

TRANSPORT NETWORKS AND THEIR USE: HOW REAL CAN MODELLING GET?

Richard E Allsop
Centre for Transport Studies
UCL
rea@transport.ucl.ac.uk

Abstract

The context for modelling of transport networks and their use is set out. Such modelling is based on the representation of the transport system by nodes and links, and the characteristics of this representation that are needed to reflect the principal realities of the system are outlined. Characteristics of the use of the system that need to be reflected are described in terms that relate to movement of people, but are also relevant mutatis mutandis to the movement of goods. Purposes of modelling in this context are set out and the evolution of transport network modelling is described, starting from the basic traffic assignment model and discussing its generalisations and extensions in the search for greater realism - first in steady state modelling for fixed demand, and then considering variable demand and time-dependence. Further progress towards appropriate realism is seen as requiring communication and cooperation between modellers and the users of the models.

1 Introduction

The study of transport is concerned with the transport system, the purposes it serves and the external effects of the system and its use. The transport system comprises transport infrastructure, vehicles, equipment, how they work in practice (including how people operating them or using them at the time behave), and how their operation and use are managed - by means of design, information, regulation, control and pricing.

Transport is not an end in itself: it is the means by which people are able to engage in activities that require people themselves and material goods to be in different places at different times. It is a means of overcoming physical separation for the purpose of economic, social cultural and personal activity. It is therefore necessary to consider the transport system in relation to the activities it serves - their nature, timing and location. A framework for doing so is shown in Figure 1 (adapted from Allsop [3]).

The prevailing desires for activity interact with the prevailing transport system through the choices made by users of the system, to produce a prevailing pattern of movement and associated elements of the direct costs of movement. Perceptions of these costs in turn feed back into users' choices in a process of short term equilibration.

Direct costs of the pattern of movement are of course by no means its only consequence. Movement and the provision of the transport system in which it takes place have many other economic effects, and widespread and pervasive social and environmental effects which are too familiar to need spelling out here.

The direct costs of movement and these wider effects, especially dissatisfaction with the ones people do not like, give rise to longer term feedback in the form of motivation to change the transport system itself on the one hand and motivation to alter the mix, location and timing of activities, including where to live, on the other.

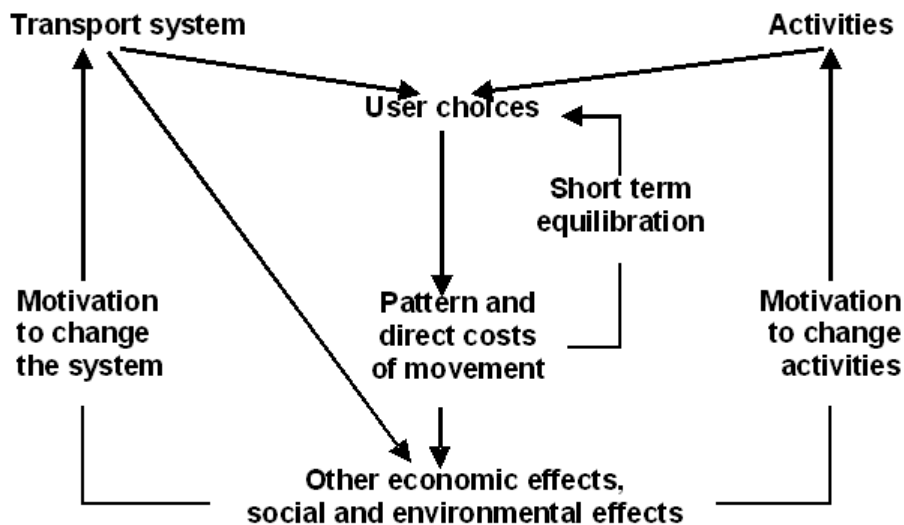


Figure 1: A framework for the study of transport

Transport networks are thus part of a system which is on the one hand physically and technologically complex and strongly influenced by physical geography, and on the other hand has people both as crucial elements in its operation and as one of the principal recipients of its impacts, for better and for worse - and the people in question are not (as is the case with some other complex systems) some trained personnel or clearly identified clientele, but the population as a whole.

The physical parts of the system are subject to the laws of mathematical and physical science, and the human parts to the laws of economics and to the law of large numbers in the probabilistic characterisation of whole populations notwithstanding the unpredictability of the behaviour of any one individual.

2 Transport networks...

Even in prehistoric times, transport by land tended to be concentrated on well-defined tracks, and in earliest history this tendency was reinforced by the construction of roads and bridges. The building of ports led to a similar tendency, as far as navigational capability and weather allowed, in maritime transport. The formalisation of land ownership and rights of way, and the development of inland waterways, railways and modern road transport, have defined land transport networks ever more precisely. Modern shipping in crowded waters navigates so-called motorways of the sea, and the requirement for air traffic control constrains civil aviation to corridors in the sky.

For the purposes of modelling, it is therefore natural for transport systems to be represented by networks of links along which traffic - in the form of people, animals, vehicles or craft - is modelled as flowing, and nodes at which flows of modelled traffic merge or diverge, or may enter or leave the model network.

It is the role of links and their incidence with the nodes, together with quantities and functions associated with the links, to represent:

- the topology of the transport system being modelled (though not usually its geography or topography - the innovative triumph of Harry Beck's iconic map of the London Underground (see for example Garland [12]) lay in letting go of these; and
- the operational characteristics and local management of traffic on each length of road, track or other right of way.

The relevant topology comprises for each link:

- the direction of travel along the link;
- from which upstream links or other sources traffic can enter the link; and
- to which downstream links or other sinks traffic can leave the link. Enough nodes and links must be defined to
- enable all movements that are both physically possible and permitted by regulations in the real system to take place in the model;
- prevent any movement that is physically impossible or prohibited by regulations in the real system from taking place in the model; and
- enable all costs of any kind experienced by each user of the system to be reflected correctly in the costs of traversing the links that make up their modelled route.

To achieve this requires skill, ingenuity and rigour in network building that is easily forgotten by those whose main concern is rightly with the theoretical analysis of the model network and movement within it. The operational characteristics of each link may comprise those of the following that are relevant to the system being modelled and the technique and purpose of the modelling:

- a maximum average rate at which traffic can pass along the link;
- a maximum average rate at which traffic can leave the link;
- a maximum accumulation of traffic that can be accommodated on the link;
- rules and parameters of any mechanism for control of movement along the link;
- parameters of the mechanism by which traffic moves along the part of the link where it is unimpeded by waiting traffic (perhaps based on a relationship between speed and concentration of traffic);
- rules and parameters of a mechanism by which traffic waits if necessary to leave the link and then does so (which may depend on how and how much traffic leaves certain other links);
- relationship between the time taken to traverse and leave the link and the associated operating costs on the one hand, and the rate at which traffic enters the link and the accumulation of traffic on the link on the other;
- the prevailing rates at which traffic enters the link from each upstream link or other source;
- the prevailing rates at which traffic leaves the link for each downstream link or other sink;
- the rules determining the amount that each unit of traffic is charged for passing along the link; and
- the way in which information is provided to users and potential users of the link.

Such characteristics need to be known in order to model traffic using the link in any given scenario concerning demand for use of the transport system being modelled. Many of these characteristics are expressed in a link performance function (Patriksson [17]) relating the perceived losses incurred in traversing the link to the rates at which traffic passes along this and other links.

Other operational characteristics describe the effects of use of the link and are concerned with interpretation of the modelled use in relation to purposes for which the modelling is being carried out. Some examples are:

- the rate of occurrence of traffic accidents on the link;
- the nature and amounts of noise generated by traffic on the link;

- the nature and amounts of exhaust emissions generated by traffic on the link;
- the nature and degree of visual intrusion resulting from traffic on the link; and
- the revenues accruing to the operating organisation and the exchequer from charging for use of the link.

Many of the characteristics are random variables or descriptions of random processes. What is used in modelling is often a single parameter (usually a mean) and sometimes also relevant variances and covariances, but deciding which parameters to use and how to estimate their values requires understanding of the randomness concerned and sometimes a good deal of analysis of the random process. When a steady state is being modelled, everything is in some appropriate way averaged over time. In dynamic modelling, evolution of the real system over part of a day (perhaps between a few minutes and a couple of hours), or over a sequence of days, is modelled, but with short-term (eg second-to-second) variability averaged out.

3 . . . and their use

The crucial role of people in the operation of the transport system has already been reflected in small part implicitly in some of the operational characteristics of links, notably in mechanisms for moving along and leaving a link. But this is only the tip of an iceberg.

In what follows, use of transport systems is discussed in terms of people making journeys. Some of it goes over *mutatis mutandis* to goods being moved, but with the important difference that goods once consigned can be moved as the carrier within reason sees fit, subject only to the law and to achieving secure and prompt delivery. Travellers, in contrast, remain active participants in the transport system throughout their journeys. And whereas intelligence imparted to packets of data in telecommunications systems may be known completely by the system operator, travellers in transport systems are individual and often apparently idiosyncratic humans whose perceptions, preferences and behaviour are hard to plumb, and even harder to influence.

In modelling the use of a transport system represented by a network of nodes and links, travellers are typically represented as having available to them certain points of entry to and exit from the network, each represented by a node. In many scenarios, where most or all of the transport system concerned is being modelled, these points represent the origins and destinations of the travellers' journeys, and their whole journeys are being modelled. In other scenarios, where only part of the system is being modelled, while some of the entry and exit nodes may represent origins and destinations of journeys, others may represent intermediate points at which travellers enter or leave the modelled part of the system, so that only parts of these travellers' journeys are modelled.

Consider first a number of travellers represented as entering the network at a certain entry node and leaving it at a certain exit node. Unless the network is unusually small and simple, they will have available to them in principle a large, probably countably infinite, number of routes through the network, each route consisting of a succession of links identifiable from the topology of the network. Routes which contain closed loops are easily excluded, but the number remaining is still large, and even when attention is confined to routes which are in some sense plausible choices in terms of network topology and the operating characteristics of the links, these may still be quite numerous. The travellers are modelled as being distributed over some set of these routes.

Modelling of the number of travellers and their distribution over the relevant routes can be helped by knowledge of or assumptions about:

- the operational characteristics of all the links in the light of the routes used by these and all other travellers;
- use by operating organisations of management tools such as provision of information, traffic control and charging for use

- travellers' knowledge and perceptions of the operational characteristics and system management;
- travellers' perceptions and relative valuations of elements of the losses they incur in making their journeys; and
- travellers' preferences concerning conditions of travel (for example between means of transport or between a longer but relatively uncongested route and a shorter but congested one).

Now consider how the numbers of travellers between the various entry nodes and exit nodes arise. Let the number of entry and/or exit nodes be Z in all. The a $Z \times Z$ matrix formed by the numbers of travellers from each entry node to each exit node is called a demand matrix.

In one commonly modelled type of scenario the demand matrix is taken as a given and fixed input to the modelling. In this case it does not matter whether the entry and exit nodes represent origins and destinations or intermediate points on journeys. In more realistic scenarios where the demand matrix is dependent upon the conditions of travel encountered in the network, it is clearer how to address this dependence when the entry and exit nodes all represent origins and destinations than when some of them represent intermediate points.

When whole journeys from a certain origin to a certain destination are being modelled, then it is possible to think in terms of a demand function for those journeys, so that the number of travellers making those journeys is a decreasing function of perceived cost, an aggregate indicator of the losses a traveller incurs in order to make the journey. A corresponding supply function, derived from the operating characteristics of the links on the various relevant routes, expresses the perceived cost as a typically increasing function of the number making the journey. The two functions together determine the estimated number making the journey. This process should proceed concurrently for all origin-destination pairs so as to reflect not only travellers' tradeoffs between travelling and not travelling, but also their tradeoffs among the advantages to be gained and losses to be incurred by travelling to each one of the available destinations.

No similar approach is available when some of the entry and exit nodes represent intermediate points on travellers' journeys, because the origins and destinations of the journeys entering and leaving at such points are not represented, and journeys starting at unrepresented origins or finishing at unrepresented destinations need not enter or leave the modelled part of the system at any one particular point of entry or exit or indeed at all.

4 What the modelling is for

Modelling of transport networks and their use has two main purposes: to estimate

- features of an existing transport system and its use that are difficult or expensive to observe or measure; and
- features of a transport system and its use in circumstances that do not yet exist - either because the system itself has not yet been built or includes as yet unimplemented alterations to an existing system, or to explore unprecedented patterns of use of an existing system - or for a combination of these reasons.

Modelling of readily observable features of an existing situation also has an important but incidental purpose - to help to establish the credibility of the modelling itself.

One principal context for modelling is the management of an existing system - perhaps to augment routine data collection in the periodic monitoring of performance, but more importantly to inform the operation of on-line management tools such as traffic-responsive signal control and driver information systems on the roads.

The second principal context is that change is contemplated, either to deal with dissatisfaction with existing conditions or to accommodate alterations in the scale, location or mix of activities served by the system. Modelling an existing system can help in the diagnosis of current problems and the development

of proposals for change. Modelling of the system with proposed changes in place can help in deciding just which, if any, of the proposed changes to implement, and in just what ways.

All these purposes require interpretation of the outputs of modelling in the light of sound understanding of the scope and limitations of the modelling itself. In particular, a balance has to be struck in the modelling process between desires for realism and coverage on the one hand and tractability and reliability on the other. In professional terms, this calls for a high level of communication between modellers, who are often mathematicians, and the users of the models, who are typically a mixture of planners, engineers, economists, interested citizens and decision-makers.

5 The evolution of transport network modelling

The initial motivator for modelling of transport networks and their use was the need of industrialised societies to adapt, and especially to adapt their cities, to the widespread ownership and use of cars. This need became clear in the 1940s and 1950s concurrently with the advent of the digital computer, which made large-scale numerical modelling practicable, and the science of mathematical programming which came to underpin it. The first attempts at transport network modelling related to urban road networks and amounted to commonsense heuristic numbercrunching by people tackling a new practical problem with new tools and little awareness of the kinds of mathematics that could help them. Mathematical insights developed in parallel, drawing upon relevant and concurrent fundamental developments in mathematical programming and probability theory, and upon earlier classical mathematics, but it was well into the 1960s before theory and practice really began to converge.

The core of the mathematical development was the traffic assignment problem. This was formulated initially as the steady-state allocation to routes of a specified demand matrix of vehicle movements between points of entry to and exit from a modelled road network in which capacity exceeds the specified demand, the perceived cost of traversing each link is the same for all vehicles and a mathematically suitable known function of the flow along that link only, and the routes taken are to be such that between each pair of entry and exit points all used routes have equal cost and no unused route has a lower cost than the used routes. This formulation will be referred to here as the core problem. An encyclopaedic history of the massive attention given to this problem and substantial generalisations of it over nearly half a century is provided by Patriksson (1994), so no particular contributions will be quoted here.

One immediate benefit from the application of mathematical thinking to this problem was the clarification of the distinction between a user equilibrium that would result from travellers choosing their routes each to minimise the perceived cost of their own journey in the context of the choices being made by others, and a system optimal allocation of travellers to routes in order to minimise the total perceived cost of all their journeys. An early consequence of the formulation of the core problem was the derivation from the link performance functions (in the form required for that formulation) of the link-specific toll that should be charged for traversing each link in order for the user equilibrium to be system optimal in terms of costs other than the tolls charged.

The principal lines of generalisation and extension of this problem in the interests of realism and coverage have been:

- dependence of link costs on flows on other links as well as on the link concerned
- different classes of vehicle in the road network
- public transport networks as well as road networks and the choice between public transport and private vehicle
- relevance to traffic management and adaptation of the transport system being modelled
- differences among travellers in their perception of cost
- dependence of demand upon perceived costs of journeys and characteristics of areas represented by points of entry and exit

- variation of demand over the period being modelled within the day, including temporary overloading
- variation of perception of cost over a succession of days, including learning from experience
- estimation of a probability distribution of the pattern of travel rather than a determinate set of flows on routes or links.

For the core problem, exact conditions for existence, uniqueness and stability of solution have long been established through its formulation as equivalent convex optimisation and variational inequality problems, and progressively more effective algorithms for calculating the solution, even in very large networks, have been developed. One of the greatest advances in this last respect came relatively recently (Bar-Gera 1999) and drew upon earlier work in telecommunications network modelling. The range of generalisations and extensions for which similar completeness of theory and capability of implementation have been achieved is, however, so far limited.

In terms of realism, ingenuity in network building, the analysis of movement of traffic on roads and in public transport systems, and the investigation of external effects of traffic movement have largely achieved quantitative descriptions of the network characteristics listed in Section 2, including plausible algebraic forms for link performance functions. Many ways in which operating organisations can use management tools such as provision of information, traffic control and charging for use are largely understood, and a good deal of empirical evidence has been assembled about characteristics of travellers as listed in Section 3. The realism of modelling therefore depends mainly on the extent to which these understandings can be reflected in the mathematics of the modelling.

6 Extension and generalisation - steady state and fixed demand

Pathbreaking as were the formulation and successful analysis of the core problem, the path soon led into challenging terrain. Elasticity of demand was addressed at once and is considered in Section 7, and time-dependence, which became an issue somewhat later, is discussed in Section 8. Some challenges that soon arose even in the context of steady state modelling for fixed demand are discussed first.

6.1 Link flows and route flows

Early analysis of the core problem was in terms of the total flow on each link, and these flows corresponding to the user equilibrium were soon shown to be unique, but it is easily seen that when multiple routes are used between more than one pair of entry and exit points, the route flows may not be unique - and indeed they usually are not. This may be of limited consequence if the main concern of the model user is to manage the existing traffic in the system, but as application is broadened to influencing the conditions experienced and choices made by travellers between particular entry and exit points, it becomes relevant to know how they are distributed among the routes they use, or might well use. Solving the core problem does no more in general than identify the routes of least cost which they might be using. Early reformulation in terms of route flows rather than link flows has been productive in other ways, but has not overcome this limitation of the core problem. Bar-Gera [5] investigates entropy-maximising allocation between parallel route segments of equal cost within the solution to the core problem where this allocation is not determined by the solution itself.

6.2 Non-separability of link performance functions

The requirement in the core problem that the perceived cost on each link should be a function of the flow on that link only was soon challenged in two respects. One is that at all road junctions except the most free-flowing grade-separated ones, costs on the links representing most or all approaches depend, often heavily, on the flows on links representing other approaches. The other is that when several demand matrices, representing different kinds of traffic, are to be allocated simultaneously to the same network

(or strictly speaking to multiple copies of a network), the costs for each kind of traffic on links used by more than one kind are dependent on the flows of all kinds on that link. Uniqueness of the solution to such generalisations of the core problem has been shown to depend on properties of the Jacobian matrix of the link cost functions which the matrix often does not possess when the cost functions are realistic.

Experience of road traffic suggests that situations in which there is more than one substantially distinct relevant user equilibrium in the real system are in practice rare, if they exist at all. Drivers are not accustomed to bulletins advising that traffic this morning has pattern B, whereas yesterday, on an otherwise pretty similar day when the roads were in the same state it had quite distinct pattern A. Non-uniqueness of solutions to a model by no means implies that none of the solutions is relevant, but it does call for care in identifying which is the relevant one.

6.3 Availability of different means of transport

Network building for public passenger transport systems, whether road-based, rail based or mixed, calls for different ingenuity from road traffic network building, but is not fundamentally different. Nor does the fact that buses and trams influence costs for users of other road vehicles and vice versa cause difficulty other than non-separability of cost functions. One feature of public transport systems that has to be kept in mind is that response of service providers to increased demand on a route by increasing the frequency of the service (as distinct from enlarging the vehicles) can lead to decreasing link cost functions, which implies potential non-uniqueness of user equilibria. Subject to this possibility, traffic assignment to a suitably built multimodal network need not differ in principle from assignment to a single mode network, in that choice between means of transport can be seen simply as a consequence of choice of route, as was recognised by Abdulaal and LeBlanc [?], and travellers without the use of a car can be assigned solely to the public transport network.

In practice, however, modelling of multimodal systems has usually proceeded by separate estimation of choice of means of transport by travellers between a given origin and destination and choice of route by the users of each means of transport, despite the fact that this may result in different perceived costs by car and by public transport even for users who have the choice of either.

6.4 Traffic assignment in network design and traffic management

To the extent that user equilibrium is regarded as providing a realistic estimate of how choices of route by users of a system will determine the link flows resulting from a given demand matrix, link flows estimated in this way (and route flows if these can also be estimated) are relevant to traffic management and to design of adaptations of the system. Analysis to assist the latter is referred to in this context as network design. This was first recognised explicitly in the context of road traffic signal control by the author (Allsop [2]), who proposed that traffic signal timings should be chosen to optimise some indicator of performance of the system under the pattern of traffic that the timings themselves would induce. Calculation of such timings was formulated as the upper level of a bilevel problem in which user equilibrium assignment formed the lower level. Initial searches for these timings were confined to timings that were optimal in the sense usually understood in traffic signal control - namely optimal for the prevailing pattern of traffic regarded as fixed. It was soon recognised by Dickson (1981, but circulated in 1977), however, that the required timings would not be optimal in that sense. This distinction is illustrated by Bell and Allsop (1998) and calls for a change of approach in signal control practice.

The scope of the bilevel approach, with traffic assignment as the lower level, in network design and traffic management has broadened as ingenuity in adapting transport systems has increased and the range of tools for traffic management has grown, notably to include road user charging. Its scope is widely recognised in the use of traffic assignment in the comparison and appraisal of proposed traffic management schemes (see e.g. Allsop 1996). This offers many opportunities for systematic optimisation at the upper level, optimisation with respect not just to total travel cost, as in the classic system optimum, but with respect to any of a range of criteria reflecting the modelled externalities of use of the transport system as well as cost to its users.

Another use of the bilevel approach which is relevant to traffic management is to provide an estimate, often on-line, of the demand matrix and route flows from traffic counts, often from automatic sensors, on selected links of the network. In this case the lower level is again traffic assignment, and the upper level is adjusting the demand matrix to fit the assigned flows to the observed ones (e.g. Bell and Grosso 1998).

6.5 Less than completely informed travellers

The core problem presumes that the travellers are completely informed in that for each route from their entry point to exit point all travellers have the same perception of the current cost of using that route. One way of relaxing this very strong assumption is to suppose that for each route, the perceptions of its cost by the travellers who might use it have a known probability distribution, and that each individual traveller will use the route which in their own particular perception has least cost. This leads to a user equilibrium, again unique in link flows, known as the stochastic user equilibrium - rather misleadingly because it is in fact the determinate solution of a convex optimisation problem formulated in terms of separable mean link costs and expectations of the relevant random variables, namely the minimum route costs between entry and exit points.

Solution of this problem is made difficult by the need to consider routes, potentially very large numbers of them, explicitly, and by the implications of the form of the probability distributions of their perceived costs. Assuming identical and independent Weibull-Gumbel distributions for their random components leads to tractable logit allocations of travellers to routes but is unrealistic in neglecting correlations between perceptions of overlapping or otherwise similar routes. A multivariate normal distribution of perception errors allows for these correlations, and with variances proportional to the means of the perceived costs leads to a probit allocation of travellers to routes. An estimate of this allocation is needed to provide estimates of the minimum route costs in order to solve the convex optimisation problem iteratively. At first it was thought that these minimum costs could be estimated only by Monte Carlo simulation, but approximate numerical methods have been found (Maher and Hughes 1998) with the help of a standard approximation to the distribution of the lesser of two normally distributed variates (Clark 1961), and this should open the way to viable use of probit stochastic user equilibrium models in practice.

7 Steady state demand dependent upon route costs

The initial formulation of the core problem recognised that it extended immediately to demand for travel between each pair of entry and exit points that was a suitable function only of the cost of travel between them by the used routes. This followed by connecting the exit point back to the entry point by a link on which the cost function is the negative of the inverse of the demand function. The resulting flows then correspond to the balance between supply and demand for travel between the pairs of entry and exit points, which is meaningful if these points represent the origins and destinations of the journeys being modelled, as distinct from intermediate points on journeys the rest of which are outside the scope of the model.

This elegant outcome takes no account of the cross-elasticities between demand for travel between different pairs of entry and exit points. One way of achieving this is by combining the assignment model with a model of the demand matrix in terms of the whole matrix of costs of used routes between the origins and destinations. Probably the most widely used and deeply understood such model is the gravity model with exponential cost function (e.g. Erlander and Stewart 1990). One of its several derivations is as the solution of a convex minimisation problem subject to given total numbers of travellers leaving each origin and arriving at each destination and a known matrix of costs of travel between origins and destinations. Its one parameter can be calibrated from an exogenous estimate of the total cost of all the journeys modelled. Evans showed (1976) that by combining the objective functions and constraints of this problem and the core problem, a demand matrix consistent with the travel costs resulting from

assigning the matrix according to the core problem, together with the corresponding link flows, are the solution of a single convex optimisation problem.

Such a combined model replaces the known trip matrix of the core problem by known row and column sums for the trip matrix. A further step might be to allow these sums in turn to depend upon some measure of accessibility of the origins and destinations in terms of costs of travel and the numbers of potential destinations and origins respectively, perhaps within some overall constraint on the total amount of travel. This would begin to incorporate the effect of conditions of travel upon the locations of the activities it serves.

8 Time dependence in route choice

Steady state modelling is potentially realistic for situations in which conditions in the transport system being modelled remain broadly similar for a good deal longer than the duration of the longest of the modelled journeys, and for conditions that are repeated over a long enough calendar period for most travellers to be presumed familiar with them. But on the one hand many of the greatest difficulties in transport are associated with peak periods over which demand first rises markedly, often temporarily exceeding the capacity of the system, and then declines to a level which allows the consequences of temporary overload to clear, and on the other hand changes in the transport system and in the lives of its users mean that there are usually at least some users who are relatively new to the conditions of travel that they are facing. This has led to modelling of within-day and day-to-day time dependence in travel choices.

8.1 Within-day time dependence

An initial approach to the modelling of the traffic flows resulting from time-varying demand is to divide the modelled period into timeslices in each of which the demand is taken to be steady and a corresponding steady-state assignment is modelled, in which where demand exceeds capacity some links are overloaded accordingly. On each such link, the excess traffic is accumulated and added to the demand for the next timeslice, entering the network by the link on which it has accumulated. This process has proved effective in practice, but does not address the modelling of variation over time fundamentally, nor does it lead clearly to any way of incorporating choice of departure time by travellers in the light of the variation in travel conditions over time.

Continuous modelling of variation in conditions in the transport system over a modelled period of the day requires the consideration of accumulation of traffic on links as well as flow along them. For road traffic, it turns out that requiring vehicles to leave a link in the same order in which they enter it and that travel time of a vehicle along a link should not be affected by the entry of subsequent vehicles places quite severe limitations on the form of appropriate models of the movement of traffic through a link. When these are satisfied, a necessary condition for dynamic equilibrium between routes between any pair of entry and exit points can be derived (Heydecker and Addison 1996). Application of this principle and related analysis allow dynamic equilibrium route flows to be calculated incrementally in time.

Variation of conditions in the transport system with time of day also implies that for a traveller with freedom of choice when to travel, there are more and less favourable times, in terms of travel conditions, at which to set out on the journey. Since most travellers have some choice of departure time it is relevant to consider this choice along with choice of route. In modelling the morning peak period the values attached to leaving home later and to arriving at work or place of study close to a preferred time have long been considered along with the cost of the journey. Recent analysis of this kind in the context of dynamic modelling (Heydecker and Addison 2005) has shown how these values, as functions of time, influence the dynamic equilibrium route flows and enable them to be calculated.

8.2 Day-to-day time dependence

All assignment models assume considerable knowledge on the part of travellers about the perceived costs of the routes available to them, and one mechanism by which such knowledge may be obtained is by learning from day to day. Attempts have therefore been made to model choices of route over a succession of days on which, for example, a traveller's expectation of the cost of a route is adapted from their previous day's expectation on the basis of that day's experience. This has the potential to extend modelling from the estimation of equilibria to the representation of equilibration, and leads to the modelling of traffic assignment as a stochastic process (Watling and Hazelton 2003) in which the probability distribution of each route flow on a particular day is obtained from the corresponding distributions on previous days by a transition matrix in a process which converges to an equilibrium probability distribution. Where the starting point for this process is a change in the modelled system, to which users then adjust, such a model provides an estimate of not only the new equilibrium resulting from the change, but also the trajectory by which it is approached. In this type of modelling, it is not a determinate equilibrium allocation of traffic to routes that is estimated, but the probability distribution of the allocation.

Modelling of this kind requires a model of the process of learning by each traveller, which may not be taking place every day, but rather only when triggered by some event or experience of conditions on the journey that are outside the range to which the traveller has become accustomed. Statistical issues in such modelling and its calibration from observable data have been addressed, for example, by Jotisankasa and Polak (2007).

9 Concluding remarks

Much applied mathematics is a compromise between elegance and tractability of the mathematics and relevance and usability of the application - and modelling of transport networks is no exception. In this context the fruitfulness of the compromise depends, as Patriksson wrote (1994) of work on a particular modelling development, 'on cooperation between operations researchers and users' - users, that is, of the wider techniques and procedures to which the models are intended to contribute.

Mathematical modellers and users of the models have both contributed much to the search for appropriate realism in models of transport networks - realism that keeps the modelling consistent with the fundamental operational and behavioural characteristics of the transport system whilst allowing the power of mathematical thinking and techniques to be brought to bear productively. But the realities of those operational and behavioural characteristics still present many challenges to the priorities of mathematicians, and much that mathematicians have achieved has still to be fully appreciated by the potential users of the resulting models. Getting more real still calls for clearer talking and harder listening on the part of both the modellers and the users.

References

- [1] Abdulaal M and L J LeBlanc (1979) Methods for combining modal split and equilibrium assignment models. *Transportation Science* 13 292-314
- [2] Allsop R E (1974) Some possibilities for using traffic control to influence trip distribution and route choice. *Proceedings of the 6th International Symposium on Transportation and Traffic Theory* 345-373. Amsterdam: Elsevier
- [3] Allsop R E (1980) Transport studies and the quality of life. *Environment and Planning A* 12 339-356
- [4] Allsop R E (1996) Introduction to modelling for the design and evaluation of area-wide traffic management schemes Short Course: *Modelling the Effects of Area-Wide Traffic Management*, June 1996. London: PTRC Education and Research Services Ltd

- [5] Bar-Gera H (1999) Origin-based algorithms for transportation network modeling. Technical Report #103. Research Triangle Park NC: National Institute of Statistical Sciences
- [6] Bell M G H and R E Allsop (1998) Introduction. In M G H Bell (ed) *Transportation Networks: Recent Methodological Advances*. Oxford: Elsevier
- [7] Bell M G H and S Grosso (1998) The path flow estimator as a network observer. *Traffic Engineering and Control* 39 540-550
- [8] Clark C E (1961) The greatest of a finite set of random variables. *Operations Research* 9 145-162
- [9] Dickson T J (1981) A note on traffic assignment and signal timings in a signal-controlled road network. *Transportation Research* 15(B) 267-271
- [10] Erlander S and N F Stewart (1990) *The Gravity Model in Transportation Analysis - Theory and Extensions*. Utrecht: VSP
- [11] Evans S P (1976) Derivation and analysis of some models for combining trip distribution and assignment. *Transportation Research* 10 37-57
- [12] Garland K (1994) *Mr Beck's Underground Map*. Harrow Weald: Capital Transport Publishing
- [13] Heydecker B G and J D Addison (1996) An exact expression of dynamic traffic equilibrium. In J-B Lesort (ed) *Transportation and Traffic Theory* 359-383 Oxford: Pergamon
- [14] Heydecker B G and J D Addison (2005) Analysis of dynamic traffic equilibrium with departure time choice. *Transportation Science* 39 39-57
- [15] Jotisankasa A and Polak J W (2007) A framework for travel time learning and behavioural adaptation in route and departure time choice. *Transportation Research Record* 1985 231-240
- [16] Maher M J and P C Hughes (1998) Recent developments in stochastic assignment modeling. *Traffic Engineering and Control* 39 174-179
- [17] Patriksson M (1994) *The Traffic Assignment Problem - Models and Methods*. Utrecht: VSP
- [18] Watling D and M L Hazelton (2003) The dynamics and equilibria of day-to-day assignment models. *Networks and Spatial Economics* 3 349-370