This is a repository copy of *An Investigation of Area Speed Flow Relationships by Micro-Simulation Single Links*.

White Rose Research Online URL for this paper: http://eprints.whiterose.ac.uk/2152/

---

**Monograph:**

Working Paper 428

---

**Reuse**
See Attached

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
This is an ITS Working Paper produced and published by the University of Leeds. ITS Working Papers are intended to provide information and encourage discussion on a topic in advance of formal publication. They represent only the views of the authors, and do not necessarily reflect the views or approval of the sponsors.

White Rose Repository URL for this paper:
http://eprints.whiterose.ac.uk/2152/

Published paper
AN INVESTIGATION OF AREA SPEED-FLOW RELATIONSHIPS BY MICRO-SIMULATION SINGLE LINKS

A. D. MAY AND S. P. SHEPHERD

ITS Working Papers are intended to provide information and encourage discussion on a topic in advance of formal publication. They represent only the views of the authors, and do not necessarily reflect the views or approval of the sponsors.
Abstract

This working paper is the first in a series relating to the EPSRC funded project, "The Definition of Capacity in Urban Road Networks: The role of Area Speed-Flow Relationships". It defines the measures collected and the methods for data collection in terms of space-time diagrams. The four measures are collected by simulating different levels of demand. Two cases are presented, the first simulates constant demand throughout each simulation period, the second simulates a demand profile which stays uniform as peak demand increases.

The paper concentrates on single link models with only one OD pair and only one route. Two equivalent networks are presented. The first represents a single link where the capacity of the link is maintained. The second is a purely hypothetical network where the capacity is reduced to zero as demand increases. This is achieved by a loop configuration so that the departure rate is affected by the queue blocking-back on itself.
1. Introduction

1.1 Objectives

The objectives of the research are :-

i) to investigate the interaction between vehicle-kilometres and vehicle-hours in a network as individual junctions become saturated;

ii) to develop improved area speed-flow relationships which reflect the relationships in i);

iii) to use the relationships in i) and ii) to explain the process by which networks reach capacity;

iv) to assess the significance, for the evaluation of road pricing policies, of the use of the improved relationship in ii).

This paper describes the data collection process employed in the simulation model NEMIS (Mauro, 1991) and the four basic measures used to assess the performance and supply of the network under different levels of demand. NEMIS is a micro-simulation model developed in Italy by Mizar Automazione. It simulates the movement of individual vehicles from origins to destinations through a network in increments of one second. The model can represent the behaviour of vehicles at signalised and non-signalised junctions using lane changing and gap acceptance models combined with a car following law.

The results from two single link networks will be discussed in detail. The single link networks represent the simplest types of network possible in that they have only one OD pair with only one possible route, however they provide an extremely useful insight into the pitfalls of the data collection process.

1.2 Background

The relationship between flow, speed and concentration is fundamental to the understanding of congestion, and of the increased delay which is generated as flow increases. Speed-flow relationships for individual links and junctions are well accepted (even though based on limited data), and are used extensively to evaluate the benefits of investment in additional capacity.

The process by which increases in flow lead to increases in delay is also fundamental to the justification of road pricing; the concern being that each additional driver adds to the delay, and hence imposes a cost on others, of which he is unaware (Smeed, 1964). However, this effect is of greatest concern in urban road networks, where the delays on individual links are not independent, and where the process by which delay increases as demand increases is much less clear. Some attempts to assess road pricing have used conventional network-based transport planning models, but these typically do not treat the process of queuing accurately, and therefore do not reflect the way in which queues block upstream junctions, and hence deny access to parts of the network. Current research at Middlesex University is investigating this process more fully (Wright and Abbess, 1992).

An alternative, more frequently used in the analysis of road pricing, is to use either a single link model or an area speed-flow relationship, in which the relationship between vehicle-km/h and vehicle-h/h, or vehicle-km/h and network speed (expressed as (vehicle-km/vehicle-h)) is used to describe the supply of road capacity. The literature uses
the terms speed and flow in this context very loosely. We use flow to mean vehicle-km/h, and speed to mean (vehicle-km/vehicle-h) in what follows. The parameters can alternatively be expressed in pcu rather than vehicles. However, other authors have used other definitions and have often failed to explain how they have aggregated data from links to networks.

The earliest work in the UK on such relationships was developed for the Smeed Report, and its sequel Better Use of Town Roads (MoT, 1967) by Thomson (1967a, 1967b). He developed a relationship between speed and flow which was linear at higher flows but convex at lower flows. In his earlier work he produced separate relationships for main and minor roads, with the latter relationship having a steeper slope. In parallel, Wardrop (1968) produced a relationship between speed and flow for Central London, which included terms describing link travel time and junction delay, and which was derived partly theoretically and partly from empirical data. The relationship, which included terms for average carriageway width, frequency of junctions, and average proportion of effective green time, was virtually linear through Thomson's observed points, but its slope increased at speeds below 15 km/h to reach a value of capacity some 15% above the flow accommodated at 15 km/h. Further theoretical work was conducted by Wright (Wright, 1975). In parallel, the series of urban congestion surveys in 13 towns and cities over the period 1963-1976, and a separate set of surveys in London, were being analysed to produce similar relationships (Marlow and Evans, 1978; Duncan et al, 1980). Their relationship was linear between speed and the increase in flow above off peak conditions. It also included a term in junction frequency for central areas, and one in density of development for non-central areas. Similar work was carried out for road pricing studies in Bogota (unpublished) and Hong Kong (Harrison et al, 1986). The Hong Kong work was assisted by a theoretical unpublished review by Wright, in which he argued that Wardrop's was the most appropriate relationship of those available, but that it was affected by the propensity for drivers to reroute, which helped to increase the capacity of the network.

Duncan's results were incorporated into TRRL's London Area Model (LAM) (Oldfield, 1993) as a linear relationship between speed and the change in vehicle-km/h. This model was used for the transport strategy analysis conducted for the London Planning Advisory Committee (May and Gardner, 1990). However, the relationship had to be modified by introducing an arbitrary increase in slope below a specified speed; without this, the model predicted large increases in traffic in central and inner London with only small speed reductions which was not credible.

A relationship similar to that used in LAM has since been employed in the START model (Bates et al, 1991), which has been used for integrated transport studies in Edinburgh, Bristol, Luton-Dunstable, London and Merseyside (May, 1993). In this case each zone is defined as having three area speed-flow relationships, for inbound radial traffic, outbound radial traffic and orbital traffic, and a separate linear relationship is specified for each, based primarily on survey data.

A major problem with the development of such relationships from empirical data is that average network speeds of less than 15 km/h are rarely observed. It is not clear whether this is due to demand responses, with drivers reluctant to travel at lower speeds, or to network response, with queuing restricting access to an area, and hence regulating vehicle-km within it. Either of these processes could lead to the commonly experienced process whereby low speed conditions gradually spread in both space and time, without becoming significantly worse at any already congested location. Equally, the theoretical
derivations of the kind developed by Wardrop do not reflect the effects of queues on upstream links and hence on delay and throughput. As noted earlier, work for LPAC modified the linear relationship in this low speed range to improve model performance, but there is no justification empirically or theoretically for doing so.

The issue is an important one, firstly because of the widespread use of linear area speed-flow models for road pricing analysis, secondly because the shape of the relationship in the range in which junctions become saturated will significantly affect the estimated benefits of road pricing, and thirdly because the implications of the model can easily be misinterpreted. The last two points are amply illustrated in a recent paper by Evans (Evans, 1992) who uses a linear area speed-flow relationship, in the form of a single link model, to assess the relative benefits of road pricing and bus priority. Using a linear relationship he demonstrates that the benefits from road pricing only become substantial at speeds significantly below those currently experienced, even in central London. However, he also experiments with a truncated relationship in which speed falls to zero once flow reaches a defined capacity, and a simple bottleneck formulation, in which speed is unchanged until capacity is reached, and then falls to zero, and demonstrates that the benefits, for conditions beyond the discontinuity in these relationships, are greater than for similar levels of speed in the simple linear case.

Fundamentally, also, he argues that it is possible to achieve zero speed at a non-zero level of throughput. Hills (1993) disputes this, argues that it is implausible for throughput to be achieved when traffic is stationary, and suggests an alternative analysis based on trips rather than flow. Strategic models (as opposed to single link models) do conduct their analysis on the basis of trips, but still need a relationship between speed and flow in specific areas. Underlying the debate between Evans and Hills is a difference of view on the way in which congested urban networks operate and, crucially, whether low (or zero) speeds are in practice attainable within defined areas. This in turn exposes a lack of understanding of about what is meant by capacity in such circumstances, and of how networks perform as 'capacity' is approached.

1.3 Definitions

For this project network speed is defined as vehicle-km/vehicle-hours and flow is defined as vehicle-km/hour.

Bates (1994) analysed the difference of opinion between Evans and Hills and suggested that two types of functions exist. The first is a network performance measure based on the engineering approach to capacity. The second type of function being a supply measure based on economic analysis. Four network measures were suggested, the first being the performance measure, the other three being supply based measures:

1) **Speed (veh-km/veh-h) vs Actual Flow (veh-km/h) : Performance**

   This is a speed flow performance curve and is calculated from flow which actually occurs on the physical network or within an area/zones of the network.

2) **Speed (veh-km/veh-h) vs Demanded Flow (veh-km/h) : Supply**

   Here the demanded flow (veh-km/h) for a given time period is defined as the demand for trips in that generating sub-period multiplied by the vehicle-km/h achieved at the lowest
level of demand simulated. Thus the different demand levels are simulated in NEMIS by increasing the vehicle generating factor for the whole OD matrix, the resulting demand in vehicle-km/h is calculated assuming the same route choice as results for the lowest demand level simulated.

3) **Average Time/km (veh-h/veh-km) vs Demanded Flow (veh-km/h) : Supply**

This is the inverse of measure 2) and gives the supply curve for a network or zone.

4) **Actual Flow (veh-km/h) vs Demanded Flow (veh-km/h) : Supply**

This measure gives an indication of whether the demanded flow can be sustained.

2. **Data Collection**

Two approaches have been developed for collecting the above measures from NEMIS simulations. The approaches are described in relation to the space-time diagrams shown in figure 1. The space-time diagram is a simplistic representation of a road or link on the Y-axis with time increasing on the X-axis. The arrows represent vehicle speeds or trajectories. This representation although simple is valid for discussing data collection methods for single links.

2.1 **The Time Slice Approach**

The first approach is the time slice approach which is based upon dividing the space-time domain into equal rectangles ABEF defined by the link length and time period AB. The flow and vehicle-hours are collected for each such rectangle and the rates are calculated by simply dividing by the period AB. The approach collects vehicle-km and vehicle-hours which occur on the physical network only; no external queues are included at high demands. This gives the speed flow performance curve (measure 1) and is in fact describing the state of the network for a given time slice.
2.2 The Tracking Approach

The second approach, the tracking approach, is used for the other three supply measures. The approach tracks vehicles through the network and aggregates the data according to the generating sub-periods. The approach collects the vehicle-km and vehicle-hours associated with the quadrilateral ABCD in figure 1 for the uncongested case and ABC'D for the congested case. Here AD represents the speed of vehicles at the start of the period AB, and BC' the (lower) speed once congestion has set in.

To calculate the vehicle-km and vehicle-hour rates associated with the demand in period AB a divisor is used. The divisor is defined as \((AB+DC)/2\) for the uncongested case and \((AB+DC')/2\) for the congested case. Note that for the uncongested case the demanded flow rate is achieved and the divisor is actually equal to the period AB, whereas for the congested case the flow is achieved at a rate lower than demanded and the divisor is greater than the period AB.

2.3 Time Slices

Each simulation may be split into a number of generating sub-periods or time slices. The time slice approach is well behaved for any number of time slices, however the tracking approach becomes more complicated as the demand level is increased and may suffer from missing data as the simulation is ended.

Consider the simulation period to be 3600 seconds, split into two generating sub-periods each of 1800 seconds. Figure 2 shows a space time diagram for the single link network for a number of possible demand levels (1-6) held constant through both generating slices.

Note that any excess demand in this case is assumed to be in an external queue and hence displaced to enter the physical network in a later time slice. In this process, external queues from slice 1 may displace those in slice 2 to later time periods.

Considering the first generating sub-period :-

Case 1

This encompasses all cases where the flow is uncongested and the tracking divisor is the average of AB+DC and is equal to 1800 seconds. This gives similar results to the time slice approach.
Case 2

This is where the flow generated can enter the network in slice 1 but causes congestion. It seems logical that the flow rate is less than that demanded in the first 1800 seconds and so the divisor, the average of \( AB+DC' \), is greater than 1800 seconds.

![Space-Time diagram for various demand levels](image)

Figure 2. Space-Time diagram for various demand levels

Case 3

This is the first case where the demand cannot enter in slice 1. For the three supply measures the extra vehicle-hours in the external queue are collected and the flow occurring in \( AB'C''D \) is collected. The divisor used is the average of \( AB'+DC'' \).

Case 4

This is an extension of case 3 only the flow collected is limited by the end of the simulation period at 3600 seconds.

Case 5 and 6

In these cases the demand is so great that the flow is displaced beyond the end of the simulation and the divisor used tends to 3600 seconds. This is similar to the performance curve for the total simulation with the addition of the time spent in external queues. The flow (veh-km) collected is the same as for the total performance (time slice) flow.
Case 6 is a special case whereby gridlock has formed. This will only occur for networks where the capacity reduces to zero as demand is increased. Networks in which a capacity is reached and maintained will eventually service all demand so that the worst case is something like case 5 where the excess demand is queued up outside the network. In fact for this latter type of network, where capacity is maintained, the speed lines within the network become parallel and only the time spent in the external queue will increase with demand. That is to say that once in the physical network the time taken to exit is the same for all demands above capacity.

The second generating sub-period.

There is once again no problem for the time slice performance measures. However for the tracking approach, this second time period is bounded by the end of the simulation for all demand levels. Furthermore the flow demanded in slice 2 is displaced in time to the right of the lines for cases 2-6. The divisor is defined as the average of (3600-B)+(3600-C) or equivalent for C'C" etc. This means that in the extreme case 6 the divisor is (1800+0)/2=900.

Whether the results from tracking the last time slice are valid or not will be discussed later, but in general it depends upon the type of network. For networks which maintain a flow the results are not valid because the demand is simply displaced to a later period. For gridlocked networks the results are valid as the flow tends to zero and the demand will never be met so that recording vehicle-hours and no flow in the extreme case predicts a speed of zero (which happens to be right).
3. Networks and Demand Profiles

3.1 Single Link Networks

Two types of single link network have been coded in NEMIS and simulated for 20 levels of demand. The networks are shown in figure 3 and are designed to be equivalent in terms of length and signal capacity but differ in that in case b) the single link loops round upon itself. This is obviously a hypothetical network but it represents a network whereby the capacity is reduced by blocking-back as the demand is increased. The single link network in case a) represents a network where capacity will be maintained.

![Diagram of single link networks](image)

a) Single Link Network

b) Equivalent Single Loop Network

Figure 3. Single Link Networks

3.2 Demand Profiles

The networks were simulated for two separate demand profiles depicted in figure 4. The first case is for a flat one hour peak for each demand level. Each curve is made up from 20 simulations, each simulation representing a different peak hour demand. The peak hour was divided into 4 time periods each 15 minutes long.
In the second case the flat one hour peak is preceded by a half-hour warm-up period and followed by a half-hour cool-down period. For the results presented here the demand level in the warm-up and cool-down periods was set as the lowest level of demand simulated within the peak. This warm-up and cool-down demand level was held constant whilst the peak demand level was increased. The simulations are again divided into 15 minute periods. Note that the peak hour now covers slices 3 to 6 compared to 1-4. The case 2 profile data is collected using the tracking approach until the vehicles from slice 6 have all been serviced or until the end of the simulation (3600 seconds).

Figure 4. Demand profiles used for simulations

4.0 Results

Figures 5-12 show the four measures for the peak hour slices for the single link (type a) networks for demand cases 1+2 described above. The case 1 measures are titled "Single Link Network..." and case 2 measures are titled "Single Link Profile Results...". In general the two measures for case 1+2 are displayed side by side.

Figures 13-20 show the same measures in a similar manner for the single loop network (type b). The results will be discussed in terms of difference in network and difference in demand profile.

4.1 Single Link vs Single Loop

Irrespective of the demand profile used to generate the curves there are obvious differences between the two types of network.
4.1.1 Speed vs Actual Flow Performance Curves

Figures 5, 6, 13, 14. For these curves the demand is increasing from one point to the next in a clock-wise manner. Both networks exhibit a relatively free-flow regime until the speed drops to about 50 km/h, although the break point is at a slightly lower demand for the loop network. Both network speeds then drop more rapidly with increases in demand down to a speed of about 10 km/h at flows of around 2800 veh-km/h for the single link and 2300 veh-km/h for the single loop.

Any further increases in demand produce a slight loss in capacity for the single link until a speed of around 8 km/h and a flow of 2600 veh-km/h is reached and maintained for further increases in demand. This point corresponds to the actual physical network being full during the time slice with a constant departure rate and hence a constant speed or time to traverse the physical link.

However any further increase in demand for the single loop network produces a loss in capacity due to blocking-back and the curve bends back to the origin where a gridlock has occurred i.e. all vehicles are queued in the loop and cannot exit due to blocking-back. It must be pointed out that this network is purely hypothetical and that the blocking-back model incorporated within NEMIS does not allow for a gridlock to be "unlocked" by driver behaviour mechanisms which may alleviate the situation in a real network.

The profile results where the network is allowed to warm up and cool down at the lowest demand level for half an hour either side of the peak give very similar curves to the one hour flat simulations.

4.1.2 Speed vs Demanded Flow

Figures 7, 8, 15, 16. The demanded flow defined previously is basically a vehicle generation factor multiplied by the veh-km/h for the lowest demand tested. The approach used to collect the vehicle-km and vehicle-hours is the tracking approach for each generating sub-period. The vehicle-hours now include any time spent in a queue external to the physical network or link when the demand is high. This has the effect of producing lower speeds as the external queue increases; even the single link speed would tend towards zero if there were an infinite external queue.

Figure 7 shows that the speed drops for later time slices, slice 4 actually shows a zero speed at high demands. This is actually misleading as the flow has not yet occurred, it has simply been displaced beyond the end of the simulation period as described in case 5 section 2.2 above. The profile method overcomes the displacement problem to some extent and slice 6 (figure 8) can be compared to slice 4 (figure 7) showing a more gradual trend to a zero speed as demand is increased. The flow displacement problem still exists but only for very high demands where flow is displaced beyond the two-hour simulation period.

The single loop results (figure 16) show a trend towards zero speed which is actually a valid result due to the formation of gridlock. Note that for both networks the first time slice is acting as a warm-up period and so produces higher speeds than for the other slices. The profile method produces a very similar warm-up slice, obviously the degree to which the network is "warmed up" depends upon the choice of demand in the first half-hour.
4.1.3 Average Time/km vs Demanded Flow

Figures 9,10,17,18. The graphs show the 4 time slices from the peak hour for both networks. Slice 4 and to some extent slice 3 for the single link network suffers from the flow displacement problem in that as the veh-km is displaced beyond the simulation period then the external queuing time is divided by a flow which is tending to zero merely because the simulation has to end somewhere. This displacement problem is reduced for the profile results but it still occurs (see figure 10).

The separate curves for each time slice show the time dependent nature of the problem once the demand exceeds the capacity and a queue forms, basically the queue will grow with time until the demand is reduced (unless gridlock has occurred). The profile results are grouped more tightly for both networks and the mid-peak results are represented by adding slices 4+5 together. The single link times/km (figure 10) increase linearly with increasing demand and give a reasonable supply curve, as do the single loop times/km which increase sharply as gridlock sets in (figure 17).

4.1.4 Actual Flow vs Demanded Flow

Figures 11,12,19,20. This measure is the only measure which is not independent of the definition of the divisor in the tracking approach (see section 2). All these figures show how actual flow equals demanded flow up to a point, then for the single link case the flow levels off to a capacity and for the single loop network the flow drops to zero as gridlock occurs.

Figure 11 shows the process of flow displacement beyond the simulation period for slices 3+4. Figure 12 confirms that the profile method has reduced the problem but that it still occurs for slice 6.

The jumps in the single loop flows (figures 19+20) are due to large changes in the divisor, as specified for the tracking approach, as gridlock occurs. This problem does not occur for other measures as the flow is divided by the hours or vice versa and so the measure is independent of the divisor.

5. Conclusions

A method for collecting speed flow measures by micro-simulation has been developed and applied to single link, single OD networks. Two basic approaches for collecting data have been developed. The first time slice approach give the speed flow performance curve for the physical network, i.e. it gives the state of the network in a particular time slice. The second tracking approach follows vehicles through the network aggregating the data according to the origin generating period. This approach uses the concept of external queues to give supply measures when demand exceeds capacity and the physical network becomes full.

Two types of single link network have been identified, type a) where a capacity is reached and maintained, and type b) where the capacity drops to zero as demand is increased and gridlock occurs. It is thought that realistic networks will produce curves in between these two hypothetical cases depending on the network topology.

Two approaches to the definition of the demand profile were simulated. The problem of flow displacement and missing data due to ending the simulations was identified for the tracking approach under high demands. The cool-down period of the profile approach
reduced the magnitude of this problem but did not prevent it completely. Hence the profile approach can give better results for the last peak hour time slice.

The other time slices were very similar for both approaches, probably due to the fact that the warm-up and cool-down demand level was the lowest level simulated. The second profile approach would enable any demand profile to be tested but without real data the assumption that off-peak is a fraction of the peak demand level for each simulation seems fair. In fact the assumption made for the flat peak approach starting from an empty network was that the network would warm up and settle down by the second time slice, i.e. so long as the warm-up demand is less than the peak demand then the second peak time slice should be the same as if the network was not warmed up previously.

Both approaches however showed the time dependent nature of the supply curves in congested conditions. The question of which slice or curve to use is not obvious as it depends upon how long the demand level is sustained and perhaps on the demand level in the cool-down period. For the rest of this project the peak demand is always assumed to be one hour long and the speed flow measures will be represented by adding together the mid-peak slices i.e. slices 4+5 for the profile results (as shown) or slice 2+3 for the flat peak results.

The current methods have been adapted for use on grid networks. The networks have been divided into zones as in the START model. Within each zone three types of links are specified for data aggregation, inbound, outbound and orbital. To save time in simulations the one-hour flat peak approach is used as the mid-peak results will be similar to those produced by a profile approach. The results for the grid network under 6 different OD patterns will be discussed in the next working paper in the series along with problems associated with the definition of demand in central zones and the tracking approach.
6. References


Harrison W J, Pell C, Jones P M and Ashton H (1986). Some advances in model design developed for the practical assessment of road pricing in Hong Kong. Transportation Research 20A.


Ministry of Transport (1967). Better use of town roads. HMSO.


Wright C (1975). An explanation of some aspects of the behaviour of congested road traffic in terms of a simple model. Transportation Research 9(5).

Fig 5
Single Link Network: Time Slice Approach
Speed vs Actual Flow Slice 2

Fig 6
Single Link Profile Results: Time Slice
Speed vs Actual Flow (mid-peak 4+5)

Fig 7
Single Link Network: Tracking Approach
Speed vs Demanded Flow

Fig 8
Single Link Profile Results: Tracking
Speed vs Demanded Flow
Fig 13
Single Loop Profile Results: Time Slice
Speed vs Actual Flow (mid-peak 4+5)

Fig 14
Single Loop Network: Time Slice Approach
Speed vs Actual Flow Slice 2

Fig 15
Single Loop Profile Results: Tracking
Speed vs Demanded Flow

Fig 16
Single Loop Network: Tracking Approach
Speed vs Demanded Flow
Appendix A: Area Speed Flow Publications


Technical Notes


To Be Written

WP SATURN: Summary of Saturn work (joint) Barcelona abstract?