AGENT-BASED TRANSPORT SIMULATION VERSUS EQUILIBRIUM ASSIGNMENT FOR PRIVATE VEHICLE TRAFFIC IN GAUTENG

P J FOURIE

Planning Support Systems, CSIR Built Environment, P O Box 395, Pretoria, 0001

ABSTRACT

This paper compares the performance of an agent-based microsimulation technology (MATSim) with that of an established equilibrium assignment model (EMME/2) for the private vehicle traffic demand description from the Gauteng Transportation Study 2000, commissioned by the Gauteng Department of Roads and Public Works. MATSim and EMME/2 produce similar volumes on high capacity links, but closer investigation reveals that MATSim agents also favour secondary routes to avoid congestion. MATSim produced a dramatically better distribution of travel times compared with reported values from the 2003 National Household Travel Survey.

INTRODUCTION

The CSIR Urban Dynamics Laboratory investigates the viability of next-generation modelling technologies for urban planning decision support. We are currently modelling urban form evolution in UrbanSim (Waddell, 2002), a suite of open source Python modules aimed at capturing the effects on overall urban development resulting from the interactions between transport and land-use.

Accurate transport modelling forms an essential complement for UrbanSim to function. Basically, UrbanSim (re-)locates people and jobs and develops parcels of land on an annual basis in reaction to transport system performance, as predicted by a transport demand model. These relocations are translated into a new transport demand, which is fed back into the transport model. The newly predicted transport system performance is then taken as input to predict the urban development response for the following year. This process is repeated for the required prediction horizon.

The question this paper aims to answer is which transport demand model to use in conjunction with UrbanSim. We limited our selection to MATSim and EMME/2, as interfaces to UrbanSim have been developed for these two model systems. MATSim is an agent-based transport simulation system, using queue simulation to perform dynamic traffic assignment, while EMME/2 performs a static assignment using a numerical optimisation process (see following section for background on both systems).

The key performance indicators that influence our choice of transport model system are listed below:

Speed of execution: Currently, equilibrium assignment is favoured in many UrbanSim applications because of the speed of performing an equilibrium-based assignment. Even though the last number of years have seen great improvements in execution speed for simulation assignment, performing large-scale simulations still typically require at least an order of magnitude more time to run than is the case for EMME/2.

Proceedings of the 29th Southern African Transport Conference (SATC 2010) 16 - 19 August 2010
Proceedings ISBN Number: 978-1-920017-47-7 Pretoria, South Africa
Produced by: Document Transformation Technologies cc Conference organised by: Conference Planners

Level of disaggregation: An agent-based approach could be used to give individualised impressions of transport system performance. Great efforts are made to produce accurate representative populations of study areas in UrbanSim, and all that resolution is lost when this individualised demand is aggregated into trip flows for use in equilibrium assignment models.

Monetary cost: An EMME/2 license is prohibitively expensive, and limits the number of concurrent users that can work on an UrbanSim project, whereas MATSim is free and open-source.

Predictive power: Finally, and most importantly, we seek accurate predictions of transport system performance in terms of link volumes, travel times and other measures of generalised transport cost. Generally, link volumes are of more interest to the transport planner than system users (firms, developers, households), so we place a relatively higher premium on the other measures for the purposes of UrbanSim, which models the system user response to generalised transport cost. Currently we take peak hour private vehicle travel time to be our best indicator of generalised cost, but this picture will have to change in future to take account of all relevant measures (e.g. distance-based tolling, vehicle operation cost, public transport accessibility) to produce generalised costs across all modes.

Earlier investigation established the viability of using the MATSim platform in the South African metropolitan context (Fourie and Joubert, 2009). But we need to be sure that its capabilities at least match those of equilibrium-based assignments in predicting essential transport system performance indicators such as link volumes and travel times.

Previous work by Gao et al. (2010) showed how MATSim's performance compares with EMME/2 in the Greater Toronto and Hamilton Area Network. They found that speed and link volume outputs produced by MATSim are comparable to those from EMME/2 while MATSim produced more realistic travel times. In this paper, MATSim's performance is compared with that of EMME/2 for Gauteng to see if their findings extend to the South African metropolitan context. To this end, their methods were replicated as far as possible, and the use of their code in the demand conversion and post-simulation analyses is acknowledged.

This study uses travel demand and network data from the Gauteng Transport Study for 2000 (GTS2000) EMME/2 model, commissioned by the Gauteng Department of Transport, Roads and Public Works. The EMME/2 demand description for private vehicles departing during the morning peak hour is transformed into a set of MATSim agent plans. These are simulated on a network graph derived from the EMME/2 model's network description and volume delay functions. Predicted volumes for both models are then compared with actual counts for 20 pairs of counting stations on strategic network links recorded during 2001. Predicted travel times are compared with those reported in the 2003 National Household Travel Survey (NHTS).

BACKGROUND

This section gives a short description of the two transportation forecasting models used in this study. The interested reader is referred to Gao et al. (2010) for an extensive comparison of MATSim and EMME/2, as well as the methodological details of synthesising a MATSim population from an EMME/2 demand description.

EMME/2

EMME/2 is a traditional four step transportation forecasting model, operating on the principles of trip production and attraction, trip distribution, modal split and route assignment. This paper is mostly concerned with the performance of the last step of the procedure, route assignment, where trip flows between origin-destination pairs are routed through the network.

EMME/2 solves this traffic assignment problem (TAP) using the Franke-Wolfe algorithm to achieve user equilibrium after a user specified number of iterations or measures of convergence. User equilibrium is achieved when no user in the network can decrease travel time or generalised travel cost by shifting to a new path (e.g. Florian, 1976, Ortuzar & Willumsen, 1994, Wikipedia contributors, 2009).

MATSim

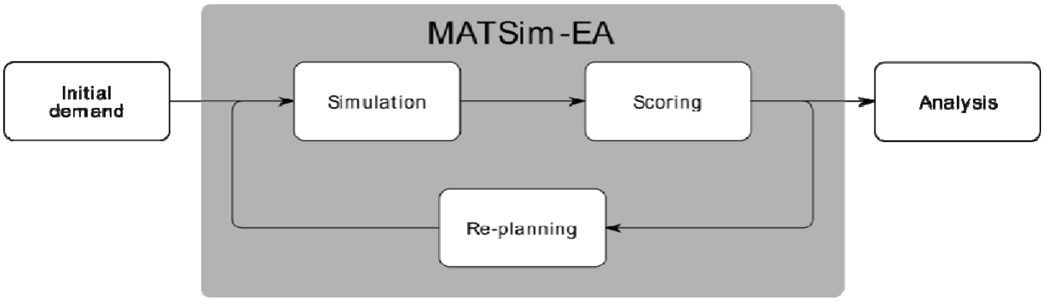


Figure 1 Schematic illustration of MATSim's transport demand modeling process. The shaded area indicates the evolutionary "engine" of MATSim, which simulates system learning and adaptation. (Source: Rieser, 2009).

MATSim is an agent-based transport demand modelling framework that operates on the basis of individual agent plans; a plan being a schedule of activities, their locations in the study area and the travel legs connecting them.

Figure 1 illustrates the basic principle of operation. An initial demand of a full day plan of activities for each agent is generated and executed in a mobility simulation. Plans are scored after the simulation step and, based on the score, agents adapt their plans in response to conditions that arose during the simulation.

A MATSim simulation converges to a state analogous to the user equilibrium through a process of systematic relaxation (see, for instance, Balmer, 2007). Such convergence is achieved through adapting and deriving a set of feasible plans for each agent from their original initial plan. As these sets of plans grow to a limiting number, bad performing plans are discarded. Consequently, each agent's set of feasible plans improves with increasing iterations. Feasible new plans can be derived from existing ones by changing activity timings, locations, re-routing travel legs between activities, changing transport modes connecting activities or dropping activities from the activity schedule altogether. Following the method of Gao et al. (2010), the only re-planning strategy allowed in this study was re-routing, as the EMME/2 demand description implicitly assumes agents to depart simultaneously at the start of the morning peak hour.

The current implementation of MATSim favours a queue simulation, which models network links as first-in-first-out queues (Rieser, 2009). Vehicles entering each link cannot leave that link for the next before a certain time has passed, equal to the free flow travel time for the link. Only a limited number of vehicles can leave the link in each time step, determined by the flow capacity of the link. A qualifying vehicle can then also enter the following link if that link has available storage capacity; a maximum queue length determined by the physical length of the link divided by the average interspacing of vehicle centres during congestion (usually taken to be 7.5m).

The interested reader is referred to Balmer et al. (2006) for a detailed description of the MATSim framework, and to Cetin (2005) for a detailed description of the queue simulation algorithm.

METHOD

For purposes of comparison, the EMME/2 network description and peak hour private vehicle travel demand from the GTS2000 model was transformed into a MATSim network and agent plans file.

Network conversion

An EMME/2 network description takes the form of directed links between nodes. The cost of traversing each link is determined by a volume delay function (VDF), which gives travel time as a function of link capacity fraction. The most widely used VDFs, such as those used in the GTS2000 model, take the form prescribed by the US Bureau for Public Roads (1964), and can be stated as:

$$t(v) = t_0\left(1 + \left(\frac{v}{c}\right)^{\alpha}\right)$$

where t_0 is the free-flow travel time across the link, v is the traffic volume on the link and c its capacity. The scaling parameter α determines the rate of change in travel time with increasing volume fraction.

A MATSim network description is also a directed graph of links connecting nodes, but rather than using a VDF, the time-dependent cost of traversing links are determined by the dynamic simulation of network congestion. Link parameters serve as constraints during the queue simulation, as described in the previous section. Each link in the MATSim network requires the following parameters: link length, link free speed, number of lanes and total link flow capacity. Link length and free speed determine free-flow travel time (minimum time for an agent to traverse a link), link length and number of lanes determine queue storage capacity. Deriving these parameters from the GTS2000 EMME/2 link VDFs is an obvious and trivial task.

Demand conversion

The morning peak hour private vehicle demand from the provided GTS2000 EMME/2 model was taken as input for both models. This demand amounts to 605,329 vehicle trips between 899 traffic zones. For purposes of comparison the EMME/2 demand description had to be converted into an agent population whose activity schedules translate into a peak hour transport demand that approximates the EMME/2 demand matrix. This process proceeded as follows:

Population synthesis

The number of trips originating in each zone was determined by rounding off the row total for each zone to the nearest integer (EMME/2 allows for fractional trip flows between zones). This number was taken to be the number of private vehicle drivers living within that zone, assuming that all morning peak hour trips are home-based. We do not consider individual demographics in this comparison and all agents are assumed to have identical activity patterns (leave home at 06h00 for work by car and return nine hours later).

The EMME/2 assignment procedure assumes that all trips originate and terminate at zone centroids. MATSim, on the other hand, performs an activity-oriented, agent-based simulation assignment, and requires a synthetic population of agents with day activity schedules to execute on the network. Activity locations are points in Euclidean space that are associated with links in the network, and not just with zone centroids. This is an important difference in comparison with equilibrium transport assignment: besides allowing for a better analogy to physical activity locations and their association with actual transport network links, it is also necessary to distribute activity locations in space to satisfy network capacity constraints. If all agents in a traffic zone had to depart from exactly the same location (zone centroid), it would cause an unnatural wave of congestion to propagate from each centroid.

We generated the synthetic population living within each zone and randomly assign their home activity locations within a circle around the zone centroid, as can be seen from Figure 2. This circle has a radius of 70% the distance to the nearest neighbour zone centroid. Each home location is then associated with the network link nearest to its coordinate.

Work location assignment

Each agent is assigned a work location zone based on their home location zone, using the EMME/2 origin-destination demand matrix to approximate conditional probabilities. The probability that an agent from a particular zone will work in any other zone is taken as the trip flow to the destination zone divided by the total trip flow from the origin zone. Agents' work locations are scattered around corresponding centroids in the same way as for home locations.

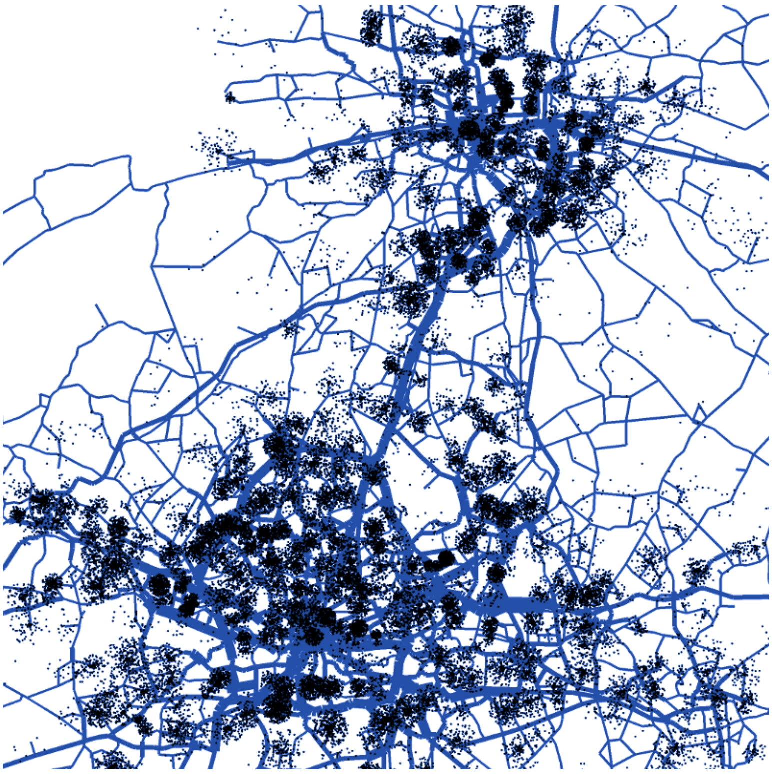


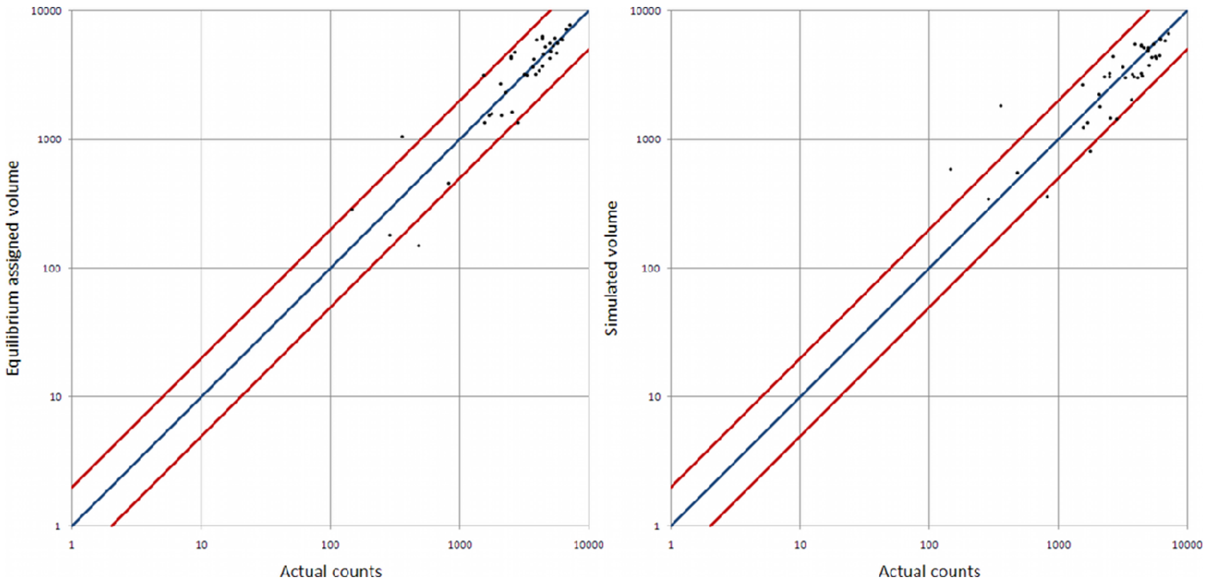
Figure 2 Agent home locations on a section of the GPS2000 network, showing the centres of Pretoria and Johannesburg. Agent home locations are scattered in a circle around each zone centroid.

Plans file generation

Once each agent has been assigned a work location, their activity schedules are compiled. It is assumed that all agents depart from home at six in the morning, as count station data reveal 06h00 – 07h00 to be the peak hour for private vehicle transport demand. Each agent is assumed to spend nine hours at work before returning home. The return trip home is taken as further validation of our transport demand description; if agents produce realistic traffic volumes on the return trip home, our confidence in the demand description increases.

RESULTS

Counts comparisons



(a) **(b)**

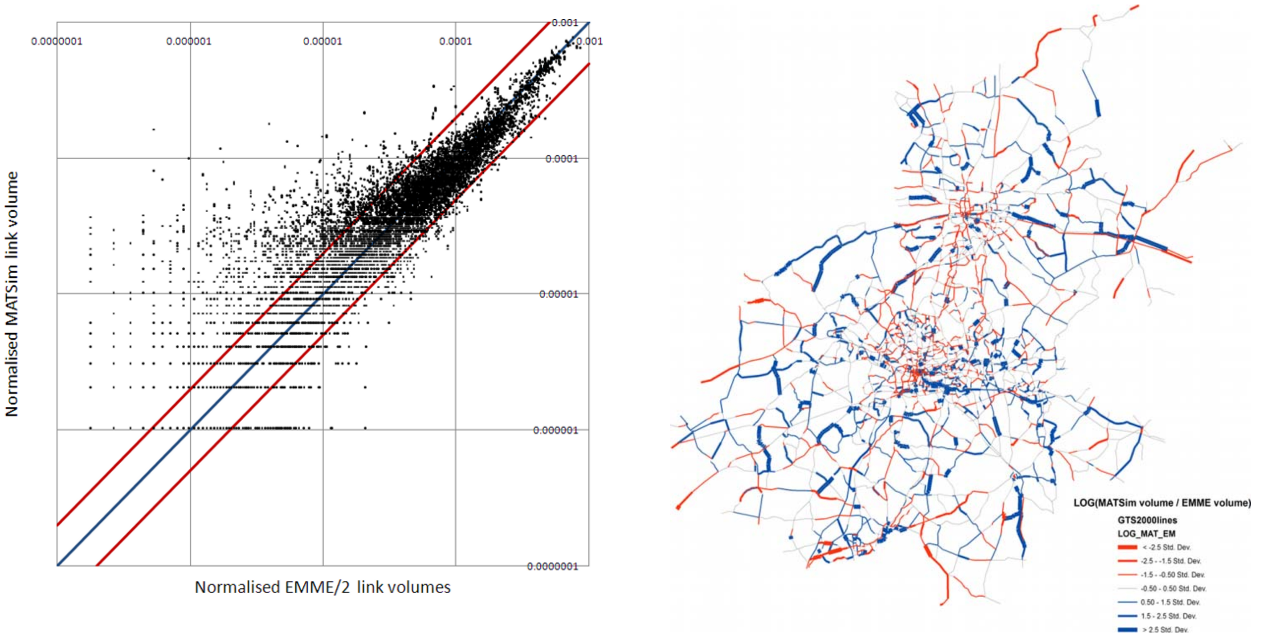
Figure 3 Comparison of predicted volumes for EMME/2 (a) and MATSim (b) versus traffic counts from 2001 on a selection of major links in Gauteng networkduring the morning peak hour. Red lines indicate 2:1 and 1:2 ratios of volume:count.

After running the MATSim simulation for a hundred iterations, and performing the EMME/2 assignment until reaching equilibrium after 19 iterations, the predicted volumes were compared with actual count station values on strategic network links for 06h00 – 07h00 on a typical Wednesday in 2001, as shown in Figure 3. In general, MATSim and EMME/2 produced similar volumes on this selection of links, with the mean relative error of the EMME/2 assignment being 10.6% over actual recorded values for 2001, while MATSim is slightly higher at 12.9%.

In absolute terms, EMME/2 volumes are on average 202 vehicles/hr more than recorded counts for the selection of 20 pairs of counting stations. MATSim volumes are 196 vehicles/hr less than recorded counts. The root mean squared error (RMSE, an indication of error spread) for the equilibrium assignment is 905 vehicles/hr, compared to 934 vehicles/hr for the simulation assignment.

Figure 4 shows a comparison of link volumes produced by MATSim and EMME/2 for all links on the network. In the case of MATSim volumes, only those produced between 06h00 and 07h00 (the AM peak hour) were considered. For purposes of comparison, all link volumes were normalised against the total AM peak hour traffic volume. This step was necessary because the total traffic volume predicted by EMME/2 was 11,215,170 vehicles, compared to 9,730,560 for MATSim, as congestion in the MATSim simulation caused travel times in excess of 60 minutes.

Figure 4 shows that MATSim and EMME/2 apportion comparable volumes to high-capacity links, but that agreement between the two datasets diminishes with decreasing link capacity. MATSim agents clearly seek out available capacity on secondary routes in order to avoid congestion.



(a) **(b)**

Figure 4 Scatter-plot comparison of peak-hour volumes on all links, normalised with respect to the total peak-hour link volume across the entire network for each simulation (a), and a network map comparing the logarithm of volume ratios (b): line thickness indicates magnitude, red indicates normalised EMME volume > MATSim and blue vice-versa. MATSim agents generally produce lower volumes than EMME/2 on high-capacity links, and its agents seek out available capacity on secondary routes to avoid congestion.

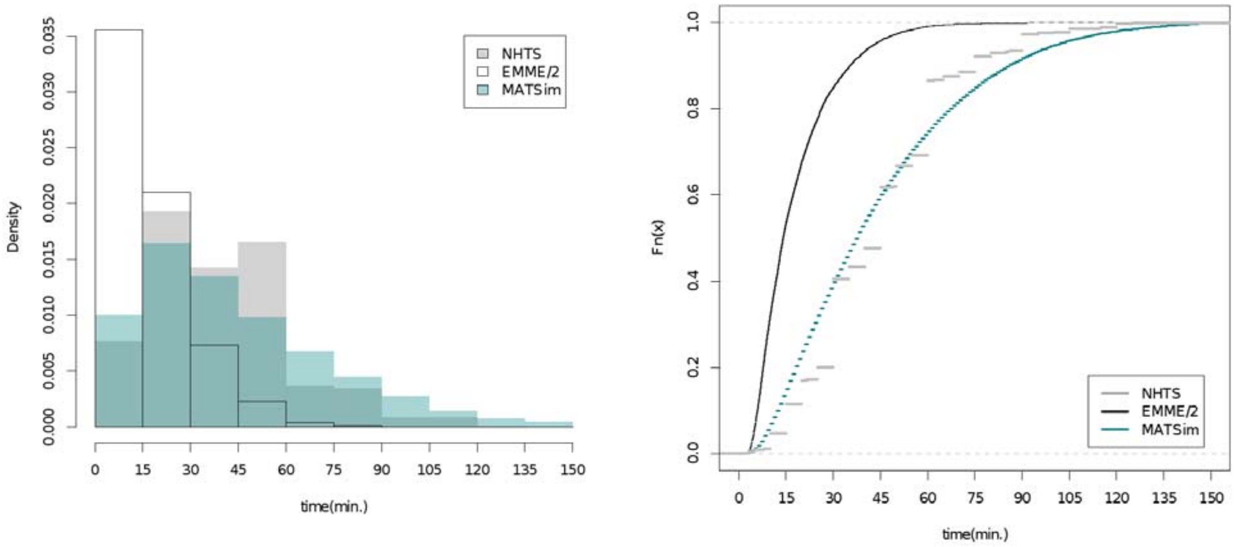
<u>Travel times</u>

Travel times produced by both models was compared with reported values from the 2003 National Household Travel Survey (NHTS) as shown in Figure 5. The NHTS travel time distribution was reconstructed by isolating the private vehicle driver person records with reported departure times between 06h00 and 07h00 and weighting the frequency of their reported travel times. This sample of 1,762 records' weights add up to a total of 562,808; a number which is of a similar order of magnitude than the trip volume predicted in the EMME/2 demand matrix.

As we deal with reported travel times, the peaks around 30 and 60 minutes are to be expected, as people are bound to round up or down to these salient quantities. The average weighted travel time reported during the NHTS comes to 44.98 minutes, with a standard deviation of 25.96 minutes. In comparison, EMME/2 predicts an average travel time of only 19.23 minutes, standard deviation 12.47 minutes, while MATSim agents report an average travel time of 45.46 minutes, standard deviation 32.12 minutes for the morning peak.

A two-sample Kolmogorov-Smirnov test (R Development Core Team 2008) with NHTS data give test statistics of D_n = 0.6474 and D_n = 0.1784 for comparison with EMME/2 and MATSim respectively. The Kolmogorov-Smirnov test statistic is a minimum distance estimation between empirical cumulative density functions, so a value of D_n = 0.0 means perfect agreement.

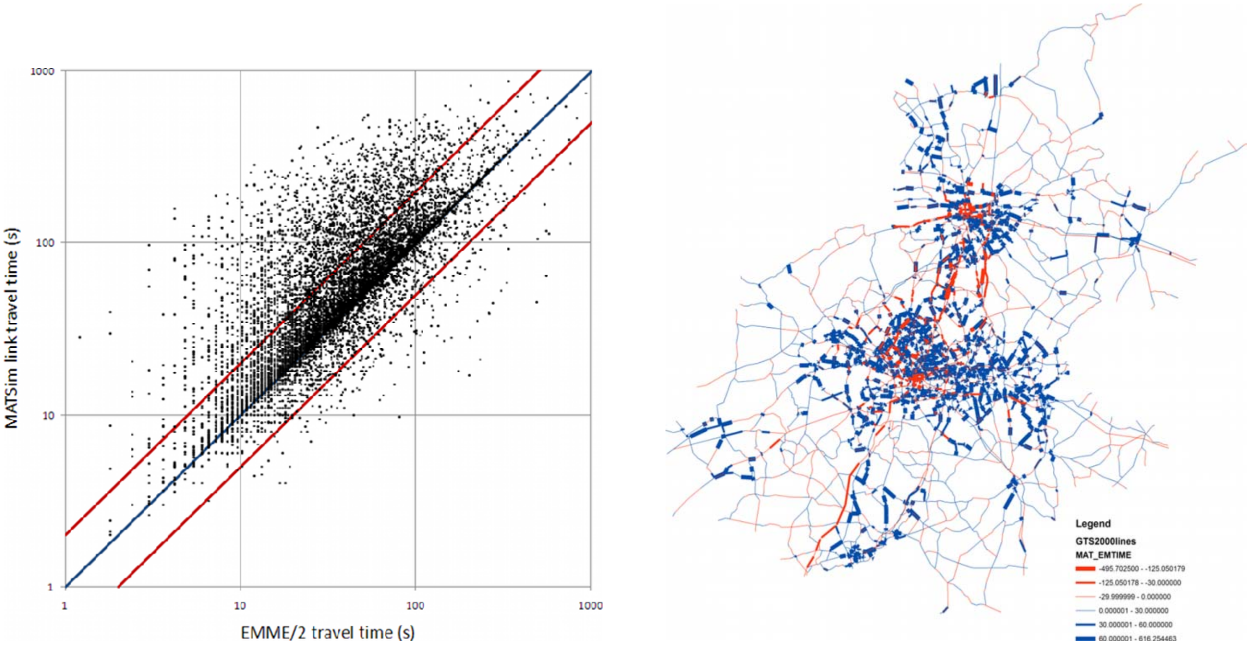
Figure 6 confirms that MATSim produces longer average travel times on a far greater proportion of network links, while Figure 7 suggests why this is the case. This snapshot of Johannesburg CBD shows how small pockets of congestion sporadically occur in the network purely due to the dynamics of agents moving through queues or waiting to enter queues. These delays eventually add up, as they do in reality, to give longer travel times than what the equilibrium model predicts.



(a)

(b
)

Figure 5 Probability density histograms (a) and empirical cumulative distribution functions (b) of NHTS, EMME/2 and MATSim travel times.



(a)

(b)

Figure 6 Comparison of MATSim and EMME/2 link travel times on all links. MATSim travel times are the average travel time recorded for each link during the peak hour. (a) shows a scatter plot comparison, while (b) shows the difference between MATSim and EMME link travel times, with line thickness denoting magnitude, blue denoting MATSim > EMME and vice-versa for red.

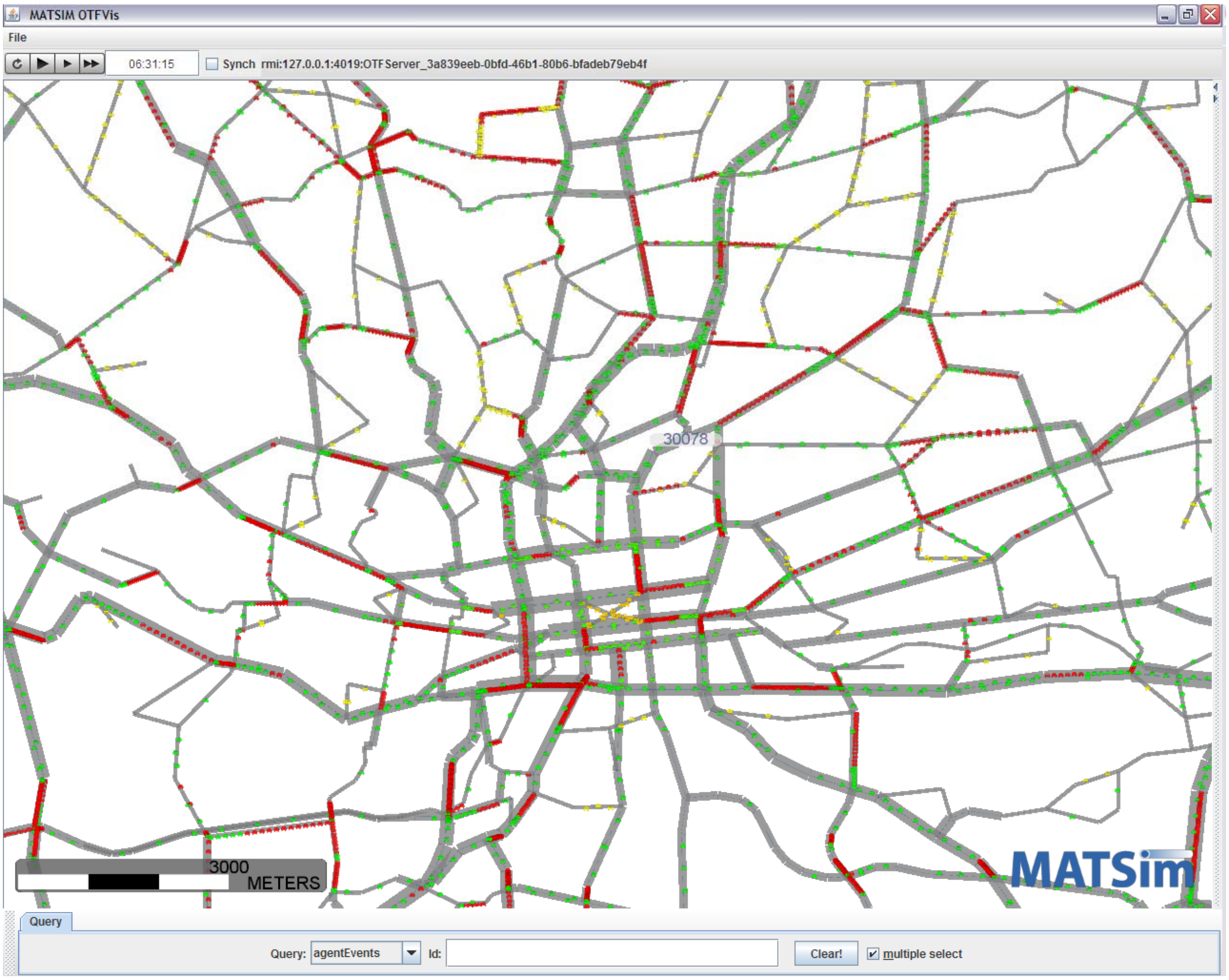


Figure 7 Simulation snapshot of Johannesburg CBD halfway through the morning peak hour. Yellow and green car icons denote agent vehicles moving at link free speed, while red icons are vehicles on congested links.

CONCLUSION

As we only execute the private vehicle demand for the peak hour, ignoring spill-over from earlier departures, public transport and freight traffic, faster travel times than those found in reality are to be expected. But our conclusion is that MATSim produces more realistic travel times than the EMME/2 model, as MATSim takes account of the time-dependent effect of congestion in the network, whereas an EMME/2 assignment assumes, in a sense, that traffic flow occurs on all links simultaneously.

The effects of congestion also have dramatic effects on network utilisation, with MATSim agents seeking out available capacity on secondary routes to produce much higher volumes than what EMME/2 predicts. In fact, an analysis of volumes on all network links show that MATSim and EMME/2 differ radically with respect to their picture of network utilisation outside of the set of high capacity links.

The improved estimation of travel time produced by MATSim carries computation time cost, as executing 100 iterations of a 10% sample requires 2 hours and 13 minutes on a 2.4 GHz Intel Core 2 Duo processor, compared to the less than five minutes required to perform the required 19 iterations for EMME/2 to achieve convergence on the same platform. However, a single commercial EMME/2 seat costs in excess of at least R100,000, whereas MATSim is an open source free software package. Several developments are also underway to improve MATSim execution speed. Tweaking an EMME/2 model to produce correct travel times takes a considerable amount of expert 'calibration', whereas MATSim produces realistic travel times with no intervention.

The computational cost of microsimulation therefore needs to be weighed against the need for more accurate travel time based on the requirements of a particular application. For instance, in the case of our application of UrbanSim, inter-zonal travel time is the single most important indicator of accessibility that feeds back into the way people and jobs organise spatially over time. In such an application, MATSim should at least be considered during the latter stages of scenario development to provide a 'reality-check' for the influence of more realistic travel times on overall urban evolution.

In an energy or emissions analysis application, MATSim should be the clear application of choice, as accurate travel times and the influence of 'cold starts' are important factors affecting energy consumption and emissions. As MATSim maintains a full record of each individual agent's experience in the road network, it is possible to accurately calculate the emissions and fuel consumption of a heterogeneous fleet of vehicles, assigned to agents based on their individual characteristics. Emissions results can then be analysed against those characteristics, so for instance a profile of emissions per capita vs. household income can be generated. As we move into an era of increasing carbon footprint awareness, it can be expected that the role of microsimulation in accurate accounting and prediction will become more prominent, hopefully affording better funding opportunities to develop a better understanding of transport behaviour.

REFERENCES

Balmer, M. (2007). *Travel Demand Modeling for Multi-Agent Transport Simulations: Algorithms and Systems.* PhD thesis, ETH Zürich

Balmer, M., Axhausen, K. and Nagel, K. (2006). *Agent-Based Demand-Modeling Framework for Large-Scale Microsimulations.* Transportation Research Record: 125–134.

Bureau of Public Roads (1964). *Traffic Assignment Manual.* U.S. Dept. of Commerce, Urban Planning Division, Washington D.C.

Cetin, N. (2005) *Large-Scale parallel graph-based simulations.* PhD thesis, Swiss Federal Institute of Technology (ETH) Zürich, Switzerland.

Fourie, P. J., and Joubert, J. W. (2009). The first agent steps in agent-based transport planning. In *28th Annual Southern African Transport Conference, South Africa* (pp. 6-9).

Florian, M. A. (1976). *Traffic equilibrium methods*. Springer.

Gao, W., Balmer, M. and Miller, E.J. (2010). *Comparisons between MATSim and EMME/2 on the Greater Toronto and Hamilton Area Network*. Paper 10-1274, Proceedings of TRB 89th Annual Meeting (January 10-14, 2010)

Ortuzar, J. D., and Willumsen, L. G. (1994). *Modelling transport*. Wiley New York.

R Development Core Team (2008). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rieser, M. (2009). Adding Transit to an Agent-Based Transportation Simulation: Concepts and Implementation. Draft of PhD thesis on November 8, 2009, Technical University of Berlin, Germany.

Waddell, P. (2002). *UrbanSim: Modeling urban development for land use, transportation, and environmental planning.* Journal: American Planning Association, 68(3), 297-314.

Wikipedia contributors (2009). *Route assignment.* Wikipedia, The Free Encyclopedia. Wikipedia, accessed 17 Feb. 2010. Retrieved from http://en.wikipedia.org/w/index.php?title=Route_assignment&oldid=334904095