Land-use/Transport Interaction Models

Past and Future

A. G. Wilson*

1. Introduction

Analysing the development of cities and regions has been one of the great social science tasks of this century. At its core are urban and regional modelling techniques, used to develop the theory which represents our understanding. The heart of the task is to understand transport/land-use interactions, which are obviously highly complex. Ideally, a good understanding is needed of urban and regional evolution in different epochs, over different time periods, with different technologies, in different cultures and places. The difficulty is shown by attempts to do this for ancient times (for example, by Jacobs, 1970). In this paper, the focus will be on the second half of this century, when the need has arisen from the twin problems of traffic congestion and pressures for land development — sometimes driven by a belief that the development of this infrastructure can be planned and controlled to generate 'good' cities. Recently, concerns with the environment — especially pollution — and with sustainability have come increasingly to the forefront (Department of the Environment, 1994; and Royal Commission on Environmental Pollution, 1994).

The modellers' task is to predict how people and organisations will live in 'good' and 'sustainable' cities; how the infrastructure will, or should, grow; and how activities and traffic flows are, where appropriate, best managed, priced and regulated. The understanding that can be achieved represents the basic science of cities and regions. Here, much of the interest derives from the utility of that science in planning. This in turn depends on predictive capability. As will be seen later (in subsection 2.3), the discoveries of urban and regional dynamics, and the properties of dynamic models, show that the ability to predict in detail is strictly limited. Predictive capability means, first, whether a model can reproduce the current (or a historical) situation as represented by the data on that situation. We might call that "reproductive" capability.

* University of Leeds. The author is grateful to Professors Michael Beesley, Michael Batty, Martin Clarke, Tony May, Stan Openshaw, and Huw Williams, and an anonymous referee, for extensive comments on an earlier draft of the paper. The responsibility for any shortcomings remains the author's.
Second, it will be necessary to identify the circumstances where accurate prediction is possible. Typically, where reproductive capability is good, it will be possible to predict the short-run in some circumstances. This is valuable for certain kinds of planning. It will also be possible in some cases to predict types of development, even when the detail is beyond reach. This will also produce valuable insights for planning. In yet other cases, it will be possible for planners to control some of the more ‘difficult’ variables (from a modeller’s point of view) and this will make predictions more reliable. In what follows, ‘predictive capability’ will be used to mean ‘the predictive capability which is possible in the circumstances’; these circumstances will be elaborated as appropriate. Predictive capability provides, in principle, ‘what if’ forecasting power for use in a variety of planning situations.

This paper reviews over thirty years of the history of transport/land-use modelling. A full treatment would now be a multi-volume enterprise, but it seeks to give a flavour of that history, attempting to give as much attention to what has not been achieved as to what has. It focuses on contemporary issues and the future, for which the research agenda remains interesting and exciting. In this spirit, the understanding — the science — which has been created through modelling, the applications and the future are considered in turn in Sections 2, 3 and 4.

2. The Modelling Achievement
The following subsections give an account of the core models and their origins, by describing their underlying theory, and the advances in mathematics which from the late 1970s onwards have made it possible to address more fully the issues of dynamic modelling. From the perspective of this Journal, the economic basis of modelling is of the greatest importance; but to achieve a reasonable historical understanding, this has to be understood in the context of approaches from different disciplines. Ultimately, these different approaches should be integrated into a unified multidisciplinary approach under the umbrella of something like urban analysis — but this remains a difficult task.

2.1 The core models and their application
The history begins in the 1950s, with the start of the great American urban transport studies. They were engineering tools, but had their impact in urban analysis. The first paper in the first volume of Papers in Regional Science was by Carroll (1955), describing transport modelling in one of the earliest studies: the Penn-Jersey Study, centered on Philadelphia. This study was particularly significant; it embraced land-use/transport interaction, generating some of the earliest and most important work in this field, and was a stimulus for academic development of the subject in the Departments of City Planning and of Regional Science at the University of Pennsylvania (for example, in relation to residential location modelling, see Harris, 1962, 1964; and Herbert and Stevens, 1960).

In the mid-1960s, there were at least three distinct contributions to the development of a model which has come to play an archetypal role in the canon of land-use models: the
urban retail model, developed by Harris (1962, 1964), as part of a transport study; Huff (1964) from a marketing perspective; and Lakshmanan and Hansen (1965), again from a transport perspective. This period of development is fully reviewed in Boyce et al. (1970). A further, independent, breakthrough was the development of a relatively simple — but powerful and elegant — integrated urban model by Lowry (1964), providing a framework for all engaged in the task of building a comprehensive urban model.

**Interaction models**

The transport and land-use perspectives generated, in essence, two kinds of interaction model, predicting the number of trips between any pair of zones in a city for a purpose, of which the archetypes were the journey to work and the journey to shop. The essence of each is now presented.

The transport model was the so-called ‘doubly-constrained’ model:

\[ T_{ij} = A_iB_jO_iD_j\exp(-\beta c_{ij}) \]  \hspace{1cm} (1)

where \( T_{ij} \) is the number of trips between any pair of zones \( i \) and \( j \) of a city; \( O_i \) is the number of trip origins in zone \( i \) (say, roughly the number of resident workers for modelling the journey to work); \( D_j \) is the number of trip destinations to zone \( j \) (say, proportional to the number of jobs); and \( c_{ij} \) is a measure of the cost of travel from \( i \) to \( j \). The parameter \( \beta \) measures the ‘strength’ of the impedance. \( A_i \) and \( B_j \) are balancing factors to ensure that

\[ \sum_j T_{ij} = O_i \] \hspace{1cm} (2)

and

\[ \sum_i T_{ij} = D_j \] \hspace{1cm} (3)

Hence

\[ A_i = 1/\sum_j B_j D_j \exp(-\beta c_{ij}) \] \hspace{1cm} (4)

and

\[ B_j = 1/\sum_i A_i O_i \exp(-\beta c_{ij}) \] \hspace{1cm} (5)

The retail model was singly constrained:

\[ S_{ij} = A_iO_iW_i\exp(-\beta c_{ij}) \] \hspace{1cm} (6)

where \( S_{ij} \) is the flow of retail expenditure from residential zone \( i \) to shops in \( j \); \( O_i \) is the total amount originating at \( i \) (and so related to population, income, and so on); \( W_j \) is a measure of the attractiveness of the shops in \( j \); \( c_{ij} \) is as in the transport model; and \( \beta \) again is a parameter. \( A_i \) is a single balancing factor, calculated as

\[ A_i = 1/\sum_j W_j \exp(-\beta c_{ij}) \] \hspace{1cm} (7)

to ensure that

\[ \sum_j S_{ij} = O_i \] \hspace{1cm} (8)

In each of the models, the relationship between transport and land use is explicit: the \( O_i \)s, the \( D_j \)s, and the \( W_j \)s will all be functions of land use; and the \( c_{ij} \)s, which would include travel-time in a measure of ‘generalised cost’, are measures of the effectiveness of the transport system to ‘deliver’ journey times. The models are potentially rich. The \( O_i \)s, \( D_j \)s, and \( W_j \)s can all be presented as functions of land use (via the trip generation model); the \( c_{ij} \)s can be functions of congestion (via the trip assignment submodel); and the parameter \( \beta \) can be broken down in various ways into a richer parameter set. Much empirical work is needed to calibrate and test the models. They can be made to work well in the sense that
they can be made to fit existing situations well. Their ability to function predictively depends on the quality of variables which are the exogenous inputs to the models.

**Location models**

The transport models, however, are essentially still interaction models alone; the land-use variables are fixed and given — or can only be changed exogenously over time. This is acceptable if they can all be planned, but the obvious question was: "Is it possible to model them?" The breakthrough for land-use modelling was the recognition that by dropping one of the constraints, as in the retail model, it was possible to calculate a variable like

\[ D_j = \sum_i S_{ij} \]  \hspace{1cm} (9)

— the total retail expenditure attracted into shops in zone \( j \) — which, of course, is a locational variable. It will be seen later that this insight provided the basis for a shift to dynamic modelling. Even so, the vector of structural variables (related, say, to the size of retail centres), \( \{ W_j \} \), is still being taken as given exogenously.

**Uses of the models**

As seen above, the origins of the model were application-driven: how could they help to ensure the delivery of best-value transport systems within an 'optimal' urban structure? The methodology for this kind of planning involved an articulation of cost-benefit analysis first in the transport field. The origins lay in the evaluation of major water resources projects in the United States (see Maass et al., 1960).

This methodology had to be recast for the transport system, and a crucial element in this was the measurement of the value of time to transport users. Key papers were published by Foster and Beesley (1963) in relation to the proposed Victoria Line in London, and by Beesley (1965). They were significant not only in providing the basis for an evaluation methodology, but also because perceived time cost was identified as a component of what came to be called generalised cost in transport models. The \( c_{ij} \) term became a sum of elements which included the different components of travel time, each of which could be weighted in the light of model calibration.

More generally, the question of performance indicators for evaluating alternative plans was raised. There were many approaches, not all rooted in good theory. (For a review of the early history, see Clarke and Wilson, 1987a; this issue will be considered further in Section 4.)

In approaching the question of 'optimal' urban structure, there are essentially two ways of proceeding. The first is to generate alternative plans, testing them against whatever battery of performance indicators can be developed — ideally using the tools of cost-benefit analysis. The second, more ambitiously, is to seek procedures for developing optimal plans through defining an appropriate objective function. This was attempted in the early days of land-use modelling by Schlager (1965). The transport problem, because of the network complications, is more difficult; a procedure was sketched in Wilson (1983) but, to the author's knowledge, has never been implemented. The generation and testing of alternative plans, therefore, became the norm: this is the design element of land-use/transport planning.
2.2 Theoretical developments

Equilibrium models

Many of the early-1960s models were fashioned by a research community functioning as engineers, planners and economists engaged on practical studies, mostly for government, directly or indirectly. The modelling methodologies that evolved did not fit easily into any of the established disciplines. Slowly, however, these disciplines began to absorb or respond to them, and in many cases to put their own stamp on them. This generated a number of competing approaches to the problems of land-use/transport interaction. Their differences can be understood in terms of system characterisation (for example, treating space continuously or in terms of discrete zones); or in terms of the typical underpinning theories of the various disciplines (for example, utility maximising approaches in economics). These approaches are more complementary than competing. For example, virtually all theories need to find a way of solving the aggregation problem: how to link theories of individual behaviour with a system model, say of flows on a transport network. With hindsight, it can be seen that virtually all the early approaches were concerned with equilibrium models, though this comment can only be fully understood in the context of dynamic models, (considered below). There is an implicit assumption of a fast dynamic — that the system being modelled, if disturbed, will return to equilibrium very quickly.

The following review is organised mainly in terms of disciplines, though it begins with the entropy-maximising methodology, which is essentially supra-disciplinary. A variety of approaches emerges; most of them are rooted in particular disciplines. It can be argued that an interdisciplinary integrated approach is the most desirable. However, the traditions make it difficult to force an effective integration. In giving a flavour of the different narratives, the issue of integration is kept in mind as the argument develops.

Arguments by analogy: the entropy-maximising approach

The original spatial interaction model was built on analogy with the Newtonian gravity model, with origins in the works of Carey (1858); Ravenstein (1885); Lill (1891); Young (1924); Stewart (1942); and Zipf (1946) (see Erlander and Stewart, 1990, for a review). However, a model based on a strict gravity analogy simply did not work. The $A_i$ and $B_j$ factors in equation (1) were introduced essentially as fudge factors to ensure that it did.

These were recognised as versions of the partition functions that occur in statistical mechanics. This led to an alternative framework (Wilson, 1967, 1969). This methodology showed that an interaction pattern was the most probable state consistent with a set of constraints — essentially the land-use pattern — and a parameter which measured ease of travel — the $\beta$-parameter in equation (1). The so-called entropy-maximising methodology could be seen as a statistical averaging procedure which worked for large populations of loosely interacting entities — in this case people travelling in cities. The entropy term can also be seen as adding a degree of dispersion relative to what might otherwise be seen as an optimisation problem. The basic spatial interaction model of equations (1), (4), and (5), for example, is a suboptimal version of the transport problem of linear programming (Evans, 1973). The retail model could be derived from the same methodology.
This is a high-level methodology which can be applied in any discipline where a large number of units, relatively weakly connected, occur, whether particles in a gas or people in a city. Its power stemmed from the fact that it was easy to modify and generalise for a great variety of circumstances (Wilson, 1970). Its reproductive power was good. In this sense, it provided the basis for a number of the submodels for a general urban model. It was possible to take the Lowry (1964) model, for example, and to improve its submodels (Wilson, 1971, 1974, 1981a).

Economic models
Many productive strands of work by transport and urban economists have contributed to land-use/transport modelling (see Berechman et al., 1996). Economic models essentially turn on the theory of consumers' behaviour, or the theory of the firm, involving, in some form or other, the maximisation of a utility function or profit. Six contributions will be distinguished here: the bases of cost-benefit analysis; the economic interpretation of the doubly-constrained interaction model and its implications for land use; the logit model; and three forms of land-use theory — a continuous space version, a mathematical programming discrete-space one, and an econometric model.

The work by Foster and Beesley (1963), and Beesley (1965), on the value of time was an essential element in facilitating the development of cost-benefit analysis in the transport planning field, and later research showed how their ideas could be incorporated into land-use/transport models in powerful ways.

The conventional doubly-constrained model of equations (1), (4), and (5) has had many derivations; the entropy-maximising approach was often criticised because it did not have an economic interpretation. This was later rectified. However, in one of the earliest contributions on land-use interaction in this Journal, Neuberger (1971) gave a full analysis of how user benefit measures could be derived, and how these connected to land-use issues. Neuberger and Wilcox (1976) extended this argument to generating benefit measures resulting from land-use changes based on transport models.

The logit model arises from assumptions that transport behaviour is determined by the maximisation of a utility function that takes into account the costs of transport. Dispersion is added to the model through a random variable to represent the distribution of preferences. An assumption has to be made about the distribution of this random variable. It is possible, by making an appropriate assumption about dispersion, to generate a model that is equivalent to the entropy-maximising model. This approach has a number of advantages, possibly the most important being that a precise measure of consumers' surplus can be derived, which can be used as the benefit measure in transport cost-benefit analyses.

The various terms of the model can be given an underlying economic interpretation. When the model is calibrated to fit a particular situation, for example, the functional forms that turn up, and the parameter values, can be interpreted in terms of the form and parameters of utility functions. A full discussion of these models and associated measures of consumers' surplus — which complete the derivation of appropriate benefit measures — is provided by Williams (1977). An earlier paper in this Journal was an important
contribution; that is, Champenowrne et al. (1976). These analyses continue to provide the theoretical bases for cost-benefit analysis.

The microeconometric approach emphasises preference and choice. A further microeconomics approach was that associated with an activity-travel framework, which emphasises interpersonal and environmental constraints as precursors to choice and has an affiliation with the work of Lancaster (1965) in economics. The practical synthesis of these two approaches, though long anticipated, remains incomplete. For a discussion of applications and progress of the activity-travel approach, see the recent reviews by Axhausen and Garling (1992) and Fox (1995).

Economists have made many approaches to the problem of modelling urban land use. Richardson's general review (1977) remains outstanding. The urban economist's task is to show how preferences for, say, housing, in relation to other costs and benefits within a utility function, determine land-use patterns. Key works were those by Alonso (1960, 1964). He saw that, building on von Thunen's (1826) ideas about agricultural location, the task could be accomplished through the concept of bid rent. What rents would consumers be prepared to bid at a location for each possible level of utility that can be achieved? Alonso was able to show that such a market could be cleared and that an equilibrium solution exists.

Alonso's analyses were carried through with a continuous space representation of the city, which makes any generalisation away from the assumption of a monocentric city difficult. To solve this problem, Herbert and Stevens (1960) used a discrete zone representation akin to that used in transport modelling, and formulated the bid-rent maximisation problem as a linear programming problem. Arguably this suffers as a realistic representation in the same way that the transport problem of linear programming suffers as a representation, say, of the journey to work. It can be generalised, likewise, by adding an entropy term to create dispersion, as in Senior and Wilson (1974). By writing the entropy-maximising model as a non-linear mathematical programming model, and interpreting the dual, it was also possible to show that the equivalent of the balancing factors in equations (4) and (5), which turn up in the revised model, can be interpreted as functions of rents (see Wilson and Senior, 1974).

Another approach to solving the monocentric city problem is econometric. Beesley and Dalvi (1974) use a von Thunen-Alonso framework to construct variables which can then be used in a regression model to estimate journey-to-work lengths from employment centres and residential areas; they apply this to London Census data for 1951 and 1961. Another example is provided by McCarthy (1977); his econometric model was also rooted in an Alonso framework. His focus was on estimating the effect of residential location on modal choice.

**Geographic models**

Geography theorists had long been concerned with land use, but until the 1960s hardly at all with spatial interaction. Von Thunen's (1826) theory of agricultural land use, as seen earlier, provided the basis of Alonso's work. Weber (1909) built mainly micro-level models of industrial location. Residential location was left in the main to descriptive rather
than mathematical models (Hoyt, 1939; Harris and Ullman, 1945). Services were dealt with through theories of settlements, one based on catchments and geometry (Christaller, 1933), one on economics (Losch, 1940). Spatial interaction did not really go beyond the gravitational analysis of Zipf (1946). An excellent geographical synthesis at the start of the model-building period was by Haggett (1965). It is possible to use the techniques of spatial interaction modelling to integrate these approaches, and to offer new insights (see Wilson, 1989, building on Birkin and Wilson, 1986a,b; and Wilson and Birkin, 1987). The argument relies on extensions to dynamics and is developed below. An emerging school of thought in geography uses new techniques to generate ‘optimal’ interaction models to fit large datasets well, irrespective of the theory — a kind of ‘statistical’ analysis of the possibilities. Methods of neural computing have been used to generate the models (Openshaw, 1988, 1992, 1993).

Sociology and Social Geography

Sociology’s contribution to urban modelling was more through statistical analysis than mathematics — for example, through factor analysis, and the notion of factorial ecology. A good example is provided by Timms (1971). The geographers also were enthusiastic exponents of this method (see, for example, Berry and Rees, 1969). It could be argued that whatever insights can be gained from a factorial ecology, they would be better if there were adequate arrays of performance indicators from a model-based analysis. However, this is still to be fully delivered. For a current account, see Bertuglia et al. (1994).

Many authors anchored in social theory attacked what they saw as the positivist underpinnings of mathematical modelling, perhaps initiated by Harvey’s (1973) Social Justice and the City. An explicit and influential attack was Sayer’s (1976); in effect arguing that positivist approaches neglected the social context, and that mathematical modelling was “positivist” and therefore inadequate or worse. While it can be argued that the full breadth of the attack was misguided — Marxist economists, for example, had no hesitation in using mathematical modelling — the analysis does point to the need for an understanding of deeper underlying structures that might be at the core of urban evolution.

2.3 Enabling Technologies

Progress in modelling development has depended on mathematical capability. Probably no opportunities have been missed for lack of it; but, as noted earlier, the expansion of the number of professional users of, and researchers on, models has not taken place as rapidly as might be desirable. Almost certainly, progress with certain kinds of problem now depends on solving new kinds of mathematical problems. Operational research has made important contributions because mathematical programming has been particularly important in the integration of different approaches to theory. In Italy, for example, much of modelling development has been under the OR umbrella. There is also a long and valuable tradition of formulating certain families of planning problems as optimisation problems — particularly the so-called location-allocation (see, for example, Scott, 1971, for an early account). However, progress with the mathematics of nonlinear systems, initiated by Thom (1975), has been the major stimulus for more recent developments.
Integrated approaches

Most equilibrium models can be formulated as mathematical programming problems. Both the primal problem and its dual can be used to elucidate the differences between the various approaches, and help in interpretation. For example, when an entropy-maximising model is so formulated, similarities to certain economic models enable an economic interpretation for some of the terms to be discovered. Thus, for spatial interaction models to be effective, generalised cost measures should reflect the loads on the networks; this usually implies an internal iteration to seek to achieve this balance. However, by embedding all the component models within a mathematical programming framework, a theoretical integration can be achieved (see, for example, Boyce, 1978). Williams (1997) noted that this equilibrating task was only taken seriously in practice from the 1980s onwards.

An even more general framework is provided by dynamic modelling. This provides new insights, for example, for nonlinear programming models about the numbers and stability of equilibrium points at various points in parameter space; this in turn leads to very important insights about the nature of urban models and of the cities they purport to represent (Wilson et al., 1981). The availability of these methods dates from the mid-1970s. They facilitated a new phase of theoretical development in dynamics. The complexities revealed by dynamic models can only be resolved with the help of computing power, with corresponding graphics capabilities. This has become available through the 1980s’ dynamic models.

The core models reviewed above (subsection 2.1) are, in essence, of a comparative static equilibrium type. In so far as the dynamics have ever been considered, in the entropy-maximising context for example (Wilson, 1970, Appendix 1), the concern has been with fast dynamics and a rapid return to equilibrium for the key predicted variables \( \{T_{ij}\} \) or \( \{S_{ij}\} \). The reproductive power of the models has been good. However, many of the key variables — and particularly the land-use inputs into transport models — were taken as exogenous. Longer-run predictions will be dependent on the quality of the forecasts of these variables. One way to determine them is by the planning process. But how can they be determined in a model? It turns out that these issues are best formulated within a dynamic framework.

In the 1970s, mathematicians discovered and presented some of the interesting properties of nonlinear dynamic systems, initially under the rubric of catastrophe theory (see Thom, 1975), later in terms of the properties of large systems of nonlinear differential equations, and more recently under the misleading umbrella of chaos theory. The properties of nonlinear systems include the possibility of discrete changes in systems at critical parameter values. Initial applications in land-use/transport modelling were in terms of catastrophe theory and applied to retail structures — particularly the nature of the corner-shop-to-supermarket transition (Poston and Wilson, 1977; Wilson and Oulton, 1983), and modal choice in transport (Wilson, 1976).

Retail structure as an example

The power of the new opportunities, however, can best be presented through a develop-
ment of the retail model in equations (6) to (9) (Harris and Wilson, 1978). Similar ideas can be applied within alternative frameworks. For example, a dynamic logit model has been developed by Nijkamp and Reggiani (1991).

The set of attractiveness terms in the model, \( \{W_j\} \), is usually measured by some index of shopping centre size raised to a power, \( \alpha \). \( W_j \), therefore, is now redifined to be the size of the shopping centre, and so the vector \( \{W_j\} \) can be seen as a crude representation of urban structure. \( W_j^\alpha \) then appears as the attractiveness term in the model equations. This will suffice to illustrate an argument which can obviously be presented more generally. To create a dynamic framework for the analysis, a hypothesis is needed for the rate of change of any one \( W_j \). \( \{D_j\} \) is a model-calculated array of revenues; let \( \{C_j\} \) now be the array of associated costs for the suppliers of retail facilities. Then

\[
D_j - C_j
\]

is a measure of profit, and a suitable hypothesis might be

\[
dW_j/dt = \epsilon(D_j - C_j)
\]

for some suitable parameter \( \epsilon \). These equations are, for an N-zone system, \( N \) linked nonlinear simultaneous differential equations. Their nature can be seen by substituting for \( D_j \) from (9), \( S_j \) from (6), and \( A_j \) from (7), and replacing \( W_j \) by \( W_j^\alpha \):

\[
dW_j/dt = \epsilon(\Sigma_i \{O_j\}W_j^\alpha W_k^\alpha \exp(-\beta c_{ij})/\Sigma_k W_k^\alpha \exp(-\beta c_{ij}) - C_j)
\]

which shows that each \( dW_j/dt \) is dependent on the whole array, \( \{W_j\} \). The specification of a simple but fruitful dynamic model can be completed by assuming that costs of provision are approximately proportional to size, and so

\[
C_j = kW_j
\]

which can be substituted in (12). If this is done, then the equilibrium condition \( (D_j = C_j) \) can be written

\[
\Sigma_i \{O_j\}W_j^\alpha W_k^\alpha \exp(-\beta c_{ij})/\Sigma_k W_k^\alpha \exp(-\beta c_{ij}) = kW_j
\]

If the left-hand side of this equation is written as \( f(W_j) \), then it can be plotted for different values of \( \alpha \) and \( \beta \) against the right-hand side, which is, of course, a straight line. An analysis then shows that, typically, there is either one stable equilibrium point, \( W_j = 0 \), or two, also \( W_j = 0 \), but with an additional \( W_j \) with finite value. Then, if there is the possibility of a \( W_j > 0 \), it can be assumed that it will actually be greater than zero: some entrepreneur will identify the possibility. The set of \( W_j \)s generated when the equations are solved represents a spatial pattern, \( \{W_j\} \). At the extreme, if there is only a single non-zero \( W_j \), typically in the centre, then this is a very centralised pattern; alternatively, if all \( W_j \neq 0 \), then the pattern is very dispersed. There is, of course, a huge variety between these extreme cases.

This analysis provides new insight into the difficulty in forecasting. These kinds of nonlinear equations have many possible solutions. The specific outcome from solving the equations is likely to depend sensitively on the starting values — and on, for example, any perturbations from idiosyncratic individual behaviour — so there is an enormous range of possible futures. However, the type of pattern is likely to be much more predictable. A valuable use of a dynamic model is the potential to identify circumstances in which the pattern is likely to change radically, as in the corner-shop-to-supermarket transition mentioned above.
There is also, however, a technical difficulty in solving the equations: an analysis of the equation for each \( j \) demands that \( \{W_k\}, k \neq j \), be fixed. This is the so-called backcloth problem (which is considered further in Section 4). However, it is possible to proceed experimentally by solving equation (14) numerically in the hope that the starting values of \( \{W_k\} \) do not bias the result. The results show (as one would expect) that for high \( \alpha \) and low \( \beta \) there are relatively few non-zero \( W_j \), and vice versa. Typical plots (for a hypothetical 729-zone system) are shown in Clarke and Wilson, 1985b.

The above models may readily be interpreted and formulated as \( N \)-player dynamic games in which the current or anticipated actions of retail competitors are used to determine the future courses of action of each (see, for example, Williams et al., 1990).

**Incorporating prices in the dynamic model**

The dynamic model presented in the previous subsection focuses on shopping centre size. In the context of land-use/transport interaction models, rent is also obviously very important. It turns out to be relatively straightforward to introduce (Wilson, 1985a, 1990). Two other sets of prices should be introduced for completeness: the consumers' perceptions of the prices of goods (to facilitate the introduction of an elastic demand term), and the prices of the goods at a particular shopping centre. Let \( \pi_i \) be the perceived price at \( i \), \( p_j \) the price at \( j \), and \( r_j \) the rent at \( j \). Then \( \pi_i \) can be built into the \( O_i \) term and this need not be of further concern here; and the \( p_j \) and \( r_j \) terms can be adjusted by mechanisms analogous to that for \( W_j \) in equation (11):

\[
\frac{dp_j}{dt} = -\varepsilon_1(D_j - C_j)
\]

and

\[
\frac{dr_j}{dt} = \varepsilon_2(D_j - C_j).
\]

Equation (15) implies that retailers would reduce prices as the profits went up — but of course this may not be the same for all kinds of good, and alternatives can be introduced in a disaggregated model. Rents would increase as profits increase. The actual dynamics will depend on the relative sizes of the \( \varepsilon \)-parameters, and examples are shown in Birkin and Wilson (1989).

**Extensions and integration**

This kind of analysis can in principle be carried out for the other submodels. One way to represent contemporary understanding, therefore, is to take a Lowry model and rewrite it, incorporating these developments. This is done in essence in Wilson and Bennett (1985), chapter 15. A critical element of being able to model land-use/transport interaction is to be able to represent the fact that the supply of land is not infinite. Lowry, in his original model, built in these constraints explicitly.

As already seen, equilibrium models are essentially concerned with fast dynamics and are appropriate in cases where forecasting is for a relatively short time into the future. The dynamic models tackle those structural variables governed by the slow dynamics. The two kinds of model can in principle be integrated; there have been a number of approaches to this task. Perhaps the most ambitious has been led through international collaboration from Italy. The results are published in a series of volumes (Bertuglia et al., 1987; Bertuglia et al., 1990; Bertuglia et al., 1994).
The ideas presented above for the retail model were fully worked out for other sectors during the 1980s (see Wilson and Birkin, 1983, for agriculture; Birkin and Wilson, 1987a,b, for industry; Clarke and Wilson, 1983, for residential location; and Clarke and Wilson, 1985b, for land use). These models represent model-based replacements for the geographic classics: von Thunen, Weber, Hoyt and Harris, Ullman, and Christaller and Losch respectively.

Some broad conclusions from these theoretical ideas are:

- The mechanisms now elucidated do give a plausible account of urban dynamics.
- They demonstrate the essential nonlinearities involved. In the retail case, these arise from the presence of the denominators $\Sigma_k W_k^a \exp(-\beta c_{ij})$ in equation (14); there is a high degree of nonlinear interdependence between the $W_k$s. The effect of this term is to represent the competition at shopping centres for consumers at each residential zone in the light of the pattern of competing shopping centres, $\{W_k\}$, $k \neq j$.
- Because of the backcloth problem, and the nature of these kinds of equations, there are many possible equilibrium solutions. The evolution of the system in any particular case will be very dependent on the initial conditions.
- This reveals a fundamental point. While equations for the slow dynamics in variables such as $\{W_k\}$ can be written down, which represent urban structure, and these equations can be used to gain understanding, they cannot be used for prediction in the conventional sense. It is possible to predict, for example, when a parameter like $\beta$ decreases to a value which is critical so that a more concentrated and less dispersed retail structure will occur; but it is not possible precisely to predict what that pattern would be in a particular situation.
- It is still possible to use this kind of model in planning — because what cannot be predicted can be fixed exogenously in a plan, and it is then possible to use the models to calculate arrays of performance indicators for evaluation purposes — and, because the framework is dynamic, to test for stability and directions of change.

**Alternative approaches**

The equation systems that occur in dynamic analysis also occur in other fields. This generates insights — for example, when it is recognised that the equations (12) behave mathematically like those that occur in ecology to model multiple species systems, where the species are competing for a fixed supply of resources (see May, 1971, 1973; Maynard Smith, 1974; and Wilson, 1981). Analogously, retailers are competing for a fixed supply of consumers. Here, unlike the competing species in the ecological model, the outcomes depend critically on spatial patterns. It was May who first exposed the possibilities of oscillatory or chaotic behaviour for these kinds of systems; his ideas can be carried through to retail modelling. However, it is unlikely in practice that urban systems would approach chaotic states (and this helpfully assists in bounding the $\epsilon$-parameters). The relationship of chaos theory to spatial dynamics has been explored in this Journal by Cochrane (1992). Another school of thought developed in Belgium by building on the work of Prigogine on self-organising systems in Chemistry (see Nicolis and Prigogine, 1977; with applications in Allen and Sanglier, 1972).
The theory of nonlinear systems has been characterised in recent times as *complexity theory*, and the leading exponents of this have been located in the Sante Fe Institute. Examples of their work are to be found in Anderson *et al.* (1988); Arthur (1988, 1994); Lane (1993, 1995); Langton (1989); and Langton *et al.* (1995).

A totally different picture can be used for generating dynamic models involving a simulation of a hypothetical but large population — the so-called 'microsimulation methodology'. This was first deployed in urban geography as long ago as the mid-1970s (Wilson and Pownall, 1976), but is now becoming more popular in economics, geography, sociology and transport planning (see Duley and Rees, 1991; Gilbert, 1995; Kain and Apgar, 1985; Kain, 1987; Hancock and Sutherland, 1992; and Harding, 1990).

Finally, two other approaches concerned with change should be noted: migration within demographic modelling; and inter-regional input-output modelling. These are usually applied to systems with large areal units and provide the regional context for urban models. As these units are made smaller, however, they can become alternative frameworks for land-use/transport models. This possibility for using economists' trade theory models in relation to location theory is developed explicitly by Krugman (1991, 1993). Input-output models were used explicitly to estimate the development impacts of transport investments by Liew and Liew (1985).

*Interactions with Geographical Information Systems developments*

As seen earlier, the development of dynamic models in the late 1970s was followed by the advent of the PC and better computer graphics. These also facilitated the development of *geographical information systems* (GIS) (see Coppock and Rhind, 1991, for a historical review). There were two major consequences. First, model-based planning systems became more user-friendly, and this, in principle, facilitated application; second, the possibility of combining human and machine intelligence in tackling problems which are analytically intractable was opened up.

Traditional GIS are essentially mapping systems with added facilities for overlays among other aids. That there is a potential for constructing model cities in graphic detail can be seen from the best-selling computer game SimCity, used in teaching. Unfortunately, the underlying model which drives it is not as realistic as it could be (see Macmillan, 1996, for an analysis), but GIS potential can arise from unusual directions. In a modelling context, GIS can be used to present not only core maps but model outputs and performance indicators on a customer-driven basis. In this way, real intelligence can be added to GIS — and this leads to the concept of *intelligent* geographical information systems — or *IGIS* (see Birkin *et al.*, 1996). Both the analytical and planning consequences of these developments are likely to turn out to be particularly important.

In traditional algebraic equation-solving, it is possible to develop an analytical feel for solutions of equations. With large nonlinear systems, this is much more difficult. Solution spaces should be explored through direct exploration of computer-graphic representations of them, and new insights can be gained in this way. In essence, this is the case when there is a combinational problem: in this instance, it is the multiplicity of equilibrium solutions and the complexity of the dynamics that necessitate these methods.
A second kind of combination problem arises in planning: the complexity of the dynamics expands the richness of future scenarios. Again, it is potentially fruitful to explore them on a person-machine basis.

3. The Applications

Some university town or city planning departments did — and still do — undertake significant research programmes in urban modelling. Their history is well charted in the papers by Batty (1994) and Wegener (1994) in a recent special issue of *The Journal of the American Institute of Planners*. They take a retrospective view of another paper which became a significant element in the critique of land-use/transport models by Lee (1973). He offers his retrospective (Lee, 1994) with contributions also from Harris (1994) and Klosterman (1994). In 1973 Lee argued that models were complicated, expensive, mechanical, and did not work. Modellers, he argued, had not produced any new theory and were inadequately connected to existing theory. By contrast, the present paper argues that modelling offers a framework for carrying theory and then provides a basis for further theoretical development. Through dynamic modelling, it is now possible to make better judgements about the validity of predictions. What can be achieved with models has to be compared with the alternatives. Lee conceded in his 1994 paper that his original arguments were used “by some reactionary city planners who used it to justify their resistance to learning anything new and to vindicate their current practices” (Lee, 1994).

Modelling can provide the analytical core of the process. Planners should be able to provide the overall framework for the use of models. Planning typically involves design (the invention of alternatives); and the process of design is non-trivial (Alexander, 1964). Planning problems can be formulated as optimisation problems, but this is usually an impracticable or unrealistic approach (see May *et al*., 1995, for a contrary view). For this purpose, planners need a good understanding of the performance indicators, the key to arguing that the best alternative is chosen. To do this, planners need to have a good understanding of the policy framework, inevitably engaged with the political process.

If the main brief of the planners is to recommend the ‘shape’ of cities, then it is usually left to the engineers to design, build and manage the transport systems. Engineers, therefore, can use models as design tools: for predicting loads to be carried using assignment programs; as the basis of network optimisation routines; and as the basis of evaluation methodologies since, typically, they will have concerns with project appraisal. They can use models to optimise with respect to management policies, pricing, and regulation. Less concerned with model development as such, they have played a major role in the use of models, as the dominant practitioners in recent years.

In the set of linked models that represent land-use/transport interaction, it is probably fair to say that two sub-models, transport and retail, have become ‘standard’. Their deployment is highly professional and successful when they are deployed alone. ‘Retail’ can be interpreted very broadly to relate to any consumer services — public and private — and so extends into sectors like education and health.

In the transport case, the transport models of the 1960s and 70s have been refined to
a considerable degree. Many of these refinements have been in response to practical requirements to represent transport systems at different levels of detail and complexity (as, for example, in the case of SATURN — see van Vliet, 1982 — which may be used in both traffic management and strategic transport planning contexts). Additionally, the traditional four-stage model (trip generation, distribution, modal split, and assignment) is now often divided into different combinations of sub-models to reflect specific approximations in different applications’ contexts. Sometimes, to facilitate development and operation, the model has been applied as a fixed-trip matrix in conjunction with an assignment model. In certain cases, this has proved problematic. For example, in the standard cost-benefit methods used to appraise road schemes in the UK, there was no allowance for the generation of new traffic. This was changed by the conclusions of SACTRA (Standing Advisory Committee on Trunk Road Assessment, 1994); these conclusions were accepted by the Department of Transport, which has now commissioned extensive research into generated traffic. One aspect of this is traffic generated by ‘unanticipated development’ (Williams, 1997, private communication). Many of the applications in the UK have been led by a perceived need to integrate as many aspects of transport planning as possible (May, 1991).

Many of the most fruitful applications of the retail model have been in the private sector. This has been facilitated by the growth of companies, as described in Birkin et al. (1996), which covers model applications in retailing, health, water resources, the motor industry, and financial services. Particular attention should be drawn to the ability of the models to provide the basis for better performance indicators — even in terms of very simple traditional ideas such as catchment areas of retail centres (Clarke and Wilson, 1987b).

It does not seem to be the case that there is any standard, widely applied general model that addresses the land-use/transport interaction issue. The task of exploiting the insights arising from the new generation of dynamic models largely remains to be undertaken. However, Wegener (1994), in a wide-ranging review, estimates that there are about twenty groups around the world which have maintained this research impetus, mainly academically driven but with many working in applied contexts. There have been significant applications in the United States — mainly led by Putman (1983, 1991) using essentially an extended Lowry model with entropy-maximising sub-models — most recently driven, as noted earlier, by the requirements of air quality legislation (Still, 1996). This requires pollution impact studies to be carried out in the context of major transport proposals, and this in turn demands the rigour of land-use/transport modelling. In Europe, there have also been significant applications, notably by Echenique and his colleagues (Echenique et al., 1990). These activities have also been extended to South America, notably by de la Barra (1989). Several; European academics have continued to develop comprehensive models. Their work was also reviewed by Webster, et al. (1988) and Paulley and Webster (1991) as part of the ISGLUTI Project (which also incorporated the US experience through Putman). There are further details in papers by Mackett (1980, 1990, 1993), Wegener (1986a,b), and Nakamura et al. (1983).

The number of applications of general models has been much lower than in the case
of transport and retail models. The first — and perhaps the most important, if least recognised reason for this — is that there is now an understanding that while the general model is an exciting concept, our knowledge of its complex dynamics means that its application in planning is far from routine. Second, it remains true that the planning community has not developed a sufficiently serious interest and expertise in these models to be in a position to use existing knowledge (see Still, 1997, for an analysis of this). Third, analytically-based planning became unpopular from the mid-1970s onwards — possibly because it was more comfortable to face the very difficult political decisions associated with cities in a haze rather than with maximum clarity. (Though May, 1991, argues that some responsibility for this lies with the model-users who tended to present results as a blueprint rather than as a framework. See also May and Roberts, 1995.) There are exceptions where clearer needs have been identified (as, for example, in relation to air quality in the United States). Fourth, the pressures on public expenditure have probably meant that neither central nor local government has wanted to ‘afford’ modelling expertise. Overall, the decline in interest in the public sector since the optimism of the 1960s is well charted by Batty (1989, 1994). It may be argued that this would have been more worthwhile expenditure than much that has been undertaken. For example, in health, much more value could have been extracted from enormous expenditure on database development, had the associated strategies been underpinned by model-based analysis. (See Clarke and Wilson, 1985a, for an example of a modelling application in the health field.) Moreover, the new generation of models is significantly cheaper.

4. The Future of Modelling

4.1. Model Development

A position has been reached where there is a set of models that can be linked to provide a general and comprehensive model, within which it is possible to embed the questions and issues of land-use/transport interaction and planning.

If the knowledge so gained is adequately combined within a theoretical and modelling framework, and if enough research is undertaken to achieve a sound knowledge of functional forms and parameter values, then it is possible to say that the means exist to understand land-use/transport interaction. A theoretical framework can be assembled that integrates these different perspectives, although this has not been fully articulated (see Wilson, 1995, 1996).

The reproductive capabilities of the models are good. But the nature of the system — and the models that represent it — is such that it is not possible to predict land use in the way it might have been envisaged at the time of the Penn-Jersey Transportation Study. However, as has been argued, it should be possible to predict the type of cities that would evolve under particular conditions, and this at least provides warnings about what to plan for. It should be possible to interpret the history of urban development in various situations, and the knowledge gained from this kind of empirical work would be
Land-use/Transport Interaction Models

A. Wilson

invaluable in a planning context. (The models have even been applied in ancient history in a context which is relevant to archaeology: see Rihill and Wilson, 1987a,b, and 1991.)

Modellers and their associates in academia and in planning should continue their work and find ways of undertaking large-scale empirical development of the systems currently available. The difficulty is that this needs substantial funding. Yet it would be much cheaper to do this now than it was in the 1950s and 60s, because of more powerful and less expensive computing technology, and the existence of communications networks which should make it easier (and, again, cheaper) to assemble the relevant data.

However, there are continually new theoretical problems to solve. It has not in general been possible to interest the professional mathematical community in the associated mathematical problems of land-use/transport modelling, not least because of the attraction of areas like theoretical physics. But, almost certainly, progress could be made with more expertise. For example, reference was made in Section 2 to the backcloth problem: the fact that a dynamic analysis of \( W_i \) involves the whole backcloth, \( \{ W_k \}, k \neq j \), which will change as \( W_j \) changes. The complexities of dynamics in the high-dimensional spaces needed to represent cities are a long way from being fully understood. (See Wilson, 1988, for a full articulation of this problem.)

More could be done to complete our understanding of the model systems available if full reconciliation were possible to link alternative approaches. Much work has been done in this respect. More may well reveal new syntheses that combine the benefits of various perspectives. This could, therefore, be a fruitful line of research. In such explorations, the aggregation problem always looms large. This is best known in economics, but it is particularly acute in land-use/transport modelling because the essence of what is represented is at a scale 'above' the micro and 'below' the macro ('bundles of trips', 'land use in a zone').

It is unlikely that this problem will be resolved in general terms; rather, it is necessary to become more skilled at selecting that 'picture' which is most appropriate to a particular analysis task, and to be able to connect that picture, up and down, to other scales as necessary. A good example from another field is provided by spatial demographic models (as, say, formulated in Rees and Wilson, 1976) and spatial economic models (Stone, 1967, 1970). At what point do migration submodels in spatial demographic models become dynamic residential location models? Recent work at different scales is exemplified by Rees (1996) and Rees and Phillips (1996).

Integrated models have always been connected to aggregative demographic and economic models. An increasing concern of planners in recent years has been the extent to which they can influence economic development, for example by attracting inward investment through transport investment. and this remains a major research area. In the UK, the topic has recently been taken up by SACTRA, which has released an extensive bibliography on the subject. Gillen (1996) reviews the American experience.

The performance indicators that can be developed from the model are almost certainly under-exploited. For example, it has been noted that most of the models predict some measure of rent, but little or no work has been done on comparing these estimates with real rents.
It has been noted that the dynamic models show that there are many alternative futures for a given set of parameter values, but the actual futures will be strongly dependent on the initial conditions at any one time, and on small perturbations, such as individual or governmental decisions. There is a potentially fruitful research area here. First, it should be possible to categorise types of future for a variety of circumstances, and this would offer considerable insight for planners of all kinds. Second, although high-dimensional spaces are involved, it may be possible to explore the alternative specific futures that could evolve within a category.

4.2. Using Models
Two points were noted earlier in different contexts: first, the success of GMAP in a wide range of application (Birkin et al., 1996); and second, the need for deeper and more extensive empirical work in an academic context. The irony is that the first is being strongly supported by companies who find the planning applications invaluable to them, while the second, which would deepen understanding and possibly spin out into the public sector, is seriously under-funded. The experience of GMAP and others shows that the work can be done. The test is to focus the methods that can be applied there in an academic context.

There is no doubt that the agenda of city and regional planning is more important (and difficult) than ever. In the last three decades, social polarisation has increased to the extent where not only are the problems very severe for those who might be considered to constitute an underclass, but some problems — such as higher crime levels — manifest themselves by spilling out into middle-class areas.

This polarisation is rooted in differential access to jobs (and, for many, no jobs at all), which must also have its origins in differential access to services, and especially education. Although there has been a small movement of middle-class residents back to the city centre, the flight to the suburbs and now, even more so, the exurbs (rural small towns and villages) continues apace. This exacerbates transport congestion, as the M25 around London illustrates on a daily basis, and remains to be fully understood and explored through land-use/transport interaction models. Town planners in the UK, who might be at the heart of this, have been driven more towards either the formalities of the planning system (who gets planning permission for what, without much analytical input) or, as noted earlier, operating as economic developers whose prime concern is to attract inward investment in competition with other cities.

It is also the case that the European and American concerns with traffic growth, congestion, and particularly pollution, are being mirrored in the UK, not least through the publication of Planning Policy Guidance 13 (Department of the Environment and Department of Transport, 1994; see also Standing Advisory Committee on Trunk Road Assessment, 1992, and Owens, 1996). These concerns, combined with the lack of resources for public investment in infrastructure to meet what is perceived to be an ever-expanding need, is continuing to shift the agenda towards trip reduction and more effective management. Model-based systems will continue to be needed to support these endeavours.
5. Concluding Comments

It seems to be the case that most of the equilibrium models developed in the 1960s and 70s have stood the test of time in offering at least reproductive power — and particularly in applications where relatively short time-scales are involved. They can be elaborated, and they can be recombined in various ways, and there will be continuing advances in technique. These models provide the economic core for policy work, particularly in cost-benefit analysis, through the developing capacity to calculate social benefits and the interpretation of model variables in terms of rents and consumers' surplus. It has also been argued that there have been major advances in modelling since the late 1980s through the application of contemporary mathematical skills that facilitate the development of dynamic models. These have still not been tested or exploited to anything like their full potential. In particular, there is much more experience with the sub-models of an integrated land-use/transport system than with the general model itself.

The agenda for the application of the models has changed. There are greater concerns with management and manipulating land use to control traffic and to reduce pollution levels. Over the next decade, it is likely that planners will have to call on a full range of instruments, including land use, investment, pricing, and regulatory policies to address the expected increases in traffic. Again, it can be argued, the models exist to test the impacts of various schemes. While the range of application is to be admired, it seems clear that there remains tremendous scope for broadening it — and, of course, much would be learned from new experience of both modelling and its application.

References


Land-use/Transport Interaction Models

A. Wilson


Land-use/Transport Interaction Models

A. Wilson


Standing Advisory Committee on Trunk Road Assessment (SACTRA) (1992): Assessing the Environmental Impact of Road Schemes. London: HMSO.


25
Young, E. C. (1924): The Movement of Farm Population. Bulletin 426, Cornell University, Agriculture Experiment Station, Ithaca, N.Y.

Date of receipt of final manuscript: June 1997