The evaluation of traffic microsimulation modelling

D. O’Cinnéide & B. O’Mahony
Traffic Research Unit, University College Cork, Ireland

Abstract

In recent years, traffic simulation software which models each vehicle on the network separately has been developed. These models are becoming more accessible due to the increases in computing power. However, the acceptance of microsimulation by traffic modellers has been far from enthusiastic. While modellers agree that conventional macroscopic software cannot represent dynamic phenomenon or many ITS applications, they are sceptical regarding the accuracy of microsimulation in modelling even conventional traffic systems. One reason is that there has been surprisingly little independent, objective evaluation of microsimulation models.

In this paper, microsimulation modelling is evaluated based on indicators for accuracy, calibration requirements and model construction time by using the Paramics program to model various small traffic networks, including the Leeds Network developed by the European Union Smartest Project which has previously been used to evaluate other microscopic models from different European countries. A number of microscopic/macroscopic comparisons are also undertaken using the Transyt and Arcady programs.

The main conclusions are that microsimulation is at least as accurate as macroscopic modelling and the only logical method of modelling dynamic effects on a traffic network. Microsimulation was found to require additional model construction time and input data for calibration. The importance of providing a list of assumptions used and changes to default parameters made with every traffic model is highlighted as is the importance of using multiple model outputs to aid model calibration and validation.

Keywords: microsimulation, Paramics, evaluation, traffic modelling.
1 Introduction

This paper describes an independent, objective evaluation of microsimulation modelling accuracy, calibration and model construction time requirements. It also considers the benefits and disbenefits of using alternative modelling software over microsimulation modelling. The Paramics program is used to evaluate microsimulation modelling under the headings of model result accuracy, calibration requirements and model construction time requirements.

2 The evaluation tool

The Paramics program, developed in the UK and widely used in the UK, the US and Australia, is used to evaluate microsimulation modelling in general. Paramics software is portable and scalable, allowing a unified approach to traffic modelling across the whole spectrum of network sizes, from single junctions up to national networks. The real time visual display of microsimulation, and Paramics in particular, are easily understood by non-experts, making the program a useful aid in public consultation.

3 Evaluation procedure

The evaluation methodology is based on that in the literature. There is a close similarity with the study networks and the evaluation methodology used by the Smartest Project [1], Choa [2], Abdulhai et al. [3] and Transport for London [4]. The networks used in these evaluations were ‘line networks’, with no route choice between any origin and destination, and with two major junctions at either end of a stretch of road. The elimination of route choice allows the more fundamental aspects of the models to be evaluated. The method of evaluation is summarised as follows:

1. Use a ‘simple’ network. One without unusual traffic phenomenon.
2. Construct the model while making comments on assumptions made, difficulties encountered in modelling, and an approximate model construction time.
3. Extract standard / simple outputs from the model where possible.
4. Compare modelling results with real data.
5. Undertake a sensitivity analysis.
6. Draw conclusions based on the similarity or dissimilarity between results and real data.

This modelling methodology was applied to a number of small traffic networks, two are described in this paper.

4 Modelling undertaken

4.1 The Leeds Network: accuracy and calibration requirement evaluation

The Leeds Network consists of a 1.4km stretch of a main arterial road called the ‘Dewsbury Road’. This is one of the main radial routes into Leeds, carrying
approximately 23,000 vehicles per day. It is also a heavily used public transport corridor, peak bus flows being in excess of 36 buses per hour.

The Leeds Network was used during the Smartest Project Transferability Test [1] to determine whether three microsimulation models developed in various locations in Europe could accurately model a UK traffic network. The network was chosen for modelling because of available traffic count, number plate OD surveys, queue length and travel time data as well as modelling results from the Dracula, Aimsun2 and Nemis microsimulation programs. The aim of the current study was to model the same network using Paramics and to compare the results with measured quantities such as queue lengths and travel times.

The Leeds Model was constructed in Paramics using three levels of input data to produce Low, Medium and High Level Models. Table 1 shows the components of input data for the three models. The aim of using these three levels of data was to determine how necessary each component was for accurate model results.

After model construction and visual interrogation of the simulation the model was then run for an AM peak hour with queue lengths and travel times recorded for the Low, Medium and High Level Models. In each case the simulation was run four times and an average of the outputs recorded.

Table 1: Input data levels for the three Leeds Paramics Models.

<table>
<thead>
<tr>
<th>Low level model</th>
<th>Medium level model</th>
<th>High level model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Road geometry</td>
<td>Road geometry</td>
<td>Road geometry</td>
</tr>
<tr>
<td>Traffic control</td>
<td>Traffic control</td>
<td>Traffic control</td>
</tr>
<tr>
<td>Vehicle O/D matrix</td>
<td>Separate car &amp; HGV Matrix</td>
<td>Separate car &amp; HGV matrix</td>
</tr>
<tr>
<td></td>
<td>Public Transport</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Measured and modelled queue lengths.

<table>
<thead>
<tr>
<th>Queue ID</th>
<th>Low (veh)</th>
<th>Medium (veh)</th>
<th>High (veh)</th>
<th>Measured (veh)</th>
<th>SD</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18.0</td>
<td>15.75</td>
<td>24.75</td>
<td>25.6</td>
<td>8.75</td>
<td>22.5, 28.7</td>
</tr>
<tr>
<td>2</td>
<td>21.47</td>
<td>17.4</td>
<td>20.95</td>
<td>9.3</td>
<td>5.17</td>
<td>11.2, 14.1</td>
</tr>
<tr>
<td>3</td>
<td>5.32</td>
<td>5.27</td>
<td>7.63</td>
<td>12.6</td>
<td>3.04</td>
<td>8.2, 10.5</td>
</tr>
<tr>
<td>4</td>
<td>5.02</td>
<td>4.95</td>
<td>4.95</td>
<td>5.2</td>
<td>2.87</td>
<td>4.4, 6.0</td>
</tr>
<tr>
<td>5</td>
<td>2.6</td>
<td>2.23</td>
<td>2.775</td>
<td>3.0</td>
<td>1.92</td>
<td>2.5, 3.5</td>
</tr>
</tbody>
</table>

Table 3: Measured and modelled travel times.

<table>
<thead>
<tr>
<th>Route ID</th>
<th>Low (s)</th>
<th>Medium (s)</th>
<th>High (s)</th>
<th>Measured (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5</td>
<td>148.75</td>
<td>142.75</td>
<td>174.5</td>
<td>164.7</td>
</tr>
<tr>
<td>6-10</td>
<td>109.25</td>
<td>109.25</td>
<td>132.75</td>
<td>159.6</td>
</tr>
</tbody>
</table>

Table 2 shows the modelled queue lengths for the three models. The measured queue lengths with standard deviations and 95% confidence intervals
are also shown in this table. Table 3 shows the measured and modelled travel
times between observer points 1-5 and points 6-10.

In general the results showed that more accurate results were obtained with
more input data. The addition of public transport data had the greatest effect on
queues and travel times, bringing the modelled averages closer to the real
measured values. Starting with the ‘low’ level modelling results, there was a
general reduction in queues and travel times with the medium level model. The
difference in these two models lay in the vehicle proportions on all links. The
medium model had all HGV’s reassigned to the major links. From the Medium
and Low-Level Model to the High Level Model there was a marked increase in
queues and travel times. The High Level Model featured public transport stops,
routes and timetables. Busses have significantly lower acceleration and, given
that the network is a heavily used public transport corridor, this had the effect of
slowing traffic flow and hence in increasing travel times and queues.

Sensitivity tests were then undertaken on the network in order to establish the
most suitable areas for model calibration. The details of the tests are not included
in this paper, but the findings are listed below.
- The effect of road gradient on stop line saturation flow is modelled in
  Paramics.
- The sensitivity tests showed that relationships such as speed-flow obtained
  from microsimulation are of similar shape to macroscopic speed-flow curves
  obtained from site surveys.
- The sensitivity tests provided a good insight into the relationships between,
demand, green time, queues, traffic flows and travel times, but not
necessarily into the best parameters to use for calibration of the Leeds Model.

Without further calibration, the results of the High Level Model were next
compared with the results obtained using Dracula, Aimsun2 and Nemis during the
Smartest Project Transferability Test [1]. Tables 4 and 5 show the queue
length and travel time comparisons between the models. The Smartest Project
deemed that the accurate representation of Queue No. 1 was critical. There is
good agreement between the observed and modelled queues by Paramics at this
approach. However, Paramics had difficulty in correctly modelling Queue No. 2;
the modelled queue was twice the measured value.

<table>
<thead>
<tr>
<th>Queue ID</th>
<th>Dracula (veh)</th>
<th>Aimsun (veh)</th>
<th>Nemis (veh)</th>
<th>Paramics High (veh)</th>
<th>Measured (veh)</th>
<th>SD</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>41.5</td>
<td>25.8</td>
<td>34</td>
<td>24.75</td>
<td>25.6</td>
<td>8.75</td>
<td>22.5,28.7</td>
</tr>
<tr>
<td>2</td>
<td>4.4</td>
<td>7.4</td>
<td>9.0</td>
<td>20.95</td>
<td>9.3</td>
<td>5.17</td>
<td>11.2,14.1</td>
</tr>
<tr>
<td>3</td>
<td>10.3</td>
<td>8.6</td>
<td>8.0</td>
<td>7.63</td>
<td>12.6</td>
<td>3.04</td>
<td>8.2,10.5</td>
</tr>
<tr>
<td>4</td>
<td>3.2</td>
<td>8.8</td>
<td>7.0</td>
<td>4.95</td>
<td>5.2</td>
<td>2.87</td>
<td>4.4,6.0</td>
</tr>
<tr>
<td>5</td>
<td>2.0</td>
<td>3.0</td>
<td>3.0</td>
<td>2.78</td>
<td>3.0</td>
<td>1.92</td>
<td>2.5,3.5</td>
</tr>
</tbody>
</table>

All four models gave queue lengths that were less than the measured values
for Queue No. 3. There was some variation in the modelled queue given by the
four models for Queue No 4. For all four models, there was very close agreement between the observed and modelled queue lengths for Queue No. 5. None of the four microsimulation models performed significantly better in modelling travel times on the network.

Table 5: Measured and modelled travel times.

<table>
<thead>
<tr>
<th>Route ID</th>
<th>Dracula (s)</th>
<th>Aimsun (s)</th>
<th>Nemis (s)</th>
<th>Paramics High (s)</th>
<th>Measured (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5</td>
<td>178</td>
<td>153</td>
<td>183.3</td>
<td>174.5</td>
<td>164.7</td>
</tr>
<tr>
<td>6-10</td>
<td>94.1</td>
<td>111</td>
<td>98.7</td>
<td>132.75</td>
<td>159.6</td>
</tr>
</tbody>
</table>

4.2 Time evaluation and macroscopic comparison

The Transyt program [5] was used to model the Leeds Network in order to compare the model construction time of microscopic simulation using Paramics and that of macroscopic modelling with Transyt. Only signalised junctions were modelled with Transyt.

The time for model construction, visual interrogation and results extraction and interpretation for the High Level Paramics model was 425 minutes. The time for data entry, calibration of right turners and results extraction and interpretation in Transyt was 140 minutes. These times are not an absolute guide to modelling times as this depends greatly on the level of expertise of the modeller.

Since model outputs are limited in Transyt, it was decided to compare the modelled queue lengths from Transyt with those from the Paramics High Level Model. Table 6 shows the queue length comparison for the two software programs. There is a good match between the two models for queues 1, 3, 4 and 5, despite the definition of queue in Transyt as Mean Max. Queue, or MMQ [5].

Table 6: Paramics and Transyt queue length comparison.

<table>
<thead>
<tr>
<th>Queue No.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measured</td>
<td>25.6</td>
<td>9.3</td>
<td>12.6</td>
<td>5.2</td>
<td>2.78</td>
</tr>
<tr>
<td>Paramics</td>
<td>24.75</td>
<td>20.95</td>
<td>7.35</td>
<td>4.95</td>
<td>2.78</td>
</tr>
<tr>
<td>Transyt</td>
<td>24</td>
<td>7</td>
<td>10</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

4.3 Conclusions from the Leeds Network

There was no major difficulty in modelling the Leeds Network in Paramics. The results obtained after visual interrogation of the model were reasonably accurate. The network was modelled using three levels of input data. These Low, Medium and High-level Models showed the effect of extra input data on model result accuracy. The addition of public transport data to build the High Level Model improved model results for queue lengths and for travel times. The addition of public transport slowed the free flow speed of traffic bringing modelled travel times closer to those observed.

The microsimulation sensitivity tests produced relationships for speed/flow, demand/queue, gradient/saturation flow and green time/flow & queue that matched the trends seen in well established aggregated relationships.
Similar to the sensitivity analysis, there are an extensive range of parameters that could be changed for calibration of the model. Guidance from the UK DMRB [6] states that “Arbitrary adjustments to measurable quantities (e.g. link length or junction geometry) should not be made”. In this study no further calibration was undertaken after a visual interrogation of the simulation.

In terms of data requirements, the Transyt model required the same amount of input data as the Low Level Paramics Model. Both models required the same level of input data to accurately represent turning radii at stop lines and the distance between junctions. The same signal data was needed by both models to represent stages, green times and cycle offsets for the three signalised junctions on the network.

Overall it was concluded that Paramics is suitable for modelling in urban congested conditions. This conclusion is based primarily on the accuracy of the modelling results obtained for flows, queues and travel times after a ‘visual calibration’ of the simulation after initial model coding and construction.

### 4.4 Wilton Roundabout

The aim of modelling the Wilton Roundabout was to determine the difference between Arcady [8] and Paramics modelling of an Irish roundabout. This evaluates the ability of microsimulation to model a situation where gap acceptance rules are important. This modelling concentrated on the fundamental relationship between entry flow and circulating flow on each approach arm.

Wilton Roundabout is located in Cork City at the intersection of Bishopstown Road (N71), Glasheen Road, Sarsfield Road and Wilton Road (N71) and is one of the main arterial routes into the city centre. At peak hours, congestion at this junction can cause queues to extend to upstream junctions.

With the junction geometry coded the traffic matrix in the model was manipulated in order to extract the entry flow / circulating flow relationship (Qe/Qc) for the Wilton and Bishopstown approach arms. Therefore the only data input for the model was measured junction geometry. The empirical Kimber formulas [7,8] that form the basis of the Arcady program were then used to derive this relationship for these approach arms. Up to this point the only input to the Paramics and Arcady models of the roundabout were the geometry and dimensions. Points on the Qe/Qc line were then measured for the two approach arms by visiting the roundabout during congested conditions, and measuring the entry and circulating flows for short time periods when there was a continual queue on the approach arms.

Figure 1 shows the Qe/Qc relationship predicted by Arcady and as modelled by Paramics along with the measured data points for the Wilton approach arm. The same trends were found for the Bishopstown approach arm. The Kimber formulas give a capacity that is almost twice that modelled by Paramics. The measured data points lie slightly above, but very close to the Qe/Qc line modelled by Paramics.

This modelling exercise demonstrates the unsuitability of using Arcady to model a large non-circular roundabout. It also highlights the large difference in model output between two programs that were run on the same geometric input.
data. The high capacities predicted by Arcady could be due to the fact that the number of circulating lanes is not parameterised. Thus Arcady could assume three lanes based on the size of the inscribed circle diameter. In Paramics the user can define the number of circulating lanes.

![Figure 1: Qe/Qc Relationships for the Wilton Approach Arm.](image)

The need for some form of model validation for even a single junction analysis has been clearly demonstrated above. The use of roundabout Qe/Qc flow counts for validation was useful, as the data points were relatively easy to obtain from counts during congested conditions.

5 Discussion

There are many methods and headings under which a microsimulation model can be evaluated. For example, a previous microsimulation model comparison study concluded that Paramics and Vissim were the two models with the greatest range of modelling ability [9]; neither model could be rated above the other. These two microsimulation models are the two main programs used in the UK and to a lesser extent in the US. Based on the literature and on the investigations outlined in this paper, it is concluded that results obtained with the Paramics microsimulation model can be applied to microsimulation in general.

Static assignment modelling techniques using macroscopic models fail to model many dynamic phenomenon such as queue spillback into upstream junctions or driver re-routeing due to congestion. From this work and from a review of the available literature, it is suggested that microsimulation is a logical method for modelling both dynamic phenomenon and also conventional traffic movements.

Consequently, microsimulation models are increasing the range of traffic phenomenon that can be modelled and are taking over from macroscopic models. At present, a macroscopic program such as Transyt is often used for traffic modelling, another program is required for pollution modelling and a third for result presentation. Microsimulation models such as Paramics can incorporate all
three of these programs. Future development of this program includes a signal optimiser, giving it the modelling capability of the Transyt program.

For practical reasons it would still be more suitable to use the single junction analysis programs such as Picady and Arcady rather than microsimulation for situations where simple outputs are required.

The advantages of microsimulation modelling over conventional macroscopic modelling are:

- The ability to model complex, non standard junction layouts
- An increased range of modelled traffic phenomenon
- An increased range of model outputs
- A graphical network builder and editor
- Mistakes in the coding can be seen in the simulation
- Improved graphical outputs
- No limit on model size
- Can model dynamic traffic phenomenon

5.1 The accuracy and calibration of microsimulation

The modelling undertaken in this investigation has shown that microsimulation is at least as accurate as conventional macroscopic modelling. However, microsimulation involves extra model construction time and extra input data for calibration and validation. In this research, Transyt model construction time was only 33% of the time required for microsimulation modelling of a three node traffic network. Thus the decision to use microsimulation modelling in this case could only be justified if outputs were required that could not be obtained from the conventional software.

Microsimulation model calibration is a key issue. It has been suggested that the wide variety of adjustable parameters in a typical microsimulation model allows the user to obtain results which suit the required answer. Thus there is a worry that microsimulation model results can be ‘fixed’. For example, in Paramics, one can change parameters in the model to make the queues shorter or longer in the simulation. Alternatively, one can change the definition of when a vehicle is in a queue, thereby leaving the simulation the same while changing the value of the modelled results.

While it is true that this form of model fixing can be used with microsimulation, it can also be applied to all other modelling software packages. For example, some practitioners consider that an advantage of the Transyt program over microsimulation is that the input saturation flows can be changed to get the required queue length. This procedure contradicts the advice of the UK Design Manual for Roads and Bridges [5], which states that arbitrary changes to measurable quantities should not be made to get a better match between results and survey data.

As all traffic modelling software can be misused in this way, there is a reliance on the integrity of traffic modellers to make reasonable assumptions in the models. It would be helpful if a list of changes to default parameters and assumptions were provided with every set of modelling results submitted. This
would make modelling more transparent. Every model run from Paramics produces a ‘configuration’ file. This file contains a list of assumptions and values for various parameters in the model.

This investigation has found that the visual calibration of the simulation immediately after model construction is the most important part of model calibration. The results obtained after a visual calibration of the Paramics Leeds Network Model were very close to the final calibrated results.

5.2 Shortcomings of microsimulation

As mentioned previously, microsimulation modelling requires more time and input data than macroscopic modelling. While model construction is not especially difficult using microsimulation, due mostly to the graphical user interface, the proper interpretation and calibration of modelling results requires an experienced modeller.

Result extraction in microsimulation can be arduous for some model outputs. Using the example of the Paramics Program, average, minimum and maximum queues can easily be obtained, however the queue at start of green at a set of signals must be extracted from a lengthy text file. However, in general, outputs can be viewed more easily in microsimulation than macroscopic models.

6 Summary

- Available literature on the evaluation of microsimulation models is limited.
- The Paramics and Vissim commercial microsimulation models appear to have the greatest modelling ability.
- Microsimulation is the only way of modelling many ITS applications and dynamic traffic phenomenon.
- For a single junction analysis requiring standard model outputs, it is still better to use single junction software such as Picady and Arcady.
- Microsimulation modelling was found to be at least as accurate as macroscopic modelling. This conclusion is in agreement with the available literature.
- Microsimulation modelling relies on the integrity of the modeller when making assumptions and when changing model parameters. All models should be accompanied by a list of assumptions made and justification for changes made to default parameters.
- The visual calibration stage was found to be the most important part of the calibration process.
- The use of multiple model outputs is very important for microsimulation model calibration and traffic or road scheme evaluation.
- It appears that microsimulation models developed in the UK are transferable to Irish traffic conditions.
References


[5] Transyt 11 TIPS file. Transport Research Laboratory, Old Wokingham Road, Crowethorne, Berks RG45 6AU, UK.


