AP-R318/08

AUSTROADS RESEARCH REPORT

Freeway Traffic Flow under Congested Conditions: Literature Review

Austroads
promoting improved road transport outcomes

Freeway Traffic Flow under Congested Conditions: Literature Review

Freeway Traffic Flow under Congested Conditions: Literature Review

First Published 2008

Freeway Traffic Flow under Congested Conditions: Literature Review

ISBN 978-1-921329-40-1

Austroads Project No. NS1204

Austroads Publication No. AP–R318/08

Project Manager

Ted Vincent, Asha Pas, Raj Ramalingham

Prepared by

James Luk, Clarissa Han, Charles Karl

Published by Austroads Incorporated
Level 9, Robell House
287 Elizabeth Street
Sydney NSW 2000 Australia
Phone: +61 2 9264 7088
Fax: +61 2 9264 1657
Email: austroads@austroads.com.au
www.austroads.com.au

Freeway Traffic Flow under Congested Conditions: Literature Review

Austroads

Sydney 2008

Austroads profile

Austroads' purpose is to contribute to improved Australian and New Zealand transport outcomes by:

- providing expert advice to SCOT and ATC on road and road transport issues
- facilitating collaboration between road agencies
- promoting harmonisation, consistency and uniformity in road and related operations
- undertaking strategic research on behalf of road agencies and communicating outcomes
- promoting improved and consistent practice by road agencies.

Austroads membership

Austroads membership comprises the six state and two territory road transport and traffic authorities, the Australian Department of Infrastructure, Transport, Regional Development and Local Government, the Australian Local Government Association, and Transit New Zealand. It is governed by a council consisting of the chief executive officer (or an alternative senior executive officer) of each of its eleven member organisations:

- Roads and Traffic Authority New South Wales
- Roads Corporation Victoria
- Department of Main Roads Queensland
- Main Roads Western Australia
- Department for Transport, Energy and Infrastructure South Australia
- Department of Infrastructure, Energy and Resources Tasmania
- Department of Planning and Infrastructure Northern Territory
- Department of Territory and Municipal Services Australian Capital Territory
- Department of Infrastructure, Transport, Regional Development and Local Government
- Australian Local Government Association
- Transit New Zealand

The success of Austroads is derived from the collaboration of member organisations and others in the road industry. It aims to be the Australasian leader in providing high quality information, advice and fostering research in the road sector.

SUMMARY

Austroads Project NS1204 (*Optimising Freeway Traffic Flow under Congested Conditions*) aims to study the freeway flow breakdown process and determine how to maximise freeway utilisation during congested conditions. The motivation for this study came from the observation that most capital cities in Australia and New Zealand are experiencing some levels of congestion on their freeway networks. Sydney, Melbourne and Brisbane are experiencing congestion between three to eight hours per day. The phenomenon of flow breakdowns on freeways, however, is still not well understood.

The possible reasons for congestion and flow breakdowns on a freeway include:

- mainline freeway flow in excess of capacity
- uncontrolled access to the freeway from on-ramps
- inadequate road geometries such as sudden lane-drop, excessive upgrades, freeway lane merge and insufficient weaving capacity
- undisciplined driving behaviour
- lack of real-time driver information to encourage better use of alternative routes or lanes
- lack of an integrated approach that encourages better use of alternative transport modes.

This report represents the first output of Project NS1204 on literature reviews. Other tasks of the project include the selection of freeway study sites, compilation of freeway data and the analysis of the data towards better understanding of flow breakdowns and will be completed as the project progresses. The contents of this report are as follows:

- basic freeway flow theory and automatic freeway control tools (Section 2)
- models for the characterisation of congested freeway flow (Section 3)
- outline of future research tasks (Section 4)
- conclusions (Section 5).

A traffic system is a complex, time-dependent system. It consists of the driver, vehicle and road infrastructure, and the interactions amongst the three components. Driver behaviour is especially difficult to control. The complexity of a traffic system is not a major concern as long as a road facility is uncongested but, with increasing congestion, certainly needs to be well understood for efficient network operations.

There have been many attempts to develop mathematical models to characterise and understand traffic behaviour. For example, the Lighthill-Whitham (1955) kinematic wave model, despite its limitations, remains useful for understanding shock waves at road bottlenecks. Even though most models lack detail and reality, traffic control systems have long been employed to control arterial and freeway traffic macroscopic relationships.

In recent years, freeways in some overseas and Australian cities have been installed with detector stations that can provide reliable traffic data at a high resolution in time and space. It is now possible to obtain accurate empirical contours of flow, speed and occupancy for the analysis of flow breakdowns at bottlenecks.

The work of Kerner (2004) has been extensively reviewed in this report. Kerner developed a model from empirical data consisting of three phases or states: free-flow, synchronised flow and moving jam. From various combinations of these phases and their transitions, different flow patterns observed on a freeway can be reproduced or predicted. An accurate prediction model is the first step in the effective implementation of freeway control tools such as ramp-metering and speed limit signs.

This project has reviewed other models of flow breakdowns including the stochastic models and models with more than three states. A general finding is that the Kerner approach has gone through a significant amount of research over ten years. Apart from providing a good platform for the analysis of flow breakdowns, it has been successfully deployed for real-time control. This project can benefit from the Kerner three-phase concept with additional input from other sources.

The project team has already scanned through the flow, speed and occupancy data from the Monash Freeway, Westgate Freeway and Western Ring Road in Melbourne to decide what freeway sites should be selected for analysis. The recommendations for subsequent stages of this project are as follows:

- A freeway route is preferable to a few isolated sites from one or several freeways so that freeway flow can be studied in a network context – what happens upstream affects downstream flow. The demand for that corridor will also be similar for a particular time of day if all sites are from the same freeway.
- Analytical modelling processes are unlikely to explain complex driver behaviour before, during and after flow breakdowns. An empirical spatio-temporal approach is recommended for congested freeway flow analysis. The two-phase approach of HCM 2000 is too simplistic and the Kerner (2004) three-phase approach appears appropriate for the identification of flow breakdowns and will be adopted in this project.
- The software MATLAB would provide a useful platform for data analysis and the development of rules and artificial intelligence in general for the identification of flow breakdowns and tracking of shock waves.

CONTENTS

TABLES

FIGURES

1 INTRODUCTION

The purpose of Austroads Project NS1204 (*Optimising Freeway Traffic Flow under Congested Conditions*) is to determine how best to extract maximum utilisation and efficiency from existing freeway systems during peak periods and severe congested conditions by:

- identifying factors that contribute to flow breakdown
- developing a technique for determining the location of critical bottlenecks
- identifying the speed of propagation of shockwaves and the extent of their influence
- identifying any influence of current freeway design features on the formation of bottlenecks.

The motivation for this study came from the observation that most capital cities in Australia and New Zealand are experiencing some levels of congestion on their freeway networks. Sydney, Melbourne and Brisbane are experiencing congestion between three to eight hours per day. The phenomenon of flow breakdowns on freeways, however, is still not well understood.

The possible reasons for congestion and flow breakdowns on a freeway include:

- mainline freeway flow in excess of freeway capacity (uncontrolled access)
- uncontrolled access to the freeway from on-ramps
- inadequate road geometries – pre-1960 and earlier freeway designs may not be suitable for congested flow conditions at present , e.g. sudden lane-drop, excessive upgrade, freeway merge lane and insufficient weaving capacity
- undisciplined driving behaviour with frequent and unnecessary lane-changing and speed changes
- lack of real-time driver information to encourage better use of alternative routes in a corridor and alternative lanes on a carriageway
- lack of an integrated approach that encourages better use of all transport modes – buses, rail, trams and cycling apart from car use.

This project will not be able to address all of these issues and will focus on the *characterisation* of freeway flow breakdowns under a range of traffic and geometrical conditions. A clear understanding of the flow breakdown process will assist in identifying the correct parameters and relationships to be used in the tools for automatic freeway control, e.g. variable speed limit signs (VSLS) and ramp metering. The use of these tools is included in the literature review in Section 2.3 but the operation and development of these tools will be addressed in other Austroads and road authority projects.

It is also important to note that freeway flow breakdown does not necessarily imply a catastrophic failure of traffic flow with vehicles moving at slow speeds. Similarly, a loss of maximum flow by itself may not necessarily imply flow breakdowns. As the project progresses and as a better characterisation of flow breakdowns is available, then breakdowns can be quantified objectively. It is likely that the definition of a flow breakdown would involve a substantial loss of maximum flow over a measurement time period and over some distance along a freeway segment.

This report provides a literature review and represents the first deliverable of the project. Its contents are as follows:

- basic freeway flow theory and automatic freeway control tools (Section 2)
- models for the characterisation of congested freeway flow (Section 3)

- outline of research tasks (Section 4)
- conclusions (Section 5).

VicRoads has an extensive network for vehicle detection in Melbourne freeways and can provide comprehensive freeway traffic data. This project will therefore make use of data from VicRoads.

The project work will also complement two other Austroads projects:

- NS1375 (*Freeway Design Parameters for Fully Managed Operations*), which has a focus on improving freeway geometric designs to minimise flow breakdowns
- NS1378 (*Best Practice for Variable Speed Limits*), which develops best practices in the use of VSLS for speed management.

2 BASIC FREEWAY TRAFFIC FLOW THEORY AND CONTROL TOOLS

A freeway is generally known as an *uninterrupted* facility because traffic on the mainline is not interrupted by a control device such as a traffic signal. Traffic flow on a freeway segment is therefore relatively easier to model analytically, especially in free-flow or uncongested conditions. This section provides the basic analytical framework for freeway traffic analysis and also briefly reviews the tools currently used for automatic freeway control. It provides the context for subsequent sections on reviews of the characteristics of flow breakdowns.

2.1 Basic Traffic Flow Model

A traffic system is characterised by three traffic parameters: *flow* (q in veh/h), *speed* (v in km/h) and *density* (k in veh/km). Flow can be measured with point sensors such as the inductive loop sensors. A pair of sensors at a known distance apart (e.g. 5 m) is necessary to determine speed. The speed so measured is called a spot speed because it is the speed at a specific location in space.

Density in vehicles per unit distance is a spatial concept and is difficult to measure using a point sensor and *occupancy* is used as a proxy density value. Occupancy is the percentage of time in a measurement period (e.g. 1 min) that a sensor is occupied with vehicles. As congestion increases, a sensor is occupied more often and both density and occupancy increases. The occupancy becomes 100% for a particular minute when a vehicle becomes stationary on top of the sensor for that minute. For a square loop of 2 m x 2 m, an occupancy in excess of 15% indicates slow moving traffic.

On a freeway segment a relationship exists amongst the three parameters as follows:

$$\text{Flow } q \text{ (veh/h)} = \text{Speed } v \text{ (km/h)} \times \text{Density } k \text{ (veh/km)}$$

or

$$q = v\,k \qquad 1$$

v is the space mean speed, or the average speed of all vehicles in a road segment where the density is determined from the number of vehicles in that segment.

Empirically, it has long been observed that speed bears a reasonable linear relationship with density before over saturation occurs, i.e. speed drops as density increases. A linear relationship between speed and density can take the following form (often known as the Greenshields model (see, e.g. May 1990):

$$v = v_f - (v_f/k_j)\,k \qquad 2$$

where v_f is the *free-flow speed* and k_j is the *jam density* so that speed is zero at $k = k_j$.

Substituting $k = q / v$ from Equation 1 into Equation 2, flow becomes:

$$v = v_f - (v_f/k_j)\,(q/v)$$

Rearranging

$$q = k_j\,(v - v^2 / v_f) \qquad 3$$

This is a parabolic relationship and there are two values of speed for each value of flow.

The critical speed can be obtained by differentiating Equation 3 and equating $dq / dv = 0$:

$$dq / dv = k_j [1 - 2v / v_f] = 0$$

Hence, the critical speed is given by:

$$v_c = v_f / 2 \qquad 4$$

Similarly, the critical density k_j can be shown to be equal to $k_j/2$. The critical flow when traffic flow starts to break down is given by Equation 1:

$$q_c = v_c k_c$$

$$= v_f k_j / 4 \qquad 5$$

with a free-flow speed of 100 km/h and a jam density of 100 veh/km (e.g. a stationary queue of passenger cars with a jam spacing of 10 m/veh), the critical flow becomes:

$$q_c = 100 \times 100 / 4$$

$$= 2500 \text{ veh/h}$$

Figure 2.1 illustrates what are commonly known as the three fundamental diagrams relating speed, flow and density and can found in most textbooks on classical traffic flow theory, e.g. May (1990) and Taylor, Bonsall & Young (2000). The relationships in Figure 2.1 have been useful for the basic understanding of traffic flow and control, especially for uninterrupted facilities such as freeways.

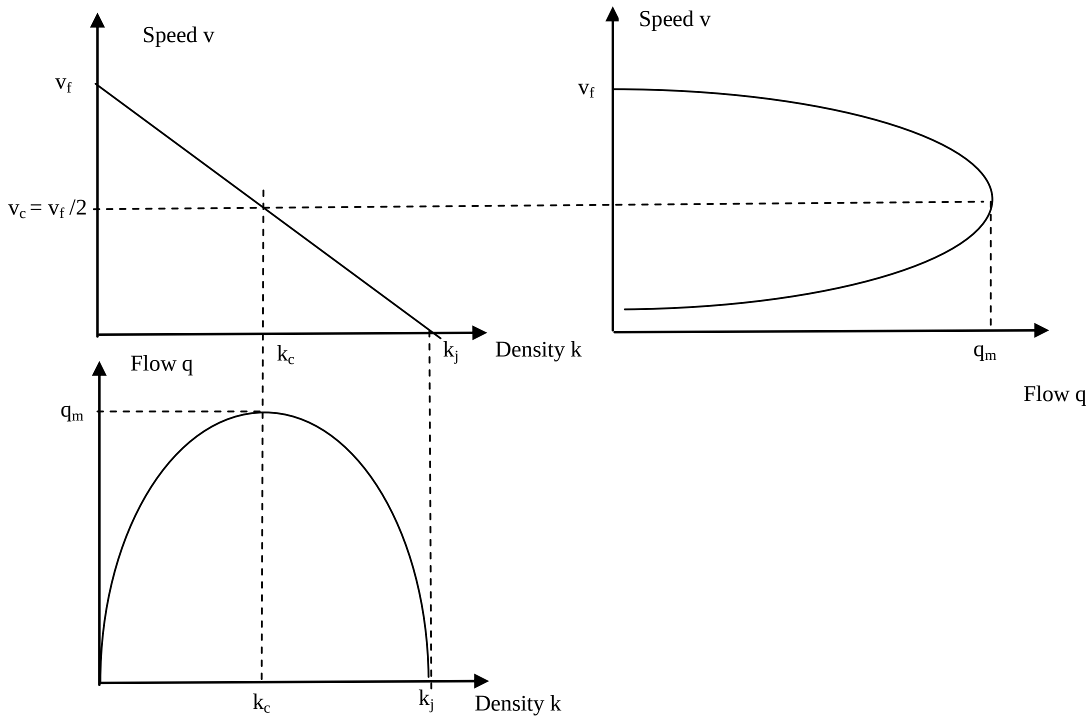


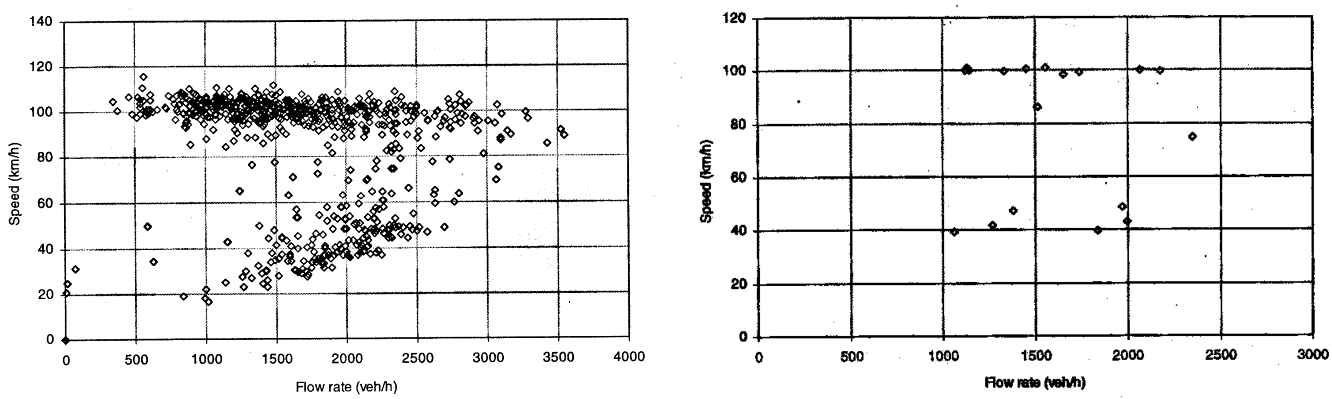
Figure 2.1: Fundamental diagrams assuming linear speed-density relationship

Traffic science is a relatively new discipline and has become popular amongst academics and researchers only after the Second World War and with the advance of computers. The collection and analysis of traffic data was often quite limited. A measurement time slice of 15 min was commonly adopted but such a long aggregate time period often masks out important traffic phenomena.

Traffic data analysis and applications now employ time slices of 20 s or even smaller with data collection carried out on 24-hour 7-day basis. This level of detail has enabled the development of intelligent transport systems (ITS) amongst road agencies in Australia. These include incident detection systems in the 1980s (Luk & Sin 1992), the calculation of travel times in Drive Time in the 1990s (Hearn et al. 1996) and more recently the monitoring of congestion in real-time promoted as part of a revised National Performance Indicators Program (Troutbeck, Su & Luk 2007).

Figure 2.2 illustrates the difference between a speed-flow curve based on data from 20 s and 15 min time slices from a pair of loop detectors in the middle lane of an inbound freeway carriageway in Melbourne (Akcelik, Roper & Besley 1999). Note the loss of information due to aggregation into 15 min data as mentioned earlier. The maximum flow or capacity using a 20 s time slice was 3,200 veh/h per lane whereas it became 2,300 veh/h per lane at a 15 min time slice. The study further identified that the occupancy at maximum flow was 25% at a 5 min time slice.

The test site in Akcelik, Roper and Besley (1999) was purposely chosen to be away from any bottleneck and their data is therefore unsuitable for analysing flow breakdowns in the context of this project. The intent of their work was also to develop speed-flow relationships in congested and uncongested regimes largely for planning purposes. Further, flow breakdown analysis will require data from more than one single detector at a bottleneck (Banks 1990, Hall, Hurdle & Banks 1992) and preferably from a network context (i.e. a whole freeway route) to identify shock waves (Sections 2 and 3 of this report).



Source: Akcelik, Roper & Besley (1999)

Figure 2.2: Speed-flow data measured at time slices of 20 s and 15 min

The concept of capacity is critical for the understanding of flow breakdowns. Capacity is unlikely to be a fixed number but a variable from a stochastic process that depends on traffic dynamics, traffic composition, weather conditions and geometric designs. It is important to note again the influence of the size of time slices on capacity. A 20 s slice shows a much higher capacity value than a 15 min slice as expected. A high capacity value of 3200 veh/h per lane raises the question of what should be a suitable size for the understanding of flow breakdowns. Traffic flow at a 20 s slice is too random to decipher a pattern and overseas work has employed a 1 min slice. These issues will be addressed as the project progresses (see also Section 3.3).

2.2 Kinematic Wave Model

An extension of the fundamental relationships is to consider speed, flow and density as functions of time (t) and space (x), and they are not independent parameters. For example, flow is a function of density k, which is a function of time t. A model that considers the traffic process in time and space is the *kinematic* wave model of Lighthill and Whitham (1955), which is more suitable for high density conditions and therefore has its place in analysing flow breakdowns.

The kinematic model assumes that high density traffic will behave like a continuous fluid (hence also called a *continuum* model). Consider the flow in and out of a short length of road Δx. The condition of *continuity* requires that if the density of vehicles has increased it must have been due to a difference in the amounts flowing in at one end and out at the other, or:

$$\frac{\partial k}{\partial t} + \frac{\partial q}{\partial x} = 0 \qquad (6)$$

where q is the flow (veh/h)

k is the density (veh/km)

x is distance (km)

t is time (h) to travel a distance of x km.

With q as a function of density k, Lighthill and Whitham developed Equation 6 further into the LW model as follows:

$$\frac{\partial k}{\partial t} + \frac{\partial q}{\partial k}\frac{\partial k}{\partial x} = 0 \qquad (7)$$

Define below a *wave speed* U that represents the speed of waves carrying continuous changes of vehicle flow in a traffic stream:

$$U = \frac{\partial q}{\partial k}$$

then

$$\frac{\partial k}{\partial t} + U\frac{\partial k}{\partial x} = 0 \qquad (8)$$

Because q = v k from Equation 1, the wave speed:

$$U = \frac{\partial (vk)}{\partial k}$$

$$= v + k\frac{\partial v}{\partial k} \qquad (9)$$

Because speed decreases with density, $\frac{\partial v}{\partial k}$ is always negative (Figure 2.1) and the wave speed U is therefore always less than the space mean speed v.

The relationship between space mean speed (v) and wave speed (U) are illustrated in the flow-density diagram in Figure 2.3, which also shows the shock wave speed (U_{SW}). The following observations can be made (Wohl & Martin 1967):

- At low densities when vehicle-to-vehicle interactions are minimal, $\frac{\partial v}{\partial k}$ is almost zero and the wave speed is similar to the space mean speed. The wave moves forward relative to the road.
- At the maximum flow and critical density, the wave is stationary. At densities higher than the critical density (k_c), the wave moves backward relative to the road.
- The wave speed changes with density according to Equation 9 and a traffic stream can have different densities at different sections of a freeway. A section of light traffic could follow a section of high density due to a decrease in lanes, an accident or on-ramp traffic. The wave in the low density traffic moves forward (relative to the freeway) at a speed faster than the wave in the high density traffic.
- When the two waves meet, a new wave called a *shock wave* will be formed. All three waves move forward for the situation shown in Figure 2.3. The shock wave speed U_{SW} is given by:

$$U_{SW} = \frac{q_2 - q_1}{k_2 - k_1} \quad 10$$

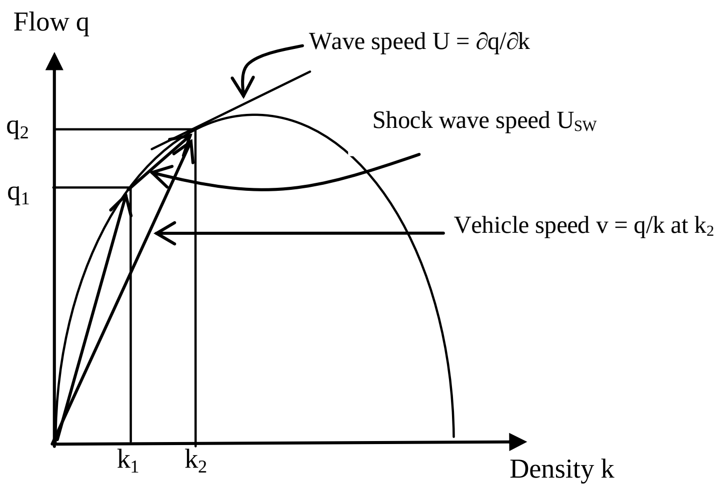


Figure 2.3: The relationship between vehicle speed, wave speed and shock wave speed

Figure 2.4 illustrates the case for a negative shock wave speed due to capacity decrease at a bottleneck (e.g. lane drop) on a freeway. Two fundamental diagrams are required. The inner diagram represents the characteristics of the bottleneck with capacity q_b less than the approach section. The approach flow q_a is larger than q_b and a complex queuing situation occurs at the entry to the bottleneck.

The density at the bottleneck entry suddenly increases from the density at C to the density at E in Figure 2.4. The wave speed at E is negative with respect to the freeway and will be reflected from the bottleneck back to the approach section. The reflected wave will meet the oncoming wave corresponding to the slope at C. A shock wave of negative speed relative to the freeway is formed. The effect of the bottleneck will be reflected along the entire approach section if the arrival flow remains constant, with a consequent loss of maintaining capacity flow (q_m).

Edie and Foote (1958) reported how shock waves were generated at an upgrade leading to the Holland Tunnel exit in New York. The shock waves propagated backward towards the tunnel entry with inefficient traffic flow. The solution was to control the entry of vehicles into the tunnel so that the entry flow did not exceed the capacity of the bottleneck section. The vehicles entered in short platoons of about 40 veh every 2 min with a 10 s gap between platoons.

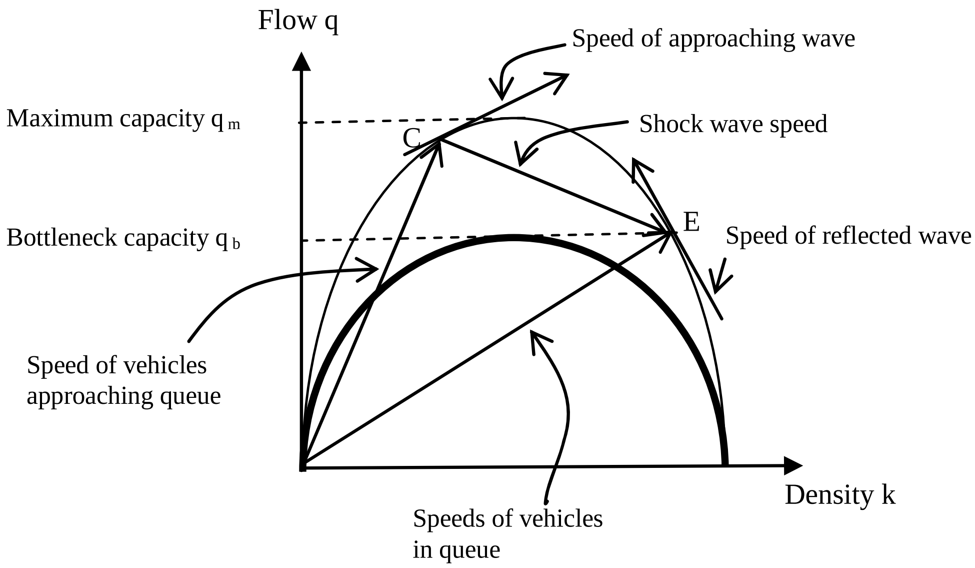


Figure 2.4: Fundamental diagrams of a bottleneck section and the approach section

The kinematic model can be solved using the finite difference (or finite element) method and has continued to be an interesting area of research (see, e.g. Leo & Pretty 1992, Michalopoulos 1988, Ngoduy, Hoogendoorn & Van Zuylen 2006, Papageorgiou 1983, Payne 1971). At the University of Queensland, Leo and Pretty were able to model the propagation of congested density upstream in a freeway lane drop situation and platoon movements in a pair of coordinated signals at very small, discrete levels of time (0.5 to 1 s) and space (about 15 m) but have not further pursued their research since the early 1990s.

The LW model is a first order model with limitations such as (Papageorgiou 1998):

- Assume that vehicle speeds can change instantaneously, i.e. large values of acceleration and deceleration rates are assumed possible at a bottleneck (E in Figure 2.4).
- Predict that the tail-end of a platoon on arterial roads will speed up to catch up with the main platoon when it is more common to observe a dispersed rear-end.
- Assume that outflow (q_b) at a freeway bottleneck is best achieved with some congestion at the bottleneck entry. This is equivalent to assuming that the outflow cannot be increased by avoiding mainline congestion, i.e. no control. The reality is that some control of a bottleneck (if possible) can improve throughput.

Second order LW models have been proposed (Daganzo 2006, Papageorgiou, Blosseville & Haj-Salem 1990, Payne 1971, Schonhof and Helbing 2007) to overcome these limitations. Kinematic models can contribute towards the understanding of freeway flow breakdowns and will be revisited as the project progresses.

2.3 Automatic Freeway Control Tools

Automatic freeway control tools aim to improve the efficiency or safety of the mainline traffic movement. They represent the more commonly employed control tools on freeways. They have been in operation in overseas and Australian cities. Control tools relevant to the identification and management of flow breakdowns are as follows:

- automatic incident detection (AID) system
- ramp metering
- variable speed limit signs (VSLSs)
- driver information systems (DISs)
- traffic lane management systems.

2.3.1 Automatic Incident Detection (AID)

A traffic event that results in the loss of freeway flow efficiency can be due to recurrent congestion or a non-recurrent incident. A freeway AID system generally aims to detect non-recurrent events (flow breakdowns) *after* they have occurred. The traffic management centre (TMC) then initiates incident management plans to resolve the incident. Ramp metering and VSLS generally aim to prevent flow breakdowns *before* they occur.

A freeway AID system identifies unusual changes in speed, flow and occupancy due to planned and unplanned events such as vehicle breakdowns and roadworks. An operator at the TMC receives a warning automatically when an incident is identified. The infrastructure of lane-by-lane, dual loop detector stations and the incident parameters of 500 m spacing and 20 s time slice was introduced initially on the South Eastern Arterial in Melbourne (now part of the CityLink) subsequent to a major incident in the 1980s. As mentioned, the infrastructure has been most useful for the development of incident detection systems, driver information systems, network performance monitoring, speed limit signs and ramp metering.

The characterisation of an incident and hence its detection in real time (Luk & Sin 1992) was based on the *time series analysis* of speed, flow and occupancy, and by comparison of these parameters between adjacent detector stations and between lanes of the same station.

Dia and Rose (1997) continued the research by employing an artificial neural network (ANN) to calibrate more accurately the detection algorithm. They collected traffic data associated with 50 incidents identified on Melbourne's freeways. The use of ANN for incident detection has continued at the Nanyang Technological University in Singapore (Mak & Fan 2004) and other academic institutions.

The characterisation of a flow breakdown is closely related to the characterisation of an incident. Some of the related issues are:

- A freeway incident is an unplanned event due to, e.g. road crashes or vehicle breakdowns on a freeway.

- Flow breakdowns could be due to unplanned events, planned events (e.g. road works), over saturation, or a range of reasons mentioned in Section 1 (merge area, lane drop, upgrade, driver behaviour, etc.). The focus of this project is on recurrent congestion and the characterisation of flow breakdowns due to factors other than over saturation.
- The characterisation of flow breakdowns can be carried out offline (and will be an offline exercise for this project). Ultimately, this process will have to be undertaken in real time to facilitate the application of ramp metering and other freeway control tools. The research into real-time ANN (or other similar techniques) is continuing but the limitation of these artificial intelligence techniques as essentially offline tools must be recognised.

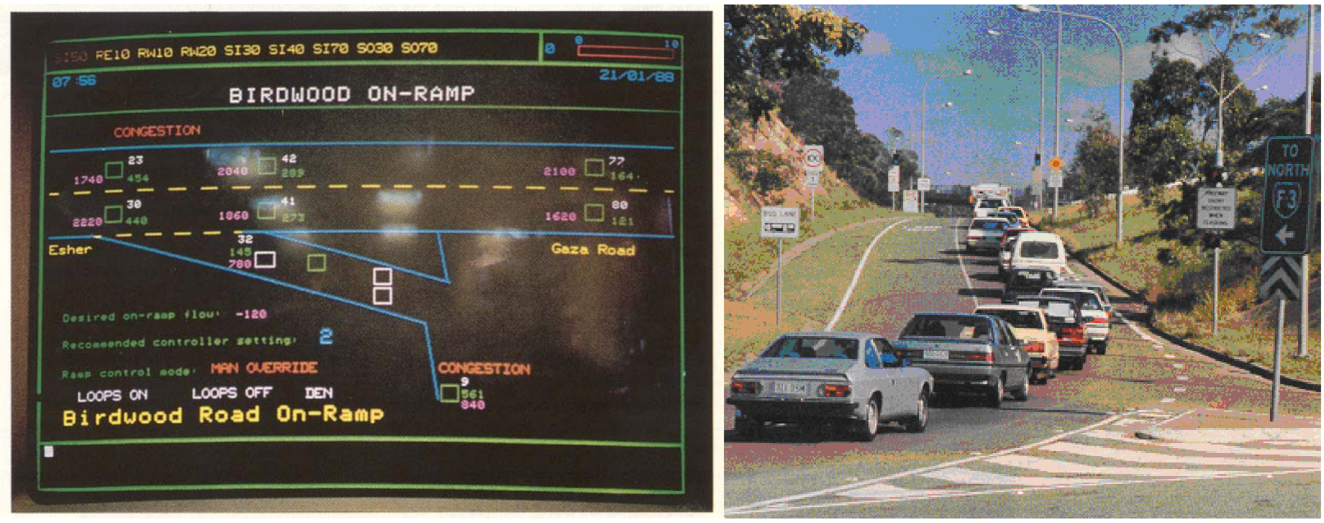
2.3.2 Ramp Metering

Ramp metering began in the US in the early 1970s (Payne, Thompson & Isaken 1973) and in the UK in the late 1980s (Owens & Schofield 1988). It reduces the on-ramp flow so that the mainline demand is maintained at or just below capacity and therefore reduces the occurrences of flow breakdowns. All motorists using the freeway will benefit because the ultimate capacity and inherent safety of the freeway is maintained (Lowrie 1996). Ramp metering also improves traffic conditions at the ramp merge point because it prevents the formation of small platoons at an on-ramp. On-ramp traffic can then merge easier and there is less interruption to the mainline flow and improved safety at the merge point.

A ramp metering signal was installed in Brisbane in 1987 to resolve a bottleneck on the South East Freeway downstream of the Birdwood Road on-ramp (Blinco 1988). The Brisbane system employed 1 min data using a single loop per lane in a detector station (speed was estimated by assuming an average vehicle length). The bottleneck capacity was first determined and was found to be 4800 veh/h but dropped to 3600 veh/h during congested periods. The bottleneck threshold could be changed by a TMC operator to account for wet weather conditions. The system then *predicted* the real-time demand at the bottleneck. Prediction was found necessary to allow time (1 to 2 min) for the on-ramp flow to be reduced. The permitted ramp flow was set equal to the bottleneck capacity less the mainline flow upstream of the on-ramp.

The system was successful in reducing the duration of congested flow at the bottleneck by 30%, with a corresponding 19% reduction in ramp flow. The benefits were however constrained by queuing from the on-ramp to the surface streets and the reality that the mainline freeway flow could not be gated to clear bottlenecks once flow breakdowns occur. The system effectively deferred the onset of congested flow.

Figure 2.5 illustrates this first generation of ramp metering in Australian cities. Plans are in hand to meter other adjacent ramps on this freeway in Brisbane in a coordinated manner.



Source: Blinco (1988)

Figure 2.5: The ramp metering system in Brisbane installed in 1987

Ramp metering is a continuing area of research in Europe (e.g. Kosmatopoulos et al. 2006, Papageorgiou, Hadj-Salem & Middelham 1997). An assessment of 20 on-ramps equipped with metering on a motorway in the Netherlands showed reductions in travel time ranging from 3 – 10% (Taale & Middelham 2000).

The Minnesota Department of Transportation switched off the entire ramp metering network in the Twin Cities area in St Paul for six weeks from October 16 to December 8, 2000 to determine the benefits of using ramp metering (which was a controversial issue in St Paul). Cambridge Systematics (2001) reported the following results:

- freeway flow decreased by 7% (14% in peak period)
- travel time increased by 22% (7% reduction in speeds)
- road crashes increased by 26%.

Apart from Brisbane, ramp metering has also recently received interest in Auckland (Auckland Motorways n.d.) and Melbourne (Gaffney 2007, Transurban 2006, VicGov 2006) and Perth because of increasing congestion on the freeways of these cities.

As mentioned, the aim of ramp metering is to maximise the mainline flow by controlling the amount of flow entering from the on-ramps without adding an unacceptable amount of delay to on-ramp traffic. The issues that need to be addressed include:

- For urban freeways, there may not be enough storage space on the entry ramps to store queuing vehicles.
- The equity issue of balancing the delay of ramp traffic and maximum throughput of the mainline flow must be recognised. This issue is not new and is similar to minimising side-street delay and maintaining good platoon progression on an arterial road.
- The best way to characterise flow breakdown in recurrent congestion and how to deploy a control tool such as ramp metering to delay or prevent a breakdown needs investigation.

- It is difficult to meter mainline traffic on a freeway in a manner similar to the Holland Tunnel in New York (Edie & Foote 1958; Section 2.2). Without such metering, any improvement in mainline flow will invariably attract more traffic. An alternative measure is freeway congestion pricing as used in Singapore (Luk 1999) when congestion tolls are varied to keep freeway speeds in the range 45 – 65 km/h, but congestion pricing has yet to be implemented as a demand management tool in Australia and New Zealand cities.

Ramp metering can be at an isolated local level and has been implemented on various on-ramps in Brisbane and Melbourne. As mentioned, VicRoads will soon be trialling also the benefit of coordinating a network of six entry ramps on the inbound direction of Monash Freeway. A coordinated metering scheme offers the opportunity of spreading entry flows more evenly over several on-ramps.

2.3.3 Variable Speed Limit Signs (VSLSs)

VSLSs have been in use for many years on motorways in the Netherlands (Jenezon, Klijnhout & Langelaar 1987) and more recently in London (Highways Agency n.d.), Melbourne, Sydney and Brisbane (Bean 2002, Herley & Lennie 2007). In Europe, these signs generally aim to reduce secondary road crashes by advising drivers upstream of a primary incident to slow down, especially in harsh weather conditions.

VSLSs have been used also for a range of purposes as follows:

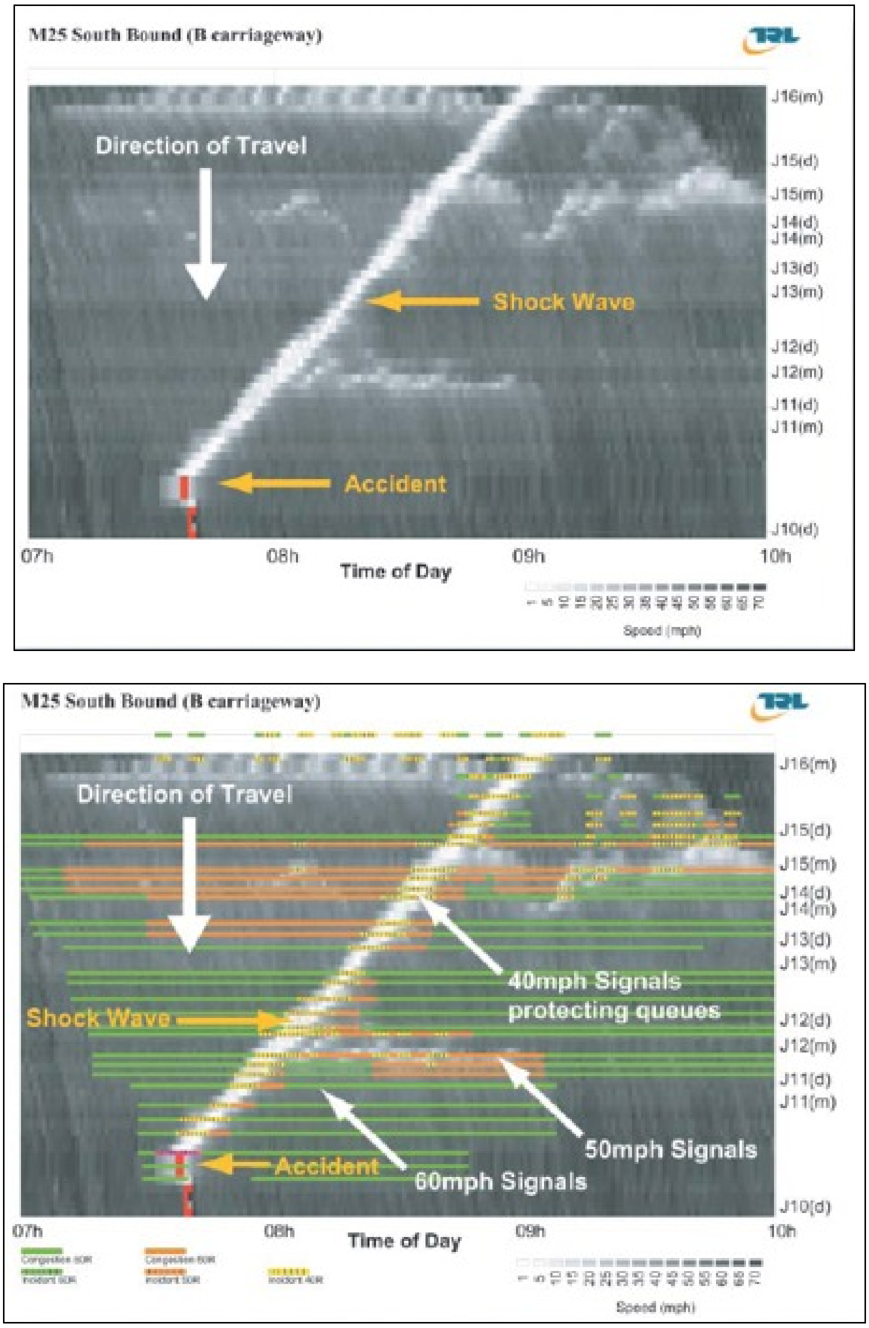
- reduce flow breakdowns
- reduce road crashes
- increase motorway efficiency
- provide general speed management during planned events such as road works or sports events.

Figure 2.6 illustrates the formation of a shock wave subsequent to an accident on the M25 in London and how lower speed limits (60 mph, 50 mph and 40 mph) were then implemented to protect the queues formed.

In the May 2007 National Variable Speed Limit Forum, Herley and Lennie (2007) reported the following statistics on Queensland motorways:

- up to eight hours of flow breakdowns
- minimum speeds in 100 km/h zones during peak hours as low as 15 km/h
- standard deviations of traffic speeds up to 25 km/h
- approximately 35% of road crashes in peak periods
- about 50% of peak period crashes being rear-end crashes
- about 20% of road crashes occurred during harsh weather conditions
- up to 50% of crashes in harsh weather conditions being rear-end crashes.

VSLSs are another useful freeway control tool to reduce flow breakdowns and road crashes and maintain efficient traffic flow. This project aims to gain better understanding of the flow breakdown process and hence better use of this tool.



Source: Highways Agency (n.d.)

Figure 2.6: Variable speed limits to avoid flow breakdowns on M25 in London

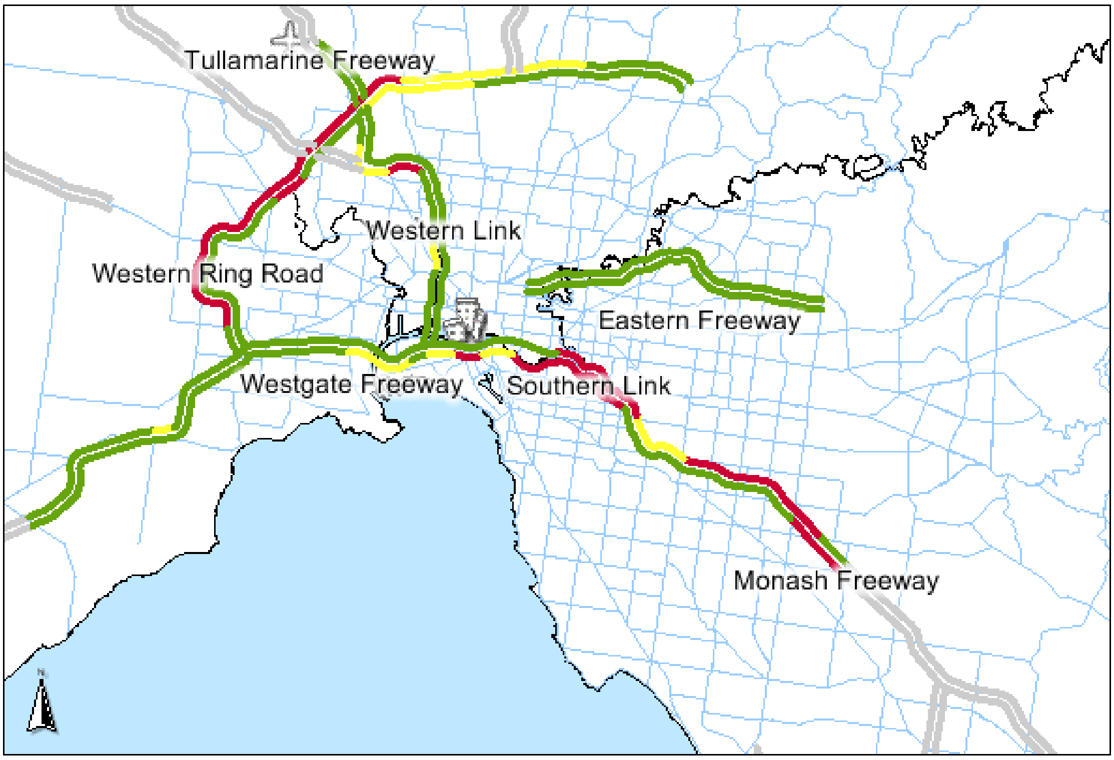
2.3.4 Driver Information Systems (DISs)

DISs can take many forms that include roadside variable message signs (VMSs), radio broadcasting, Internet access, highway radio, SMS and other personal subscribed services. An example of traveller information from a VicRoads website is shown in Figure 2.7. A DIS can provide information for a traveller to:

- be aware of road network conditions and conditions en route, e.g. the Drive Time System in Melbourne (Hearn et al. 1996)

- change travel mode before a trip starts from car mode to, say, public transport if severe congestion has occurred
- change to a different destination if car travel is considered necessary
- change routes and get off from a congested freeway if possible to avoid severe bottlenecks
- change lanes on a freeway in the case of an incident
- reduce speeds where appropriate at specific locations for safety reasons.

A VSLS is, naturally, a particular case of a DIS that advises drivers to reduce speeds on a freeway corridor to avoid road crashes (e.g. in harsh weather conditions) or flow breakdowns in recurrent congestion. DISs have yet to be fully exploited amongst road authorities to manage a road network but are expected to play more important roles as the larger cities in Australia and New Zealand become more congested.



Source: http://traffic.vicroads.vic.gov.au/trafficinfo/default.aspx?width=1024

Note: red – heavy congestion, yellow – medium congestion, green – light congestion as recorded at 5:16 p.m. on 4 April 2007;

Figure 2.7: Traveller information from VicRoads

2.3.5 Traffic Lane Management System

The capacity of a carriageway in the peak period can increase if the shoulder lane in the peak-flow direction is opened for traffic. This measure can be achieved by time-of-day using a static sign or a VMS and has been adopted in congested cities.

Further, contra-flow operation has been introduced on freeways and arterial roads where tidal flows occur. For example, more lanes are allocated to the southbound direction in the morning peak period at the Sydney Harbour Bridge and more lanes in the northbound direction in the afternoon peak. This operation was previously a manual operation many years ago but has become traffic adaptive according to prevailing conditions at the Sydney Harbour Bridge and other river crossings in Sydney (Longfoot 1984).

The traffic lane management system is essentially a road space reallocation system and offers the potential to reduce flow breakdowns on freeways. With improving driver information and other ITS technologies, traffic adaptive lane management systems will become more prevalent.

2.4 Summary

Classical traffic models have been useful for the development of relationships amongst speed, flow and occupancy (density). These relationships are *macroscopic* by nature and have provided the basic understanding of a traffic process and for application of automatic freeway control tools.

The review of basic freeway traffic flow theory has confirmed the validity of the concept embedded in the first order Lighthill-Whitham model. The model forms the basis of flow breakdowns in an analytical framework. Active research in second order models with solutions obtained by the finite difference method is on-going. However, analytical models are unlikely to be able to capture driving behaviour before, during and after flow breakdowns.

Various tools have been employed for freeway control in Australian cities. These include ramp-metering, incident detection, VLSLs, driver information tools and lane management. They are already in operation and some already make use of more traditional, macroscopic models of traffic flow.

In recent years, freeway data at a detailed level with small time slices and high density of detector stations has been available in Australian and overseas cities. There has also been renewed interest in the physics of traffic flow, especially amongst researchers in Europe. Their research will be reviewed in Section 3.

The objective of this project, as mentioned, is to make better use of freeway control tools from a better understanding of the flow breakdown process through the study of empirical spatio-temporal maps of traffic data.

3 MODELS FOR THE CHARACTERISATION OF FREEWAY CONGESTED FLOWS

The literature review on freeway congested flow continues in this section with the following contents:

- freeway congested flow model of *two states* as adopted in the US Highway Capacity Manual (HCM) (Banks 1990, Hall, Hurdle & Banks 1992, Transportation Research Board (TRB) 2000) (Section 3.1)
- *three state model* promoted by Kerner (2004), largely based on 1 min data from German motorways (Section 3.2)
- other models that include the stochastic concept of freeway capacity (Brilon, Geistefeldt & Regler 2005, Kerner 2007a, Kerner 2007b, Kerner 2007c, Schonhof & Helbing 2007) (Section 3.3).

As mentioned in Section 1, the factors that lead to *recurrent* flow breakdowns on freeways include mainline flow that exceeds freeway capacity, uncontrolled access from on-ramps, inadequate road geometries, or erratic driver behaviour when the freeway is near maximum flow. This section aims to study how various models published in the past deal with the characteristics of these breakdowns.

3.1 Two Phase HCM Model

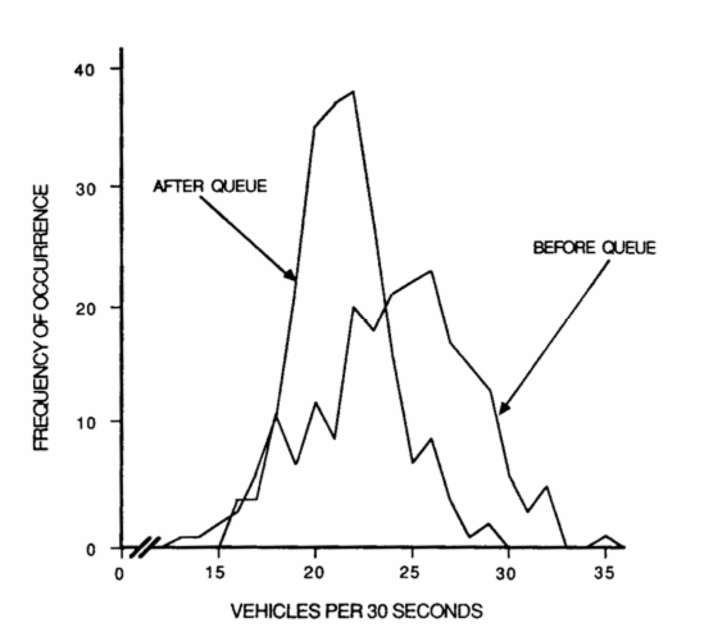
A conventional understanding of the formation of congested flow conditions is that a queue would form upstream of a bottleneck due to conditions such as lane drop, merge area, weaving section or upgrade. The trailing edge of the queue moves upstream at a rate depending on demand and capacity conditions. When the tail of this queue reaches any upstream location, freeway operation moves from the uncongested regime to the congested regime, at approximately the same flow (Section 2.1).

The HCM 2000 (TRB 2000) and the earlier 1986 edition have advocated the need to consider maximum flows or capacities of a freeway segment in two regimes or phases. Two maximum flow rates can be identified as follows:

- Maximum flow when flow is stable – this is the maximum flow before the formation of a queue at a bottleneck, i.e. the maximum *pre-queue* flow.
- Maximum queue discharge flow – this is the maximum flow after a queue is formed and is associated with a speed drop, and has been found to be less than the pre-queue maximum flow rate. A possible reason for this decrease in flow rate is *driver caution* – departures from a freeway queue require more care because drivers may not be aware of conditions downstream. This is in contrast to a start-up queue at a signalised approach where maximum flow is achieved even though different vehicles have different acceleration rates.

There have been debates on where the maximum flows should be measured. Hall and Agyemang-Duah (1991) argued that the two phases are observable only if detectors are located at some distance *upstream* of a bottleneck, and that there is only one congested regime if they are *at a bottleneck*.

In a study of a bottleneck on a four-lane freeway near San Diego (Interstate 8), Banks (1990) measured the above two maximum flow rates. The frequency distribution polygons of the counts on the fast lane are shown in Figure 3.1. The results clearly showed that there is a statistically significant difference between the two flow rates.



Source: Banks (1990)

Figure 3.1: Frequency distribution polygons of vehicle counts on the fast lane

Hall, Hurdle and Banks (1992) finalised a speed-flow diagram as shown in Figure 3.2 which is adopted in the HCM 2000 (Figure 3.3). The diagram overcomes the following issues:

- the parabolic shape in uncongested flow is no longer used; speed remains quite similar until the degree of saturation or volume/capacity ratio reaches 0.75.
- the queue discharge regime is included in the speed-flow diagram.
- two maximum flow rates are used, one for the stable, pre-queue regime and another for the queue discharge rate (which is lower than the maximum pre-queue flow rate).

As mentioned, ramp metering is useful for reducing on-ramp flow so that the mainline demand is maintained at or just below capacity and therefore reduces the occurrences of flow breakdowns and also improves traffic conditions at the merge point.

Hall, Hurdle and Banks (1992) also suggested that much more research is needed in understanding freeway congested flow. The vertical segment in Figure 3.2 or Figure 3.3 is not really a speed-flow function, but is plotted on a graph without the location axis. It is therefore important to analyse flow-breakdowns in spatio-temporal diagrams described in the following section.

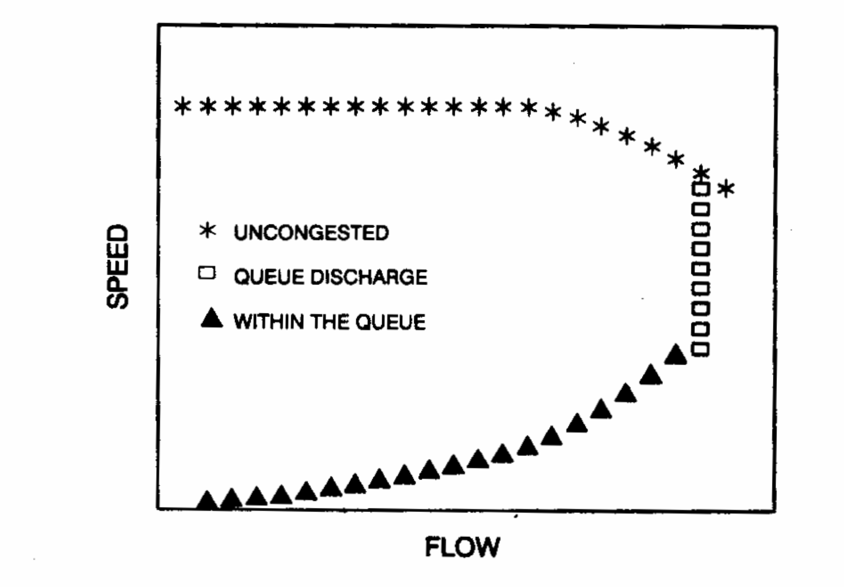


Figure 3.2: Generalised speed flow relation for a typical freeway segment

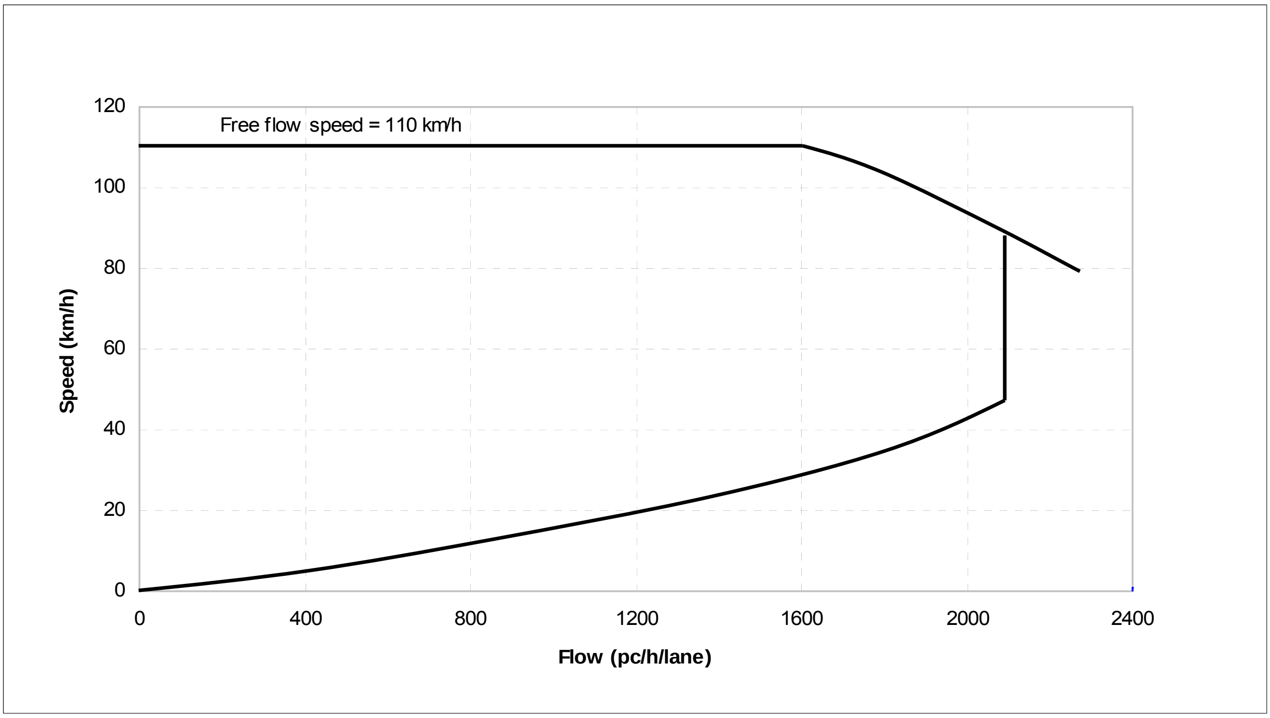


Figure 3.3: Speed-flow relationship for freeway in HCM 2000

3.2 Three Phase Model

This section reviews the three-phase model of Kerner (2004) and includes:

- definition of three traffic phases
- empirical probability nature of traffic breakdown

- engineering applications of three-phase theory (FOTO and ASDA)
- freeway control application (ANCONA)
- comments on Kerner's three-phase model.

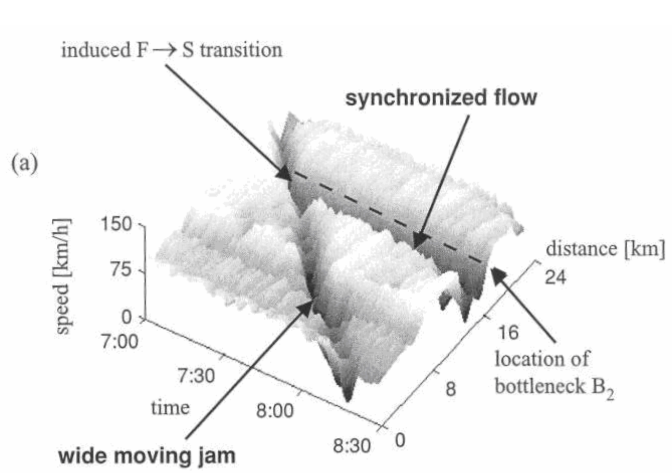
3.2.1 Three Traffic Phases

Kerner and Rehborn (1996) first proposed the classification of freeway traffic flow into three phases based on time series of flow, occupancy, and average speed. Kerner (2004) later completed the three-phase traffic theory based on earlier work. In the three-phase traffic theory, there are two traffic phases in congested traffic, *synchronised flow* and *wide moving jam,* defined as follows:

- A synchronised flow is a congested traffic state and the downstream front of this flow is often fixed at a freeway bottleneck. Within the downstream front of synchronised flow, vehicles accelerate from lower speeds in synchronised flow to higher speeds in free-flow.
- A wide moving jam is a moving jam that maintains the mean velocity of the downstream jam front, even when the jam propagates through any other traffic states or freeway bottlenecks.

The three traffic phases are therefore free-flow (F), synchronised flow (S) and wide moving jam (J).

Figure 3.4 illustrates the traffic phase definition of synchronised flow and wide moving jams (Kerner 2004). The data in Figure 3.4 came from a section of Autobahn 5-South freeway near Frankfurt, Germany. There are three bottlenecks labelled as B1, B2 and B3. Average 1 min speed data in space and time is shown in (a). A two-dimensional graph of the same data with the free-flow phase in white, the synchronised flow phase in grey, and the wide moving jam phase in black is shown in (b).



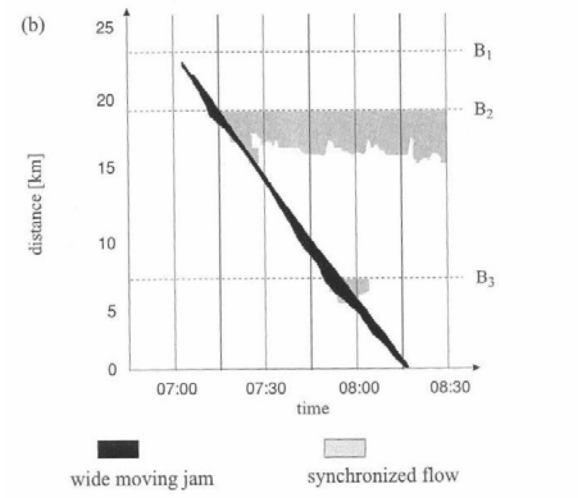


Figure 3.4: Synchronised flow and wide moving jams in congested traffic

The three-phase traffic theory explains the complexity of traffic phenomena based on phase transitions among these three traffic phases. For example, transitions can be *spontaneous* F→ S or *induced* F→ S, and their complex nonlinear spatio-temporal features. In Kerner's three-phase theory, a transition from F→ S is a flow breakdown (Kerner 2004, Kerner et. al. 2005).

An induced F→ S transition is caused by a short-term external disturbance in traffic flow. This traffic flow can be related to the propagation of a moving spatio-temporal congested pattern that initially occurs at a different freeway location. Figure 3.4 (a) shows an example of induced F→ S transition – the wide moving jam propagated through the bottleneck location B2 and induced the synchronised flow at this bottleneck.

Figure 3.5 shows an example of spontaneous F→ S transition. This breakdown phenomenon or F→ S transition is caused by an internal local disturbance (e.g. an on-ramp bottleneck) in traffic flow. There are no external disturbances in traffic flow responsible for this phase transition

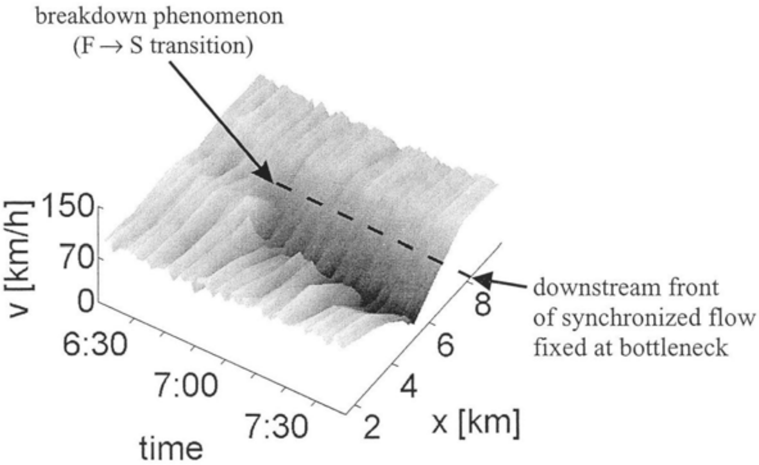


Figure 3.5: An example of spontaneous F→ S transition

The F→ S transition or breakdown phenomenon usually occurs at the same freeway bottleneck. These bottlenecks are called *effectual* bottlenecks in Kerner's model. Examples of effectual bottlenecks are the bottleneck in Figure 3.5 and B1, B2, and B3 in Figure 3.4.

Based on different combinations of traffic phases, different congested patterns are formed. Kerner studied traffic flow on the A5 freeway over a large number of days and found that the spatio-temporal structure of congestion patterns exhibits predictable features. These features can be used to forecast freeway congestion and develop effective freeway control tools such as mentioned in Section 3.

Lindgren (2005) also investigated a 30 km section of A5 freeway north of Frankfurt and found some similar traffic patterns that match Kerner's three traffic phases. In Lindgren's A5 freeway study, traffic flows were observed in which speeds across all lanes were notably lower than in free-flowing conditions, and they were more consistent across all lanes. This phenomenon was observed in congested flows upstream of the bottleneck following activation. This pattern matched Kerner's synchronised flow phase. Lindgren also revealed several occurrences of congested patterns in which a relatively short duration traffic disturbance travelled several kilometres upstream. This pattern matched Kerner's wide moving jam.

Lindgren's study represented some of the first apparent independent validation of Kerner's traffic phase findings (Lindgren 2005, Lindgren et. al. 2006). However, Lindgren also offered different analysis techniques and comments on Kerner's work are reported in Section 3.2.5.

Further, Brilon et al. (2005) also showed that three traffic flow states exist in a freeway: fluent traffic state, congested traffic state and a transient state that occurs in each breakdown and recovery of traffic flow. Their study will be examined further in Section 3.3.1.

3.2.2 Empirical Probabilistic Nature of Traffic Breakdown

Kerner (2004, 2007a, 2007c) found that the traffic breakdown exhibits a probabilistic nature. At a given flow rate, traffic breakdown at a freeway bottleneck can occur but it may not necessarily occur.

The probability for an F→ S transition, i.e. a traffic breakdown, ($P_{FS}^{(B)}$) at a bottleneck is an increasing function of the flow *downstream* of the bottleneck q_{sum} as shown in Figure 3.6 and Figure 3.7. q_{sum} is the sum of the flow on the on-ramp q_{on} and mainline upstream flow q_{in}. There is a threshold flow rate $q_{th}^{(B)}$ and a critical flow rate $q_{max}^{(B)}$. Regardless of free-flow control application there is a range when $q_{th}^{(B)} \leq q_{sum} \leq q_{max}^{(B)}$ within which traffic flow breakdowns can occur with probability $P_{FS}^{(B)} > 0$.

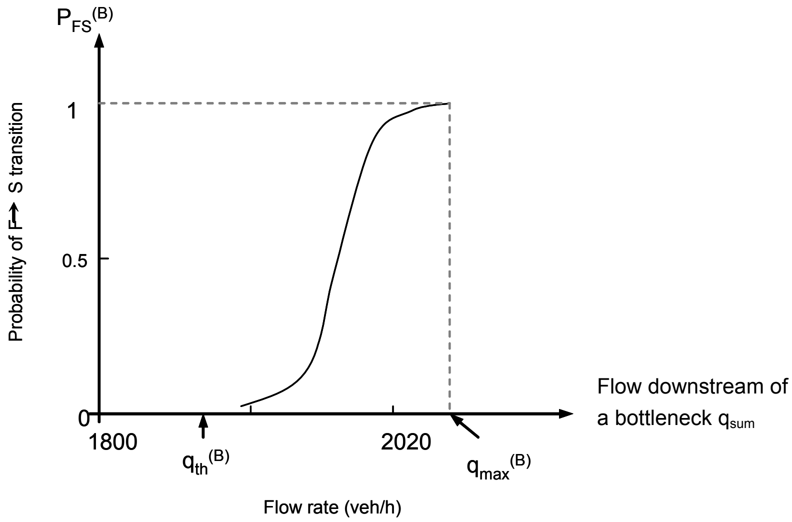


Figure 3.6: Probability of traffic breakdown

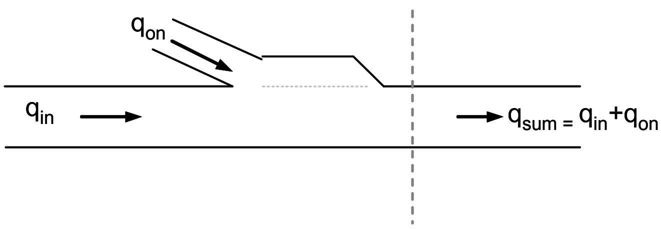


Figure 3.7: Traffic flow downstream of a bottleneck (q_{sum})

A flow breakdown, if due to a speed disturbance in free flow in the neighbourhood of a bottleneck, occurs only when the speed decreases below a critical speed. The critical speed depends on the q_{sum}. The smaller the q_{sum}, the lower the critical speed required for breakdown. The probability for traffic breakdown $P_{FS}^{(B)}$ is the probability of random critical speed disturbances appearing at the bottleneck. Disturbances with small amplitudes in free flow at the bottleneck do not lead to breakdown. However, if a random short-term speed disturbance in free flow at the bottleneck exceeds some critical values, traffic breakdown occurs.

Kerner (2007a) stated that empirical fundamental features of probabilistic traffic breakdown cannot be explained and cannot be predicted by earlier freeway flow models.

3.2.3 The FOTO and ASDA Analysis Tools

To recognise and track the spatio-temporal congested traffic patterns at freeway bottlenecks based on the three-phase theory, Kerner developed the FOTO and ASDA models (Kerner 2004, Kerner et. al. 2004).

FOTO is used to recognise the location of the upstream and downstream fronts of synchronised flow. ASDA is used to recognise the upstream and downstream fronts of wide moving jams. Both could track these fronts in time and space (Figure 3.8). In these models, artificial intelligence (fuzzy inference system) is used to classify the traffic phases based on local measurements.

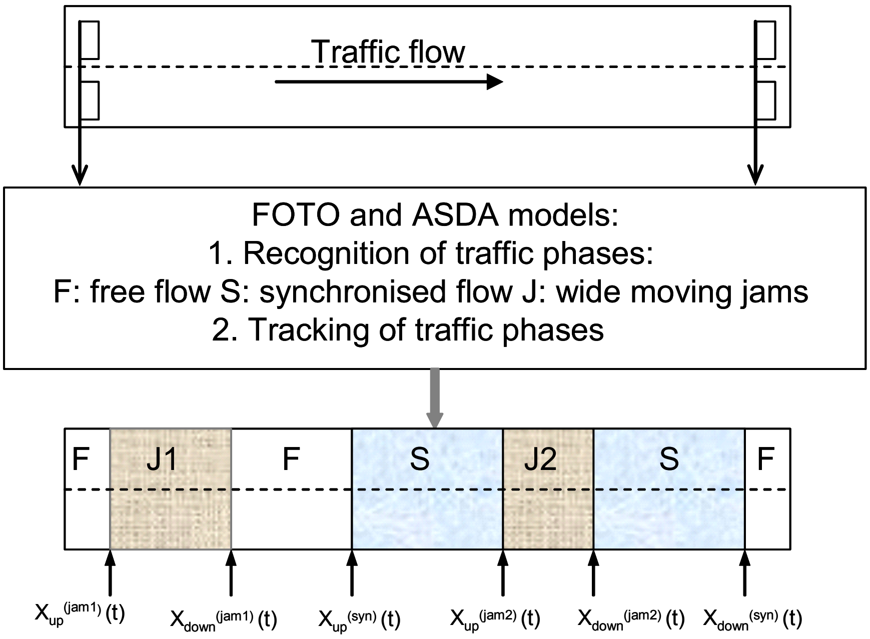


Figure 3.8: Illustration of FOTO and ASDA model approach

The FOTO and ASDA softwares have been installed in the whole freeway network in the state of Hessen in Germany with approximately 2500 double loop detectors and 1200 km of freeway network. Kerner also applied FOTO and ASDA models on American freeways in California (Los Angeles) based on 30 s stationary detector measurements of flow rates and occupancy.

Kerner compared German and USA freeway applications and claimed that the FOTO and ASDA models could recognise and track all congested patterns with similar accuracy in different road networks, and without much effort in re-calibration.

3.2.4 Freeway Control Application Based on Three-phase Theory

Kerner and colleagues further presented freeway control methods based on the features of the breakdown phenomenon and congested pattern that emerge at freeway bottlenecks (Kerner et al. 2005, Kerner 2007a, Kerner 2007c). The ANCONA on-ramp metering system is an example.

The ANCONA approach is different from earlier on-ramp metering methods that are based on the free-flow approach. The fundamental diagram in Figure 3.9 explains the theoretical background of free-flow approach. The downstream bottleneck capacity q_{cap} in this approach is related to the maximum flow of the flow-occupancy diagram. Traffic congestion upstream of the bottleneck occurs only when upstream flow exceeds q_{cap}. Free-flow approaches aim to maintain the free-flow at the bottleneck by keeping the downstream occupancy close to a chosen optimal occupancy. This optimal occupancy should be less than the critical occupancy o_{cr} (when maximum flow q_{cap} occurs; Section 2). This theory is in contradiction with the probability of traffic breakdown in the three-phase theory - the probability of traffic breakdown is very close to one if the downstream occupancy is very close to the critical occupancy (Figure 3.6).

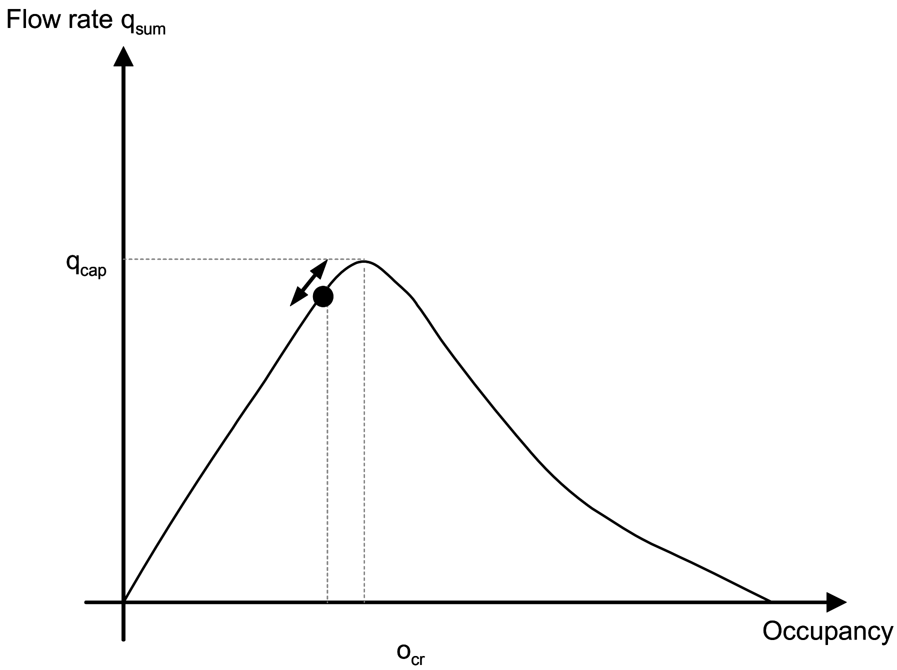


Figure 3.9: Fundamental diagram of free-flow control approach

The ANCONA approach aims to control the congested pattern localised on the mainline in a small neighbourhood of the bottleneck. The ANCONA detector is located at the upstream of an effectual bottleneck (Figure 3.10). The on-ramp metering adjusts the ramp flow q_{on} based on the average speed v_{det} measured at the feedback control detector. If v_{det} drops down to equal or below a chosen congested speed v_{cong}, which means a spontaneous traffic breakdown (F→ S transition) has occurred, the on-ramp metering then starts to reduce the ramp flow q_{on}. This will achieve a phase transition S→F and return to normal uncongested flow. When v_{det} increases to above v_{cong}, the on-ramp meter starts to allow greater ramp flow q_{on} to merge into the mainline. A F→ S transition may appear again, and the on-ramp meter reduces q_{on} again. Figure 3.11 illustrates the phase transition of this control approach.

Kerner further observed that:

- under a small enough on-ramp flow q_{on}, S→F transition can easily occur
- within the synchronised flow, the average speed is relatively high (about or more than 60 km/h)
- no wide moving jams are formed spontaneously within the high speed synchronised flow
- the synchronised flow pattern can propagate upstream, but only at the localised area

- under a very high traffic demand when congestion has to occur somewhere in the traffic network, ANCONA tries to achieve greater throughputs.

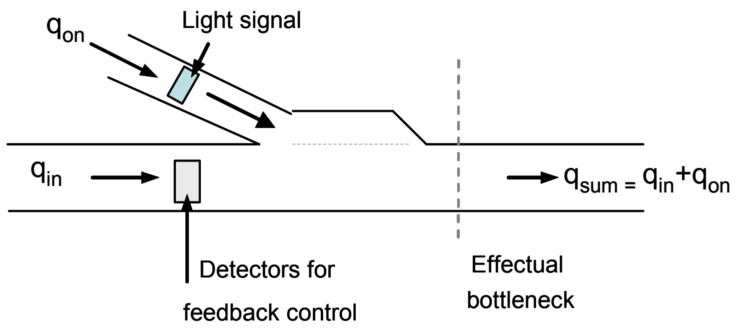


Figure 3.10: Location of ANCONA detectors

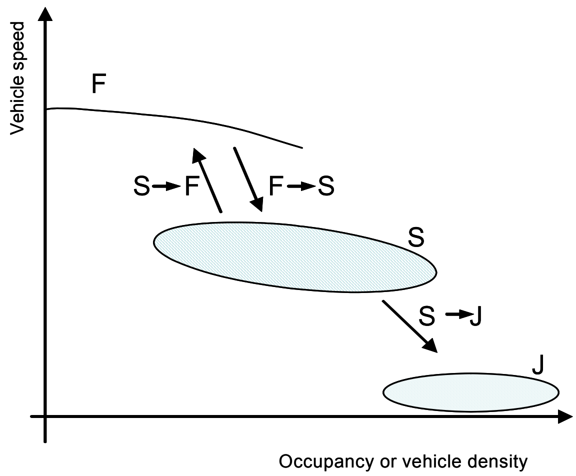


Figure 3.11: Theoretical background of ANCONA

Kerner (2007a, 2007b, 2007c) claimed the benefits of ANCONA as follows:

- prevention of upstream propagation of congestion
- greater throughputs at the bottleneck
- shorter waiting time at the light signal in the on-ramp lane.

3.2.5 Comments on Kerner's Three-phase Model

Lindgren (2005) and Schonhof and Helbing (2007) investigated the same section of A5 freeway traffic flows as Kerner did. Lindgren reviewed Kerner's work on three-phase models and commented that Kerner and colleagues had not made their data available to other researchers.

Lindgren made the following observations:

- Kerner's time series plots cannot show excess accumulation (queuing) between measurement locations resulting from bottleneck activation. Therefore, Lindgren applied a *cumulative count curves technique* that was used to complement the three-phase models to observe transitions between free flows to queued conditions and identify time-dependent traffic features of bottlenecks. Figure 3.12 shows the bottlenecks in time and space identified by Lindgren.

- Kerner and colleagues only investigated a limited number of days and bottlenecks of A5, which failed to consistently determine the spatio-temporal limits of bottleneck activation before computing characteristics such as discharge flow. Lindgren investigated 81 bottleneck activations and deactivations, where queued traffic prevailed upstream of each bottleneck and un-queued traffic was present downstream.
- Kerner suggested that there was up to a 50% capacity difference between free-flow and congested traffic, and claimed that there was significant variability of discharge flows. However, Lindgren found that high flows existed particularly in the mid and left lanes for several minutes prior to bottleneck activation. The bottleneck discharge flows in Lindgren's study were found to be essentially reproducible over several days, across all lanes and in the individual lanes.
- Kerner suggested that traffic congestion can form and traffic can self-organise without a physical bottleneck. However, it was shown that all 81 bottlenecks diagnosed in Lindgren's study, activated at a predictable location (e.g. merge, diverge, vertical curves) and appeared to be linked to particular triggers rather than to have occurred spontaneously.

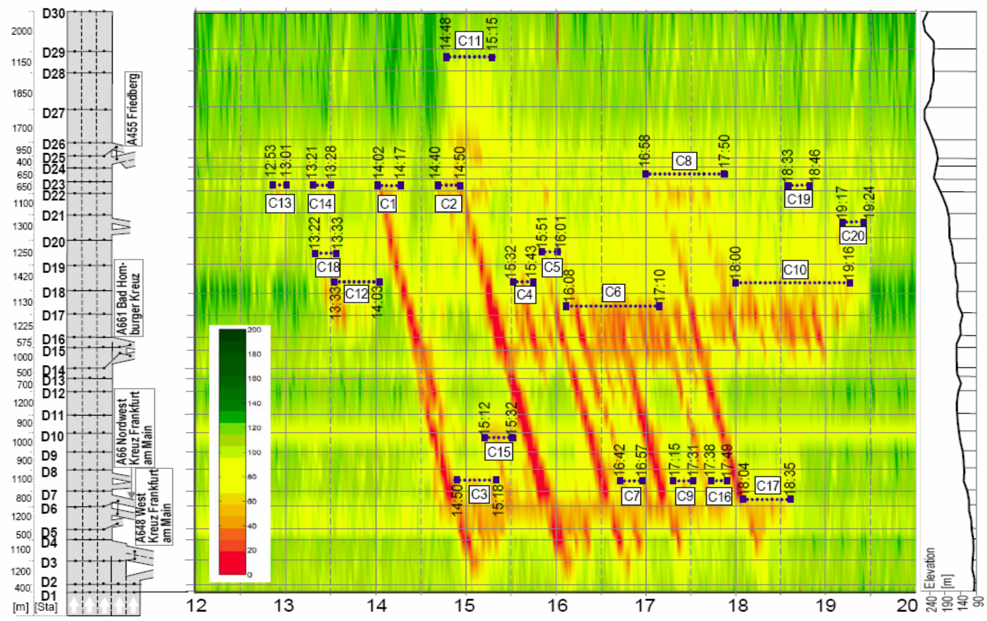


Figure 3.12: A5 speed contour diagram in Lindgren's study – 1 min data

Schonhof and Helbing (2007) introduced a *six traffic phase* model (see also Section 3.3.2) and commented on Kerner's three-phase theory as follows:

- Kerner claimed that the three-phase theory is the only theory that can describe the empirical phenomena accurately. However, Kerner has reported new spatio-temporal traffic patterns and the definition of synchronised flow is still a controversial topic.
- Kerner's criticism of models that use the fundamental diagrams is not convincing in Schonhof and Helbing's study. Therefore, it is not necessary to discard all previous knowledge accumulated in traffic modelling as suggested by Kerner.

3.2.6 Summary of Three-phase Model

In summary, it has been found that there are empirical features of phase transitions and spatio-temporal congested patterns at freeway bottlenecks that are reproducible. The earlier freeway flow models and theories cannot explain or predict these empirical features of traffic breakdowns. The literature review of Kerner and colleagues' work in the past ten years shows that the three-phase theory models provide a promising tool to analyse mainline empirical spatio-temporal congested pattern features. This spatio-temporal analysis of traffic flow includes:

- observing and measuring traffic flow parameters at many freeway locations over the course of many days, and identifying the recurrent congested locations
- finding effectual bottlenecks based on the investigation of congestion locations and freeway infrastructure plans
- analysing the predictability and reproducible features of different congested patterns for each effectual freeway bottleneck or each set of several adjacent effectual bottlenecks
- studying some specific congested patterns to identify synchronised flow patterns and wide moving jams based on the objective criteria for traffic phases in congested traffic situations.

3.3 Other Models of Flow Breakdowns

3.3.1 Stochastic Concept of Traffic Capacity

Brilon et al. (2005) studied 5 min data on the freeways around the city of Cologne, Germany and found that the concept of stochastic capacities seems to be more realistic and more useful than traditional use of single value capacity. Their empirical analysis shows that the distribution of freeway capacity fits very well into a Weibull distribution (Figure 3.13). The shape parameter seems to be in a range from 9 to 15 with an average of 13 for German freeways. The overload probability (traffic breakdown) for a single bottleneck is equal to the capacity distribution function as shown in Figure 3.13. This finding is consistent with Kerner's analysis of probabilistic nature of traffic breakdown (Section 3.3.2).

The concept of randomness permits the demonstration of the capacity reducing effect of wet road surfaces (-11%) and the capacity increasing effect of traffic adaptive variable speed limits.

The study by Brilon et al. also showed that three traffic flow states exist in a freeway: fluent traffic state, congested traffic state and a transient state that occurs in each breakdown and recovery of traffic flow. These three states seem to match Kerner's three-phase theory but the definitions of the phases are slightly different.

The stochastic concept of capacity reveals that the optimum degree of saturation for a German freeway is around 90%. If the degree of saturation increases further, the risk of a breakdown becomes too high, so that the efficiency of freeway operation must be expected to be lower than a saturation of 90%.

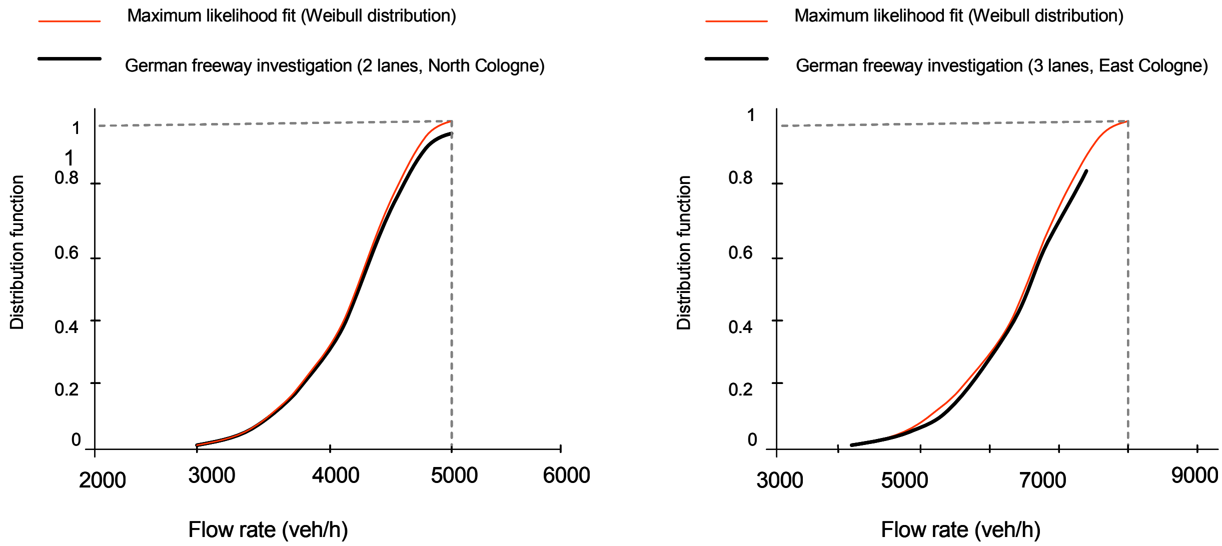


Figure 3.13: Capacity distribution functions for two German freeway sections – 5 min data

3.3.2 Six Traffic State Model

Schonhof and Helbing (2007) investigated 1 min date for the same section on the A5 freeway as Kerner and Lindgren. They interpreted traffic flow by six states: free traffic (FT), pinned localised cluster (PLC), moving localised cluster (MLC), stop-and-go waves (SGW), oscillating congested traffic (OCT) and homogeneous congested traffic (HCT). The most frequent states at the investigated freeway are the PLC and OCT states. HCT occurs mainly after serious accidents with lane closures or during public holidays. An adaptive smoothing method is used to identify the different traffic states. This method interpolates and smoothes traffic date from successive freeway sections, taking into account the propagation speeds of perturbations in free and congested traffic.

Schonhof and Helbing (2007) found that the congested traffic states identified by this model are in good agreement with prediction of some second-order macroscopic traffic models and some microscopic car-following models.

3.4 Summary

This section reviews the three-phase model and other relevant traffic models. These traffic state models classify traffic flow into different state regimes and provide a useful way to illustrate congested freeway flow patterns. Freeway data at a detailed level with small time slices and high density of detector stations is becoming more available in recent years. It is now possible to investigate the time and space features of flow breakdowns in spatio-temporal diagrams such as speed contours described above. After the spatio-temporal pattern features are understood, additional analysis of nonlinear pattern features could be performed using Lindgren's cumulative count curve method.

4 PROJECT TASKS AND TIMELINE

The inception of this project took place in June 2007. This section provides an outline of future tasks beyond the current project task of literature review. The various project tasks agreed with Austroads are as follows:

- Task 1: Literature review and presentation – this review consists of traditional and current spatio-temporal techniques for analysing flow breakdowns.
- Task 2: Progress report 1 – this stage involves the development of an experimental design for data collection, the compilation of data and the identification of sites suitable for flow breakdown analysis. The sites are most likely from Melbourne freeways because of the large number of two-loop detector stations available.
- Task 3: Progress report 2 – the tool for the analysis of flow breakdowns will be developed in this task and will follow the recommended approach as a result of the literature review in Task 1.
- Task 4: Progress report 3 – the analytical tool will be applied to the sites included in the experimental design in Task 2.
- Task 5: Draft project report and presentation – the results of analysis and the methodology will be compiled and presented.
- Task 6: Final report and presentation.

Table 4.1 shows the planned timeline for these tasks.

Table 4.1: Project timeline

Task	Description of work	Month
-	Project inception	June 2007
1	Literature review	July 2007
2	Progress report 1	Aug 2007
3	Progress report 2	Oct 2007
4	Progress report 3	Dec 2007
5	Draft project report & presentation	Feb 2008
6	Final project report & presentation	Mar 2008

Sections 3.1 to 3.4 of this report constitute the output for Task 1. Meetings have been held with VicRoads staff regarding the current M1 ramp metering project and the retrieval of data from VicRoads' databases.

As mentioned, inadequate road geometries such as sudden lane-drop, excessive upgrades, freeway lane merge and insufficient weaving capacity would cause flow breakdowns. Inadequate geometries are some of the issues to be addressed in this project, and will also be pursued in greater detail in Austroads Project NS NS1375 (*Freeway Design Parameters for Fully Managed Operations*).

5 CONCLUSIONS

A traffic system is a complex, time-dependent system. It consists of the driver, vehicle and road infrastructure, and the interactions amongst the three components. Driver behaviour is especially difficult to control. The complexity of a traffic system is not a major concern as long as a road facility is uncongested but, with increasing congestion, certainly needs to be well understood for network operations. The phenomenon of flow breakdowns on freeways is, however, still not well understood.

There have been many attempts to develop mathematical models to characterise and understand traffic behaviour. For example, the Lighthill-Whitham (1955) kinematic wave model, despite its limitations, remains useful for understanding shock waves at road bottlenecks. Even though most models lack detail and reality, traffic control systems have long been employed to control arterial and freeway traffic macroscopic relationships.

In recent years, freeways in some overseas and Australian-New Zealand cities have been installed with detector stations that can provide reliable traffic data at a high resolution in time and space. In other words, it is now possible to obtain accurate empirical diagrams in time and space. These spatio-temporal diagrams provide a good framework to analyse freeway congested flow and breakdowns at bottlenecks. The work of Kerner (2004) has been extensively reviewed in Section 3. Kerner developed a model from empirical data and it consists of three phases or states: free-flow, synchronised flow and moving jam. From various combinations of these phases and their transitions, different flow patterns observed on a freeway can be reproduced or predicted. An accurate prediction model is the first step in the effective implementation of freeway control tools such as ramp-metering and speed limit signs.

This project has reviewed other models of flow breakdowns including the stochastic model of Brilon et al. (2005) and the six-state model of Schonhof and Helbing (2007). A general finding is that the Kerner approach has gone through a significant amount of research over ten years. Apart from providing a good platform for the analysis of flow breakdowns (which is the concern of this project), it has been successfully deployed for real-time control. This project can benefit from the Kerner three-phase concept with additional input from the cumulative count curve method of Lindgren (2005).

The project team has already scanned through the flow, speed and occupancy data from the Monash Freeway, Westgate Freeway and Western Ring Road in Melbourne to decide what freeway sites should be selected for analysis. The recommendations for subsequent stages of this project are as follows:

- A freeway route is preferable to a few isolated sites from one or several freeways so that freeway flow can be studied in a network context – what happens upstream affects downstream flow. The demand for that corridor will also be similar for a particular time of day if all sites are from the same freeway.
- Analytical modelling processes are unlikely to explain complex driver behaviour before, during and after flow breakdowns. An empirical spatio-temporal approach is recommended for congested freeway flow analysis. The two-phase approach of HCM 2000 is too simplistic and the Kerner (2004) three-phase approach appears appropriate for the identification of flow breakdowns and will be adopted in this project.
- The software MATLAB would provide a useful platform for data analysis and the development of rules and artificial intelligence in general for the identification of flow breakdowns and tracking of shock waves.

REFERENCES

Ackelik, R, Roper, R & Besley, M 1999, *Fundamental relationships for freeway traffic flows*, research report ARR 341, ARRB Transport Research, Vermont South, Vic.

Auckland Motorways (n.d.), Ramp signalling, viewed 20 August 2007, <http://www.aucklandmotorways.co.nz/rampsignalling.html>

Banks, JH 1990, ‘Flow processes at a freeway bottleneck’, *Transportation Research Record*, no.1287, Transportation Research Board, Washington, DC, pp.20-8.

Bean, S 2002, ‘A dynamic variable speed limit system’, *World Congress on Intelligent Transport Systems, 9th, Chicago, 2002,* ITS America, Washington, D.C.,10pp.

Blinco, R 1988, ‘Monitoring operations on the South East Freeway’, *Queensland Roads*, vol.23, no.3, pp.26-30.

Brilon, W, Geistefeldt, J & Regler, M 2005, ‘Reliability of freeway traffic flow: a stochastic concept of capacity’, *International Symposium on Transportation and Traffic Theory, 16th, College Park, Maryland, July 2005*, Elsevier, Oxford, pp.125-44.

Cambridge Systematics 2001, *Twin Cities ramp meter evaluation*, final report prepared for Minnesota Department of Transportation, Cambridge Systematics, Oakland, California.

Daganzo, ZF 2006, ‘In traffic flow, cellular automata = kinematic waves’, *Transportation Research*, vol.40B, no.5, pp.393-406.

Dia, H & Rose, G 1997, ‘Development and evaluation of neural network freeway incident detection models using field data’, *Transportation Research*, vol.5C, no.5, pp.313-31.

Edie, LC & Foote, RS 1958, ‘Traffic flow in tunnels’, *Highway Research Board Proceedings*, vol.37, Highway Research Board, Washington, DC, pp.334-44.

Gaffney, J 2007, ‘Monash-CityLink-west gate upgrade project: building a business case for the deployment of integrated its operations’, *World ITS Congress, 14th, Beijing, Oct 2007*, (submitted for presentation).

Hall FL & Agyemang-Duah K, 1991, ‘Freeway capacity drop and the definition of capacity’, *Transportation Research Record*, no.1320, Transportation Research Board, Washington, DC, pp.91-8.

Hall, FL, Hurdle, VF & Banks, JM 1992, ‘Synthesis of recent work on the nature of speed-flow and flow-occupancy (or density) relationships on freeways’, *Transportation Research Record*, no.1365, Transportation Research Board, Washington, DC, pp.12-7.

Hearn, B, Ramsay, E, Catchpole, J & Luk, JYK 1996, ‘Evaluation of the VicRoads Drive Time system’, *ARRB Transport Research Conference, 18th, 1996, Christchurch, NZ*, vol.18, no.7, ARRB Transport Research, Vermont South, Vic., pp.325-42.

Herley, T & Lennie, S 2007, ‘Variable speed limits and lane control signs: site selection criteria for Queensland motorways’, *National Variable Speed Limits Forum, Brisbane, March 2007*, 12pp.

Highways Agency (n.d.) The M25 controlled motorway, viewed 6 July 2007, <http://www.highways.gov.uk/knowledge/14334.aspx>

Jenezon, JH, Klijnhout, JJ & Langelaar, HCG 1987, ‘Motorway control and signalling’, *Traffic Engineering & Control*, vol.28, no.6, pp.349-55.

Kerner, BS 2004, *The physics of traffic*, Springer, Berlin.

Kerner, BS 2007a, 'On-ramp metering based on three-phase traffic theory - part I', *Traffic Engineering & Control,* vol.48, no.1, pp.28-35.

Kerner, BS 2007b, 'On-ramp metering based on three-phase traffic theory – part II', *Traffic Engineering & Control,* vol.48, no.2, pp.68-75.

Kerner, BS 2007c, 'On-ramp metering based on three-phase traffic theory – part III', *Traffic Engineering & Control,* vol.48, no.3, pp.114-20.

Kerner, BS & Rehborn, H 1996, 'Experimental properties of complexity in traffic flow', *Physical Review E*, vol. 53, no. 5, pp. R4275-8.

Kerner, BS, Rehborn, H, Aleksic, M & Haug, A 2004, 'Recognition and tracking of spatio-temporal congested traffic patterns on freeways', *Transportation Research Part C,* vol.12, no.5, pp.369-400.

Kerner, BS, Rehborn, H, Haug, A & Maiwald-Hiller, I 2005, 'Tracking of congested traffic patterns on freeways in California', *Traffic Engineering & Control,* vol.46, no.11, pp.380-5.

Kosmatopoulos, E, Papageorgiou, M, Bielefeldt, C, Dinopoulou, V, Morris, R, Mueck, J, Richards, A & Weichenmeier, F 2006, 'International comparative field evaluation of a traffic-responsive signal control strategy in three cities', *Transportation Research*, vol.40A, no.5, pp.339-413.

Leo, CJ & Pretty, RL 1992, 'Numerical simulation of macroscopic continuum traffic models', *Transportation Research,* vol.26B, no.3, pp.207-20.

Lighthill, MJ & Whitham, FB 1955, 'On kinematic waves II – a theory of traffic flow on crowded roads', *Royal Society Proceedings (Series A)*, vol.229, pp.317-45.

Lindgren, RVF 2005, *Analysis of flow features in queued traffic on a German freeway*, Ph.D Dissertation, Department of Civil and Environmental Engineering, Portland State University, Portland, 147pp.

Lindgren, RVF, Bertini RL, Helbing, D & Schonhof M 2006, 'Toward demonstrating predictability of bottleneck activation on German Aotubahns', *Transportation Research Record*, no.1965, pp.12-22.

Longfoot, JE 1984, 'Control system for three lane tidal flow bridges', *ARRB Conf., 12th, 1984, Hobart*, part 5, Australian Road Research Board, Vermont South, Vic., pp.155-66.

Lowrie, PR 1996, 'Freeway ramp metering systems', *Institute of Transportation Engineers Regional Conference, Melbourne, 1996*, publication PP-047, Institute of Transportation Engineers, Washington, D.C., 10pp.

Luk, JYK 1999, 'Electronic road pricing in Singapore', *Road and Transport Research*, vol.8, no.4, pp.28-40.

Luk, JYK & Sin, FYC 1992, 'The calibration of freeway incident detection algorithms', *Road Engineering Association of Asia and Australasia Conference, 7th, Singapore*, vol.1, Road Engineering Association of Asia and Australasia Conference, Kuala Lumpur, Malaysia, pp.347-55.

Mak, CL & Fan, HSL 2004, 'Analysis of incidents on Singapore's Central Expressway', *Road & Transport Research*, vol.13, no.1, pp.51-62.

May, AD 1990, *Traffic flow fundamentals*, Prentice Hall, New Jersey.

Michalopolous, PG 1988, ‘A numerical methodology for analysing traffic flow at congested surface streets’, *ARRB Conf., 14th, 1988, Canberra*, part 2, Australian Road Research Board, Vermont South, Vic., pp.173-85.

Ngoduy, D, Hoogendoorn, SP & Van Zuylen HJ 2006, ‘Continuum traffic model for freeway with on- and off-ramp to explain different traffic congested states’, *Transportation Research Record,* no.1965, pp.91-102.

Owens, D & Schofield, MJ 1988, ‘Access control on the M6 Motorway: evaluation of Britain's first ramp metering scheme’, *Traffic Engineering & Control*, vol.29, no.12, pp.616-23.

Papageorgiou, M 1983, *Applications of automatic control concepts to traffic flow modelling and control*, Springer-Verlag, Berlin.

Papageorgiou, M 1998, Some remarks on macroscopic traffic flow modelling’, *Transportation Research*, vol.32A, no.5, pp.323-9.

Papageorgiou, M, Blosseville, J-M & Hadj-Salem, H 1990, ‘Modelling and real-time control of traffic flow on the southern part of Boulevard Peripherique in Paris: part I: modelling’, *Transportation Research*, vol.24A, no.5, pp.345-59.

Papageorgiou, M, Hadj-Salem, H & Middelham, F 1997, ‘ALINEA local ramp metering: summary of field results’, *Transportation Research Record*, no.1603, pp.90-8.

Payne, HJ 1971, ‘Models of freeway traffic and control’, *Simulation Council Proceedings*, vol.1, pp.51-61.

Payne, HJ, Thompson, WA, Isaken, L 1973, ‘Design of a traffic responsive control system for a Los Angeles freeway’, *IEEE Transactions on Systems, Man and Cybernetics*, vol.3, pp.212-24.

Schonhof, M & Helbing D 2007, ‘Empirical features of congested traffic states and their implications for traffic modelling’, *Transportation Science*, vol.41, no.2, pp.135-66.

Taale H & Middelham, F 2000, ‘Ten years of ramp metering in the Netherlands’, *International Conference on Road Transport Information and Control, 10th, April 2000,* IEE pub. no. 472, Institution of Electrical Engineers, London, pp. 106-10.

Taylor, MAP, Bonsall, PW & Young, W 2000, *Understanding traffic systems: data analysis and presentation*, Ashgate, Aldershot, UK.

Transurban 2006, *Monash-City Link-West Gate upgrade project*, brochure, Transurban Limited, Melbourne,

Transportation Research Board 2000, *Highway Capacity Manual 2000*, TRB, Washington, DC.

Troutbeck, RJ, Su, M & Luk, JYK 2007, *National performance indicators for network operations*, report RCNS1207-6, Austroads, Sydney, NSW.

VicGov 2006, *$1 billion boost to Monash West Gate Corridor*, Media release, May 17, The Office of the Premier, Victorian Government, Melbourne, Vic.

Wohl, M & Martin, BV 1967, *Traffic system analysis*, McGraw-Hill, New York.

INFORMATION RETRIEVAL

Austroads, 2008, **Freeway Traffic Flow under Congested Conditions: Literature Review**, Sydney, A4, 42pp, AP-R318/08

Keywords:

freeway, speed, flow, bottleneck, congestion, efficiency, occupancy, detector, time, space, contour, map

Abstract:

This report provides a literature review on the analyses of the flow breakdown process on a freeway. It begins with the basic freeway flow theory and automatic freeway control tools now in operation. A traffic system is a complex, time-dependent system. The factors that contribute to flow breakdowns on a freeway include: mainline freeway flow in excess of capacity, uncontrolled access to the freeway from on-ramps, inadequate road geometries, undisciplined driving behaviour, and lack of real-time driver information to encourage better use of alternative routes or lanes. There have been many attempts to develop mathematical models to characterise and understand traffic behaviour. The Lighthill-Whitham model has been useful, despite its limitations, for understanding shock waves at road bottlenecks. Even though most models lack detail and reality, traffic control systems have long been employed to control arterial and freeway traffic with macroscopic relationships. Models for the characterisation of congested freeway flow are described and the Kerner three-phase model is reviewed in some detail.