RELATIONSHIPS BETWEEN THE MEMBERS OF AN EXTENDED FAMILY OF TRAFFIC NETWORK EQUILIBRIA

Michael A P Taylor
Professor of Transport Planning and Director, Transport Systems Centre
School of Geoinformatics, Planning and Building
University of South Australia, City East Campus
Adelaide South Australia 5000
Australia
(fax) +61 8 8302 1880
(email) MAP.Taylor@UniSA.edu.au

abstract: This paper examines the use of static equilibrium assignment models as models of congested traffic networks that may be used framing transport policies for environmental and energy management, travel demand management, traffic calming, congestion management, road pricing, and land use-transport interactions. To make static assignment models really useful for these broader level policy analyses requires a more flexible model definition, and this is the principal aim of this paper. In addition, the paper seeks to establish relationships between the resulting family of equilibrium assignment models that may then find use in the evaluation and comparison of alternative transport policies.

1. INTRODUCTION

Traffic assignment models have a central place in transport planning, for they provide necessary information on the traffic and congestion loads to be borne across a network and how those loads may vary depending on network configuration, design standard, control regime and travel demand distribution. Static models, such as the user equilibrium (minimum individual travel time) model, have been used in transport planning for many years. In more recent times, there has been growing interest in dynamic assignment models, for use in network simulation and with route navigation systems and advanced traveller information systems (ATIS).

This paper focuses on static equilibrium assignment models because of their relevance in transport policy analysis, that is as models of congested traffic networks that may be used to examine the effects of policy alternatives in environmental and energy management, travel demand management (TDM), traffic calming, congestion management and road pricing, and land use-transport interactions. To make static assignment models really useful for these broader level policy analyses requires more flexible definition of the models, and it is this broader definition that is the principal aim of this paper. In addition, the paper seeks to establish relationships between the resulting family of equilibrium assignment models that may then find use in the evaluation and comparison of alternative transport policies.

2. THE BASIC NETWORK FLOW PROBLEM

The network flow problem is a basic problem in transport planning. Traffic assignment models to solve this problem for road networks require the following inputs:

(1) a network description, where the network is a connected graph of nodes and links;

(2) an origin destination matrix $\{T_{ij}\}$ of trips from origins i to destinations j in the network, which describes the travel demand for the network. On occasions there may be a set of origin-destination matrices, split in terms of factors such as trip type or vehicle class. There may also be separate matrices for time of day, particularly in studies where trip timing or peak spreading is important, and

(3) descriptions of the physical and traffic-carrying characteristics of the network links (e.g. road type, number of lanes, free flow travel time or speed, relationship between link volume and travel time) and (possibly) the network nodes (e.g. intersection geometry and control type, traffic signal setting where appropriate, turn penalties and bans as necessary).

The network assignment problem is one of selecting a specific strategy to allocate the trips from the O-D matrices to routes through the network, and thus accumulate the flows on the network links and through the network nodes. The principal complication arises from the interaction between traffic volume and travel time (or cost) on network elements. Route choices are assumed to be based on travel times and/or costs between origins and destinations; these times and costs change as traffic volumes on the network build up - the phenomenon of traffic congestion - and thus an iterative mechanism has to be established.

2.1 Continuity of Flow

The generic solution to the network assignment problem is based on continuity of flow considerations. A given level and pattern of travel demand in a given time period has to be loaded on to the network. We may write the generic network flow solution as the following sets of continuity of flow equations.

$$T_{ij} = \sum_{r} X_{rij} \qquad \forall\ i, j \tag{1}$$

$$q(e) = \sum_{ijr} \delta_{eijr} X_{rij} \qquad \forall\ e \tag{2}$$

and

$$q(e) \geq 0 \qquad \forall\ e \tag{3}$$

where

δ_{eijr} = 1 *if and only if* e is in path r from i to j,
= 0 otherwise

T_{ij} is the number of trips from origin i to destination j, q(e) is the volume on link e and $X_{rij} \geq 0$ is the number of trips using path r between origin i and destination j.

2.2 The Equilibrium Family of Flow Solutions

Any flow pattern satisfying the above constraint equations (1)-(3) is a feasible solution to the network flow problem. A number of alternative solutions, satisfying different performance standards set in terms of the resulting travel times or 'costs' for journeys made through the network, can be defined. Each of these alternatives can be ascribed to a particular strategy or policy for organising or representing the travel pattern Further, these different network flow solutions may be evaluated separately or compared with each other, thus providing a means to assess and rate alternative policies. This equilibrium family of flow solutions may be described by defining a set of alternative objective functions for the network, based on a set of route choice principles.

3. THE WARDROP-JEWELL PRINCIPLES

For planning and evaluation purposes traffic assignment and route choice modelling is perhaps most successfully undertaken by the formulation of the assignment-route choice problem as a mathematical programming problem. Three basic principles are in common usage: Wardrop's first and second principles, and Jewell's principle. These three principles may be seen as individual members of a family of traffic network equilibria. Equilibrium assignment involves the solution of a mathematical programming problem in which the objective function is non-linear but the constraints are linear. The constraints ensure the conservation of flows in the network, and are set to ensure that the travel demand is satisfied. The objective function represents the strategy adopted by drivers in selecting the routes for their journeys.

The assignment principles defined by Wardrop (1952) and Jewell (1967) may be used to define a family of equilibrium assignment models. Different members of this family will have different objective functions in the mathematical programming formulation for the equilibrium assignment problem. The theme of this paper is to develop an understanding of the relationships and similarity between various individual equilibrium solutions as members of a family of equilibrium solutions, representing different planning and traffic control objectives.

3.1 Wardrop's First Principle

Under this principle, the journey times on all of the routes used for travel between an origin and a destination will be equal at the equilibrium point, and will be less than those times which would be experienced on any other route. No individual driver can gain an advantage by a unilateral change of route. This strategy implies that each driver seeks a route minimising that individual's travel time given that all drivers are attempting to implement a similar strategy for themselves. The strategy is one of individual travel time optimisation, and there is no cooperation between drivers. Indeed, the situation is one of competition between drivers, who are all seeking the best outcomes for themselves independently of each other. It may be regarded as a competitive solution point under the mathematical theory of games.

3.2 Wardrop's Second Principle

This principle is concerned with the overall minimisation of the travel task represented by the total travel time (vehicle-hours of travel, VHT) in the network. In this case drivers will select their routes to produce the minimum VHT which is necessary for the travel demand to be satisfied, i.e. for all of the trips in the O-D matrices to reach their destinations. This model is described as *system travel time minimisation*. The solution to this problem implies a degree of cooperation between drivers to attain this result, and may be seen as a Pareto solution in the theory of games. Although the total VHT will be less than that arising in the user travel time minimisation, some individual drivers may encounter much longer travel times than the minimum available to them. Should such drivers decide to improve their own situations, then the system-wide optimum solution will be lost, and there is no incentive (other than the ideal of cooperation for the overall gain of the community) for them not to do so. The consequence is that this solution is unstable. It does, however, define a datum in terms of the 'best' distribution of flows that could occur if the overall minimisation of 'travel effort' (e.g. VHT) were to be achieved. Other solutions (e.g. for user travel time minimisation) may be compared to it on those grounds.

3.3 Jewell's Principle

This principle may be seen as a generalisation of the two Wardrop principles, each of which can be seen as a special case of Jewell's principle: that the assigned flow pattern should optimise some overall economic objective for the network. This objective may be the minimisation of travel time, either by individuals (Wardrop's first principle) or for the system as a whole (Wardrop's second principle). In addition, other definitions of economic objective may be chosen, where such objectives are sensitive to the level of congestion. These would include objectives such as minimum fuel consumption, vehicle operating costs, generalised cost of travel, or pollutant emissions. The optimisation problem for any of these objectives may be defined as either a user minimisation problem or a system minimisation problem. Congestion pricing equilibrium solutions can also be obtained, when marginal link costs can be derived from average link costs. Further solutions to be included in the family of traffic network equilibria include solutions for networks with route guidance information available, in which different proportions of 'advised' and 'unadvised' drivers are on the road.

4. TRAFFIC CONGESTION

Traffic congestion presents a common if not inevitable facet of traffic activity in urban areas. The spread, duration and intensity of congestion, the processes that lead to it, and the consequences of it are of special concern in urban policy making and transport planning.

4.1 What is Congestion?

If knowledge about congestion and its extent and intensity is important, then the first consideration is to define just what congestion is. Congestion is an integral part of a transport system, but its specific definition and identification are not immediately obvious. A number of different definitions of traffic congestion and the observed phenomena associated with it were reviewed by Taylor (1992). On the basis of this review three recurrent ideas that occurred in the various definitions of congestion were identified:

(1) congestion involves the imposition of additional travel costs on all users of a transport facility by each user of that facility;

(2) transport facilities (e.g. road links, intersections, lanes and turning movements) have finite capacities to handle traffic, and congestion occurs when the demand to use a facility approaches or exceeds the capacity, and

(3) congestion occurs on a regular, cyclic basis, reflecting the levels and scheduling of social and economic activities in a given area.

The following definition of congestion can be proposed for use in traffic studies: 'traffic congestion is the phenomenon of increased disruption of traffic movement on an element of the transport system, observed in terms of delays and queuing, that is generated by the interactions amongst the flow units in a traffic stream or in intersecting traffic streams. The phenomenon is most visible when the level of demand for movement approaches or exceeds the present capacity of the element and the best indicator of the occurrence of congestion is the presence of queues'. This definition recognises that the capacity of a traffic systems element may vary over time, e.g. when traffic incidents occur.

Thus congestion may always be present in any part of a transport system, but that the level of congestion may have to exceed some threshold value to be recognised. The threshold may be context-specific, for instance owing to the occurrence of incidents such as breakdowns, road works, or road crashes. Peak periods are recognised as prone to congestion, it must also be recognised that congestion can occur at other times, due to different traffic management regimes in place off-peak, or due to traffic incidents or unusual local traffic generating activity.

The investigation of any traffic planning or traffic management strategy requires the determination and possible subsequent monitoring of the level of congestion. Thus there is a need to collect and analyse data on congestion. Several measures can be used, and although the definition of traffic congestion would suggest that delay time and queue length are essential parameters, they are almost certainly not sufficient measures. The set of factors reflecting the level of congestion includes: delay times, the equitable distribution of delays between road users and traffic streams, reliability of travel time and travel costs, queue management, incident management, excessive energy consumption and additional emissions of air and noise pollution.

For strategic transport planning purposes a satisfactory definition of the level of congestion on a network component (e.g. a route, link or intersection turning movement) is the excess travel time incurred by a traveller when traversing that network component. Excess travel time is the additional travel time over and above the free flow travel time (c_0), which is the minimum amount of time required to cover the component - thus the excess travel time corresponds to the 'system delay' pertaining to the component under the given traffic conditions (Taylor *et al*, 1996). Further, we assume that travellers may be able to trade-off excess travel time (or indeed total travel time) for other components of the overall cost of travel on a trip. This requires the concept of a generalised cost of travel for a trip. The economic basis for the trade-off is illustrated by the theory of road pricing (e.g. May *et al*, 1996), and this trade-off is explored later in this paper.

4.2 Congestion Functions

A congestion function describes the relationship between the amount of traffic using a network element and the travel time and delay incurred on that element. The total travel time to traverse a network element is directly related to the traffic volume using that element. As volume increases, so delay, and hence travel time, increases. The rate of increase in travel time accelerates as volume approaches the capacity of the element. For most transport planning applications the network link is the typical level at which congestion functions are applied, but for traffic engineering applications function for lanes and movements may be more appropriate.

Traffic movement along a link in a network may be seen as consisting of two components. The first component is cruising, with traffic moving along the link largely uninterrupted (except for the possibility of side friction, say due to vehicle parking manoeuvres). Travel along the link may also be punctuated by points of interruption, say pedestrian crossings, bus stops and, most importantly, road junctions. For example, the junction at the downstream end of the link may dictate the traffic progression along the link. Movement through the interruption points can be handled using the methods for intersection analysis and queuing theory. What is also needed, particularly for urban areas or other places where congestion is expected, is a composite relationship that can include the two components simultaneously. A 'congestion function' (or 'speed-flow relation') may be used to describe the relationship between link flow and speed or travel time on a network link or road section containing a set of network elements. The most convenient way to represent a congestion function is in terms of the travel time on a link, in which c is the travel time on the link when it is carrying traffic at a flow rate of q, and the function includes a set of parameters that describe the physical and environmental characteristics of the link and the sources of interference impinging on the flow (e.g. opposing traffic, or parked vehicles). The function will start with a finite travel time (c_0) at zero flow, and the actual travel time then increases with volume. One useful form for a congestion function is the modified Davidson function (Tisato, 1991):

$$
\begin{aligned}
c &= c_0\left(1 + J\frac{x}{1-x}\right) && x < x_0 \\
c &= c_0\left(1 + J\frac{x_0}{1-x_0} + J\frac{x-x_0}{(1-x_0)^2}\right) && x \geq x_0
\end{aligned}
\qquad (4)
$$

where x = q/C is the 'volume-capacity ratio' (or 'degree of saturation'), and x_0 is a user-selected proportion, usually in the range (0.85, 0.95) as discussed by Taylor (1984).

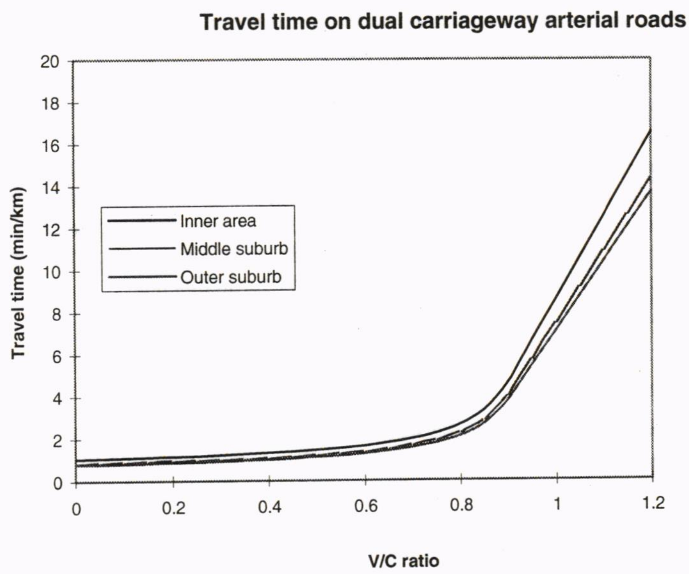

Figure 1: Davidson congestion functions for urban arterial roads in Australia

4.3 Fuel and Emissions Functions for Network Elements

Taylor (1996) indicated how a congestion function could be combined with fuel consumption and pollutant emissions models to generate link-based fuel and emissions functions for use in transport network modelling, in which fuel usage and emissions are related to link volume-capacity ratios. Figure 2 shows a set of fuel consumption functions, for unleaded petrol, corresponding to the congestion functions of Figure 1. Figure 3 shows the corresponding functions for carbon monoxide (CO) emissions. These functions were generated using the method described by Taylor (1996).

4.4 Generalised Cost of Travel

The concept of a generalised cost of travel has been widely used in transport planning, especially in relation to the analysis of modal choice. One particular formulation of generalised cost was proposed by Wigan (1976) for network studies of the interaction of road-based components of travel cost. This formulation is

$$g_A = \left(A + Bu + \frac{H}{u^2} \right) x_s + m \tag{5}$$

where g_A is the generalised cost of travel over a route or link of length x_s, u is the unit travel time on the link (time taken per unit distance), m is road toll, congestion charge or other direct out-of-pocket money cost incurred on the route or link, and A, B and H are parameters characterising the

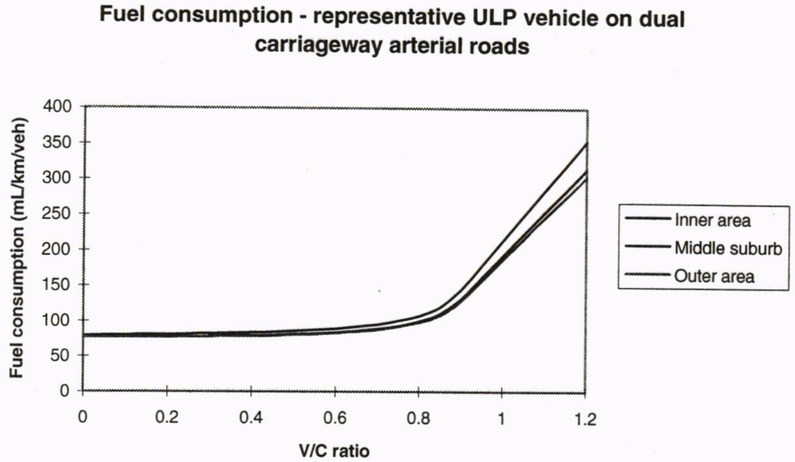

Figure 2: Fuel consumption (unleaded petrol) functions for urban arterial roads in Australia

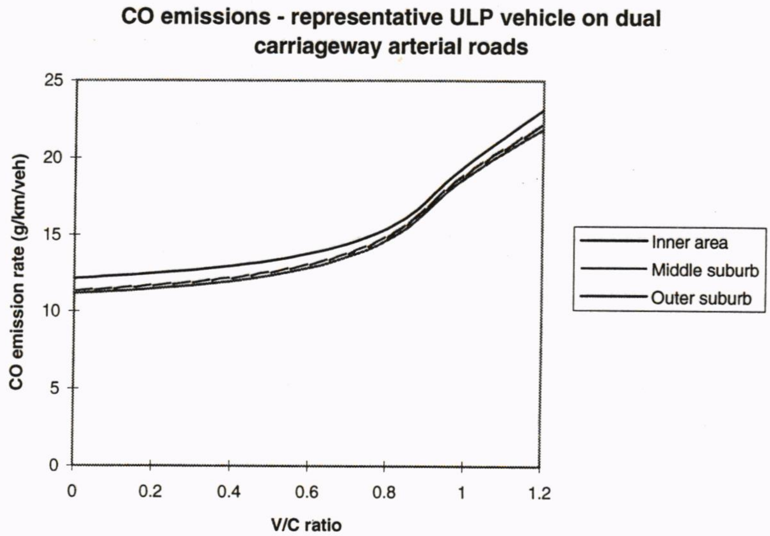

Figure 3: Carbon monoxide emissions from unleaded petrol, urban arterial roads in Australia

network and traveller behaviour. Parameter A represents the operating cost of the vehicle per unit distance while B can be taken as the traveller's valuation of travel time. Both A and B can be split into components: $A = A_t + A_c$ and $B = B_m + B_t + B_c$ from which a further breakdown of the cost function into policy related components can be made:

(a) $A_t + B_t u$ indicates the cost per unit distance due to taxation on fuel, tyres and oil;

(b) $A_c + B_c u$ indicates the operating cost per unit distance net of all taxation, and

(c) $B_m u$ is the money cost equivalent of travel time per unit distance

The term H/u^2 in equation (5) is associated with the fuel consumption and emissions of vehicles and can be taken as an approximation of the more detailed models for fuel and emissions introduced in section 4.3.

One of the applications of generalised cost of travel - and its inherent assumption that travellers may trade-off between travel time and money costs - is in the consideration of road pricing.

4.5 Road Pricing

Congestion provides a natural but partial restraining mechanism on travel demand. The additional costs (delays, queuing and inconvenience) resulting from congested conditions can act as a form of deterrent to the generation of further travel demand. However, there is widespread belief amongst transport planners that the congestion 'price' of itself is inefficient as a demand management tool. Individual drivers may not be fully aware of the true costs that they impose on other travellers and the transport system on the basis of congestion delays alone. Some other pricing signal is required to this end. Assuming that travellers will respond to a composite generalised cost (i.e. a total travel cost containing components from travel time, travel distance, out-of-pocket expenses, fuel cost, wear and tear, etc.) by trading-off the different cost components in their travel decision making, the further step is to impose a congestion tax, toll or road pricing charge on travellers in an intelligent, selective fashion (e.g. for travel on some parts of a network at some times of day)). There is a resurgence of interest in this topic, largely as a result of the new technological capabilities for vehicle identification and modelling (e.g. Jones and Hervik, 1992).

The economist's conceptual model for a congestion price is that of the demand-supply equilibrium and the relationship between the average travel cost on a link and the marginal cost. The average cost (G_l) is less than the marginal cost for all positive link volumes. A congestion charge could be imposed on motorists to enable them to meet their full marginal costs, this means the imposition of a congestion charge ΔG on each vehicle. The marginal travel cost on a link is g_m where

$$g_m = \frac{\partial G_T}{\partial q} = \frac{\partial (c(q)q)}{\partial q} \tag{6}$$

and G_T is the total travel cost on the link, given by $G_T = cq$. It then follows that for the Davidson function (c(x)) defined by equation (4), the marginal cost is given by

$$g_m = c_0\left(1 + \frac{2Jx}{(1-x)^2} - \frac{Jx^2}{(1-x)^2}\right) = c(x) + \frac{c_0 Jx}{(1-x)^2} \qquad x < x_0$$

$$g_m = c_0\left(1 + \frac{2Jx_0}{(1-x_0)^2} - \frac{J(2x-x_0)}{(1-x_0)^2}\right) = c(x) + \frac{c_0 Jx}{(1-x_0)^2} \qquad x \geq x_0 \tag{7}$$

Equation (7) enables the 'congestion tax' or 'road price' (ΔG) to be identified explicitly, given that $g_m(x) = c(x) + \Delta G$.

5. POLICY QUESTIONS AND OPTIONS

The main value of the proposed family of equilibrium assignment models may be seen in the area of transport policy evaluation. Consider a set of alternative policies that seek to optimise different characteristics of network travel performance, such as minimum overall travel (i.e. minimum overall VHT), or minimum total travel cost, overall delay, fuel usage or pollutant emissions, or peak hour trip spreading, or optimised transport/land use plans for given levels and distribution of travel demand. Each of these policies can be represented by a particular equilibrium assignment model, and the resulting network flow distributions can be compared with each other and with the individual travel time minimisation flow pattern (which can be taken as a macro-level simulation of existing travel and thus as a common starting point). The resulting solutions thus indicate the relative transport performance between the alternative policies, and the degree of similarity or difference between them can be examined.

The Wardrop principles may be treated as meeting different economic objectives for network travel, if travel time is taken as one possible alternative measure of travel cost. Thus they may be seen as particular cases of Jewell's assignment principle that the ultimate pattern of flow in a network will satisfy some explicit economic objective. For instance, direct substitution of generalised travel cost, fuel consumption or pollutant emission functions for a link travel-time volume congestion function yields assignment models that can generate traffic patterns corresponding to minimum fuel use or minimum pollution generation. Generalised cost functions including travel time, fuel consumption, pollutant emissions, tolls and road user charges, vehicle operating costs etc can also be proposed and solved. For example, a generalised travel cost function can be employed to study the effects of road user charges and tolls on the distribution of traffic in a network. Likewise, the composite fuel and emissions functions of the type described above can be used to derive system- and user- optimum traffic patterns for fuel consumption and pollutant loads. Another example is an equilibrium assignment model that includes variability of travel times as a factor for consideration in route choice. This situation might pertain to commuters who are required to be at a specified destination (e.g. their workplace) by a given time of day. The question then confronting those commuters is how to minimise the risk of arriving late (or indeed of arriving either too late or too early).

Other models may be developed by further applications of Jewell's principle. In the case of road pricing systems, for instance, use of marginal travel costs (e.g. equation (7)) rather than average travel costs (equation (4)) for all links in the network provides an appropriate network equilibrium model including individual choice in the presence of a perfect road pricing regime. Other cases of road pricing implementations may be more interesting? For example, what if the road pricing is only imposed on a subset of the links (e.g. in a downtown area or regional activity centre), and not across the whole network? In addition, there is the question of how a practical road pricing system might be implemented? Technological developments not withstanding, it seems unlikely that a perfect, real-time road pricing system, in which marginal costs are adjusted continuously in response to traffic flow variations, can be readily employed. Some simplified systems of imposing the 'congestion charge' ΔG are more likely to be used.

May *et al* (1996) described a number of alternatives for applying road pricing to real networks:

(1) road pricing based on charges for usage of road space, perhaps in a specified sub-area - the road pricing zone', which corresponds to the application of marginal travel costs on the road links in that area;

(2) cordon based road pricing, in which drivers are charged for entering the road pricing zone. The congestion charge is thus a fee imposed for traversing the links that feed in to the road pricing zone. This is the Singapore model;

(3) travel distance-based road pricing, in which a per unit distance charge is levied on each vehicle travelling along the links in the road pricing zone, and

(4) travel time-based road pricing, in which a per unit time charge is levied on each vehicle travelling along the links in the road pricing zone.

Of these alternatives, travel-time based road pricing is perhaps closest to the theoretical pricing regime but can be seen as inequitable. For instance, if drivers are held up in the road pricing zone due to some traffic incident inside (or even outside) the zone which is beyond their control or influence, then the charge they incur may be seen as unreasonable - an element of individual choice has been removed. Distance-based road pricing is then more equitable and can still be used to differentiate between different levels of travel activity in the road pricing zone. It is also easier to provide drivers with advance warning of the charges they will incur. This ease of advice is even more apparent for cordon-based pricing, but this system cannot differentiate between long and short journeys within the road pricing zone, nor for trips which are completely internal to the zone.

Route guidance is normally seen as a dynamic problem for which static models of traffic assignment may seem inappropriate. However, from the point of view of the transport planner or policy maker the advantage of a route guidance system may lie more in the resulting ability to influence the distribution of traffic load in a network, and an equilibrium model may then have a role. Specifically, the effect of the proportion of drivers receiving and using route guidance information will be of interest in determining appropriate transport policy and strategies for system operation.

6. FAMILY MEMBERS - MEET THE Z's

Using the Wardrop-Jewell principles and the various definitions of travel time, cost, fuel usage and emissions described previously we can define a family of equilibrium assignment models, in which each model represents a particular strategy or policy for the operation of the road transport network. The individual models form a family because they have a common structure, as constrained non-linear optimisation problems, through the continuity of flow constraint equations (equations (1)-(3)) and the general form of their objective functions. Each model has its own variation on the objective function (Z), but all of the models include the same constraint equations and lie in the same decision space.

6.1 Models including Inelastic Travel Demands

Inelastic travel demand is defined as a fixed O-D matrix. The starting point for assignment modelling is the well-known user equilibrium model for fixed travel demand, which is an expression of Wardrop's first principle (individual travel time minimisation). This model formulation provides a useful macroscopic simulation of travel on a metropolitan network. It is written as the following non-linear optimisation problem,

$$Z_{W1} = \min\left\{\sum_e \int_0^{q(e)} c_e(x)dx\right\} \tag{8}$$

subject to the continuity of flow constraints of equations (1)-(3).

The equivalent system-wide travel time minimisation problem, the network flow pattern satisfying Wardrop's second principle, may be written as a similar optimisation problem with objective function

$$Z_{W2} = \min\left\{\sum_{e} q(e)c_e(q(e))\right\} \tag{9}$$

with the same conservation of flow constraints.

6.1.1 Generalised travel cost models

Utilising Jewell's principle, a number of other equilibrium assignment models can be derived for both individual traveller or system-wide objectives. For example, direct substitution of link fuel nd emissions functions for the congestion function $c_e(q)$ would yield assignment models that could generate traffic patterns corresponding to minimum fuel use or minimum pollution generation. Generalised cost functions including travel time, fuel consumption, pollutant emissions, tolls and charges, road pricing, etc can also be proposed and solved. For instance, the user equilibrium assignment problem for overall travel cost based on Wigan's generalised travel cost function (equation (5)) is

$$Z_{JU} = \min\left\{\sum_{e} \int_{0}^{q(e)} g_A(e,x)dx\right\} \tag{10}$$

subject to the continuity of flow constraints of equations (1)-(3), while the equivalent system-wide travel cost minimisation problem is

$$Z_{JS} = \min\left\{\sum_{e} q(e)g_A(e,q(e))\right\} \tag{11}$$

also subject to the constraint equations (1)-(3). Road tolls and general charges (e.g. vehicle operating costs are included in the models defined by equations (10) and (11). Models to study the network traffic effects of road pricing models are better treated separately, especially when alternative road pricing regimes are contemplated.

6.1.2 Road pricing models

Equation (6) indicated the relationship between the marginal cost of travel and the average cost of travel on a link. Substitution of the marginal cost of travel in the objective function for the user-minimisation equilibrium assignment (equation (8)) shows that it is equivalent to solving the system-wide travel time optimisation problem (equation (9)). This is indicated below, starting with equation (12) which is the user equilibrium objective function based on marginal travel costs (equation (6)).

$$Z_{RP} = \min\left\{\sum_{e} \int_{0}^{q(e)} g_m(e,x)dx\right\} \tag{12}$$

Consider the integral in the righthand side of equation (12). Given the definition of marginal travel cost in equation (6), this can be written as

$$\int_0^{q(e)} g_m(e,x)dx = \int_0^{q(e)} \frac{\partial(c_e(x)x)}{\partial x} = \left[c_e(x)x\right]_0^{q(e)} = c_e(q(e))q(e)$$

from which

$$Z_{RP} = \min\left\{\sum_e \int_0^{q(e)} g_m(e,x)dx\right\} = \min\left\{\sum_e c_e(q(e))q(e)\right\} = Z_{W2}$$

Again, the continuity of flow constraints (equations (1)-(3) apply. Thus the road pricing solution yields the system-wide minimum travel time (VHT) distribution of traffic, as long as marginal travel costs are applied on all links in the network. In the case where road pricing might only be applied to a subset of the roads (e.g. the central business district) whereas other links remained in their 'normal' state, then the road pricing solution is found by using a composite objective function (e.g. derived from equation (12) for a user equilibrium formulation). Different models apply to the four road pricing schemes examined by May *et al* (1996) and described above. In the following discussion these models are, for purposes of brevity, described as the Leeds models. Assume that the full set of network links then consists of three subsets:

(1) links w which lie wholly within the road pricing zone;
(2) links *l* which cross the cordon around the RP zone, and
(3) all remaining links e, which are external to the road pricing zone

The first Leeds road pricing model, that of charging for the use of road space in the road pricing zone, is based on the application of marginal costs to all links within that zone. In this model the marginal cost functions apply to those links inside the road pricing cordon area and the average cost functions applies to all other links. This model is thus represented by the objective function

$$Z_{M1} = \min\left\{\sum_e \int_0^{q(e)} c_e(x)dx + \sum_l \int_0^{q(l)} g_m(l,x)dx + \sum_w \int_0^{q(w)} g_m(w,x)dx\right\} \tag{13}$$

subject to the continuity of flow constraints (equations (1)-(3)).

The second Leeds model is cordon-based road pricing, in which drivers are required to pay a fixed charge (m_l) when they cross the cordon line to enter the road pricing zone. In this case the corresponding objective function is

$$Z_{M2} = \min\left\{\sum_e \int_0^{q(e)} c_e(x)dx + \sum_l\left[\int_0^{q(l)} c_l(x)dx + q(l)m_l\right] + \sum_w \int_0^{q(w)} c_w(x)dx\right\} \tag{14}$$

The third Leeds model is that of travel distance-based road pricing. In this case each vehicle is charged for the distance travelled in the road pricing zone, and the objective function Z_{M3} is thus given by

$$Z_{M3} = \min\left\{\sum_e \int_0^{q(e)} c_e(x)dx + \sum_l\left[\int_0^{q(l)} c_l(x)dx + q(l)\eta_l L_l\right] + \sum_w\left[\int_0^{q(w)} c_w(x)dx + q(l)\eta_w L_w\right]\right\} \tag{15}$$

where L_w is the length of link w and η_w is the charge per vehicle per unit distance in the road pricing zone.

The final Leeds road pricing model is for time-based road pricing. In this case each vehicle is charged at a rate of ϕ_w for each unit of time that it spends in the road pricing zone. The corresponding objective function may be written as

$$Z_{M4} = \min\left\{\sum_e \int_0^{q(e)} c_e(x)dx + \sum_l \left[\int_0^{q(l)} c_l(x)dx + q(l)c_l(q(l))\phi_l\right] + \sum_w \left[\int_0^{q(w)} c_w(x)dx + q(l)c_w(q(w))\phi_w\right]\right\} \quad (16)$$

6.1.3 Perceived travel time models

A further member of the equilibrium assignment family can be constructed to include considerations of travel time variability in route choice decisions. This situation might pertain to commuters who are required to be at a specified destination (e.g. their workplace) by a given time of day. The question then confronting those commuters is how to minimise the risk of arriving late (or indeed of arriving either too late or too early).

Variability in travel times can be caused by a variety of circumstances, including regular condition-dependent variations (e.g. cyclic fluctuations in levels of demand by hour of day or day of week), irregular condition-dependent variations (e.g. accidents, breakdowns or weather conditions) and random variations. Focussing on the latter two causes, Herman and Lam (1974) indicated that, if travel times on individual links were independent random variables, the standard deviation of the distribution of link travel times σ was proportional to the square root of the mean travel time (c),

$$\sigma = \gamma\sqrt{c} \quad (17)$$

where the constant γ is known as the 'travel time variability ratio'. Richardson and Taylor (1978) tested this model with data collected in Melbourne, Australia and found a relationship between travel time variability and the level of congestion. Travel time variability can enter into individuals' route choice decisions. For example, when making departure time decisions, travellers may implicitly apply a measure of travel time composed of the mean travel time plus a certain number of standard deviations, i.e. $c + k\sigma$ where the constant 'k' may depend on trip purpose and the individual's response to the risk of arriving late. Another factor that can be incorporated in a perceived cost model is the impact of tolls and other charges, possibly as a trade-off for travel time variability? By combining these components a general expression can be written for the perceived travel time (c_p) on a link, yielding the expression

$$c_p = c + k\gamma\sqrt{c} + \frac{m}{\lambda} \quad (18)$$

where m is the money cost involved in using the link and λ is the value of time (money units per unit time). Using equation (18) as the expression for perceived link travel time, the following user equilibrium assignment model results, which accounts for travel time variability:

$$Z_p = \min\left\{\sum_e \left[q(e)\int_0 \left(c_e(x) + k\gamma_e\sqrt{c_e(x)}\right)dx + \frac{1}{\lambda}q(e)m_e\right]\right\} \quad (19)$$

subject to the continuity of flow constraints set by equations (1)-(3). This model permits study of both the effects of travel time variability and tolls for specific routes, given knowledge of the γ factors for different road classes (Richardson and Taylor, 1978).

6.1.4 Fuel and emissions models

The individual assignment models represented by the specific objective functions given above each represent an optimisation of some specific performance measure of traffic network performance. It follows that it is possible to compare the network states between models and to obtain values of the other performance measures from each specific model. This may be used to compare the effectiveness of different policies. For instance, given the flow pattern corresponding to any of the above traffic assignment models, the total fuel consumption and emissions generated can be estimated using link fuel/emissions relationships of the types shown in Figures 2 and 3.. In addition, assignment models for direct optimisation of fuel usage or pollutant emissions can also be generated, using the fuel emissions functions as link cost functions. These models might be used to investigate network flow patterns where (say) drivers adopt route choice strategies to minimise fuel consumption, or a community seeks a flow pattern minimising the overall rate of pollutant emissions. The objective function Z_E for individual minimum fuel consumption would be

$$Z_E = \min\left\{\sum_e L_e \int_0^{q9e)} E_{se}(x)dx\right\} \tag{20}$$

where L_e is the length of link and E_{se} is the fuel consumption rate per unit distance. The objective function for minimum total fuel consumption or total pollution emissions across the network (Z_X) would be

$$Z_X = \min\left\{\sum_e E_{se}(q(e))q(e)L_e\right\} \tag{21}$$

6.1.5 Route guidance

Route guidance is of great interest in contemporary transport research, for example, see Mahmassani and Chen (1991), Watling and Van Vuren (1993) and Emmerlink *et al* (1995). A particular question that is raised is the actual value of route guidance information for the community as a whole, especially under conditions of recurrent congestion. Other issues emerge about the propensity of drivers to accept and use route information, and the technologies to be used in providing information in real-time. A likely scenario is that traffic streams on a network where information is provided may well be divided into two sets of drivers: those using the supplied information and those unaware of or ignoring it. From the transport policy viewpoint, one possible advantage of route guidance information is the opportunity to direct the network flow pattern towards a system-optimisation objective, such as minimum VHT (Wardrop's second principle) or a similar objective such as minimum overall fuel consumption or pollutant emissions. Given the assumption that the static equilibrium assignment flow pattern from Wardrop's first principle represents a reasonable macro-level simulation of traffic behaviour on a network in the absence of route guidance information, the issue becomes one of steering the flow pattern towards the system-wide objective. In this case, the proportion of drivers making use of the information is an important parameter in determining the actual state of the traffic system.

A planning model can be assembled, involving a combination of the user minimisation equilibrium assignment and a system-wide optimum. Assume that a proportion of drivers (ω) make use of route guidance information in choosing their routes. This information is directed towards achieving the minimum VHT in the network (i.e. the objective function Z_{W2} given by equation (9), although alternative objectives such as Z_{JS} (equation (11)) or Z_X (equation (21)) are also possible candidates. These are the 'informed' drivers. The remaining (1 - ω) proportion of drivers can be assumed to adopt an individual travel time minimisation strategy in the face of the route choices made by each other and by the informed drivers. The following composite model then emerges:

(1) for the overall travel demand represented by the O-D matrix T, determine the system-wide flow pattern corresponding to the chosen planning objective, say $q_{W2}(e)$ as resulting from the equilibrium assignment model with objective function Z_{W2} (equation (9)). Scale the link flows, $\omega q_{W2}(e)$, to obtain the volumes of informed drivers on each link e of the network.

(2) Then compute the user travel time equilibrium assignment for the 'uninformed' drivers, given $\omega q_{W2}(e)$ as the base link flows. This requires assignment of the uninformed travel demand represented by the scaled O-D matrix (1 - ω)T to the network, yielding the uninformed link flows $q_U(e)$ on the network. These flows are determined by solving the assignment problem defined by

$$Z_{Uw} = \min\left\{\sum_e \int_{wq_{W2(e)}}^{wq_{W2}(e)+q_U(e))} c_e(x)dx\right\} \tag{22}$$

subject to the constraint equations (1)-(3). Alternatives to the objective of equation (22) might be used, involving, for example, perceived travel times or generalised costs.

(3) Compute the total link flows q(e),

$$q(e) = \omega q_{W2}(e) + q_U(e) \qquad \forall\ e \tag{23}$$

to provide the flow distribution on the network when a proportion of ω drivers use route guidance information. Then compare this solution with the other equilibrium models.

6.1.6 Summary

This discussion has defined a number of alternative equilibrium assignment solutions, each of which can be regarded as a member of a family of similar solutions. The properties of the individual members, rated in terms of network parameters such as mean travel time, distance, speed, delay, fuel consumption and pollutant emissions, may be compared to test for the differences between them and to rate the success or otherwise of particular transport policies (for which specific objective functions can be formulated). Some initial indication of the comparisons between alternative policies is given later in this paper. A basic characteristic of the model family described above is that it is for fixed travel demand. Policies concerning travel demand management or land use-transport interaction might well be directed at changing the pattern of travel demand, in time and/or space, and so consideration also needs to be given to models incorporating elastic travel demands.

6.2 Models involving Elastic Travel Demands

Elastic demand equilibrium assignment models are of growing relevance and importance (Hills, 1996). Two forms may be considered. The first is for time-elastic travel, where the total travel demand over an extended time period is fixed in space, and some drivers can choose their departure times in particular intervals within the overall time period. The second model is where the total number of trips is fixed, but travellers have the ability to set their origins and destinations in response to congestion on the network. This is a long-run space-elastic model, with applications in land use-transport interaction studies. With the growing interest in trip timing decisions, especially for peak spreading and travel demand management considerations, an obvious future step is the integration of the time-elastic and space-elastic models into a single elastic-demand model.

6.2.1Time-Elastic Travel Demand

Traffic models accounting for trip timing and peak spreading behaviour are useful for studies of the impacts of travel demand management policies and time-dependent road pricing systems (e.g. see D'Este (1985) and Alfa (1989)). Matsui and Fujita (1996) derived a model for individual travel time minimisation including departure time choice, which fits closely to the equilibrium assignment model family described previously. The model is based on the following assumptions:

(1) the total travel demand for a given period of time and represented by the O-D matrix $\{T_{ij}\}$ may be split into two components, a set of fixed-time O-D matrices $\{H_{ij}^{n}\}$ for travellers constrained to depart in a given time interval n, and a time-variable O-D matrix $\{F_{ij}\}$ representing travellers free to choose a departure interval, and

(2) those travellers able to select a departure interval have a probability P_n of choosing to depart in time interval n given by a multinomial logit model.

Matsui and Fujita (1996) applied their model to Toyota City, Japan, to study the impacts on congestion and journey times of a flexitime system and a road pricing system (in which a fixed charge was levied for travel on the entire network in one 60 minute period, 07:30-08:30, with no charge in other time intervals). They found a five per cent reduction in total VHT under the flexitime regime, and a two per cent reduction under the road pricing system.

6.2.2 Space-Elastic Travel Demand

In the case that travel demand is regarded as elastic, i.e. the trip distribution (destination choice) may vary depending on the congestion levels in the network, then an alternative model formulation is in order. The combined distribution-assignment model proposed by Evans (1976) and explained by Horowitz (1989) provides an equivalent formulation to the equilibrium assignment model, and may be solved by a similar mathematical programming approach. Such a model may be treated in identical fashion to the equilibrium assignment model for fixed travel demand. More complex models including modal choice have also been developed, e.g. Tatineni *et al* (1995).

Other recent developments in space-elastic models include the SUSTAIN model (Roy *et al*, 1996) which is a land use-transport interaction model. SUSTAIN includes user equilibrium assignment as part of a combined assignment-trip distribution-modal split model connected to a housing and employment location model. It has been used for studies of differing residential densities in very large cities (with a population of four million or more), as described in Roy *et al* (1995).

6.3 Summing Up

This paper has defined an extensive family of traffic assignment models based on the equilibrium assignment model. This family offers the capability to study the effects on the distribution of flows on a network under different transport policy and control systems. Given their common basis, direct comparisons between outputs of the different models can be made. This enhances the utility of the model family for planning and evaluation studies.

The next question is to consider just how similar or otherwise are the flow patterns that result from the different models. Comprehensive investigation of the models is not yet available, but some preliminary results are, and these are described in the next section.

7. WHAT TO DO NEXT - EMPLOYING THE Z's

Preliminary results involving comparisons between some of the individual models in the equilibrium family have been found, for a 'sketch-planning' network representing the primary road system for Melbourne, Australia. This coarse network of some 300 nodes, 1200 links and 50 zone centroids represents the principal road corridors for that city. A representation of it is shown in Figure 4. A peak period O-D matrix was developed from census journey-to-work data, with this matrix divided into four submatrices. Each submatrix represented trips with origins in different regions (inner, middle north and west, middle south and east, and outer) of the metropolitan area.

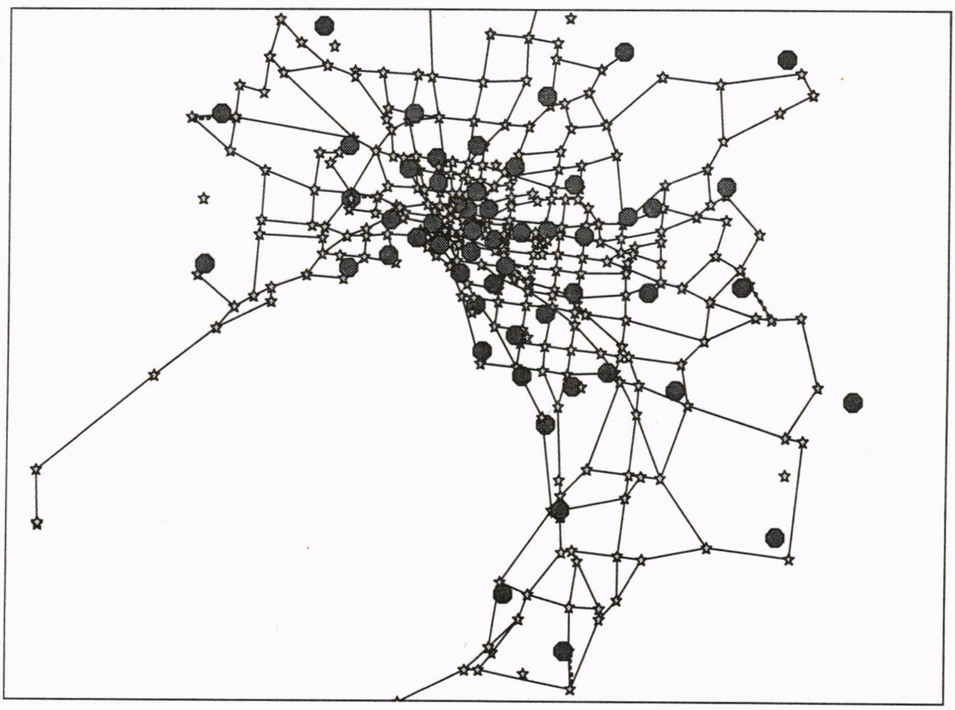

Figure 4: The Melbourne strategic road network

Five alternative assignment models were applied to the network. The first of these was individual travel time minimisation, which was taken as a simulation of the actual traffic flows on the network and was used as a datum for comparisons with four particular planning objectives:

(1) minimum total VHT;
(2) minimum fuel consumption by individual drivers;
(3) minimum total fuel consumption, and

(4) minimum total emission of pollution.
Each of these planning objectives was represented by its own equilibrium assignment model, by choosing the appropriate objective function.

Summary results for the user-optimum travel time assignment (Z_{W1}) are given in Table 4.

Table 4: Summary network travel statistics, user-optimal assignment in Melbourne, 1991 morning peak hour

Trip class	Veh-km of travel (VKT) (x 10^6)	Veh-h of travel (VHT)	Total trips	Mean speed (km/h)	Mean trip distance (km)	Mean trip time (min)
Inner	0.1237	4 629	15 228	26.7	8.1	18.2
Mid N and W	0.6452	24 031	46 369	26.8	13.9	31.1
Mid S and E	0.9663	41 222	61 208	23.4	15.8	40.4
Outer	1.2168	37 993	50 474	32.0	24.1	45.2
Total	**2.9520**	**107 875**	**173 279**	**27.4**	**17.0**	**37.4**

Relative network travel statistics are shown in Table 5, for the four alternative planning objectives compared to the simulated flow distribution. Differences of up to nine per cent may be seen in this table. The objectives of minimum VHT and minimum total pollution tend to generate higher network speeds. Minimum VHT scarcely affects aggregate VKT, the changes are in the amounts of travel time on the network, whilst minimum total pollutant emissions drives down both VKT and VHT. The two fuel consumption objectives tend towards lower network speeds, with both VKT and VHT reducing. The effect is greater for minimum total fuel consumption than for minimum individual fuel consumption. Although VKT reduces significantly for this latter case compared to the base case (individual travel time minimisation), the VHT for both these models are similar.

Some regional effects are apparent in the table, and these are most apparent for the minimum individual and total fuel consumption models.

The overall impacts of the four policies suggested by the assignment models are a possible seven per cent reduction in VHT for the minimum total VHT (and no change in VKT), a five per cent reduction in VKT for individual minimum fuel consumption, a four per cent reduction in VHT and a six per cent reduction in VKT for minimum total fuel consumption, and a seven per cent reduction in VHT and three per cent reduction in VKT for minimum total pollution generation. Note that all other factors, e.g. the network configuration and design standard, the spatial (and temporal) distribution of travel demand), and transport/vehicle technology are exactly the same in all cases.

These results are preliminary only, but they do suggest that there are some interesting differences in network performance between the alternative objective functions. Network configuration and topology may well influence the model results. For example, the Melbourne network is sparse, and may not contain a wide range of alternative routes. More work is required, using different networks at different levels of detail, to suggest possible trends in the performance of the individual assignment models, and to allow comparisons of assigned link flows for the different transport policy objectives.

Table 5: Relative network travel statistics for different assignment models in Melbourne 1991 morning peak, compared to the individual travel time minimisation model

Policy Objective	Trip Class	VKT	VHT	Mean Speed
Minimum VHT	Inner	1.00	0.98	1.02
	Mid N and W	0.99	0.95	1.05
	Mid S and E	0.99	0.91	1.09
	Outer	1.00	0.94	1.07
	Total	**1.00**	**0.93**	**1.07**
Minimum Individual Fuel Consumption	Inner	0.97	1.01	0.96
	Mid N and W	0.92	1.00	0.93
	Mid S and E	0.95	0.98	0.97
	Outer	0.95	1.01	0.95
	Total	**0.95**	**1.00**	**0.95**
Minimum Total Fuel Consumption	Inner	0.97	0.98	0.99
	Mid N and W	0.92	0.98	0.94
	Mid S and E	0.94	0.94	1.00
	Outer	0.95	0.97	0.98
	Total	**0.94**	**0.96**	**0.98**
Minimum Total Pollution Generation	Inner	0.99	0.97	1.01
	Mid N and W	0.96	0.94	1.02
	Mid S and E	0.98	0.91	1.07
	Outer	0.98	0.95	1.04
	Total	**0.97**	**0.93**	**1.04**

8. CONCLUSIONS

A family of traffic assignment models can be established, based on a common foundation and thus suitable for comparative analysis of alternative networks and transport policies. The model family starts with the well-known individual travel time minimisation model (Wardrop's first principle) and the system-wide minimum vehicle-hours of travel (Wardrop's second principle). With the addition of generalised and perceived travel cost functions and fuel and emissions relationships, the model family offers a useful means to examine the ways in which variations in vehicle fleet composition, travel demand patterns, vehicle operating costs, road user charges and tolls, and congestion levels affect network performance. Thus traffic network models sensitive to transport and land use planning objectives can be established and applied to examine the effects of alternative policies. Some preliminary results have been found, indicating that there are differences in the flow patterns resulting from the different objectives. Further investigations are needed to explore the wide variety of alternative assignment models made available in the family of equilibrium models, and to make comparisons between the resulting outputs of those models in terms of link flows and network performance parameters such as travel times, fuel consumption and pollutant emissions.

REFERENCES

Alfa, A S (1989). Departure rate and route assignment of commuter traffic during peak periods. **Transportation Research B 23B** (4), 337-344.

D'Este, G M (1985). The effect of staggered working hours on commuter trip durations. **Transportation Research A 19A** (1), 109-117.

Emmerlink, R H M, Axhausen, K W, Nijkamp, P and Rietveld, P (1995). Effects of information in road transport networks with recurrent congestion. **Transportation 22** (1), 21-53.

Evans, S P (1976). Derivation and analysis of some models for combining trip distribution and assignment. **Transportation Research 10** (1), 37-57.

Herman, R A and Lam, T (1974). Trip characteristics of journeys to and from work. In Buckley, D J (ed) **Transportation and Traffic Theory**. A H and A W Reed, Sydney, 57-85.

Hills, P J (1996). What is induced traffic? **Transportation 23** (1), 5-16.

Horowitz, A J (1989). Tests of an ad hoc algorithm of elastic-demand equilibrium traffic assignment *Transportation Research B 23B* (4), 309-313.

Jewell, W S (1967). Models for traffic assignment. **Transportation Research 1**, 31-46.

Jones, P M and Hervik, A (1992). Restraining traffic in European cities: an emerging role for road pricing. **Transportation Research A 26A** (2), 133-145.

Mahmassani, H S and Chen, P S-T (1991). System performance and user response under real-time information in a congested traffic corridor. **Transportation Research Record 1306**, 69-81.

May, A D, Milne, D, Smith, M, Ghali, M and Wisten, M (1996). A comparison of the performance of alternative road pricing systems. In Hensher, D A,King, J and Oum, T (eds), **World Transport Research 3: Transport Policy**. Elsevier, New York, 335-346.

Matsui, H and Fujita, M (1996). Modelling and evaluation of peak hour traffic congestion under flextime and road pricing systems. In Hensher, D A,King, J and Oum, T (eds). **World Transport Research 2: Modelling Transport Systems**. Elsevier: New York, 183-193.

Richardson, A J and Taylor, M A P (1978). Travel time variability on commuter journeys. **High Speed Ground Transportation Journal 12** (1), 77-99.

Roy, J R, Marquez, L O and Brotchie, J F (1995). A study on urban residential density vs transport energy consumption. **Journal of the Eastern Asia Society for Transportation Studies 1** (1), 11-24.

Roy, J R, Marquez, L O, Taylor, M A P and Ueda, T (1996). SUSTAIN - a model investigating sustainable urban structure and interaction networks. In Hayashi, Y and Roy, J R (eds). **Transport, Land Use and the Environment**. Kluwer: Dordrecht, 125-145.

Tatineni, M R, Lupa, M R, Englund, D B and Boyce, D E (1995). Transportation policy analysis using a combined model of travel choice. **Transportation Research Record 1452**, 10-17.

Taylor, M A P (1984). A note on using Davidson's function in equilibrium assignment. **Transportation Research 18B** (3), 181-199.

Taylor, M A P (1996). Incorporating environmental planning decisions in transport planning: a modelling framework. In Hayashi, Y and Roy, J R (eds). **Transport, Land Use and the Environment**. Kluwer: Dordrecht, 337-358.

Taylor, M A P, Young, W and Bonsall, P W (1996). **Understanding Traffic Systems: Data, Analysis and Presentation**. Avebury Technical Books: Aldershot.

Tisato, P M (1991). Suggestions for an improved Davidson's travel time function. **Australian Road Research 21** (2), 85-100.

Wardrop, J G (1952). Some theoretical aspects of road traffic research. Road Paper 36, **Proceedings of the Institution of Civil Engineers 1** (2), 325-378.

Watling, D and Van Vuren, T (1993). The modelling of dynamic route guidance systems. **Transportation Research C 1C** (2), 159-182.

Wigan, M R (1976). The estimation of environmental impacts for transport policy assessment. **Environment and Planning A 8**, 125-147.