Transport networks and their use: how real can modelling get?

BY RICHARD E. ALLSOP *

Centre for Transport Studies, University College London, Gower Street, London WC1E 6BT, UK

The context for modelling transport systems and their use is set out. Such modelling is based on representing the transport system by a network of nodes and links, and the characteristics of this representation needed to reflect the principal realities of the system are outlined. The characteristics of use of the system that need to be reflected are described. Purposes of the modelling are set out and its evolution is described, starting from the basic traffic assignment model and discussing its generalizations 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 the modellers and the users of models, helped perhaps by combining the advantages of analytical modelling and microsimulation.

Keywords: transport systems; network modelling; evolution of modelling; realism

1. Introduction

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

Transport is not an end in itself: it enables people to engage in activities that require people and goods to be in different places at different times. It is a means of overcoming physical separation for purposes of economic, social, cultural and personal activity. It is therefore necessary to consider the transport system in relation to the activities it serves. A framework for doing so is shown in figure 1 (cf. Allsop 1980).

The prevailing desires for activity interact with the prevailing transport system through choices made by its users, to produce a pattern of movement and associated direct costs of movement. The perceptions of these costs feed back into users’ choices through short-term equilibration.

Direct costs of movement are not its only consequence. Movement and the provision of the transport system have many other economic effects, and widespread social and environmental effects.

*rea@transport.ucl.ac.uk

One contribution of 16 to a Discussion Meeting Issue ‘Networks: modelling and control’.
Direct costs of movement and these wider effects, especially dissatisfaction with some of them, give rise to longer term feedback through both motivation to change the transport system and motivation to alter activities as well as their location and timing.

Transport is thus part of a system that is physically and technologically complex, strongly influenced by physical geography, and also has people as crucial elements in its operation and as principal recipients of its impacts; these people are not trained personnel or clearly identified clientele, but the whole population.

The physical parts of the system are subject to mathematical and physical laws, and the human parts to laws of economics and the law of large numbers in the probabilistic characterization of whole populations notwithstanding the unpredictability of individuals’ behaviour.

The transport system and its use can be modelled using a network of nodes and links and a matrix describing the demand for movement. Starting from the origins and main purposes of such modelling, this paper offers a perspective on its evolution over the last half-century. Except for authoritative texts of Erlander & Stewart (1990) and Patriksson (1994), only illustrative references are given. A few of these are the first publications of key advances, but no attempt is made thus to attribute every development discussed.

2. Transport networks...

Even in prehistoric times, transport by land was concentrated on well-defined tracks, and in earliest history this tendency was reinforced by construction of roads and bridges. Building of ports led to a similar tendency in maritime transport. Formalization of land ownership and rights of way, and development of inland waterways, railways and modern road transport have defined land transport networks ever more precisely. Modern shipping in crowded waters navigates the so-called motorways of the sea, and air traffic control constrains civil aviation to corridors in the sky.
In modelling, it is therefore natural for transport systems to be represented by networks of links along which traffic is modelled as flowing, and nodes at which flows of modelled traffic merge or diverge, or enter or leave the model network.

Links and their incidence with the nodes, together with quantities and functions associated with the links represent

— the topology of the transport system (though not usually its geography or topography) and
— the operational characteristics and local management of traffic on each length of road, track or right of way.

The relevant topology for each link comprises the following:

— the direction of travel,
— 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 in the model all movements that are both physically possible and permitted by regulations in reality,
— prevent in the model any movement that is physically impossible or prohibited by regulations in reality, and
— reflect all costs experienced by each user in the costs of traversing the links forming their modelled route.

Achieving all this requires skills in network building which are easily forgotten by those concerned mainly with the analysis of the model network and its traffic.

The operational characteristics of each link may include, as relevant to the modelling task in hand, the following:

— a maximum average amount of traffic that can traverse the link per unit time,
— a maximum average amount of traffic that can leave the link per unit time,
— a maximum accumulation of traffic that the link can accommodate,
— rules and parameters of any mechanism for control of movement along the link,
— parameters of the mechanism by which traffic unimpeded by waiting traffic moves along the link,
— rules and parameters of the mechanism by which traffic waits to leave and leaves the link (which may depend on traffic leaving certain other links),
— relationships between the cost to traffic traversing and leaving the link and the amount of traffic that enters the link per unit time and the accumulation of traffic there,
— the prevailing amounts of traffic entering the link per unit time from each upstream link or other source,
— the prevailing amounts of traffic leaving the link per unit time for each downstream link or other sink,
— the rules determining the amount charged for traversing the link, and
— how information is provided to users and potential users of the link.
Such interdependent characteristics need to be known in order to model traffic using the link in any scenario concerning demand for use of the modelled system. Many of these characteristics are expressed in a link performance function (Patriksson 1994) relating perceived losses incurred in traversing the link to rates at which traffic traverses this and other links.

Other operational characteristics describe effects of use of the link and are concerned with the interpretation of the modelled use in relation to purposes of the modelling. Examples are the nature and levels on the link of

— occurrence of traffic accidents,
— noise generated by traffic,
— exhaust emissions generated by traffic,
— visual intrusion resulting from traffic, and
— revenues accruing from charging for use.

Many of the characteristics are random variables or describe random processes. Modelling often uses a single parameter (usually a mean) and sometimes variances and covariances, but deciding which parameters to use and how to estimate them requires understanding of the randomness concerned. When a steady state is being modelled, everything is averaged over time. In dynamic modelling, evolution of the real system over part of a day, or over a sequence of days, is modelled, but with very short-term variability averaged out. Dynamic modelling embraces time dependence of the operational characteristics of links, usually with a fixed network topology.

3. ...And their use

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

The use of transport systems is discussed here in terms of people making journeys. Some of it applies to goods being moved, but goods once consigned can be moved as the carrier sees fit, whereas travellers remain active participants throughout their journeys. And whereas intelligence imparted to packets of data in telecommunications systems may be known by the system operator, travellers in transport systems are individual humans whose perceptions, preferences and behaviour are hard to plumb and influence.

In modelling the use of a system represented by a network of nodes and links, travellers are typically represented as having certain points of entry to and exit from the network, each represented by a node. In many scenarios, these points represent the origins and destinations of the travellers’ journeys, and whole journeys are being modelled. However, where only a part of the system is being modelled, some of the entry and exit nodes may represent intermediate points on some travellers’ journeys, so that only parts of these journeys are modelled.

Consider first travellers represented as entering the network at a certain entry and leaving it at a certain exit. Unless the network is unusually small
and simple, they will have available in principle a large number of *routes* through the network, each consisting of a succession of links. Routes containing closed loops are easily excluded, but the number remaining is still large, and even routes which are in some sense plausible choices may still be numerous. Travellers are modelled as being distributed over some set of these routes.

Modelling of the number of these travellers and their distribution among routes is helped by information about the following:

— operational characteristics of all links in the light of the routes used by all travellers,
— use of management tools such as information, traffic control and charging,
— corporate behaviour of commercial operators of system components like public transport and parking,
— 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 journeys, and
— travellers’ preferences concerning conditions of travel (for example among means of transport or routes).

Let the number of entry and/or exit nodes be $Z$ in all. Then, a $Z \times Z$ matrix formed by the number of travellers from each entry node to each exit node is called a *demand matrix*.

In one common type of scenario, the demand matrix is given and fixed. It does not then matter whether the entry and exit nodes represent origins and destinations or intermediate points on journeys. In more realistic scenarios where the demand matrix depends upon the conditions of travel, this dependence is easier to address when the entry and exit nodes all represent origins and destinations than when some represent intermediate points.

Whole journeys from a certain origin to a certain destination can be modelled in terms of a demand function for those journeys so that the number of travellers making a journey is a decreasing function of its *perceived cost*, an aggregate indicator of the losses a traveller incurs in making the journey. A corresponding supply function, derived from the operating characteristics of the relevant links, 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’ trade-offs between travelling and not travelling, but also their trade-offs among the advantages gained and losses incurred by travelling to each available destination.

No similar approach is available when some of the entry and exit nodes represent intermediate points on journeys, because the origins and destinations of those journeys 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 particular point of entry or exit, or at all.
4. What the modelling is for

The modelling of transport systems and their use has two main purposes as follows:

— to estimate features of an existing transport system and its use that are
difficult to observe and
— to estimate features of a transport system and its use in circumstances that do
not yet exist.

The second purpose arises when the system of interest is not yet built or includes
as yet unimplemented alterations, or when unprecedented patterns of use of an
existing system need to be explored.

The modelling of readily observable features of an existing situation also has
an incidental purpose—to help to substantiate the modelling itself. However,
capability to reproduce an existing situation should be supplemented by tests of
ability also to reproduce observed effects of changes in the situation.

One principal context for modelling is management of an existing system—
perhaps to augment data collection in monitoring of performance, but more
importantly to inform the operation of online 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 or to accommodate alterations in activities served by the system.
Modelling an existing system can help in the diagnosis of current problems and
development of proposals for change. Modelling of the system with proposed
changes in place can help in deciding which, if any, changes to implement, and how.

All these purposes require an interpretation of the outputs of modelling based
on the understanding of its scope and limitations. A balance is needed between
desire for realism and coverage and the importance of tractability and reliability.
This calls for a high level of communication between the modellers, often
mathematicians, and the users of the models, typically a mixture of planners,
enGINEERS, economists, interested citizens and decision makers.

5. The evolution of transport network modelling

The initial motivator for the modelling of transport systems and their use was
the need of industrialized societies to adapt to the widespread ownership and use
of cars. This need became clear in the 1940s and 1950s concurrently with the
advent of digital computing, making large-scale numerical modelling practicable,
and the science of mathematical programming which came to underpin it. Early
network modelling in transport related to urban roads and amounted to common
sense heuristic number crunching by people tackling a new practical problem
with new tools and little awareness that mathematics could help them.
Mathematical insights developed in parallel, drawing upon concurrent inno-
vations in mathematical programming and probability theory, and upon earlier
classical mathematics, but it was only during the 1960s that theory and practice
really began to converge.

Phil. Trans. R. Soc. A (2008)
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 and exit in a modelled road network with capacity exceeding the specified demand, the perceived cost of traversing each link the same for all vehicles and a mathematically suitable link performance function of the flow along that link only, and routes to be taken such that between each entry and exit pair all used routes have equal cost and no unused route has a lower cost. This formulation will be referred to as the *core problem*. An encyclopaedic history of this problem and substantial generalizations of it over nearly half a century is provided by Patriksson (1994), and hence no particular contributions will be quoted here.

Applying mathematics to this problem quickly clarified the distinction between a *user equilibrium* resulting from travellers choosing their routes each to minimize the perceived cost of their own journey in the context of the choices being made by others, and a *system optimal* routering of travellers to minimize the total perceived cost of all their journeys. An early consequence of the formulation of the core problem was the derivation of the link-specific tolls charged for traversing each link that make the resulting user equilibrium system optimal in terms of costs other than the tolls charged. The required link toll follows from the corresponding link performance function and the resulting link flow. Dependence on the latter means that this toll can at best be approximated in practice.

The principal lines of generalization and extension of the core problem are as follows:

— dependence of each link cost on flows on other links,
— different classes of vehicle in the 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 system being modelled,
— differences among travellers in perception of cost,
— dependence of demand upon perceived costs of journeys and characteristics of areas represented by entry and exit points,
— variation of demand over the period being modelled, including temporary overloading,
— variation of perception of cost over a succession of days, including learning from experience, and
— estimation of a probability distribution of the pattern of travel rather than a determinate pattern.

For the core problem, exact conditions for existence, uniqueness and stability of solution have long been established through its formulation as equivalent convex optimization and variational inequality problems, and progressively more effective algorithms for calculating the solution in large networks have been developed. A great advance in this last respect came relatively recently (Bar-Gera 2002) and drew upon earlier work in telecommunications network modelling. The range of generalizations and extensions for which similar completeness of theory and the capability of implementation have been achieved is, however, limited.

The ingenuity in network building, analysis of movement on roads and in public transport systems, and the investigation of external effects of traffic have largely
achieved realistic quantitative descriptions of the network characteristics listed in §2, including algebraic link performance functions. Many ways in which operating organizations can use management tools are largely understood, and much empirical evidence has been assembled about traveller characteristics listed in §3. The realism of modelling therefore depends mainly on the extent to which these understandings are reflected mathematically.

6. Extension and generalization: steady state and fixed demand

Path-breaking as were the formulation and analysis of the core problem, the path soon became challenging. Elasticity of demand was quickly addressed as discussed in §7, and time dependence became an issue somewhat later, as discussed in §8. Some challenges arising in the context of steady-state modelling for fixed demand are discussed first.

(a) 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 when multiple routes are used between more than one entry and exit pair, the route flows may clearly not be unique—and they usually are not. This may be of limited consequence if the model user is managing existing traffic, but as application is broadened to influencing the choices made by travellers between particular entry and exit points, it becomes relevant to know how they are distributed among the available routes. Solving the core problem does no more than identify their routes of least cost. Reformulation in terms of route flows rather than link flows has been productive in many ways, but has not overcome this limitation of the core problem. Bar-Gera (2006) reports a recent implementation of entropy-maximizing allocation between parallel route segments of equal cost within the solution to the core problem.

(b) Non-separability of link performance functions

The requirement in the core problem that perceived cost on each link should be a function of flow on that link only was soon challenged in two respects. First, at all road junctions except the most free-flowing ones, costs on links representing most or all approaches depend, often heavily, on flows on links representing other approaches. Second, when several demand matrices, representing different kinds of traffic, are to be allocated simultaneously to the same network (or multiple copies of a network), the cost for each kind of traffic on a link is dependent on the flows of all kinds on that link. The uniqueness of the solution to such generalizations of the core problem depends on properties of the Jacobian matrix of the link cost functions which the matrix often does not possess when the cost functions are realistic.

The experience of road traffic suggests that for more than one substantially distinct relevant user equilibrium to exist in practice is rare, if it happens at all. Drivers are not accustomed to bulletins advising that traffic today has pattern B, whereas yesterday, in otherwise pretty similar circumstances, it had quite distinct pattern A. The non-uniqueness of solutions to a model by no means implies that none of the solutions is relevant, but does require identification of the relevant one.
(c) Availability of different means of transport

Network building for public transport systems, road-based, rail-based or mixed, calls for different ingenuity from road network building, but is fundamentally analogous. And buses and trams influencing costs for other road vehicles and vice versa causes only non-separability of cost functions. The response of service providers to an increased demand by increasing frequency of service (as distinct from enlarging vehicles) can lead to decreasing link cost functions (Morlock 1979), implying a potential non-uniqueness of user equilibria. But 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 route choice, as recognized by Abdulaal & LeBlanc (1979).

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 different means of transport even for users who have the choice of either.

(d) Traffic assignment in network design and traffic management

Insofar as user equilibrium provides realistic estimates of how choices of route by users of a system determine the link flows resulting from a given demand matrix, link flows so estimated (and route flows where also estimated) are relevant to traffic management and to adaptation of the system. An analysis to assist the latter is referred to as network design. This was first recognized explicitly for road traffic signal control by Allsop (1974) who proposed that signal timings be chosen to optimize some indicator of performance of the system with the traffic pattern that the chosen timings would induce. The calculation of such timings was formulated as the upper level of a bi-level problem with user equilibrium assignment as the lower level. Initial searches for these timings sought timings that were also optimal in the sense usually understood in traffic signal control—namely optimal for the prevailing pattern of traffic regarded as fixed—but Dickson (1981, but circulated in 1977) recognized that the required timings would not be optimal in that sense. This distinction is illustrated by Bell & Allsop (1998) and implies change in signal control practice.

The scope of this bi-level approach 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 has been widely recognized in the comparison and appraisal of proposed traffic management schemes (e.g. Allsop 1996). This offers opportunities for systematic optimization at the upper level, with respect not just to travel cost, as in the classic system optimum, but to any criterion reflecting modelled externalities of use of the system as well as cost to its users.

Another long-standing relevant use of the bi-level approach is to estimate the demand matrix and route flows from traffic counts on selected links of the network. 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. More recent implementations, like that of Bell & Grosso (1998), can be used online with data from automatic traffic sensors.
(e) Less than completely informed travellers

The core problem presumes that the travellers are completely informed in that for each route from entry to exit all travellers have the same perception of the current cost of using that route. This strong assumption is implausible because travellers differ both in knowledge about routes and in their preferences. It can be relaxed by supposing that for each route, perceptions of its cost by potential users have a known probability distribution, and that each traveller will use the route they perceive to have least cost. This leads to an 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 optimization problem in terms of separable mean link costs and expectations of the minimum route costs between entry and exit points.

This problem is made difficult by having to consider the potentially large numbers of routes explicitly, and by implications of the probability distributions of 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 routes. A multivariate normal distribution of the random components allows for these correlations, and 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 optimization problem iteratively. It was thought that these minimum costs could be estimated only by a Monte Carlo simulation, but approximate numerical methods have been found (Maher & Hughes 1998) which use an approximation to the distribution of the lesser of two normal variates (Clark 1961), and this should open the way to practical use of probit stochastic user equilibrium models.

7. Steady-state demand dependent upon route costs

The initial formulation of the core problem extended immediately to demand for travel between each entry and exit pair that was a suitable function only of the cost of travel between them by the used routes. This followed by connecting the exit back to the entry by a link whose cost function is minus the inverse of the demand function. The resulting flows then balance supply and demand for travel between each entry and exit pair, which is meaningful if these points represent the origins and destinations of the journeys being modelled.

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

This replaces the known trip matrix of the core problem as input 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 travel conditions upon locations of activities.

8. Time dependence in route choice

Steady-state modelling is potentially realistic where conditions in the system being modelled remain broadly similar for a good deal longer than the duration of the modelled journeys and are repeated over enough weeks for most travellers to be presumed familiar with them. But many difficulties in transport are associated with peak periods over which demand rises temporarily above the capacity of the system, before falling enough to allow the consequences of temporary overload to clear. Moreover, changes in the transport system and the lives of users mean that usually at least some users are new to the travel conditions they face. This leads to modelling of within-day and day-to-day time dependence in travel.

(a) Within-day time dependence

One approach to modelling traffic resulting from time-varying demand is to divide the modelled period into timeslices in each of which demand is steady and a corresponding steady-state assignment is modelled, and where demand exceeds capacity some links are overloaded. On each such link, excess traffic accumulates and joins the demand for the next timeslice, entering by the link on which it accumulated. This is effective in practice, but neither addresses the modelling of variation over time fundamentally nor extends clearly to incorporating travellers’ choices of departure time according to the time-varying travel conditions.

The continuous modelling of variation in travel conditions over a period of the day has to consider accumulation of traffic on links as well as flow along them. For road traffic, requiring both that vehicles leave a link in the order in which they enter it and that travel time of a vehicle along a link should not be affected by subsequent vehicles, limits severely the form of the appropriate models of traffic traversing a link. With an appropriate model, a necessary condition for dynamic equilibrium among routes between any entry and exit pair can be derived (Heydecker & Addison 1996). The application of this condition and related analysis allow dynamic equilibrium route flows to be calculated incrementally in time.

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

(b) Day-to-day time dependence

All assignment models assume knowledge by travellers about the perceived costs of available routes, and one mechanism for obtaining such knowledge is by learning from day to day. The 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 have therefore been modelled. This enables modelling to extend from the estimation of equilibria to the representation of equilibration, and leads to the modelling of traffic assignment as a stochastic process (Watling & 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 with convergence to an equilibrium probability distribution. Where this process starts from a change in the modelled system to which users adjust, such a model estimates not only the new equilibrium resulting from the change but also the trajectory by which it is approached. This type of modelling estimates not a determinate allocation of traffic to routes but rather the probability distribution of the allocation.

Modelling of this kind requires a model of the process of learning by each traveller, which may be taking place every day, or only when triggered by some new event or unusual experience on the journey. A number of studies have addressed statistical issues in such modelling and its calibration from observations in laboratory experiments, a recent example being that of Jotisankasa & Polak (2007).

9. Concluding remarks

The foregoing account of network modelling in transport has discussed a pursuit of greater realism through development of analytical techniques based on the properties of link performance functions, demand functions, random variation and associated optimization algorithms. It has not discussed parallel developments in microscopic simulation of transport systems, as enabled by the rapid increase in computing power and speed. Microscopic simulation, for example, vehicle-by-vehicle simulation of road traffic, is beguiling in its scope for realism in detail, and visualization of the modelled system. Carried out rigorously, with thorough testing against theory and empirical data of each component of the network simulation model and their interactions, transparency concerning the underpinning assumptions and sufficient model runs for each scenario to distinguish real findings from between-runs variation, microscopic simulation of transport networks has a great potential. However, it lacks the capability of analytical modelling in principle not only to determine a solution representing conditions to be expected in a given scenario but also to determine rigorously the quantitative properties of that solution and the range of validity of the modelling.

Phil. Trans. R. Soc. A (2008)
Fortunately, potential users need not be faced with deciding either for an analytical model or for a microsimulation model. It should be possible for the two approaches to join forces in using microsimulation to extend the capability and realism of analytical modelling by using high-quality simulation to estimate tractable analytical descriptions of features of the transport system where no such descriptions have yet been yielded by theory or observation. In this way, the scope for realism offered by microsimulation could be combined with the power and rigour of analysis in a new generation of transport network models.

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

Mathematical modellers and users of the models have both contributed much to the search for appropriate realism in network models of transport systems—realism that keeps the modelling consistent with the fundamental operational and behavioural characteristics of the transport system, while 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 both their parts.

References


