than any other randomization technique. The main issue is that an activity centroid does not imply that activities are located closer to the centroid. It is likely that the actual trip origins and destinations are located along major arterial roads at the boundary of each zone.

**Centroid Connectors**

EMME/2 uses centroid connectors to allow people to access the network from activity centroids. The connectors link zonal centroids to nearby nodes on the network. As mentioned above, MATSim disperses activity locations by assigning them Cartesian coordinates in the vicinity of the centroids. With activities starting away from the centroids, the centroid connectors become unnecessary. However, when converting the network, a decision was made to retain the connectors and let MATSim treat them as normal network links. This is mainly due to the way MATSim models trip starts. MATSim does not model the time it takes for a person to travel from starting location to the link. People may wait to enter a link if the link is congested, but otherwise, they start right on the link closest to their homes. In a city block bounded by four roads one on each side without centroid connectors, all trips within the block start on one of these four links. The simulation would treat people as if they were all living on these boundary roads. Retaining the centroid connectors, on the other hand, would help model people living within the block. Even after the simulation disperses people around the zonal centroid, a portion of them would still be close to the centroid connectors. When they start travelling, they will first spend time driving within the block before getting onto the real network. Therefore, the results should be more realistic and credible than using a network with all centroid connectors removed.

**Number of Lanes and Capacity**

In MATSim, links can have multiple lanes and are unidirectional. The network visualizer displays links with their width proportional to the number of lanes on the links. There is an attribute that defines capacity of the link, which is different from the EMME/2 definition of capacity per lane. Capacity in EMME/2 is not an absolute maximum. EMME/2 uses capacity and volume delay function to determine travel time on the link. After road capacity has been reached, vehicles are still allowed to use the road, although the travel time may become prohibitive. In contrast, link capacity in MATSim is the absolute
physical capacity of the link. Since the simulation is queue based, no more vehicles can fit onto the link when the capacity has been reached and the effect of congestion will propagate backwards to other connected links. Another implication of this modelling approach is that gridlock may occur in heavily congested areas. In this case, the simulation would have to remove some of the vehicles in order for the simulation to continue. Alternatively, it can also be configured to force vehicles onto a fully congested road, ignoring the hard capacity constraint.

Since the EMME/2 capacity is the full capacity of the network, TASHA uses household expansion factors to amplify the O-D flows generated based on 5% of the real GTA population. A MATSim simulation with full population is computationally infeasible on the computers used for this research. Therefore, to use the 5% population, MATSim network capacity needs to be reduced. This reduction factor was specified in the simulation configuration file and was applied only during the actual simulation. The reduction in storage capacity is simulated by amplifying the car length. After experimenting with different values, a flow capacity factor of 0.05 and a storage capacity of 0.1 were used.

**Turn Restrictions**

EMME/2 network has turn restrictions at intersections. A program in the MATSim Toronto playground can implement turn restrictions in the MATSim network. (ManeuverCreation.java) Simulating turn restrictions is achieved by creating multiple versions of the links that lead to the intersection, and connect them according to the rules specified. This essentially creates an imaginary interchange at the intersection. A network with turn restrictions was used for the research that compares EMME/2 and MATSim performance. (Gao, 2009) For this research, however, adding turn restrictions changes the link IDs, and makes emission calculations very complicated as the emission calculator would have to distinguish real links from imaginary ones. Therefore, turn restrictions were not implemented here.

**4.2.2 Schedules**

**Plans**

Activity schedules are called plans in MATSim. The plan file is in XML format, and different XML schemas are defined by DTD files located in the DTD directory. The list of variables required varies depending on the input file version specified in the simulation configuration file. Current MATSim runs use input schema version 4, which requires person id, activity type, x and y coordinates of the activity location, link, activity start and end time, and mode. Values for these variables are either extracted directly from TASHA output or converted using methods previously described. (Figure 4-4) Trips in TASHA have
various attributes associated with trip origin and destination. MATSim plans, on the other hand, lists changes in the person’s schedule. Thus, the main challenge during this conversion process is to change the trip based records from TASHA into event based plans in MATSim.

Optional inputs such as initial route choice or a choice set of schedules for each person can be added to the input file. They can be used to model the person’s prior knowledge of the network or different plans they have in memory. MATSim can be configured to output plans after each iteration. The output is essentially a plan file with route and schedule choice. Thus, it is possible to use this file to resume simulation at the corresponding iteration.

MATSim developers in Berlin are planning to add additional information to the plan file in the future. Fully integrating a rescheduling process into MATSim would require additional data similar to those required by TASHA scheduler. The “plan” file will eventually become a “population” file that includes both people’s schedules and additional personal and household attributes.

**Modes**

To obtain information on mode of travel, output from TASHA scheduler needs to be processed by Merge_tt as well as the mode choice model before converting into MATSim compatible file format. There is on-going research to improve the mode choice module in MATSim. Currently, MATSim handles only auto trips. There is a simple build-in mode choice program that considers auto and non-auto trips. The program simply duplicates the plan file, creating a copy of each trip, but flagged as non-auto trips. MATSim calculates the score of these non-auto trips based on factored auto travel time and specified non-auto travel cost. The resulting mode share depends on how many of the original (auto) and copied (non-auto) plans are chosen by the agents.

MATSim expects schedules to be home-based trip chains. However, when using mode choice output from TASHA, selecting all trips assigned with drive mode will lead to broken trip chains. This is mainly due to the GO access and egress portion of the chain, which involve driving even though it is not identified as drive mode. This issue was encountered again for the emission modelling component of this research, and a program was written to extract car schedules from personal travel schedules.
4.3 MATSim Simulation Configurations

MATSim simulation parameters can be modified within a configuration file. The parameters are organized into different categories that manage different aspects of the model, such as input/output paths, number of iterations, network capacity modifiers, and values for scoring function parameters. Parameters not listed in the configuration file will take on the default values set in the MATSim code. Figure 4-6 shows the layout of a configuration file used for a standard TASHA-MATSim run. To ensure MATSim has the most recent information on file format and parameter definitions, the simulation will access DTD file in the online database for format specifications. The complete configuration file used is presented in Figure A-6 in the Appendix.

```xml
<?xml version="1.0" encoding="utf-8"?>
<!DOCTYPE config SYSTEM "http://www.matsim.org/files/dtd/config_v1.dtd">
<config>
  <module name="global">
    Global parameters such as time format and coordinate system
  </module>
  <module name="network">
    Network filename and path
  </module>
  <module name="plans">
    Plan filename and path
  </module>
  <module name="controller">
    Specify output directory and number of iterations
  </module>
  <module name="simulation">
    Specify start time and end time of simulation, link capacity factors, snapshot period, etc
  </module>
  <module name="planCalcScore">
    Specify scoring function parameter values
    Define activity types, priority, and typical duration
  </module>
  <module name="strategy">
    Specify the probability of performing various replanning and rerouting strategies
  </module>
</config>
```

Figure 4-6. MATSim Configuration File Layout
4.4 MATSim to TASHA Data Conversion

4.4.1 Event File

For each iteration of the simulation MATSim generates a file that stores all the events occurred during the simulation period, including the type of the event, the agent it is associated with, and the time of its occurrence in seconds. An event is defined as a change in the agent’s status, such as entering or exiting a link. This method of storing information is different from other software that directly provides information about each agent’s trip. To extract trip information such as trip duration, one must search through the event file to identify the event of the agent leaving the origin link and the event of arriving at the destination link, and then take the difference in time. Many other files generated by MATSim at the end of the simulation run are in fact based on the Event File. MATSim automatically post-processes these data to give more meaningful statistics such as trip start time distribution or link usage statistics. Extracting additional information from the Event File can be tedious and computationally intensive for a full scale simulation. The number of records far exceeds the limit Access or Excel can handle. (A TASHA-MATSim run using 5% GTA population generates approximately 20 million events.) As a result, it is necessary to write small data processing programs to extract information from the Event File. The code is preferably to be either developed within the MATSim environment using JAVA, or written in C# in order to be compatible with the future version of TASHA. The former requires a general knowledge of the various event handlers in MATSim, as many standard procedures such as sorting and searching events have already been coded into the program.

4.4.2 Link Statistics

LinkStats is an analysis module that extracts link information from the Event file. For each link in the network, the program attaches link attributes, and calculates maximum, average, and minimum link flows and speeds for every hour. The calculations are based on enter and leave link events, and link characteristics specified in the network file. The output, unfortunately, is in a rather inconvenient format, with more than two hundred fields for each link record. For this research, LinkStats was used at the early stage to make sure link flows are reasonable. The analysis program for emission calculations avoids using LinkStats output, and extracts network information directly from either the network file or the Event file.
4.4.3 Constructing the Travel Time Matrix

Travel time for each trip can be easily computed using departure and arrival events in the Event file. However, since TASHA still relies on inter-zonal travel times for peak and non-peak hours, output from MATSim needs to be aggregated both temporally and spatially. Temporal aggregation is relatively trivial, the hour of the trip is determined using the start time. This method is consistent with that used in TASHA-EMME/2, and has the same problem that a trip will be assigned to the hour of the start time even if a majority of it takes place in the next hour.

Aggregating travel times spatially is a more complicated procedure that requires mapping of links to traffic zones. One option is to query travel time for each trip with household and activity locations in order obtain inter-zonal travel time. However, since not all possible zone pairs have trips between them, this method results in an incomplete travel time matrix. Therefore, a link-based method was implemented instead. This method requires a mapping between links and zonal centroids. A custom module Link2ZoneMapping.java in the Toronto playground handles the mapping procedure. The algorithm is similar to the one used during simulation that connects trip starting locations to a link on the network. Multiple links are mapped to one zone, depending on the distance between the midpoint of the link and the zonal centroid. The mapping results are then used by the travel time calculator EventsToTTMatrix.java that calculates average zone-to-zone travel time by hour. Intra-zonal travel times can also be calculated as long as there are at least two links mapped to the same zone.

The mapping results show that all zones have been assigned at least one pair of links, with the exception of zone #209, which is the Toronto Island. Since there are only two long links connecting the island centroid to zones to nodes in Harbour Front, these links are assigned to zones in Harbour Front instead. To maintain the completeness of the travel time matrix, these links need to be mapped back to the island centroid. This problem has to be manually corrected at the moment until improvements to the mapping procedure are made.

Although zone #209 is an extreme case, it does indicate that this link-to-zone mapping approach becomes less reliable when there are very few links associated with a zone. Each link is assigned to only one zone even though people starting at that link may come from different zones. Especially for large suburban zones with only a few links at its boundary, it is unrealistic to map a boundary link to only one zone. However, for these suburban zones, it is likely that there is less congestion and the travel time on adjacent links can still provide good estimation so that the average inter-zonal time does not deviate significantly from reality.
Chapter 5. TASHA-MATSim Modelling Results

Three TASHA-MATSim simulation runs were performed, each with 50 MATSim iterations. The first two were standard runs with TASHA output directly feeding into MATSim. The last run had the TASHA output processed by the Car Allocation Model in order to prepare for emission calculations. Another set of simulation was performed using a different configuration file with Time Allocation Mutator enabled. This section first summarizes the various statistics produced by build-in analysis modules in MATSim, followed by a discussion of the evolution of TASHA schedule, mode choice, and travel time through the iterations. Finally, the effects of the Time Allocation Mutator on travel time, trip length and duration are analyzed.

5.1 Score Convergence

Figure 5-1 shows the score statistics obtained during the 50 MATSim iterations. Each person can remember up to five schedules, and for each new schedule evaluated, the one that has the lowest score will be removed from the choice set. For each iteration, the plot shows the average of people’s best, worst, and average score, as well as the one chosen for that iteration. Each MATSim iteration took about eight minutes to complete, leading to a total time of approximately six hours for a MATSim run with 50 iterations.

![Score Statistics](image.png)

Figure 5-1. MATSim Simulation Score Statistics
The plot shows that even after 50 iterations, the average of people’s best scores has not completely converged. Complete convergence could be achieved after 300 iterations. The score first improved quickly, but then the increase after 40 iterations became very small. Since only the rerouting algorithm was used to change people’s schedules, there were very limited possibilities for further improvement after the first few trials. Therefore, it is sufficient to use 50 iterations for this research.

The worse score decreased during earlier iterations as people added more schedules to their choice set. Once most people had already attempted five different schedules, the average of their worst score started to improve as they started to remove the schedule with the lowest score. However, again due to the limited ways to improve their schedule, the average of people’s lowest scores did not improve much before converging.

5.2 Distributions

5.2.1 Trip Departures and Arrivals

Since people cannot change activity times through MATSim iterations, the trip departure and arrival distributions after 50 iterations are still identical to those initially generated by TASHA. The distribution in Figure 5-2 shows that a large number of trip departures occurred at the beginning of the hour or half hour, especially during the afternoon peak. This issue was not observed in the TASHA-EMME/2 model because the O-D flows were aggregated by hour. The plot in Figure 5-2 is in 5 minute intervals, so that the peak in the afternoon appears to be a lot larger than the one in the morning. The large departure peaks can also be observed in the en route plot, which becomes chaotic during the afternoon peak. Finally, a small peak occurs at the final hour of the simulation as TASHA forces everyone to return home at the end of the day.

From a simulation point of view, these sharp peaks in travel demand make the simulation unable to allocate everyone onto the network in one time step. This results in longer waiting time and thus lower score for the trip. The lower score does not influence people’s choice of schedule because it does not affect the relative comparison between different routes when the departure time is fixed. Another problem caused by these sharp peaks is the sudden increase in congestion on the road, especially for short links. When it takes too long for a vehicle to enter a link, MATSim assumes that vehicle is lost and removes it from simulation. With flow capacity factor at 0.05 and storage capacity at 0.1, approximately 1% of the vehicles were lost during simulation, mostly at peak hours. This should not significantly affect the average hourly zone-to-zone travel time.
5.2.2 Trip Length and Duration

Using build-in modules in MATSim, the average trip length and duration by hour for each TASHA-MATSim iteration were computed and plotted in Figure 5-3 and Figure 5-4, respectively. The results indicate that there is very little difference between iterations. The only noticeable difference occurs at 6am where the average trip length produced by the first TASHA run using EMME/2 base travel time is slightly longer. The plots for the following runs are almost identical. Average trip length starts at a peak of 25km at 4am, gradually decreases throughout the day until reaching the afternoon peak, where another high of 17km can be observed. This suggests that the few people who travel very early in the morning have to do so because their destinations are far away. The afternoon peak also reflects long return home trips from downtown Toronto to suburban areas. However, average trip length in the afternoon is reduced by the large number of short shopping and leisure trips.
A similar trend can be observed for the average trip duration plot, with the exception of long trips in early morning having much shorter duration due to lack of congestion. Average trip duration peaks at 36 minutes in the morning and 25 minutes in the afternoon. The plots do show some sensitivity to changes in zone-to-zone travel time for schedules in the morning peak at 7am. Average trip duration decreases when the zone-to-zone travel time increases, indicating a possible shift in trip start time or trip distribution.
### 5.2.3 Mode Split

Figure 5-5 shows the output of the TASHA mode choice model in the form of hourly mode split expanded by household expansion factors. Comparing to results from TASHA-EMME/2, the general trend in trip counts by mode is similar throughout the day.

#### Whole Day

<table>
<thead>
<tr>
<th>Run</th>
<th>Unknown</th>
<th>Auto</th>
<th>Tran</th>
<th>Walk</th>
<th>Ride</th>
<th>Psg</th>
<th>GoAcc</th>
<th>GoEgr</th>
<th>GoTrn</th>
<th>Number of Trips</th>
</tr>
</thead>
<tbody>
<tr>
<td>MATSim Iter 1</td>
<td>3.20%</td>
<td>58.09%</td>
<td>17.71%</td>
<td>3.55%</td>
<td>8.56%</td>
<td>7.87%</td>
<td>0.09%</td>
<td>0.09%</td>
<td>0.84%</td>
<td>9777900</td>
</tr>
<tr>
<td>MATSim Iter 2</td>
<td>3.20%</td>
<td>58.06%</td>
<td>18.18%</td>
<td>3.61%</td>
<td>8.46%</td>
<td>7.42%</td>
<td>0.12%</td>
<td>0.12%</td>
<td>0.83%</td>
<td>9759996</td>
</tr>
<tr>
<td>MATSim Iter 3</td>
<td>3.20%</td>
<td>58.05%</td>
<td>18.22%</td>
<td>3.58%</td>
<td>8.47%</td>
<td>7.40%</td>
<td>0.12%</td>
<td>0.12%</td>
<td>0.85%</td>
<td>9760870</td>
</tr>
<tr>
<td>EMME/2 Iter 3</td>
<td>3.19%</td>
<td>57.95%</td>
<td>18.02%</td>
<td>3.55%</td>
<td>8.52%</td>
<td>7.82%</td>
<td>0.16%</td>
<td>0.16%</td>
<td>0.63%</td>
<td>9773003</td>
</tr>
</tbody>
</table>

#### AM Peak

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Unknown</th>
<th>Auto</th>
<th>Tran</th>
<th>Walk</th>
<th>Ride</th>
<th>Psg</th>
<th>GoAcc</th>
<th>GoEgr</th>
<th>GoTrn</th>
<th>Number of Trips</th>
</tr>
</thead>
<tbody>
<tr>
<td>MATSim Iter 1</td>
<td>3.38%</td>
<td>59.00%</td>
<td>17.78%</td>
<td>4.32%</td>
<td>1.85%</td>
<td>12.21%</td>
<td>0.28%</td>
<td>0.02%</td>
<td>1.15%</td>
<td>2522024</td>
</tr>
<tr>
<td>MATSim Iter 2</td>
<td>3.39%</td>
<td>58.83%</td>
<td>18.45%</td>
<td>4.40%</td>
<td>1.85%</td>
<td>11.58%</td>
<td>0.37%</td>
<td>0.02%</td>
<td>1.12%</td>
<td>2521141</td>
</tr>
<tr>
<td>MATSim Iter 3</td>
<td>3.38%</td>
<td>58.86%</td>
<td>18.43%</td>
<td>4.42%</td>
<td>1.85%</td>
<td>11.54%</td>
<td>0.36%</td>
<td>0.02%</td>
<td>1.15%</td>
<td>2520392</td>
</tr>
<tr>
<td>EMME/2 Iter 3</td>
<td>3.37%</td>
<td>58.70%</td>
<td>18.19%</td>
<td>4.35%</td>
<td>1.85%</td>
<td>12.18%</td>
<td>0.52%</td>
<td>0.03%</td>
<td>0.82%</td>
<td>2522530</td>
</tr>
</tbody>
</table>

#### PM Peak

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Unknown</th>
<th>Auto</th>
<th>Tran</th>
<th>Walk</th>
<th>Ride</th>
<th>Psg</th>
<th>GoAcc</th>
<th>GoEgr</th>
<th>GoTrn</th>
<th>Number of Trips</th>
</tr>
</thead>
<tbody>
<tr>
<td>MATSim Iter 1</td>
<td>3.19%</td>
<td>57.16%</td>
<td>17.87%</td>
<td>3.89%</td>
<td>9.06%</td>
<td>7.61%</td>
<td>0.03%</td>
<td>0.17%</td>
<td>1.01%</td>
<td>3711244</td>
</tr>
<tr>
<td>MATSim Iter 2</td>
<td>3.20%</td>
<td>57.12%</td>
<td>18.37%</td>
<td>3.97%</td>
<td>8.92%</td>
<td>7.14%</td>
<td>0.02%</td>
<td>0.23%</td>
<td>1.03%</td>
<td>3699040</td>
</tr>
<tr>
<td>MATSim Iter 3</td>
<td>3.21%</td>
<td>57.11%</td>
<td>18.40%</td>
<td>3.90%</td>
<td>8.93%</td>
<td>7.13%</td>
<td>0.03%</td>
<td>0.23%</td>
<td>1.06%</td>
<td>3700202</td>
</tr>
<tr>
<td>EMME/2 Iter 3</td>
<td>3.19%</td>
<td>57.01%</td>
<td>18.25%</td>
<td>3.89%</td>
<td>8.97%</td>
<td>7.52%</td>
<td>0.02%</td>
<td>0.33%</td>
<td>0.80%</td>
<td>3706167</td>
</tr>
</tbody>
</table>

---

*Figure 5-5. Mode Split by Hour of Day*

*Figure 5-6. Mode Share Percentages for the Whole Day and Peak Periods*
With the small changes in inter-zonal travel time, the mode share percentages are very stable throughout iterations, and are very close to the results from TASHA-EMME/2. A relatively large deviation can be observed for mode involving GO Rail since the associated access/egress times are not updated.

5.2.4 Travel Time Matrix

Figure 5-7 shows the zone-to-zone travel time calculated using the link-to-zone mapping procedure in MATSim. The average peaks at 53 minutes in the morning and 50 minutes in the afternoon. The averages for the peak periods are even smaller due to the sharp decrease in travel time in adjacent hours. These average travel times are lower than those estimated by EMME/2, which predicts 55 minutes and 53 minutes for the morning and afternoon peak period, respectively. The off-peak travel time, however, is higher, averaging around 41 minutes compared to the EMME/2 average of 37 minutes. This is expected because the free-flow travel time in EMME/2 is computed using the actual free-flow speed, which should be the minimum possible travel time between zones.

The purpose of comparing average travel times from the two models is to show the difference of the travel time matrices feeding back to TASHA. Since most of the origin-destination pairs do not have any actual trips between them, for comparing the performance of MATSim and EMME/2, travel times weighted by number of trips should be used. This was done by Gao (2009), and the results suggest that MATSim travel times are more sensitive to the effect of congestion. Weighted travel time in MATSim is higher than in EMME/2, even for the peak periods. However, the minimal differences in mode splits imply again that TASHA modechoice model lacks sensitivity to travel times.

![Figure 5-7. Travel Time Distribution](image-url)
5.3 Visualization

MATSim simulation results can be visualized using the build-in visualizer OTFVis. The visualizer uses the event file to produce snapshots of the system at specified intervals. Agents are represented by little car icons on the links. The color of the icons ranges from green to yellow to red, respectively indicating different levels of congestion: free flow traffic, mild congestion, and saturated links. The spacing between icons on the links is also an indication of travel speed. Figure 5-8 shows the network at 8:35 in the morning. A large number of links are congested, including many arterial roads in downtown, a large section of the Gardener Expressway along Lakeshore west of Bathurst St., and Don Valley Parkway south of Eglington Ave. Hourly snapshots are presented in the Appendix, and they are compared to emission calculation results plotted by ArcGIS.
5.4 Effects of the Time Allocation Mutator (TAM)

TAM is a build-in module in MATSim that modifies activity start time. It is often used in conjunction with the rerouting algorithm and other more sophisticated replanning modules to diversify modifications made to schedules throughout MATSim iterations. With the use of TASHA as an external provider of schedules, TAM was removed from the list of scheduling strategies. The major concern is that TASHA would be unaware of the changes made by TAM, and thus modify schedules based on inaccurate travel time information. However, as shown in Section 5.2, the synthesized activities in TASHA have their start times concentrated at specific times such as the beginning of an hour or half hour. This causes the queue based simulation to produce longer travel time because people are waiting to be put onto the network. MATSim developers from Zurich suggested using TAM to add minor shifts in start time in order to smooth out the start time distribution. It is important to make only small, infrequent modifications to start time (5 to 10 minutes, 10% possibility of choosing this module) since TASHA is not aware of these changes.

Figure 5-9. Score Statistics for MATSim Simulation with Time Allocation Mutator
Comparing to the score statistics obtained previously, (Figure 5-1) the score plot in Figure 5-9 suggests that it takes much longer for the scores to converge. It takes about 100 iterations for the scores’ rate of increase to drop to a level comparable to that obtained previously. Due to the added flexibility in start time, people have more means to avoid congestion and shorten their trip duration. As a result, the score converges to a higher value using the same initial schedule as before. Moreover, the average of people’s worst scores, which previously did not improve after 50 iterations, has now improved greatly.

With TAM, the percentage of lost vehicles is reduced from 1% to approximately 0.5%. As shown in Figure 5-10, the trip departure distribution has smoothened considerably, and it is clear that the arrival distribution follows the departure distribution throughout the day. From the en route plot, it appears that while the morning peak stays at 8am, trips in the afternoon have shifted to a later time.

Figure 5-10. Distribution of Trip Departures and Arrivals at Iteration 100 with TAM Enabled
As shown in Figure 5-11, the shift in departure time caused by TAM has altered travel time on the network. The afternoon congestion period has shifted to a later time. The changes in the morning, however, are much smaller. This is expected since the departure peaks in the afternoon were much larger before, and the effect of TAM on overall average travel time should be directly related to how much the peaks have been “trimmed”.

![Figure 5-11. Effect of the Time Allocation Mutator on Zone-to-Zone Travel Time](image)

Average trip length has increased in the morning at 5am and also at hours in the late afternoon. (Figure 5-12) This is due to the decrease in the frequency of applying the rerouting strategy, resulting in longer, less efficient routes. In addition, from the various plots presented in this section, it is clear that the effect of shifting departure time is much more effective in improving the score than attempting a new route. Thus, even if the rerouting strategy had been selected, the resulting plan would be easily bested by one produced by the Time Allocation Mutator, and would likely be discarded later.

With TAM, the average trip duration decreases and has a smoother distribution throughout the day. A noticeable shift can be seen even in the morning peak period, where no extreme departure peaks were observed before. This means that TAM has reordered people’s departure times so that more of them can avoid congestion. As described before, TASHA is unaware that the travel time matrix sending back from MATSim is no longer based on the original input schedules. More importantly, it is difficult to determine and validate the right amount of optimization provided by TAM. Shifting trips away from
peak hour will definitely improve travel time and thus the score (Figure 5-13), but the results can no longer represent the observed trip distributions. Therefore, for the emission modelling, TAM was not used in order to be consistent with previous research using EMME/2. To use TAM and other schedule optimization techniques, more research is required in the calibration of parameters representing the disutility of deviating from scheduled departure and the utility of travel time savings.

Figure 5-12. Trip Length Comparison

Figure 5-13. Trip Duration Comparison
Chapter 6. Modelling Auto Emissions

This part of the research attempts to improve the previous average speed emission model by exploiting the agent-based output generated by the TASHA-MATSim model. This new type of agent-based output allows emissions on the network to be calculated without losing linkage to each household agent. This enables emissions on the network to be tracked back to those who are producing them, allowing for analysis by household location, or various personal attributes. Exhaust, start, and evaporative (soak) emissions are calculated. Since the output is event based, the exhaust emission can be separated into two components, with one produced by vehicle travelling at free-flow speed, and the other produced by idling on congested road. This makes exhaust emissions more sensitive to the effect of congestion.

There are three major steps involved: first is to convert travel schedules into car usage schedules using the Car Allocation Model; the second step is to construct look-up tables for different types of emission factors; finally, a calculation program estimates auto emission based on vehicle usage information and emission factors. A complete flow chart of this portion of research is presented in Figure A-5 in the Appendix. The following sections describe the procedures in detail.

6.1.1 Data Preparation for the Car Allocation Model

As previously described in Section 3.3, the Car Allocation Model provides the missing link between personal schedules and vehicle emissions, namely car usage information. Inputs for the Car Allocation Model are derived from personal trip chain information predicted by the modechoice sub-model within TASHA. Since emissions are generated whenever a vehicle is being used, the input should include all auto trips generated by TASHA. The difficulty lies in extracting auto trips that are imbedded in a non-pure-auto chain. Due to the large number of records involved, a data cleaning program (DataCleaner.java) was written to remove non-auto trips and chains. The program is designed to track location of cars during the day. For example, a person drives to work, performs a shopping activity after work by walk or transit, then comes back to the work location to get the car and drive back home. The program will ignore the non-drive components of the trip chain. The result would be as same as if the person had stayed with the car all day.

Theoretically, the same treatment could be applied to trips that access to and egress from GO-Rail stations. This would require looking up the access and egress station number. Although this information is currently available in the output of Merge_tt from TASHA, it cannot be used to locate cars parked at stations. This is because Merge_tt chooses access and egress stations separately using a logit model.
With this method, stations with similar utilities become indifferent and there is no guarantee that the egress station of the egress trip is same as access station of the access trip. Omitting the access and egress GO-Rail trips should not significantly affect emission estimations. First of all, this mode accounts for only a small percentage of the total number of trips. Secondly, since most people using GO-Rail live in suburban areas and the main motivation for drive-access GO-Rail is to avoid congestion in downtown, the auto portions of these trips most likely take place far away from downtown and contribute little to the overall pollution.

Results from the Car Allocation Model indicate that a few of the auto trips assigned by TASHA are infeasible due to conflicting requests for car usage by different household members. These conflicts are caused by inconsistencies between MATSim travel times, and the more aggregated travel times used by TASHA scheduler. Since the Car Allocation Model uses hourly inter-zonal travel times from MATSim, the end time of an earlier trip may occasionally be later than the start time of the next trip. The Car Allocation Model is capable of handling such cases when the difference is small. However, on rare occasions where one extremely long travel time delays multiple trips, the Car Allocation Model would consider the trip infeasible and remove it from the schedule in order to preserve other trips. This in turn breaks the home-based trip chain MATSim requires as input. To prevent this problem from occurring, an additional function was added to the data cleaning program before the Car Allocation Model to detect extreme travel times and replace them with aggregated times used by TASHA scheduler.

6.2 MATSim Simulation

Outputs from the Car Allocation Model are first integrated into TASHA schedules by linking the household, person, and trip ID in Access. The schedule converter in MATSim was modified to use household-person-car as agent ID, instead of household-person. As a result, MATSim simulation now treats each person-car combination as an agent. With this special ID, emission calculation results can be accumulated by person or by car. This is useful when emission factors for different vehicle types are available. Other than the modification made to the agent ID, the simulation is identical to the previous TASHA-MATSim runs. The same simulation configuration file with only rerouting algorithm enabled was used. The MATSim output contains most of the information required for looking up emission factors, with the exception of link type, which is stored in the MATSim network file and can be retrieved using link ID.
6.3 Emission Calculations using MATSim Outputs

6.3.1 Emission Factors

Four types of emissions are calculated: exhaust, idling, start, and soak. Exhaust and idling emissions account for vehicle emissions on the roads, whereas start and soak occur at activity locations. Emission factors extracted from Mobile 6.2C are categorized by various dimensions such as road type, emission type, or hour of the day, and are placed in their respective files in comma-delimited format. The layout of each look-up table is presented in the table below. Emission factor files must strictly follow this format since the calculation program does not use a header or any other index to identify these emission factors. The starting hour of the day is set to 4am to comply with TASHA time.

<table>
<thead>
<tr>
<th>Exhaust Emission</th>
<th>Hour of the Day (Starting from 4am)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freeway HC</td>
<td>speed 2.5 – 10mph, 2.5mph increments</td>
</tr>
<tr>
<td></td>
<td>speed 10 – 65mph, 5mph increments</td>
</tr>
<tr>
<td>CO</td>
<td>speed 2.5 – 10mph, 2.5mph increments</td>
</tr>
<tr>
<td></td>
<td>speed 10 – 65mph, 5mph increments</td>
</tr>
<tr>
<td>NO\textsubscript{x}</td>
<td>speed 2.5 – 10mph, 2.5mph increments</td>
</tr>
<tr>
<td></td>
<td>speed 10 – 65mph, 5mph increments</td>
</tr>
<tr>
<td>CO\textsubscript{2}</td>
<td>speed 2.5 – 10mph, 2.5mph increments</td>
</tr>
<tr>
<td></td>
<td>speed 10 – 65mph, 5mph increments</td>
</tr>
<tr>
<td>Arterial</td>
<td>Same Categories as Above</td>
</tr>
<tr>
<td></td>
<td>Emission Factors (gram/mile)</td>
</tr>
<tr>
<td>Local</td>
<td>Same Categories as Above</td>
</tr>
<tr>
<td></td>
<td>Emission Factors (gram/mile)</td>
</tr>
<tr>
<td>Ramp</td>
<td>Same Categories as Above</td>
</tr>
<tr>
<td></td>
<td>Emission Factors (gram/mile)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Start Emission</th>
<th>Hour of the Day (Starting from 4am)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HC</td>
<td>soak duration 1 – 30 minutes, 1 minute increments</td>
</tr>
<tr>
<td></td>
<td>duration 30 – 60 minutes, 2 minute increments</td>
</tr>
<tr>
<td></td>
<td>duration 60 – 720 minutes, 30 minute increments</td>
</tr>
<tr>
<td>CO</td>
<td>Same Categories as Above</td>
</tr>
<tr>
<td></td>
<td>Emission Factors (grams)</td>
</tr>
<tr>
<td>NO\textsubscript{x}</td>
<td>Same Categories as Above</td>
</tr>
<tr>
<td></td>
<td>Emission Factors (grams)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Soak Emission</th>
<th>Hour of the Day (Starting from 4am)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HC</td>
<td>soak duration 1 – 60 minutes, 1 minute increments</td>
</tr>
<tr>
<td></td>
<td>Emission Factors (grams/hour)</td>
</tr>
</tbody>
</table>

*Figure 6-1. Emission Factor File Format*
Before using these emission factors for emission calculations, various statistical analyses have been performed to check the validity of these factors. For HC, CO, and NO₅ exhaust emissions, the emission factors are expected to be sensitive to vehicle speed and temperature. Therefore, the factors are plotted against the fifteen speed categories (Figure 6-2) and time of the day (Figure 6-3). The results indicate that the magnitude of emission increases exponentially when the vehicle travels at very low speed. There is also a small difference between travelling on a freeway and travelling on an arterial road. The other two road types, local and ramp, are not considered since Mobile6.2C assumes a constant speed for travelling on those two types of road.

Figure 6-2. HC Exhaust Emission at 7AM

Figure 6-3. HC Exhaust Emission Factor by Hour and Speed
Figure 6-3 shows the emission factors by hour of the day. It indicates that the highest emission rate occurs at 3pm. This is due to the temperature profile Mobile6.2C uses as input, which has the highest temperature of the day occurring at 3pm. The relationship between emission factor and vehicle speed can also be visualized in this figure, by examining the spacing between the plots.

While the HC emission factors monotonically decreases as speed increases, those for CO and NOx behave differently. For example, Figure 6-4 shows a three dimensional plot of the exhaust emission factors for CO. Comparing to those of HC, a similar variation in magnitude throughout the day can be observed. However, the emission factors reach a minimum at an intermediate speed of approximately 30mph, then slowly increase again for higher speed. The rate of increase appears to be dependent on the time of day. This observation proves that it is necessary to categorize emission factors into these dimensions in order to capture the non-linear variations in magnitude, thus justifying the need of having a large emission factor look-up table.

Figure 6-4. 3-D Plot of CO Exhaust Emission Factor
Comparing to exhaust emission factors, start and soak emissions factors are easier to manage since they are not dependent on the roadway type or travelling speed of the vehicles. Start and soak emissions are shown in Figure 6-5 and Figure 6-6, respectively. Both factors vary for different hours of the day. Similar to exhaust emission factors, this is mainly due to the difference in ambient temperature. The color gradient shown in the graphs ranges from red for high temperature to purple for low temperature. The start emission is affected by the duration of soak the vehicle has experienced prior to start. For soak emission, its relationship to soak duration is not linear. The soak duration is capped at 60 minutes, after which the vehicle is assumed to have reached ambient temperature.

![Figure 6-5. HC Start Emission Factors](image1)

![Figure 6-6. HC Soak Emission Factors](image2)
### 6.3.2 Calculation Procedures

The table below shows the calculation steps for each type of emission. The class structure of the program can be found in Section 3.4.

<table>
<thead>
<tr>
<th>Event Type</th>
<th>Variables Updated / Recorded</th>
<th>Values Calculated*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leave Home</td>
<td>- update chain number</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- reset trip number</td>
<td></td>
</tr>
<tr>
<td>Departure</td>
<td>If this is the first departure**:</td>
<td>If previous event is Arrival:</td>
</tr>
<tr>
<td></td>
<td>- record all event attributes***</td>
<td>- soak duration = Departure time – Arrival time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- calculate soak emission at time of Arrival:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• duration is capped at 1 hour</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• if soak finished in the next hour, separate soak duration into two components</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• look up soak emission factors (grams/hour) for each component of the duration</td>
</tr>
<tr>
<td></td>
<td></td>
<td>➢ Soak Emission = Σ emission factor x duration</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Calculate start emission at time of Departure:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• check duration category</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• obtain hour of day using Departure time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• look up start emission factor in grams</td>
</tr>
<tr>
<td></td>
<td></td>
<td>➢ Start Emission = emission factor</td>
</tr>
<tr>
<td>Enter Link</td>
<td>- record all event attributes</td>
<td></td>
</tr>
<tr>
<td>Leave Link</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arrival</td>
<td>- update all event attributes</td>
<td></td>
</tr>
<tr>
<td>Stuck</td>
<td>- update lost vehicle counter</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**
* All values calculated are then amplified by the household expansion factor.
** Since start emission calculation requires information on the previous arrival event, each person’s first departure is recorded and then “playback” at the end to calculate overnight soak duration.
*** Event attributes include: Time, CarID, LinkID, Event Type

*Figure 6-7. Emission Calculation Procedures*
While the procedures used for calculating start and soak emissions are identical to the ones used in the TASHA-EMME/2 model, exhaust emission is calculated differently. TASHA-EMME/2 uses average link speed to model exhaust emission. However, to account for the effect of congestion, the ideal input data for exhaust emission should have detailed vehicle dynamics including acceleration, deceleration, and timely update on actual travelling speed. Such information can only be obtained from a microscopic traffic simulation model. The traffic information produced by the queue based MATSim simulation lies in between the two extremes. From the event file, one can extract the actual link travel time and compare it to the theoretical free-flow travel time. The difference between the two can be used as an estimate on the degree of congestion. Therefore, emission on the road is separated into two components, with the exhaust emission produced by vehicle travelling at free-flow speed, and idling emission produced by vehicle stuck in congestion. In term of emission factor, the idling emission factor is just the exhaust emission factor at the lowest speed category. This method is a small step towards adding congestion sensitivity into exhaust emission calculations. The emission results are expected to be greater than those obtained from traditional average speed model, due to the nonlinear relationship between the emission factor and speed.

The first departure of the day for each car requires special treatment, since the start emission calculation depends on the duration of the overnight soak. As a result of TASHA being a 24 hour simulation, simply assuming a long overnight soak period may be insufficient as people who return home very early in the morning (before 3am) may soon start another trip to work. To solve this problem, the program stores all first departures in memory. Once all other events, including all the last arrivals, have been processed, the program then works through the list of first departures to finish the calculations.

As mentioned in Section 3.4, due to the large number of events, the program has to accumulate the desired outputs as soon as each value is calculated. Currently, the program produces seven output files: emission by household, by person, for the entire network, and by link for each of the four emission types.
Chapter 7. Emission Modelling Results and Analysis

Emission modelling results can be analyzed in many different ways depending on the scope of research and the dimensions available during calculation. The following is an incomplete list of possible dimensions that emission results can be plotted against. This section starts with the most aggregated results, then moves on to emission by agent-based attributes. Finally, the hourly link emissions are plotted using ArcGIS.

- **Emission**: type of emission, type of pollutant
- **Trip**: time, activity type, trip purpose
- **Person**: age, occupation, gender
- **Household**: household location, dwell type, number of cars
- **Network**: link, region, road type

### 7.1 Total Emissions

Total daily emission for the entire network was computed. The network produced a total of 63.95 tons of HC, 1362 tons of CO, 75.62 tons of NO\(_x\), and 25.81 kilotons of CO\(_2\). Exhaust emission has the largest percentage share in all four cases. (Figure 7-1)

Figure 7-2 compares these results to those using EMME/2. Since EMME/2 uses average speed to calculate emission, the current emission calculation program was modified to follow the same procedure and produce another set of outputs. Comparisons between these three sets of emission values indicate that using MATSim resulted in higher emission estimates. The difference is approximately 7% between MATSim average speed and EMME/2 for CO, NO\(_x\), and CO\(_2\). Separating travel into free-flow and idling components resulted in slightly higher emission estimate for all four compounds. This is expected due to the nonlinear emission profiles. CO\(_2\) emission factors are not affected by speed. Therefore, the increase for CO\(_2\) should be purely due to the differences in simulation. Although trip length has been shown to be slightly shorter in MATSim than EMME/2, (Gao, 2009) the emission calculation is based on time of actual enter and leave link events, and thus bypassing the inaccurate trip length calculations. All parts of the trips occurring on the network are accounted for,
including the intra-zonal portion not captured by EMME/2. As a result, emission estimates using MATSim are slightly higher. Another observations is that the difference in HC emission is much larger, over 16%, due to the relatively large exhaust emission factor at low speed making the effect of idling more apparent.

Figure 7-1. Emission by Type for each Pollutant (HC, CO, NO\textsubscript{x} in tons, CO\textsubscript{2} in kilotons)

<table>
<thead>
<tr>
<th></th>
<th>Emission in tons</th>
<th>% Difference from EMME/2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HC</td>
<td>CO</td>
</tr>
<tr>
<td>EMME2</td>
<td>36.33</td>
<td>1103</td>
</tr>
<tr>
<td>MATSim av. Speed</td>
<td>42.30</td>
<td>1180</td>
</tr>
<tr>
<td>MATSim Exhaust + Idling</td>
<td>43.16</td>
<td>1248</td>
</tr>
</tbody>
</table>

Figure 7-2. Exhaust Emission Comparisons between Different Modelling Methods
7.2 Emissions by Hour

Results from the emission calculation program were aggregated by hour. Start emission has very short duration so the program assumes that it occurs at the time of departure. Soak begins at the time of arrival and the program is able to properly handle the case where the soak period extends to the next hour. For exhaust and idling emissions, the hour is determined using the time vehicle entered link. This estimation is considered to be sufficient, since the time spent on most links are much shorter than one hour, and the difference between emission factors from adjacent hours is very small.

Hourly emissions were plotted in three graphs: the first one shows the total emission by hour (Figure 7-3); the second one separates hourly emission by type; (Figure 7-4) the last figure shows the percentage share of different emission types. (Figure 7-5) The graphs have been generated for all four compounds, but only the ones for HC are presented here since it includes all four types of emissions, and the trends observed are common for all four compounds.

![Figure 7-3. Total HC Emission by Hour](image)

Figure 7-3 shows that total emission peaks at 8am and 5pm, producing 6.29 and 6.57 tons of HC respectively. There is a drastic decrease in emission in the middle of the day between the two peak periods. From Figure 7-4, it can be observed that this decrease is mostly due to the decrease in exhaust and idling emission. There is very little idling emission between the peak periods when there is less congestion. Although start and soak emissions also peak during the peak hours, they remain significant throughout the day. Start emission is slightly higher in the morning than in the afternoon, since first
start of the day produces more emissions than starting again for subsequent trips. Soak emission is expected to follow the trend of start emission. This is difficult to observe at the current aggregation level, however. With the percentage plot, it can be seen that while a majority of the emission is still caused by vehicles cruising on the road, idling emission is very significant at peak hours and accounts for about 20% of morning peak emissions and 12% of the afternoon peak emissions.

Figure 7-4. Different Types of HC Emission by Hour

Figure 7-5. Percentage Share of Different Types of HC Emission by Hour
### 7.3 Emissions by Person and Household

A major advantage of using agent-based traffic simulation for emission calculation is the ability to link emissions to the people responsible for them. Emission by person was plotted against various person attributes. Figure 7-6 shows the amount of HC exhaust emission per person by age. Disregarding the unstable plot at the end due to lack of data, the graph suggests that people produce less emission as they get older. This is mainly due to people’s time of travel shifting away from peak hours as they age. (Figure 7-7) Decrease in average trip length may also contribute to the emission reduction. In contrast, the average daily start emission per person remains constant. (Figure 7-8)

![Figure 7-6. HC Exhaust Emission per Person by Age](image1)

![Figure 7-7. Percentage of Trips by Hour for Different Age Groups](image2)
The similarity in start and soak emissions per person between male and female drivers suggests that these two types of emissions are not strongly affected by time of travel and trip frequencies. The difference is much more pronounced in exhaust emissions. The average exhaust emission per person for male drivers is significantly higher. Analysis of the input schedules shows that the percentage of trips made by male drivers at peak hours is higher than that of female drivers. This increases the daily emission average by male drivers. Another possible factor is that male drivers are likely to drive further distances. There are more male drivers starting in early morning than female drivers, and those early morning work trips have been shown to be much longer in length.
In addition to age and gender, emission per person was also plotted against occupation type. People in different occupation groups all have similar start and soak emissions. Exhaust emission is considerably lower for people without work, and higher for people in professional, management, and technical positions. The amount of idling emission varies significantly, implying that people in different occupation groups travel at different time periods of the day. Overall, the result is the combined effect of multiple possible factors, including trip start time, work location, home location, trip purpose, etc. It is difficult to isolate the effect of each factor.

Figure 7-10. HC Emission per Person by Occupation Type

G: General Office/Clerical
M: Manufacturing/Construction/Trades
O: Not Employed
P: Professional/Management/Technical
S: Retail Sales and Service
The amount of emissions produced by each household was plotted in a histogram. (Figure 7-11) Measured by weighted average, each household each day produces 46.3 grams of HC, 986 grams of CO, 54.7 grams of NO\textsubscript{x}, and 18.7 kilograms of CO\textsubscript{2}. The number of households producing higher emissions follows an exponential decay. When there are five or fewer vehicles in the household, there is a strong correlation between emission and the number of vehicles. (Figure 7-12) The relationship is not linear and an additional vehicle provides a smaller increase in household emission. A correlation cannot be observed when the household owns a large number of vehicles, partially due to lack of data, and also likely due to underutilization of the vehicles.

Figure 7-11. Household HC Emission Histogram

Figure 7-12. HC Emission per Household by Number of Vehicles
Household location is also a contributing factor in the amount of emissions produced. People living in suburban areas are more likely to drive and drive further distances. ArcGIS software was used to link the GTA zone boundary file with household emission results. The data used for producing the graph were weighted by household expansion factors, and emission per household was plotted instead of zonal totals in order to eliminate bias due to different zone sizes. As shown in Figure 7-13, households located in zones further away from downtown produce more emission in general. Most of the households located within the highway ring produce less than 50 grams of HC a day. Newly developed residential areas such as Markham have substantially higher emission per household.

Figure 7-13. Weighted Average HC Emission per Household by Household Location
7.4 Link Emissions

Link emissions were also plotted in ArcGIS. The emission amount was divided by link length in order to remove the bias towards longer links, and produce smoother color transition along the link. The color scale was generated using data from a particular hour of the day (8am for exhaust and idling, 5am for start and soak) and then used for all other hours. The division of ranges were based on Jenk’s Natural Breaks suggested by ArcGIS. This method classifies data by minimizing the sum of squared differences between class members and class means. The ranges obtained were further modified through rounding and slightly decreasing the limit of higher ranges in order to obtain better colour distribution.

Figure 7-14 and Figure 7-15 respectively show the link-based CO exhaust and idling emission between 7am and 8am. Exhaust mission is much higher on the highways and arterials, and extends outwards towards regional centres. Idling emission was calculated independent of link length, and therefore not weighted by it. Comparing to exhaust emission, idling emission is more localized and appears on highways and also on roads closer to downtown.

Figure 7-14. CO Exhaust Emission
Figure 7-15. CO Idling Emission

The Appendix includes a complete set of hourly HC emission plot for each emission type, as well as a set of hourly snapshots from MATSim simulation. Comparing the two sets of plots shows that the peak in link exhaust emissions corresponds to the peak in daily travel. Emission in the afternoon peak is much higher than in the morning due to higher ambient temperature. Another factor is that there is less congestion but higher traffic flow due to more spread-out trip departures. Idling emission is high only during peak hours and decreases sharply for other hours of the day. This is consistent with the plot for network emission. (Figure 7-4) Start and soak emissions were also plotted on the links. Both types of emissions remain high throughout the day. The large number of work trips coming into Toronto can be observed on the plots: a large hot spot of soak emission can be observed in the downtown core in the morning. In the afternoon, a similar hot spot of start emissions originates from downtown.
Finally, whole day link exhaust emission plot and household emission plot were merged together in Figure 7-16. It is clear that there is high amount of emission occurring within the City of Toronto, but the heavy polluters live in suburban areas. This type of analysis is made possible by the agent-based traffic assignment from MATSim.

Figure 7-16. Household and Link Emission Overlay
Chapter 8. Future Work

8.1 Improving TASHA-MATSim Interactions

A large number of data conversion problems were encountered during integration, most of them were resolved by making trade-offs that sacrifice the advantages of microsimulation for model stability and compatibility. This is mostly due to the imbalance in model precision between TASHA and MATSim. TASHA still relies heavily on aggregated data. In fact, a large portion of the inputs is still generated by EMME/2. Possible improvements for the next version of TASHA include more sensitivity to travel time and a recalibrated mode choice model. Database management can be improved by merging similar datasets and possibly adopting a more robust file format.

On the traffic assignment side, MATSim was forced to uniformly reduce network capacity. A better representation of capacity is required to account for sample agents with different weightings. In the short-term, more experimentation with different flow and storage capacity combinations is required to better understand how the capacity-reduced network responds to congestion. Car allocation information is currently coded into the schedule by sharing the agent ID slot with person and household ID. Future expansion of the plan file should be able to handle this information as a separate attribute. This would also allow the Car Allocation Model, which is currently outside TASHA-MATSim, to be implemented within the MATSim environment. Other TASHA-MATSim configurations mentioned in Section 4.1 could also be tested. In the long-term, extending the traffic simulation to include other modes should significantly improve the current model, as TASHA would no longer be relying on aggregated outputs from EMME/2.

Unlike the fully automated process in TASHA-EMME/2, interactions between TASHA and MATSim involve tedious data cleaning and format conversions. A controller program should be implemented to streamline the process. This is very possible since most of the conversions are done by custom computer programs rather than Access queries.
8.2 Improving the Emission Model

Currently, the scope of the emission model focuses only on light-duty household vehicles, and as a result the model significantly underestimates total emission. Other types of vehicles such as trucks and buses all have significant contributions to pollution. This is especially the case for busy transportation corridors such as Highway 401 with a large number of commercial freight vehicles. When emission rates and activity diary for these types of vehicles become available, the model’s capability will no longer be limited to scenario analysis based on different household travel patterns. The model will become suited for analyzing environmental aspects of other, more general land-use and transportation development policies.

As emission rates become specific for different types of vehicles, it is necessary to improve the current Car Allocation Model. More research is required on people’s vehicle preferences so that emission modelling becomes sensitive to different car allocation strategies.

Improvement to the emission model also calls for additional information to be supplied by MATSim. The time of a vehicle reaching the end of a queue is an important piece of information that will lead to a better representation of congestion level. More research is required on extracting additional vehicle dynamics information from a queue-based traffic simulation.
Chapter 9. Conclusions

This research explores the potential of microsimulation models by experimenting with the integration of TASHA and MATSim and its application in vehicle emission modelling. The results indicate that TASHA-MATSim is capable of producing meaningful outputs similar to that of TASHA-EMME/2, without losing linkage to the agents. The results, however, may be too similar to TASHA-EMME/2 results, implying that the current version of TASHA is insensitive to zonal travel time variations.

On the emission modelling side, the emission calculation program is an improved average-speed model that adds congestion sensitivity into the calculations. Based on MATSim simulation events, the emission estimates are higher than those previously obtained through EMME/2. The amount of idling emissions, especially during peak periods, indicates that the model is indeed sensitive to congestion. The agent-based output from MATSim also allows emission analysis to be based on various person and household attributes. A non-linear relationship was observed between exhaust emission rate and person’s age. The results also confirmed that despite most of the emission occurs within the City of Toronto and other regional centres, heavy polluters live in suburban areas. This demonstrated that the new modelling framework is a promising tool that provides a better understanding of how people’s behaviour affects the system. This model should be implemented within a large-scale integrated microsimulation framework to fully utilize the linkage between agents and the system.
References


Hatzopoulou M. (2008) Ph.D. dissertation at University of Toronto


Appendices
Figure A-17. MATSim-T
Figure A-18. Complete Flowchart
Figure A-19. TASHA-EMME/2 Flowchart
Figure A-20. TASHA-MATSim Flowchart
Figure A-21 Emission Modelling Flowchart
<?xml version="1.0" encoding="utf-8"?>
<!DOCTYPE config SYSTEM "http://www.matsim.org/files/dtd/config_v1.dtd">
<config>
  <module name="global">
    <param name="randomSeed" value="4711" />
    <param name="outputTimeFormat" value="HH:mm:ss" />
    <param name="coordinateSystem" value="Atlantis" />
  </module>
  
  <module name="network">
    <param name="inputNetworkFile" value="C:/workspace/matsim/input/Runs/Run01/network.xml" />
  </module>
  
  <module name="plans">
    <param name="inputPlansFile" value="C:/workspace/matsim/input/Runs/Run01/plans.xml.gz" />
  </module>
  
  <module name="controller">
    <param name="outputDirectory" value="C:/workspace/matsim/output/Runs/Run01/" />
    <param name="firstIteration" value="0" />
    <param name="lastIteration" value="50" />
  </module>
  
  <module name="simulation">
    <!-- "startTime" of MobSim (00:00:00 == take earliest activity time/ run as long as active vehicles exist) -->
    <param name="startTime" value="00:00:00" />
    <param name="endTime" value="00:00:00" />
    <param name="flowCapacityFactor" value="0.05" />
    <param name="storageCapacityFactor" value="0.1" />
    <param name="snapshotperiod" value="00:30:00" />
    <param name="snapshotFormat" value="netvis" />
  </module>
  
  <module name="planCalcScore">
    <param name="learningRate" value="1.0" />
    <param name="BrainExpBeta" value="2.0" />
    <param name="lateArrival" value="-18" />
    <param name="earlyDeparture" value="-0" />
    <param name="performing" value="+6" />
    <param name="traveling" value="-6" />
    <param name="waiting" value="-0" />
    <param name="activityType_0" value="H" />
    <param name="activityPriority_0" value="1" />
    <param name="activityTypicalDuration_0" value="12:00:00" />
    <param name="activityMinimalDuration_0" value="00:00:00" />
    <param name="activityType_1" value="W" />
  </module>
</config>

Figure A-22. MATSim Simulation Configuration File
<param name="activityPriority_1" value="1" />
<param name="activityTypicalDuration_1" value="08:00:00" />
<param name="activityMinimalDuration_1" value="00:00:00" />
<param name="activityOpeningTime_1" value="07:00:00" />
<param name="activityLatestStartTime_1" value="09:00:00" />
<param name="activityEarliestEndTime_1" value="" />
<param name="activityClosingTime_1" value="18:00:00" />
<param name="activityType_2" value="A" />
<param name="activityPriority_2" value="1" />
<param name="activityTypicalDuration_2" value="12:00:00" />
<param name="activityMinimalDuration_2" value="00:00:00" />
<param name="activityOpeningTime_2" value="" />
<param name="activityLatestStartTime_2" value="" />
<param name="activityEarliestEndTime_2" value="" />
<param name="activityClosingTime_2" value="" />
<param name="activityType_3" value="B" />
<param name="activityPriority_3" value="1" />
<param name="activityTypicalDuration_3" value="12:00:00" />
<param name="activityMinimalDuration_3" value="00:00:00" />
<param name="activityOpeningTime_3" value="" />
<param name="activityLatestStartTime_3" value="" />
<param name="activityEarliestEndTime_3" value="" />
<param name="activityClosingTime_3" value="" />
<param name="activityType_4" value="I" />
<param name="activityPriority_4" value="1" />
<param name="activityTypicalDuration_4" value="12:00:00" />
<param name="activityMinimalDuration_4" value="00:00:00" />
<param name="activityOpeningTime_4" value="" />
<param name="activityLatestStartTime_4" value="" />
<param name="activityEarliestEndTime_4" value="" />
<param name="activityClosingTime_4" value="" />
<param name="activityType_5" value="J" />
<param name="activityPriority_5" value="1" />
<param name="activityTypicalDuration_5" value="12:00:00" />
<param name="activityMinimalDuration_5" value="00:00:00" />
<param name="activityOpeningTime_5" value="" />
<param name="activityLatestStartTime_5" value="" />
<param name="activityEarliestEndTime_5" value="" />
<param name="activityClosingTime_5" value="" />
<param name="activityType_6" value="L" />
<param name="activityPriority_6" value="1" />
<param name="activityTypicalDuration_6" value="12:00:00" />
<param name="activityMinimalDuration_6" value="00:00:00" />
<param name="activityOpeningTime_6" value="" />
<param name="activityLatestStartTime_6" value="" />
<param name="activityEarliestEndTime_6" value="" />
<param name="activityClosingTime_6" value="" />
<param name="activityType_7" value="M" />
<param name="activityPriority_7" value="1" />
<param name="activityTypicalDuration_7" value="12:00:00" />
<param name="activityMinimalDuration_7" value="00:00:00" />
<param name="activityOpeningTime_7" value="" />
<param name="activityLatestStartTime_7" value="" />
<param name="activityEarliestEndTime_7" value="" />
<param name="activityClosingTime_7" value="" />
<param name="activityType_8" value="O" />
<param name="activityPriority_8" value="1" />
<param name="activityTypicalDuration_8" value="12:00:00" />
<param name="activityMinimalDuration_8" value="00:00:00" />
<param name="activityOpeningTime_8" value="" />
<param name="activityLatestStartTime_8" value="" />
<param name="activityEarliestEndTime_8" value="" />
<param name="activityClosingTime_8" value="" />
<param name="activityType_9" value="R" />
<param name="activityPriority_9" value="1" />
<param name="activityTypicalDuration_9" value="12:00:00" />
<param name="activityMinimalDuration_9" value="00:00:00" />
<param name="activityOpeningTime_9" value="" />
<param name="activityLatestStartTime_9" value="" />
<param name="activityEarliestEndTime_9" value="" />
<param name="activityClosingTime_9" value="" />
<param name="activityType_10" value="S" />
<param name="activityPriority_10" value="1" />
<param name="activityTypicalDuration_10" value="12:00:00" />
<param name="activityMinimalDuration_10" value="00:00:00" />
</module>

<module name="strategy">
<param name="maxAgentPlanMemorySize" value="5" />
<param name="ModuleProbability_1" value="0.8" />
<param name="Module_1" value="ChangeExpBeta" />
<param name="ModuleProbability_2" value="0.2" />
<param name="Module_2" value="ReRoute" />
</module>
</config>

MATSim Simulation Configuration File Continued
Figure A-23. MATSim Simulation Snapshots (4:00 – 7:00)
MATSim Simulation Snapshots (8:00 – 11:00)
MATSIm Simulation Snapshots (0:00 – 3:00)
Figure A-24. Hourly HC Exhaust Emission (Page 1 of 6)
Figure A-25. Hourly HC Idling Emission (page 1 of 4)
Hourly HC Idling Emission (page 4 of 4) (Insignificant Emission after 7pm)
Figure A-26. Hourly HC Start Emission (page 1 of 6)
Hourly HC Start Emission (page 3 of 6)
Hourly HC Start Emission (page 6 of 6)
Figure A-27. Hourly HC Soak Emission (page 1 of 6)
Hourly HC Soak Emission (page 3 of 6)
Hourly HC Soak Emission (page 6 of 6)