21st Australasian Transport Research Forum
Adelaide, September 1997

A Model for Evaluating the Effectiveness of Incident Management Systems

Neil Prosser
PPK Environment and Infrastructure

Abstract:

In most urbanised areas of Australia traffic flow on major arterial roads and freeways is at, or close to, capacity for extended periods and traffic incidents of even a small duration can cause major delays. A model for estimating delay caused by incidents is presented that allows evaluation of different incident scenarios using established deterministic queuing formula. An incident is assumed to cause a major, temporary reduction in roadway capacity and if the variation of both capacity at the incident site and arrival flow upstream of the incident are known, then the model estimates the time for the resulting queue to dissipate and total delay. The model is used to examine the variation in incident impact for different incident characteristics including duration, severity and time of day. The use of the model to assess benefits of deploying incident management resources is demonstrated by estimating delay savings arising from reduced incident duration and traffic demand upstream of the incident.

Contact Author:

Neil Prosser
PPK Environment and Infrastructure Pty Ltd
PO Box 248
Concord West NSW 2138

telephone: (02) 9743 0333 fax: (02) 9736 1568
email: rustsyd@ozemail.com.au

Introduction

Congestion in urban areas incurs a major cost to the community in terms of delays, accidents and environmental impact. Flow on major roadways is often close to capacity for extended periods during both weekdays and weekends. Any loss in capacity due to incidents, such as vehicular accidents and breakdowns can lead to extensive delays to traffic upstream of the incident site.

To address the phenomenon of 'non-recurrent' congestion, transport agencies across the world are developing incident management systems that attempt to ameliorate the impact of incidents using advanced technology. Various mechanisms, including loop detectors, closed-circuit TV cameras and manual reporting, are used to detect and verify incidents. Confirmed incidents are dealt with via deployment of measures such as emergency services, motorist information and real-time traffic control. Effective identification and use of available incident management resources requires an understanding of the likely benefits of their application to different incident scenarios. A major benefit to implementation of incident management plans is the saving in delay to traffic upstream of an incident site, and this aspect of incidents is the subject of this paper.

A large proportion of incidents in urban areas result in a major reduction in capacity at the incident site. During periods of moderate flow, such capacity reductions are often sufficient to cause queues to develop upstream of the incident site as demand temporarily exceeds capacity. Under such conditions the effect of an incident continues well past the time to clear the roadway, as time is required for the queue to dissipate. Commonly used measures of the impact of incidents are the time required to clear the queue and the total delay to vehicles due to the incident.

This paper presents an approach to estimating the total delay arising from a particular incident occurring under known flow conditions. Previous work, for example Federal Highway Administration (1986), Morales (1986) and Transportation Research Board (1990), has developed analytical techniques for determining total delay due to an incident assuming constant flow conditions and discrete changes in capacity at the incident site. The technique developed in this paper extends the analysis to the situation where both the arrival rate upstream of an incident and the capacity of the roadway at the incident site vary over time. Established deterministic queueing equations are used in a simple spreadsheet program to allow examination of the impact of incidents, in terms of the time for queue to discharge and total delay, under different flow conditions. The effect of incident duration on queue dissipation time and total delay is examined in detail and methods for assessing the likely benefits, in terms of reduced delay, arising from reducing the duration and severity of incidents are examined.

Delay caused by an incident: constant arrival rate and incident capacity

Let us assume the arrival rate upstream of an incident is constant at level q, less than the capacity of the roadway without an incident, c_{max}. Further, assume that an incident leads to a temporary reduction in capacity to level c_i, for a period of duration t_i, with $c_i < q$.

Figure 1 shows the variation in queue length upstream of the incident under these conditions. The queue builds up at a rate equal to $(q - c_i)$ veh/h, with the maximum queue length, L_{max}, attained at the point where the incident site is cleared. The queue then dissipates at a rate $(c_{max} - q)$ veh/h, so that the time for the queue to dissipate after the start of the incident, t_q, is given by:

$$\begin{aligned} t_q &= t_i + \frac{(q - c_i)}{(c_{\max} - q)} t_i \\ &= t_i \frac{(c_{\max} - c_i)}{(c_{\max} - q)} \\ &= t_i f_i \end{aligned} \quad (1)$$

where $f_i = (c_{max} - c_i)/(c_{max} - q)$ could be interpreted as the 'severity' of an incident. For example, a value of $f_i = 2$ indicates that the queue associated with the incident lasts twice as long as the duration of the incident itself. The parameter, f_i, is equal to the ratio of the reduction in capacity due to the incident to the unused capacity once the incident has cleared. It is also equal to the rate of queue buildup over the rate of queue dissipation. Thus, in determining the impact of an incident it is equally important to consider both the capacity reduction due to the incident and the capacity available for queue dissipation once the incident has been cleared.

Other measures of the impact of an incident can be derived from the preceding formulations. The total delay, D, due to an incident is the area under the queue length curve in Figure 1, given by:

$$\begin{aligned} D &= \frac{1}{2}(q - c_i)\, t_i^2 + \frac{1}{2}(q - c_i)\frac{(q - c_i)}{(c_{\max} - q)} t_i^2 \\ &= \frac{1}{2}(q - c_i)\, t_i^2 \left[1 + \frac{(q - c_i)}{(c_{\max} - q)} \right] \\ &= \frac{1}{2}(q - c_i)\, t_i^2 f_i \end{aligned} \quad (2)$$

The number of vehicles delayed, N, is equal to the number of vehicles arriving during the period that the queue exists, i.e. $N = qt_q = qt_i f_i$. Thus, the average delay to vehicle arriving during the incident effect period, D', is equal to:

$$D' = \frac{1}{2}\frac{(q - c_i)}{q} t_i \qquad (3)$$

A number of important conclusions can be drawn from these formulae:

1. The **total** delay due to an incident under conditions of constant arrival flow is proportional to the duration of the incident squared so that, for example, with all other variables equal, an incident of duration 20 minutes will result in total delay four times that for an incident of 10 minutes duration.

2 Similarly, **total** delay increases with the square of the difference between arrival rate, q, and incident capacity, c_i.

3 The **average** delay to vehicles disrupted by an incident is determined by the capacity and duration of the incident, c_i and t_i, and the arrival flow, q. The 'normal' capacity of the roadway, c_{max}, only determines total delay and not average delay.

The value of the incident capacity, c_i, strongly determines the impact of an incident and depends very much on the nature of the incident, in particular how many lanes of the roadway are blocked. Previous work, such as Morales (1986) and Goolsby (1971), has suggested that the *per lane* capacity during an incident is significantly less than that during uninterrupted flow conditions due to the effects of vehicles slowing down as they pass the incident.

Delay caused by an incident: varying arrival rate and capacity

The preceding analysis was based on the assumption that both the arrival rate upstream of an incident, q, and the capacity during an incident, c_i, are constant. A more realistic scenario is where both the arrival rate and capacity vary continuously over time, as illustrated in Figure 2. In this example the arrival rate of vehicles upstream of an incident site is peaked, climbing to a level close to the normal roadway capacity, c_{max}, before decreasing.

In Figure 2, the capacity at the incident site drops sharply at the onset of the incident from c_{max} to c_i. Capacity remains at this level until the blocking vehicles are cleared, where capacity increases to an intermediate level during which, although all lanes are clear, discharge through the incident site is still constrained by the effects of 'rubbernecking'. Full capacity is restored once all evidence of the incident has been removed. There is strong evidence, for example Hunt and Yousif (1994), that capacity under the forced-flow conditions existing after an incident has been cleared is still less than that under free flowing conditions.

In general, a queue develops upstream of an incident when the arrival rate, $q(t)$, exceeds the capacity of the incident site, $c(t)$. The instantaneous rate at which the queue builds at any time t is equal to $q(t) - c(t)$. When $q(t) < c(t)$, the queue dissipates at a rate $c(t) - q(t)$. The expression for queue length at time t is given by:

$$L(t) = \int_{t_a}^{t} [q(t) - c(t)]\, dt \tag{4}$$

where t_a is the time where the arrival rate first exceeds capacity after the start of the incident.

The total delay due to an incident under these conditions is equal to:

$$D = \int_{t_a}^{t_b} L(t)dt \tag{5}$$

where t_b is the time when the queue dissipates after capacity exceeds arrival rate.

Note that when using the above equation to determine total delay due to an incident consideration needs to be given to possible delay due to recurrent congestion in the absence of any incident, where the arrival rate $q(t)$ exceeds normal capacity c_{max} for some period. In this case, the total delay due to an incident would equal the value given by Equation 5 above less delay due to recurrent congestion.

In terms of incident management, the major objective in minimising the contribution of an incident to delay is to reduce the area shaded in Figure 2 - the difference between the arrival rate and capacity curves. Reducing the magnitude of this area could be achieved through a combination of minimising the period of vehicle blockage and minimising the effects of rubbernecking and other factors that reduce the flow of traffic past the incident site. Total delay is minimised when the area under the queue length curve (equal to the difference between cumulative arrivals and departures) is minimised.

Variation of incident impact with time of day

An important corollary from the above analysis is that the impact of an incident, in terms of delay to traffic upstream of the incident site, will vary considerably depending on the flow rate during and after the incident. For most urban arterials and freeways, where flow rates vary considerably during a typical day, this implies that the impact of an incident will vary depending on the time the incident occurs.

To investigate the impact of different flow and capacity profiles on the duration of the incident effect period and total delay, a spreadsheet tool was developed that allowed specification of flow and capacity profiles as piecewise linear curves. Discrete forms of Equations 4 and 5 above were then used to estimate the variation in queue length and total delay arising from an incident for a particular capacity profile.

Figure 3 shows a possible flow profile for a major urban arterial or freeway during a typical weekday. The morning peak is higher and sharper than the afternoon peak. Also shown on the graph are the variations in queue dissipation time and total delay due to an incident of duration 15 minutes occurring at different times during the day. An incident capacity of 1600 veh/h was assumed with maximum capacity level of 4400 veh/h. Figure 4 shows the same results for incidents of 30 minute duration. Some key results from Figures 3 and 4 are summarised in Table 1 below.

Table 1 Variation in incident impact with time of day: key results

Variable	Values				
	Off Peak	AM Peak		PM Peak	
Flow, veh/h	2600	4400	(1.7)	4000	(1.5)
Incident Duration = 15 min					
t_q (minutes)	23	104	(4.5)	91	(4.0)
D (veh h)	49	705	(14.4)	461	(9.4)
Incident Duration = 30 min					
t_q (minutes)	46	142	(3.1)	154	(3.3)
D (veh h)	196	1858	(9.5)	1588	(8.1)

Note: Values in parentheses indicate ratio of AM and PM Peak values to Off Peak value.

From Figures 3 and 4 and Table 1, the following results can be drawn:

1. The variation in the impact of an incident, as measured by both the queue dissipation time and the total delay, is more marked than the flow variation. For example, although the maximum flow value during the AM Peak is 70% higher than in the Off Peak, the maximum AM Peak values of the duration of the incident effect period and the total delay are 4.5 and 14.4 times, respectively, the corresponding values in the Off Peak.

2. The occurrence times of incidents leading to maximum impact are well before flow maximums. For example, for 15 minute incidents, the worst occurrence time in terms of the queue dissipation time is about 45-50 minutes before both the AM and PM flow peaks. Flows at these times are well below capacity but are rising sharply particularly in the AM Peak period. Note that the worst incident occurrence time in terms of total delay is 15 to 30 minutes after that for the queue dissipation time.

3. As noted earlier, doubling the duration of an incident, in this case from 15 to 30 minutes, in the off-peak period where flow is constant, leads to a doubling of the duration of the incident effect period and a fourfold increase in total delay. However, for incidents occurring immediately before the morning and afternoon peaks, the increase in incident impact with increasing incident duration is less, as a major part of queue dissipation time extends to the period where flow rates are decreasing. For example, the maximum total delay for an incident of duration 30 minutes in the AM Peak period is 2.6 times the maximum delay for a 15-minute incident

This last point regarding the impact of incident duration on incidents occurring at different times of day is further highlighted in Figure 5 which shows the variation in total delay with incident duration for incidents occurring at different times of day. The flow and capacity assumptions are the same as those used in Figure 3 above.

The variation in total delay with incident duration for incidents starting at 10:30, at the start of the off peak period of near constant flow, is similar to that shown by Equation 2 earlier, i.e. total delay increases in proportion to the duration squared.

Incidents of short duration occurring at 6:45, well before the AM Peak flow, lead to relatively small values for total delay but as the duration of the incident increases, and the effect of the incident extends into the peak flow period, the total delay resulting from the incident increases sharply, at a rate higher than that for constant flow conditions.

For incidents starting at 7:30 and 8:00, even short incidents lead to considerable delay. The increase in delay with incident duration approaches a linear relationship for longer incidents.

The shapes of the total delay versus incident duration curves are better seen in Figure 6 where total delay is shown as a proportion of the total delay for a 30 minute incident. Figure 6 shows the smooth quadratic relationship for incidents occurring at 10:30 where flow is constant. The curve for incidents occurring at 6:45 is much steeper, while the curves for times of 7:30 and 8:00 approach a straight line.

The curves of total delay versus incident duration shown in Figure 5 can be used to estimate the benefits, in terms of reduced delay, of incident management systems that reduce the time to clear an incident. Table 2 shows the total delay, in vehicle hours, saved when an incident of a particular duration and occurrence time has its duration reduced by one minute. Shown in parentheses are the percentage reductions in total delay. For example, if an incident occurring at 7:30 has its duration reduced from 10 minutes to 9 minutes the delay due to that incident will be reduced by amount of 54.8 vehicle hours equal to 13.2% of total delay. Once again it can be seen that the largest benefits of reducing the duration of incidents is achieved for those incidents starting close to the peak flow period.

Table 2 Total delay (vehicle hours) savings resulting from reducing incident duration by one minute

Incident occurrence time	Incident duration (minutes)		
	10	20	30
6:45	2.7 (15.6%)	11.4 (11.1%)	72.0 (10.8%)
7:30	54.8 (13.2%)	69.2 (6.4%)	78.5 (4.3%)
8:00	36.2 (12.4%)	49.1 (6.5%)	57.9 (4.4%)
10:30	3.6 (14.9%)	8.0 (8.6%)	12.3 (6.0%)

Reduction in delay due to decreased upstream arrivals

A key mechanism to reduce the impact of an incident on upstream traffic is to reduce the demand for flow through the incident site either by diverting traffic to an alternative route or mode of travel, or encouraging motorists to postpone or cancel their trip. The approach used in this paper can be extended to consider the reduction in total delay resulting from a reduction in arrival flow upstream of an incident.

Returning to the incident scenario considered in section 2, where arrival rate is assumed constant immediately after the occurrence of an incident, assume that at some time t_2 the arrival rate drops from level q to q_2. Figure 7 shows the resulting change in queue length for the case where the reduction in arrival rate occurs after the incident is cleared but before the queue dissipates. At the reduced arrival rate, the queue dissipates at a faster rate and total delay is reduced. If the drop in arrival rate occurs before the incident is cleared, the maximum queue length will be less as well. Under these conditions, simple formulae can be developed for the reduction in queue dissipation time and total delay due to the decreased arrival rate

An important consideration in developing incident management plans is the response time for reducing demand upstream of the incident. With fast dissemination of incident information via roadside signs, and real-time control of, for example, ramp metering and traffic signals, reductions in traffic demand can occur shortly after an incident has been detected and verified.

Figure 8 shows the variation of reduced total delay with the time, t_2, at which arrival rate drops. The curve shown is for the case where normal capacity, c_{max}, is 4400 veh/h; incident capacity, c_i, is 1600 veh/h; and at time t_2 the arrival rate drops by 10% from a value of 3500 veh/h to 3150 veh/h. The rate of decrease in total delay is higher for low values of t_2. From this it can be concluded that reducing arrival rate upstream of an incident shortly after it occurs will have substantial benefits in terms of reducing the resultant delay. The magnitude of the delay saving falls off quickly as the time to achieve demand reduction increases

Conclusions

A methodology has been presented in this paper for estimating the impact of an incident, in terms of queue dissipation time and delay, under assumed variations in arrival rate and roadway capacity. Using discrete forms of derived expressions in a spreadsheet program allows examination of different incident scenarios. The methodology lends itself well to applications in developing specific incident management plans as well as assisting in strategic level decision making

The duration of an incident and its effect on capacity of a roadway are strong determinants of the impact the incident will have on upstream traffic. Under constant flow conditions, where an incident causes a reduction in capacity below the arrival rate, the time for the resultant queue to dissipate is proportional to the duration of the incident. A term for incident severity under these conditions has been proposed equal to the ratio of the reduction in capacity due to the incident and the excess capacity once the incident has been cleared. The total delay due to an incident occurring under constant flow conditions is proportional to the square of the incident duration

Extending the analysis to the situation of varying arrival rate during a typical weekday, it was shown that the impact of an incident varies considerably with its time of occurrence Incidents occurring immediately prior to peak flow periods will lead to delays much higher than at other times. The effect of incident duration on queue dissipation time and total delay was illustrated for incidents occurring at different times of day. Some examples were given of savings in total delay due to reduced incident duration resulting from faster incident response times.

Finally, an example was provided of possible reductions in the impact of incidents resulting from measures that reduce the arrival rate upstream of an incident site It was concluded that large savings in delay could be obtained if arrival rates are reduced shortly after an incident occurs - the magnitude of the savings decreases quickly as time to achieve flow reduction increases.

References

Federal Highway Administration (1986) *Quantification of Urban Freeway Congestion and Analysis of Remedial Measures* Washington DC

Goolsby, M E (1971) Influence of Incidents on Freeway Quality of Service *Highway Research Record 349*, 41-46

Hunt, J G and Yousif, S Y (1994) Traffic Capacity at Motorway Roadworks - Effects of Layout, Incidents and Driver Behaviour *Proceedings of the Second International Symposium on Highway Capacity* Sydney: ARRB Ltd

Morales, J M (1986) Analytical Procedures for Estimating Freeway Traffic Congestion *Public Roads 50 (2)*, 55-61

Transportation Research Board (1990) *Freeway Incident Management* NCHRP Report 156 Washington DC

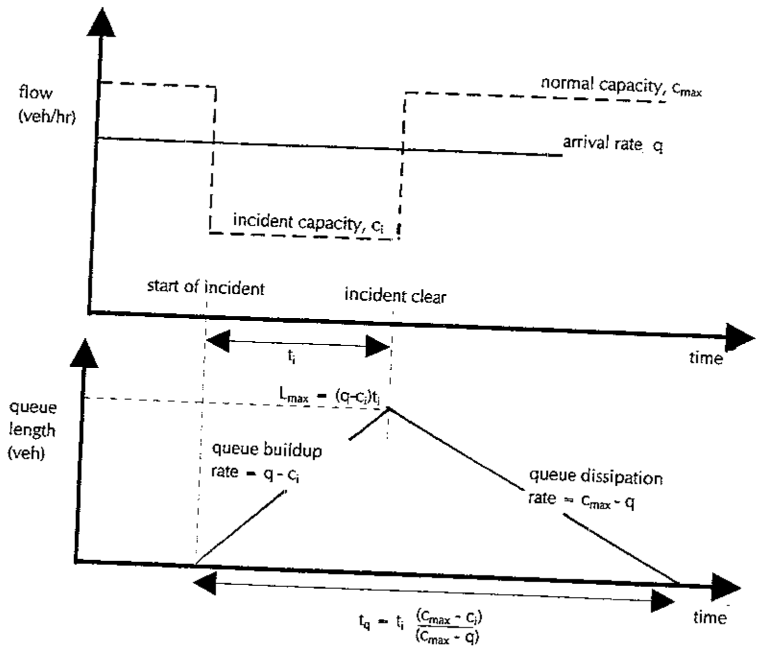

Figure 1 **Illustration of incident effect: constant arrival rate and incident capacity**

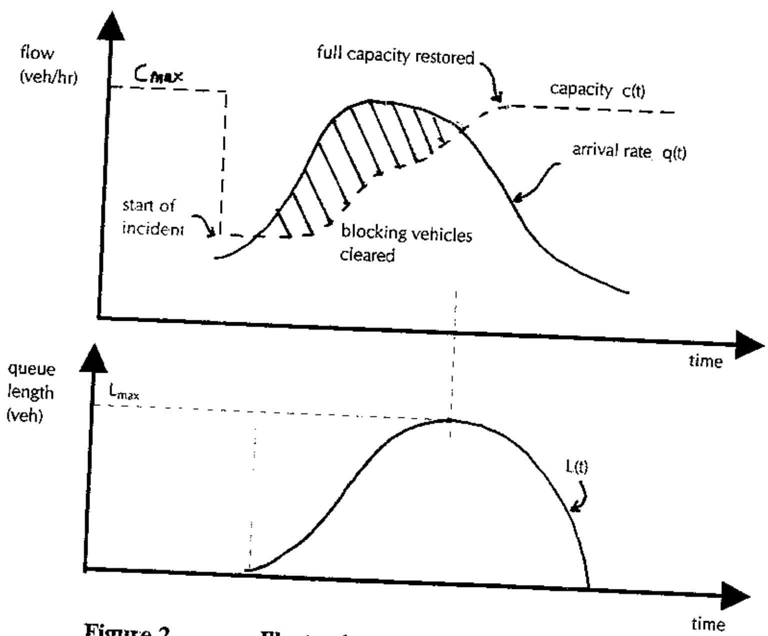

Figure 2 **Illustration of incident effect: varying 'arrival rate and capacity**

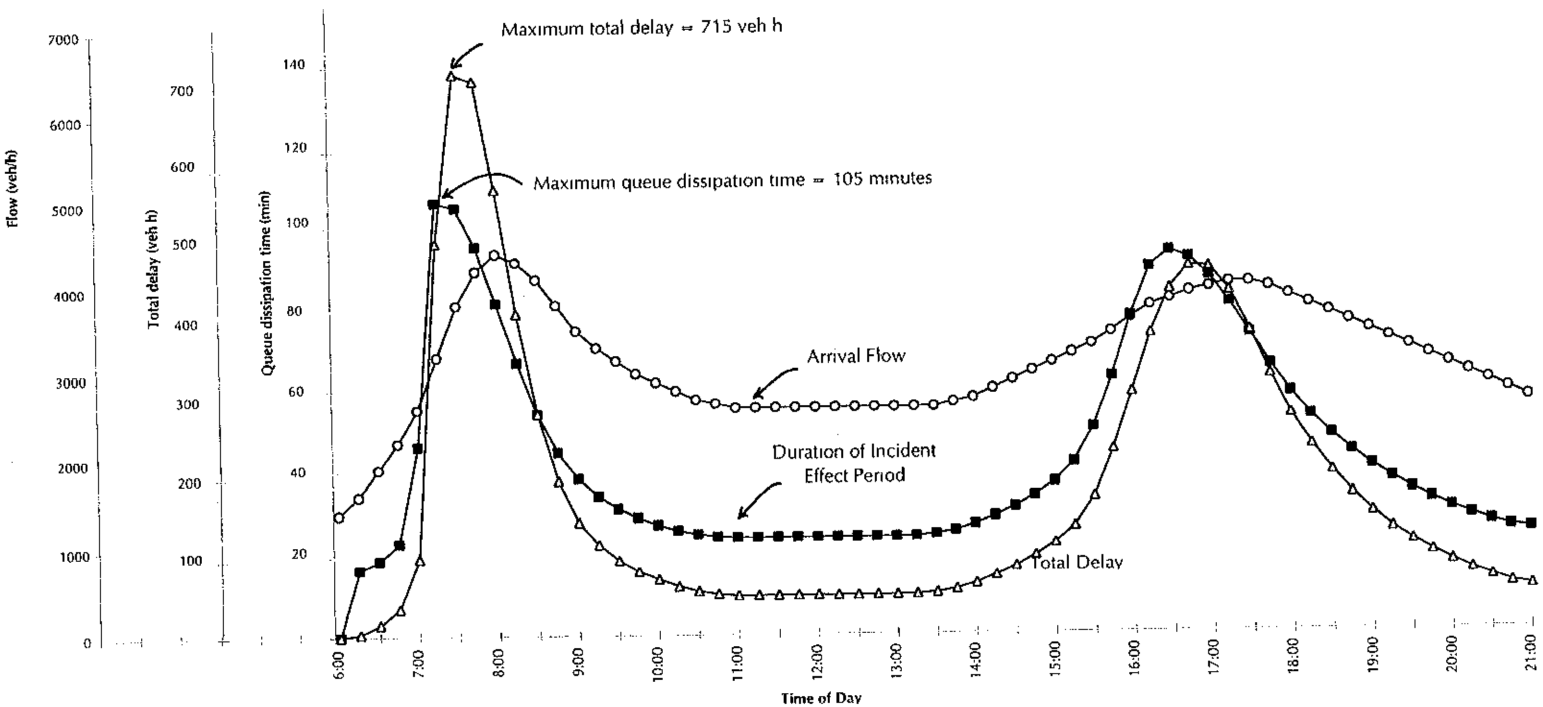

Figure 3 Variation of incident impact with time of day: incident duration = 15 minutes

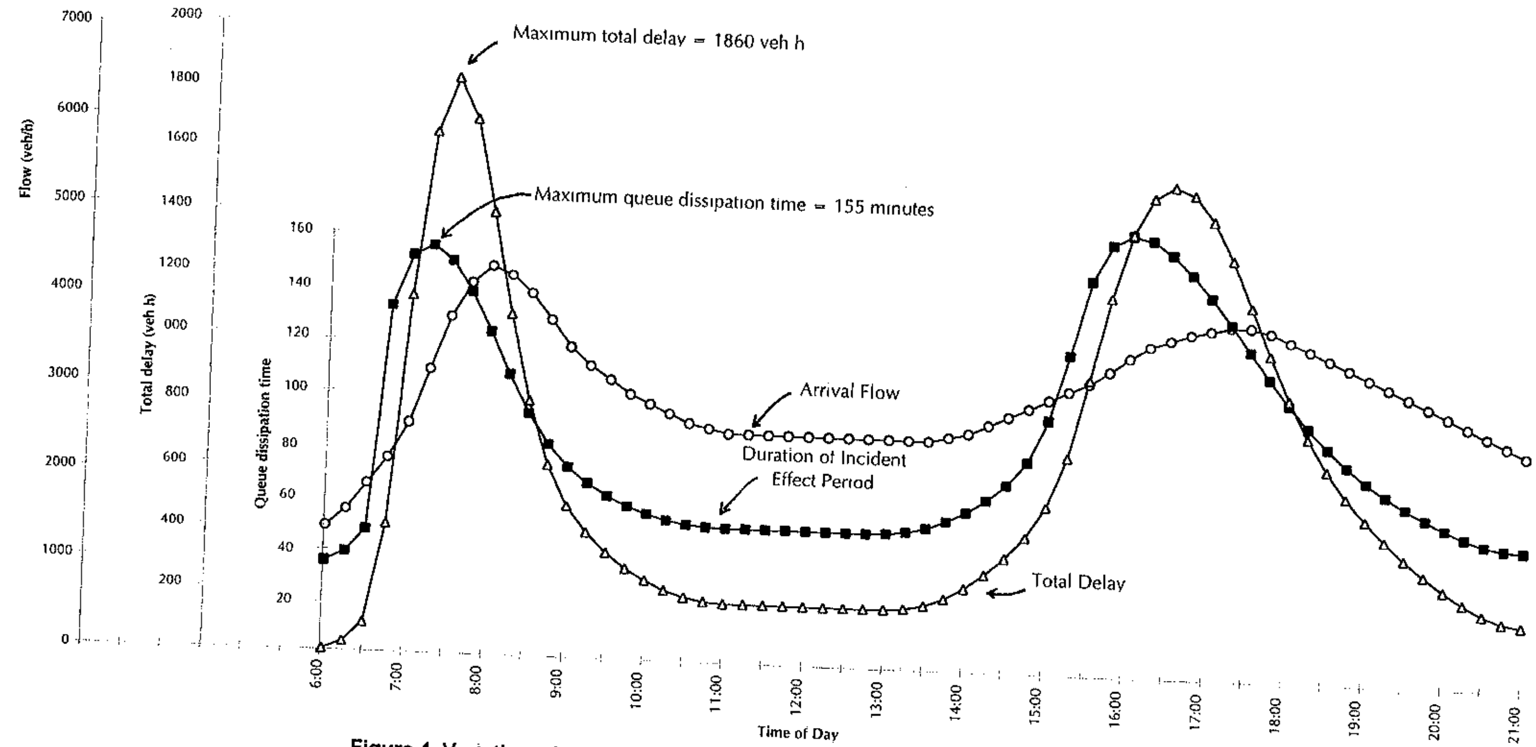

Figure 4 Variation of incident impact with time of day: incident duration = 30 minutes

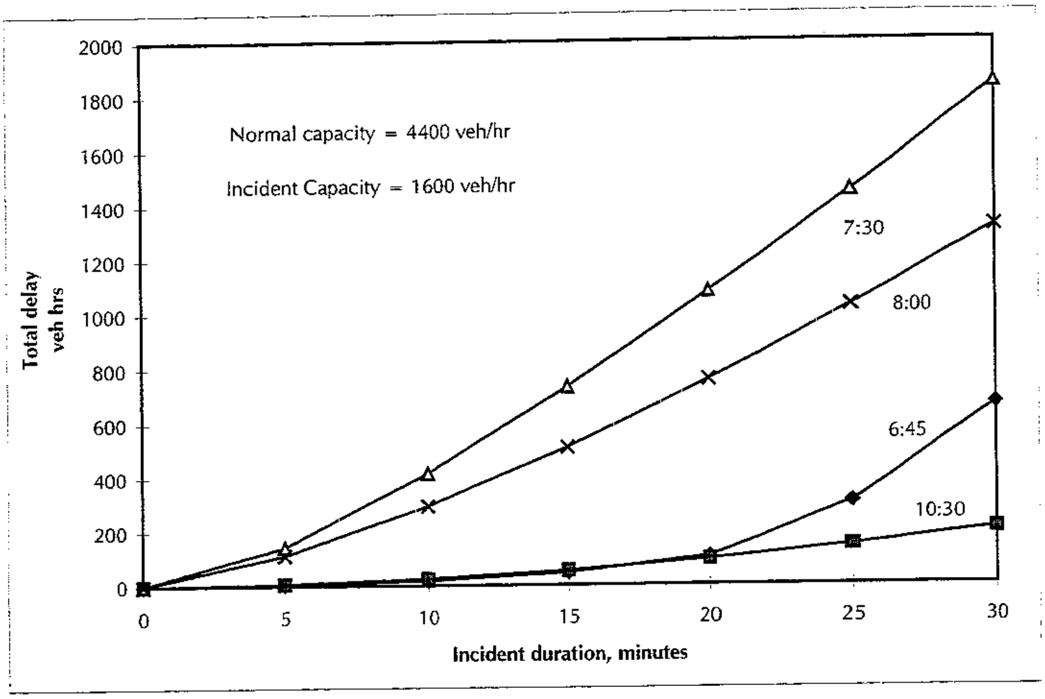

Figure 5 Variation of total delay with incident duration and occurence time

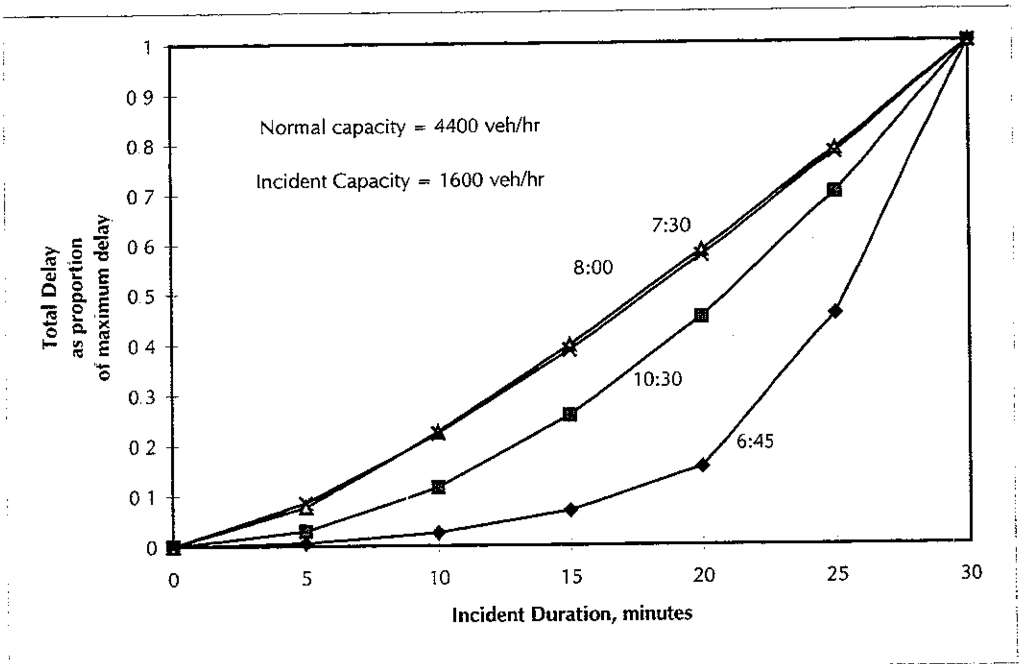

Figure 6 Variation of total delay with incident duration and occurence time: total delay as proportion of 30 minute incident total delay

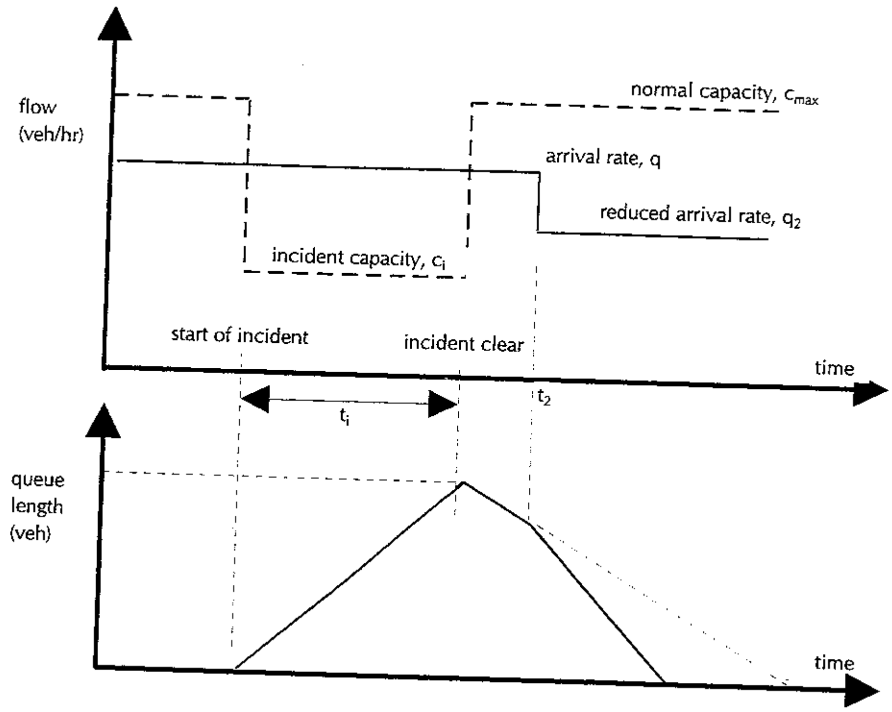

Figure 7 **Illustration of reduced incident effect due to decreased arrival rate**

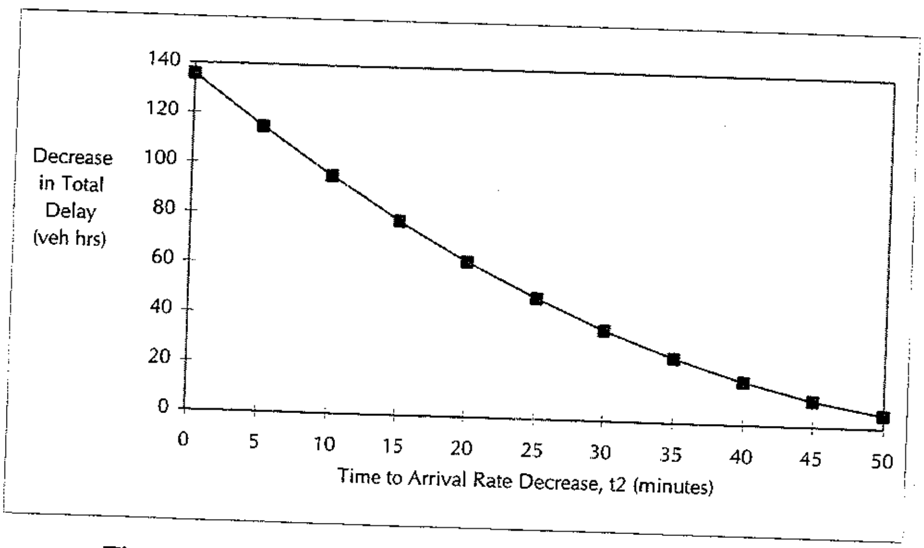

Figure 8 **Decrease in total delay versus time to decreased arrival rate**

21st Australasian Transport Research Forum
Adelaide, September 1997

Behavioural Approaches to Travel Demand Management

Bruce James
Travel Demand Branch
WA Department of Transport

Gary John
Travel Demand Branch
WA Department of Transport

Abstract:

The need to reduce the growth of car driver only trips in urban areas has been identified in transport strategies for Perth, Melbourne, Sydney and Brisbane. One of the available tools to mitigate this growth is travel demand management Travel demand management requires behaviour change interventions common in public policy areas such as health This paper explores the relevance of various social change processes to reduce car driver only trips Social change processes reviewed include planned social change, social marketing, organisational behaviour and community involvement. The lessons learnt from these reviews are then applied to two travel demand management initiatives being trialed in Perth, Western Australia

Contact Author:

Bruce James
WA Department of Transport
6th Floor
19 Pier St
Perth WA 6000

telephone: (08) 9231 2600 fax: (08) 9231 2612
email: bjames@dot.wa gov.au