Transportation systems, retail environments and pedestrian trip chaining behaviour: modelling issues and applications

Timmermans, H.J.P.; Hagen, van der, X.H.A.M.; Borgers, A.W.J.

Published in:
Transportation Research. Part B: Methodological

DOI:
10.1016/0191-2615(92)90019-S

Published: 01/01/1992

Document Version
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the author's version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
TRANSPORTATION SYSTEMS, RETAIL ENVIRONMENTS AND PEDESTRIAN TRIP CHAINING BEHAVIOUR: MODELLING ISSUES AND APPLICATIONS

HARRY TIMMERMANS, XAVIER VAN DER HAGEN, and ALOYS BORGERS
Eindhoven University of Technology, Department of Architecture, Building and Planning, Postvak 20, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

Abstract — Revitalisation processes in Western inner-cities and the construction of large-scale shopping malls in developing countries and countries with substantial economic growth require the development of models which allow the prediction of the likely effects of changes in the retail environment and the transportation network on pedestrian choice behaviour. This article gives a state-of-the-art review of existing models of pedestrian movement, discusses a few applications of these models, and identifies some important directions for future research.

INTRODUCTION: PROBLEM CONTEXT

Although the total number of studies of pedestrian behaviour is still relatively small, recent problems of urban and transportation planning lead us to believe that pedestrian movement will be a focus of increasing attention in the next decade. Two pertinent fields of application come to one's mind if one thinks about pedestrian movement: revitalisation processes in Western cities and the construction of large-scale shopping malls in developing countries and countries with substantial economic growth.

One of the most important recent problems in urban planning in many European and American countries concerns the revitalisation of their inner cities. During the last decade the position of many inner city areas has been seriously weakened as the cumulative result of a series of interrelated changes in consumer demand and methods of business supply. The process of outward migration has shifted the demand for retail facilities from the city center to suburban locations. This increasing demand in suburban locations has stimulated the construction of new out-of-town shopping centres, which in turn caused an increasing reduction of retail turnover in inner city areas. This process was further amplified by changes in the scale economies of business. Many of the larger chains looked for larger sites at lower costs at highly accessible locations, which could only be found at the periphery. This again increased the drainage of retail turnover in many inner city areas.

In response to these developments, many planning authorities have attempted to revitalise the inner cities. Their policies often include plans for new retail facilities, reflecting the assumption that urban development is triggered by elements such as in-town hypermarkets and superstores. Often plans for retail developments are integrated with plans to improve the accessibility of inner cities as exemplified by new or improved rail or road networks and parking facilities. Although revitalisation projects undoubtedly have a potentially positive effect on the position of inner city shopping areas in the regional functional hierarchy of shopping centres, they may also have a negative impact on the existing spatial retail structure within the inner city. The basic problem in this respect is to alleviate the ill effects these projects could have on the traditional commercial activity and to favour their proper integration to the built environment. An important research question therefore is whether the positive effect due to an increased attractiveness of the city centre will suffice to counterbalance the increased competition within the inner city area. Evidently, such research should preferably be complemented by macro-level research which identifies the relationships between urban change and the pattern of people movement.

While many Western countries face problems of adjustment and restructuring, many developing countries and countries with substantial economic growth experience a tendency to build new large shopping malls to satisfy the ever-increasing demand for shop-
ping goods. For example, Hagishima, Mitsuyoshi, and Kurose (1987) note that large shopping facilities have been located in many residential districts in Japan and that Japan's Ministry of Construction has recently established a system of pedestrian-space improvement. The construction of such facilities changes pedestrian flows substantially and often the surrounding transportation network is not appropriately designed. Therefore, there is a need to understand the relationships between pedestrian flows and characteristics of both the retail environment and the transportation network; such information may be a valuable input in the design process.

From a research point of view, problems like the ones discussed above require the development of a model which allows the prediction of the likely effects of changes in the retail environment and the transportation network on pedestrian choice behaviour. Such a model should depict the linkages between pedestrian movement, the characteristics of the retail structure and the transportation network. Pedestrian movements are essentially a form of multipurpose-multistop or trip-chaining behaviour: the problem is to decide in what sequence to buy a number of different goods. As such, pedestrian movement can be considered as some form of dynamic behaviour in that a consumer's choice behaviour at some stop during his trip will be influenced by both the decisions made at previous stops and the decisions to be made at future stops during the trip.

As part of an ongoing project which seeks to improve existing models of pedestrian movement and in line with the objective of this conference, the present paper gives a state-of-the-art review of existing models of pedestrian movement, discusses a few applications of these models to problems of urban and transportation planning in The Netherlands, and identifies some potentially important directions for future research. The paper is organized as follows. First, different modelling approaches will be discussed. Then, in the next section, two applications will be discussed. Finally, some avenues for future research will be identified.

MODELLING APPROACHES

Over the years a number of models of pedestrian movement have been developed in a variety of disciplines. Four different modelling approaches may be distinguished: regression models, spatial interaction/entropy-maximizing models, Markov models, and simulation models. Each of these will be discussed in turn.

Regression models

A typical example of a regression model of pedestrian movement is the model developed by Sandahl and Percival (1972). Their model is based on the assumption that pedestrian flows are characterised by the accumulation of flows around interesting objects. This implies that their model does not describe the sequence of destination choices, but rather the accumulation of pedestrians around objects. Their model can be expressed as follows:

\[ T_i = \alpha_0 + \sum_{k=1}^{n} \alpha_k X_{ik} \]

where:

- \( T_i \) is the number of pedestrians in zone \( i \);
- \( X_{i1} \) is the effective floorspace for retail facilities in zone \( i \);
- \( X_{i2} \) is the number of long-term parking lots in zone \( i \);
- \( X_{i3} \) is the number of bus routes that stop in zone \( i \);
- \( X_{i4} \) is the centrality of zone \( i \);
- \( X_{i5} \) is the number of pedestrians who cross the city centre at zone \( i \);
- \( X_{i6} \) is the number of bookstalls in zone \( i \);
- \( X_{i7} \) is the number of public places in zone \( i \);
- \( X_{i8} \) is the number of short-term parking lots in zone \( i \);
Modelling issues and applications

\[ \alpha_k, \quad k = 1, \ldots, 8 \] are parameters to be estimated;

\[ \alpha_0 \] is the intercept of the regression equation.

This model provided a reasonable description of the observed number of pedestrians in various parts of the centre of Orebo.

A similar approach was followed by Pushkarev and Zupan (1971) in Manhattan. Evidently, regression models like the one described in this section have a number of serious limitations. They do not provide any insight into the factors influencing route choice behaviour, the sequencing of visits, complementary relationships, the strength of functional relationships between streets and types of goods, the influence of the locational pattern of shops on pedestrian movements, etc. Consequently, they are difficult to use for predicting the likely effects of changes in the retail structure or the transportation network on pedestrian movement. They may be useful though to assess pedestrian safety, the need for special facilities in pedestrianised streets, etc. around objects in the city that attract pedestrians.

Spatial interaction/entropy-maximizing models

These models have been very popular for modelling all kinds of interactions. Hence, it is not surprising that many authors have applied Wilson's ideas (Wilson, 1970, 1971) to studies of pedestrian movement. A typical application can be found in Butler (1978). Her model consisted of three submodels: a trip-attraction submodel, a trip-distribution submodel and an allocation model. The attraction submodel predicts the total number of stages that starts or ends in a particular zone of the city centre. A stage is defined as the trip between two successive stops. The attraction of the first stage, disaggregated according to transport mode, is related to variables such as parking facilities, and bus stops. The remaining stops were predicted on the basis of a Hansen type accessibility measure, incorporating variables such as floorspace and number of jobs. The distribution submodel predicts the distribution of the trips across the various destinations. More specifically, the entropy-maximizing version of the spatial interaction model was used. This specification was also used in studies conducted by Ness, Morral and Hutchinson (1969) and Johnson (1972). The allocation submodel then assigns the predicted flows between the origins and destinations to the transportation network. Butler used an all-or-nothing assignment procedure, implying that each pedestrian is assumed to choose the shortest route between the origin and the destination.

The goodness-of-fit of the model for data pertaining to the city centre of Liverpool was disappointing. The attraction submodel performed acceptably only for some areas. As a consequence, the predictive ability of the distribution submodel was not very good, while in addition the data showed that many pedestrians did not choose the shortest route between origins and destinations.

A very similar model was applied in Fukuoka City, Japan by Hagishima et al. (1987). They used a production, a distribution and an allocation submodel. They started by representing the road network by a series of links and nodes. Trip generation at each node was determined by the following equation:

\[ P_j = \alpha P_{i0} + \beta C_i + 0, \]

where

- \( P_j \) is the pedestrian shopping trip generation at node \( i \);
- \( P_{i0} \) is the population at node \( i \);
- \( C_i \) is the parking capacity at node \( i \) and is assumed to reflect the number of car arrivals from external zones;
- \( 0, \) is the number of pedestrians who enter from external zones;
- \( \alpha, \beta \) are parameters to be estimated.

The nodes are divided into nodes associated with shopping facilities (set \( Z_1 \)) and those
associated with bus stops, ferry port and train station (set Z). The latter set of nodes is used to model trips to external zones. The distribution of shopping trips is modelled by a partly constrained entropy-maximizing model of the following type:

\[ T_{ij}^{(1)} = A_i^{(1)} P_j W_j \exp(-\beta c_{ij}) \quad j \in Z_1 \]

\[ T_{ij}^{(1)} = A_i^{(1)} P_j B_j Q_j \exp(-\beta c_{ij}) \quad j \in Z_2 \]

where

- \( T_{ij}^{(1)} \) is the number of pedestrians who go from node \( i \) to node \( j \) at the first stop;
- \( A_i^{(1)} \) is a balancing factor;
- \( W_j \) is a measure of attractiveness of node \( j \), operationalised as \( m^2 \) store floor space and adjusted by some constant ratio for supermarkets and large department stores;
- \( B_j \) is a balancing factor;
- \( Q_j \) is the number of pedestrians who shop in external zone \( j \);
- \( c_{ij} \) is a travel cost function between nodes \( i \) and \( j \) and operationalised in terms of weights assigned to links of the shortest route for traffic condition, pavement and street conditions reflecting the walking speed of pedestrians;
- \( \beta \) is a distance-decay parameter.

To allow for a chain, the trips between two shopping destinations are modelled as:

\[ T_{jk}^{(2)} = A_j^{(2)} r T_{ij}^{(1)} W_j \exp(-\beta c_{jk}) \quad j, k \in Z_1 \]

where \( r \) is the ratio of pedestrians who arrive at node \( j \) and continue their trip to destination \( k \). In practice, this ratio was set equal to 1.0.

Finally, the trip distribution of pedestrians \( Q_j \) who get off at node \( j \) (bus stop, train station or ferry port) and go home to node \( i \) was modelled by the following equation:

\[ T_{ij}^{(3)} = A_j^{(3)} Q_j P_i \exp(-\beta c_{ij}) \quad j \in Z_2 \]

Pedestrians were assigned to the road network according to their shortest route (all-or-nothing assignment). As in previous applications of spatial interaction models, the predictive results of the model left ample room for improvement. The Pearson correlation coefficient between observed and predicted pedestrian flows was 0.77, and observed flows were often over or underpredicted by several hundred percent.

A potential disadvantage of spatial interaction models is their failure to incorporate the basic mechanisms underlying pedestrian movement. Especially, these models lack any component to represent the sequencing of visits in any coherent fashion.

**Markov models**

The third approach of modelling pedestrian movements is the Markovian approach. Markov models have typically been used for analysing the kind and intensity of functional relations in multi-purpose trips. The assumption underlying these models is simple: only the last state occupied by the process is relevant in determining its future behaviour. That is to say, the location of the next stop is determined only by the location of the last stop. For every pair of stops the probability of choosing some destination given the choice of a particular destination at the previous stop can be empirically estimated from observed pedestrian flows and is represented by a transition matrix.

Thus, basically, these models involve an interaction matrix that describes the probability that a particular destination \( j \) will be chosen given that the previous stop was at destination \( i \). These transition probabilities are derived from observed choice patterns. The main advantage of these models lies in their ability to derive some general aspects of trip chaining behaviour by some easy manipulations of the interaction matrix. For example, the probability that a particular choice alternative will be chosen in a given number
of steps can be calculated simply by raising the matrix to a power which is equal to the number of steps. Likewise, the accessibility of the choice alternatives and the overall choice probabilities can be calculated by simple matrix operations. Examples of Markov chain models for this type of analysis can be found in Marble (1964), Horton and Schuldiner (1967), Wheeler (1972).

Notwithstanding the positive features of conventional Markov chain models, they also share some characteristics that limit their applicability. According to Thill (1988) these models produce poor results due to a set of interrelated rigorous assumptions, such as:

(a) Because the future state of the process depends only on the last occupied state it can be said that these models assume memoryless behaviour. To overcome this problem Nystuen (1967) suggested to describe the probability distribution of trip duration by a binomial process in which the decision is made to continue the trip or to go home. Another proposal is due to Kitamura (1983), who applied a Markov model the states of which were redefined in terms of stop purposes and the history of the chain, in accordance with the observation of a systematic order in the sequencing of stop purposes along trip chains. This latter solution is, according to Thill (1988), not practicable in the case of more than three or four purposes;

(b) The transition probability matrix is independent of time, characteristics of subpopulations and spatial environment. The former implies that in the case of trip purpose sequences the transition probability between a given pair of trip purposes does not vary over time regardless of the history of the trip. To relax the assumption of a stationary transition probability matrix, O’Kelly (1981) and Borgers and Timmermans (1986a), developed a time-varying Markov model. Horton and Wagner (1969) applied a Markov model, which was disaggregated for various subpopulations, while Kondo (1974) assumes that the ratios of the transition probabilities of one destination zone to the transition probabilities of all destination zones correspond to a gravity potential;

(c) The duration of stay at a stop does not vary over time. Kitamura (1983) performed statistical tests using large-scale origin-destination survey results and concluded that trip purpose linkages are not stationary within a trip chain, but are dependent on the past history. Various authors (e.g. Gilbert, Peterson, & Schofer, 1972; Lerman, 1979) proposed a semi-Markov model in which the duration of a stay is derived from a probability distribution.

Although much progress has been made over the past decades, the theoretical underpinnings of the Markovian approach are still relatively weak. Strictly speaking, Markov models are no-choice models in the sense that they do not include an explicit preference structure or choice rule. Basically, Markov models are descriptive; they do not allow predicting future behaviour given some changes in the variables which are supposed to govern the trip-chaining process. Recently, Borgers and Timmermans (1986a) have attempted to overcome this problem by assuming that trip-chaining is a form of sequential decision-making and that the choice of a destination at each step of this decision-making sequence is based on a deterministic preference structure and Luce’s choice axiom. They used O’Kelly’s time-varying Markov model of multi-stop multi-purpose trip chains, added to it an endogenous determination of interstop transition probabilities, and applied it to pedestrian movement.

More specifically, this model of pedestrian movement is based on the assumption that pedestrian movement can be represented as a multipurpose trip that consists of both intended and impulse stops. The intended stops are assumed to be the result of a multistep decision-making process of destination choice in which utilities are sequentially maximized. The impulse stops are assumed to be a function of pedestrian route choice behaviour, which, in turn, is conditional upon pedestrian destination choice. Thus, their model consists of separate submodels of destination choice, route choice and impulse stops. In particular, the model can be represented by the following set of equations. First, define the conditional probability that the \( j \)th link of the transportation network will be chosen on the \( m \)th stop for purpose \( h \), given that the \( i \)th link was chosen for purpose \( g \) on stop \((m - 1)\) as
where \( s(m) \) and \( d(m) \) denote respectively the purpose and the link associated with the \( m \)th stop. The city centre's entry points are defined in terms of the links of the network. Consequently, the transition probability matrix has a specific form (see Borgers and Timmermans, 1986a for details). The probability that a pedestrian will choose link-purpose combination \((j, h)\), given that he has chosen combination \((i, g)\) on a previous stop \( \ell \), then equals:

\[
p_{\ell+1}^{(j, h)} = \sum_{a=1}^{Z} \sum_{b=1}^{N} \sum_{c=1}^{Z} \sum_{d=1}^{N} \frac{p_{\ell}^{(j, h)}}{p_{\ell}^{(i, g)}}
\]

The total number of interactions between link-purpose combination \((i, g)\), which was visited on the \( \ell \)th stop, and link-purpose combination \((j, h)\), visited on or before the \( M \)th stop, then equals:

\[
n_{\ell}^{(j, h)}(M, \ell) = \sum_{k=1}^{M} q_{\ell}^{(j, h)}(k, \ell)
\]

The demand for retail facilities at link \( j \) in retail sector \( h \) can then be computed by:

\[
b_{\ell}^{(j, h)}(M) = \sum_{i=1}^{Z} \sum_{g=1}^{N} b_i^{(g)}(1) \sum_{k=2}^{M} q_{\ell}^{(j, h)}(k, 1)
\]

The development of a predictive model requires a submodel that predicts the derived transition probabilities. Borgers and Timmermans (1986a) assumed that the transition probabilities are a nonlinear multiplicative function of the total amount of floor space in retail sector \( g \) at link \( i \) and the distance separation between the links. In particular they assumed:

\[
p_{\ell}^{(j)}(m | m - 1) = \frac{F_i^{(g)} C_j^{(g)} \exp(c_{ij})^{-\beta_{jm}}}{\sum_{j' = 1}^{N} F_i^{(g)} C_j^{(g)} \exp(c_{ij'})^{-\beta_{jm}}}
\]

where

\( p_{\ell}^{(j)}(m | m - 1) \) is the probability that the \( j \)th link will be chosen for purpose \( g \) at stop \( m \), given that the previous stop was at link \( i \);

\( F_i^{(g)} \) is the total amount of floorspace for shop type \( g \) at link \( i \);

\( c_{ij} \) is the distance separation between links \( i \) and \( j \);

\( \alpha_{gm} \) and \( \beta_{gm} \) are purpose and stop specific parameters.

Given the predicted probability that a particular destination will be chosen at successive stops, the model then predicts the probability that a particular route will be chosen. In particular, it is assumed that route choice behaviour can be described by a multinomial logit model. Route length is used as the independent variable of this model. The route choice model is then used to calculate the number of pedestrians who pass through the various links of the network. Finally, the distribution of impulse stops across the network is predicted as a function of floorspace and the number of pedestrians that pass through the link. In particular, the following equation was used:

\[
F_i^{(g)} C_j^{(g)} = \frac{F_i^{(g)} C_j^{(g)}}{\sum_j F_i^{(g)} C_j^{(g)}}
\]
where

\[ p_i^{(g)} \] is the probability that an impulse stop for purpose \( g \) will be made at link \( i \);
\[ C_i \] is the total number of pedestrians passing link \( i \);
\( \delta_g \) and \( \theta_g \) are purpose-specific parameters.

**Simulation models**

A potential disadvantage of the modelling approaches discussed in the previous sections is that in order to be tractable the models have to be based on rather rigorous assumptions. It is often very difficult, if not impossible, to incorporate more complex theoretical notions.

Therefore, a number of researchers, working in very different disciplines, have used simulation methods for predicting pedestrian movement. Simulation methods have the potential advantage that a variety of rules can be used to simulate pedestrian movement.

Perhaps the most elaborate simulation model can be credited to Borgers and Timmermans (1986b). Their model works as follows: First, the number of goods \( I \) that is purchased by a random pedestrian is determined by drawing at random from a distribution that corresponds to the observed relative frequencies of purchases of goods. The required number of goods is obtained by drawing a random number within the range of accumulated integers that corresponds to observed purchase frequencies. The same procedure is used to identify the types of goods that are bought. This time the relative purchase frequencies, conditional upon the number of goods bought, is used as a reference distribution. The second step in the modelling process involves predicting the links where the selected goods are bought. It is assumed that the probability of buying the first good in the simulated sequence in a particular shopping street or link equals:

\[
P_{nl}^{(g)} = \frac{\left( \sum_{k=1}^{K} F_k^{(g)} \right)^{\alpha} \exp\left(-\beta \min_{r|n} \left[ \sum_{r' \neq r} d_{r'} \right] \right)}{\sum_{l'=1}^{L} \left( \sum_{k=1}^{K} F_k^{(g)} \right)^{\alpha} \exp\left(-\beta \min_{r|n} \left[ \sum_{r' \neq r} d_{r'} \right] \right)}
\]

\((n = 1, 2, \ldots, N; \ell = 1, 2, \ldots, L)\),

where;

\[ P_{nl}^{(g)} \] is the probability that a good in retail sector \( g \) will be bought at link \( \ell \) providing that the pedestrian departed from city entry point \( n \);
\[ F_k^{(g)} \] is the total amount of floorspace in retail sector \( g \) at location \( k \);
\[ \min_{r|n} \left[ \sum_{r' \neq r} d_{r'} \right] \] is the distance associated with the shortest route from city centre entry point \( n \) to link \( \ell \);
\( \alpha, \beta \) are parameters to be estimated.

Having determined the link where the first good is bought, the model proceeds by simulating the choice of the links of the remaining \((I - 1)\) goods, that is, if \( I > 1 \). Otherwise the model assumes that the consumer returns to the entry point from where he departed. This choice process is simulated using the same equation as described above. However, whereas for the first good the equation is based on the minimum distance between city-centre entry points and shopping streets, the equation for the remaining goods is based on distances between shopping streets. Together, these equations implicitly assume that pedestrians are engaged in sequential utility-maximizing behaviour.

Again, the simulation process proceeds by drawing random numbers. Each potential link receives a range of accumulated integer numbers proportional to the probabilities as indicated by the latter equation, and a link is assumed to be chosen if a randomly selected number falls within its range of integer numbers.
In the third step the model simulates the route choice behaviour of a pedestrian given information on the location of the entry point and the links where goods will be bought. It is assumed that the point of completion of this trip is the same as the point of entry. Dynamic programming techniques are used to simulate route choice. More specifically, the route choice problem was solved in terms of the generalized stagecoach problem of dynamic programming, which is applied in a series of successive steps, given two consecutive destinations in the simulated sequence of $I$ destinations.

This whole simulation process is repeated for each consumer at each entry point in a series of replications, such that the total number of replications for each city-centre entry point equals $C$. In addition, some additional quantities such as the total number of pedestrians passing through link $l$, the total number of goods in sector $g$ bought at link $l$, turnover levels and turnover-to-floorspace ratios were computed in a straightforward manner. Such quantities might prove useful in assessing the impact of alternative retail and transport policies on the viability of shopping streets within a city centre. It should be noted that the simulation model is based on the same kind of data as the time-varying Markov model. The two modelling approaches differ in terms of the submodels used to predict pedestrians' destination and route choice behaviours. Consequently, the appropriateness of the two modelling approaches for forecasting purposes are the same, at least as far as dependence on empirical data is concerned.

A somewhat less complicated simulation model of pedestrian behaviour was developed by Crask (1979). His model differs from the one outlined above mainly in that no attempt is made to incorporate the functional and spatial linkages that characterize pedestrian movement. In addition, because Crask applied his model to shopping malls with a simple layout, his model does not contain a submodel of pedestrian route choice behaviour. Finally, the attractiveness function differs. Crask also used a Hansen type accessibility function, but a choice alternative's utility or attractiveness was weighted by a term which expresses how closely a shopper's age, education, income and occupation matched those of the store's typical shoppers. To make this term operational, each respondent was asked to identify in terms of these four demographic characteristics the customers to whom particular stores appealed. These responses were then averaged to yield a profile of a typical shopper for each store. To weight the store utility for a shopper whose behaviour is simulated, the difference between each of the mean characteristics calculated for that store and the corresponding characteristic of the shopper was squared and divided by the standard deviation of that characteristic for the store. The differences were summed and the reciprocal of this sum was used to weight store utility. This procedure may be an interesting one from a theoretical perspective in that it allows segmentation, but generates additional problems when the model is used for prediction, because one would need to develop a submodel to forecast the sociodemographic characteristics of customers attracted to new stores.

APPLICATIONS

Maastricht

Maastricht is a historical city, located in the south of the Netherlands, very close to the Belgian border. Its city centre is divided into two parts by a big river. Both parts are connected by two bridges across the river; one connects the railway station to the historical heart, and there are many shops along this connection. The other bridge is located further to the north and forms no part of any commercial/retail structure. At the time of the study, one of the most important planning problems facing the urban and transportation planners of the city was that the viability of the shops between the railway station, an area called "Oud-Wyck", and the old, historical heart of the city was rapidly deteriorating. Their aim was to improve turnover levels in this part of the city centre without deterring trade too much from the other part of the city centre. This was not an easy task. At the same time the City Council wanted to develop an integrated retail/office/hotel complex along the river (very close to the problem area), plans for the construction
of a third bridge, additional parking garages in close vicinity to the historical heart, and a
new public transportation plan (which also favoured the historical heart) were already
approved by the City Council. Hence, if the area between the railway station and the
river was made too attractive, this could jeopardize the plans for the new complex.
Likewise, a too strong development of this complex and/or the historical heart could
have further adverse effects on the viability of “Oud-Wyck”. Their task therefore was to
develop all three areas in close harmony. One of the strategies was to improve the attrac-
tiveness of the shopping streets in “Oud-Wyck” and locate transportation terminals in
such a way that all streets would benefit most from pedestrian flows and strengthen
pedestrian flows between “Oud-Wyck” and the historical heart. Different scenarios were
assessed using the information that underlies Borgers and Timmermans’ (1986a, 1986b)
models of pedestrian movement.

In order to calibrate these models, a network, consisting of 88 links and 6 entry/
deptarture points was constructed first. The data for the calibration of the models were
provided by a random sample of shopping pedestrians. Pedestrians who were leaving the
city were interviewed. They were asked to mark on a map the route they had taken within
the city centre in the conduct of their shopping, the links of the network where they made
a stop, and the type of shop associated with each stop. They were also requested to
indicate which of those stops were impulse stops. The shops were classified into five
categories: groceries, clothing, department stores, markets and other.

Both the time-varying Markov model (Borgers and Timmermans, 1986a) and the
simulation model (Borgers and Timmermans, 1986b) were calibrated from these data.
We will not report all calibration results, but rather concentrate on the goodness-of-fit of
these two models as indicated by Pearson’s correlation coefficient. Of course, the correla-
tion coefficient has been subject of much criticism, but still is probably reported most in
the literature. In fact, we calculated more than 30 different goodness-of-fit measures,
some descriptive, some statistical, but the results largely seemed independent of the spe-
cific goodness-of-fit measure used. Table 1 presents the correlation coefficients between
observed and predicted arrivals at the destinations for both models. It shows that the
destination choice models perform satisfactorily as indicated by correlation coefficients
well above 0.90 for many types of shops. Evidently, the category “other” forms an
exception in this respect, but this is probably due to the heterogeneous nature of this
category. Table 1 also shows that both models perform almost equally well, but this
should not be surprising because the assumptions underlying these two models are very
similar and the models use the same data.

The results obtained for total pedestrian flows were even better. The correlation
coefficient between predicted and observed flows was 0.942 for both models, while the
correlation coefficient for the arrivals were 0.993 and 0.992 respectively. The demand per
link and type of shop was also predicted to be satisfactory: the correlation coefficient
was 0.987 for the Markov model and 0.985 for the simulation model.

The correlation coefficient which expresses the goodness-of-fit of the route choice
submodel was 0.839. The results apply to both types of models. The results of the
submodel of impulse stops were also good: the correlation coefficient ranged from 0.797

<table>
<thead>
<tr>
<th>Type of shop</th>
<th>Correlation coefficient</th>
<th>Number of observations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Markov model</td>
<td>Simulation model</td>
</tr>
<tr>
<td>Groceries</td>
<td>0.909</td>
<td>0.903</td>
</tr>
<tr>
<td>Clothing</td>
<td>0.939</td>
<td>0.942</td>
</tr>
<tr>
<td>Department stores</td>
<td>0.999</td>
<td>0.999</td>
</tr>
<tr>
<td>Markets</td>
<td>0.999</td>
<td>0.999</td>
</tr>
<tr>
<td>Other</td>
<td>0.781</td>
<td>0.639</td>
</tr>
</tbody>
</table>

Source: Borgers and Timmermans (1986a,b).
to 1,000. Finally, these submodels were combined, and the predictive results remained good. For example, all correlation coefficients for the impulse stops were greater than 0.96 and the percentage error is only 5 and 18 percent for impulse stops per type and link and total number of impulse stops per link, respectively.

Sittard

Sittard is a town located in the south of The Netherlands. It is a little smaller than Maastricht. Its city centre is characterized by a linear structure that runs from the railway station/bus terminal to the old market square. Access to this retail environment is good throughout: bus stops and parking lots are evenly distributed along this shopping street. At the regional level, the city has to compete with Maastricht and another town, Heerlen, and because those competing cities were in the process of improving the attractiveness of their inner-city shopping areas, the City Council of Sittard decided to permit the opening of a new in-town hypermarket. This hypermarket was to be located outside the existing linear retail structure, and the plan was to improve the accessibility to this new development by creating additional surplus parking spaces in its vicinity. The new hypermarket was to be linked to the existing shopping street by a series of new small shops. As a result of these developments the traditional linear retail structure would be replaced by a circuit (hypermarket market square), and it was expected that this might have negative effects on the sales levels in that part of the retail structure between the railway station and the hypermarket. Therefore, there was much pressure from the retailers who had to decide between staying in their current shops or relocating to the new shops. The federation of retailers was especially worried that some of the stores that attract pedestrians might relocate, which would further decrease the viability of this part of the city centre. The urban planners and transportation planners were faced with the problem of which policies related to parking and building programs to implement in order to minimize these possible negative effects as much as possible. It was decided to use Borgers and Timmermans' time-varying Markov model to support their decision-making process. The model was calibrated for this city and the results were used for simulating the likely impacts of possible scenarios in terms of sales levels in the various links of the retail environment. Finally, in cooperation with the city planners and retailers, the findings of the simulations were used to select the scenario that appeared most beneficial to all those involved.

The model was calibrated using the same kind of data and the same kind of measurement procedures as those described for the Maastricht case. However, because of the linear form of the retail environment, the network has more links and entry/departure points. More specifically, it consisted of 163 links and 14 entry/departure points. Also, the number of respondents was much larger; this time, 1719 respondents were interviewed. Moreover, the categorization of shop types differed slightly; that is, appliances were used rather than markets. Again, only the results pertaining to the goodness-of-fit of the models are reported in this paper. Table 2 gives the correlation coefficients for the destination choice submodel. It shows that although the results are less appealing than those obtained for Maastricht, they are still satisfactory.

Data on pedestrian behaviour were collected for a Tuesday and a Saturday. All

Table 2. Goodness-of-fit of the destination choice submodel
(Sittard)

<table>
<thead>
<tr>
<th>Type of shop</th>
<th>Correlation coefficient</th>
<th>Number of observations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tuesday</td>
<td>Saturday</td>
</tr>
<tr>
<td>Groceries</td>
<td>0.842</td>
<td>0.802</td>
</tr>
<tr>
<td>Clothing</td>
<td>0.894</td>
<td>0.895</td>
</tr>
<tr>
<td>Department stores</td>
<td>0.964</td>
<td>0.945</td>
</tr>
<tr>
<td>Appliances</td>
<td>0.739</td>
<td>0.786</td>
</tr>
<tr>
<td>Other</td>
<td>0.852</td>
<td>0.940</td>
</tr>
</tbody>
</table>

Source: Borgers, Timmermans & van der Waerden (1988)
submodels were calibrated for these two days separately. The route choice submodel performed reasonably as evidenced by a correlation coefficient of 0.749 and 0.680 respectively. The correlation coefficients which represent the goodness-of-fit of the submodel for impulse stops ranged from 0.677 to 0.987 for the Tuesday and from 0.847 to 0.954 for the Saturday.

DISCUSSION

The aim of the present paper has been to discuss existing models of pedestrian movement in terms of their theoretical underpinnings and reported goodness-of-fit. It should be emphasized that the discussion has been restricted to modelling approaches that actually have been applied to pedestrian behaviour. Models such as nested logit models that have been applied to other types of multi-purpose, multi-stop behaviour, and models of dynamic choice behaviour were not discussed in this paper. These models are discussed elsewhere (see e.g. Timmermans and Borgers, 1985; Thill and Thomas, 1987; Kitamura, 1988). This state-of-the-art review suggests that over the years a few attempts of modelling pedestrian movement have been made in a variety of disciplines. However, this review also suggests the need for substantial improvement of existing models. First, this review demonstrates that many recent developments that have improved conventional choice models have not yet been incorporated into models of pedestrian movement. Perhaps more importantly though is that the theoretical underpinnings of most models are still relatively weak. Most models combine models that have been applied in other choice contexts before and consequently often lack mechanisms that are typical for pedestrian destination and route choice behaviour as a form of dynamic travel behaviour. In particular, most models fail to incorporate mechanisms that represent the functional and locational relationships between the various stops of a pedestrian's shopping trip. Regression models describe the accumulation of demand at various links of the network. No attempt is made to model dynamic aspects that are involved. Spatial interaction models arbitrarily break down the trip chain and represent attempts of modelling resulting flows. Markov models and simulation models have more to offer in this respect, although most of these models are based on the limiting assumption of time invariant utility functions: the utility of visiting some shop remains invariant regardless of the shops that will be visited later or that have been visited previously on the same trip. Of course, Borgers and Timmermans circumvent this problem, but their models are still rather ad hoc. Their solution is to estimate stop-specific models by assuming sequential utility-maximizing behaviour; no attempt is made to formulate a more general theory and incorporate dependencies explicitly in the model.

Therefore, two research avenues need further attention in future research. First, in line with existing choice models, the various submodels could be improved by incorporating the latest developments in the original models. For example, the destination submodel could be improved by introducing context-sensitive utility functions (see e.g. Borgers and Timmermans, 1987, 1988). This would allow assessing the influence of the similarity among the choice alternatives and their relative location on choice probabilities. Likewise, Kitamura's concept of prospective utility (Kitamura, 1984) could be incorporated in the model to estimate the effect of the attractiveness of choice alternatives visited later in the chain on current choice probabilities. Also, the submodels of route choice behaviour could be improved. Most models of pedestrian movement are still based on the assumption of distance-minimizing behaviour. However, empirical findings suggest this decision rule to apply only to a minority of pedestrians. The logit model of route choice might thus be viewed as an improvement in the sense that route choice is taken to be a probabilistic rather than deterministic phenomenon. It has the associated problem that different routes are assumed to be independent. To avoid this problem, a multinomial probit model could be used to account for the covariance in alternative routes. It is also possible to model route choice as a function of the functional, morphological and aesthetic characteristics of the streets, or, alternatively, of pedestrians' evaluations of such attributes. This is not to say that such improvements would also increase the predictive success of the
model. In fact, the authors have tested many such alternative specifications in different application contexts, among which pedestrian movement, but often found the predictive validity of the model to improve only slightly.

However, improvements like these would still be based on rigorous assumptions such as sequential utility-maximizing behaviour. An alternative approach therefore would be to develop conceptual frameworks that allow different decision strategies, both temporally and spatially to exist. Some pedestrians may reveal sequential utility-maximizing behaviour, others simultaneous utility-maximizing behaviour, while still others may reveal suboptimal behaviour. In fact, empirical research (van der Hagen, Borgers and Timmermans, 1991) suggests that pedestrians may apply such different decision strategies when shopping. The local-distance minimizing strategy, underlying most models of pedestrian route choice behaviour, is only one of various possible decision heuristics pedestrians may apply. For example, Säisä and Gärling (1987) found, both in a laboratory setting and in real-world environments, that individuals tend to use global-distance minimizing strategies rather than local-distance minimizing heuristics when the latter heuristic would result in a substantially longer distance required to complete the tour. In two other projects (Gärling et al., 1986, Gärling, 1987) they found that the choice between these two heuristics is dependent upon the cognitive representation of the environment. If individuals possess some maplike mental representation of the environment, they were able to reveal simultaneous utility-maximizing behaviour, whereas otherwise they tend to minimize distance locally in a sequential decision-making process. These findings were supported by analyses of actual behaviour (Gärling and Gärling, 1988). The most frequently observed behaviour is that individuals tend to first choose the destination farthest away from their entry point and then minimize distance locally back to the point where they entered the city centre. Because pedestrians apparently use different decision heuristics, the predictive success of models of pedestrian destination and route choice behaviour might be substantially improved if different decision-making strategies could be incorporated into existing modelling approaches.

To accomplish this, it seems important to distinguish between planned and unplanned or impulse stops because it may be assumed that trip structures depend primarily on planned stops. Pedestrian destination choice behaviour will depend on the location and attributes of the kinds of shops that should be patronized. In contrast, unplanned stops relate to those visits that are decided upon during the trip. It may result from needs arising from merely passing certain shops, or the effect of items purchased during the trip on future demand.

A second factor refers to the issue of whether a pedestrian tends to always buy certain items in the same store (fixed destination) or that he/she patronizes different stores to purchase a particular item (unfixed destination). The distinction is important in that it segments pedestrians into two groups. The first group constitutes the store-loyal group. Their choice set is highly restricted. In fact, in the short term it consists of the fixed destinations only. The second group consists of those pedestrians who frequent different stores. They are not necessarily loyal to one particular store, and their choice behaviour should be modelled.

The combination of these factors gives rise to different types of pedestrians. At one extreme of the spectrum, there is the pedestrian who is involved in rational, routine-like, habitual behaviour. At the other end, there is the pedestrian who maybe goes shopping just for fun, and whose behaviour is based completely on impulse stops. It could be argued that none of these pedestrian types can be modelled from a choice perspective, because either their behaviour is totally planned or their behaviour is not based on any choice considerations at all. Perhaps the most interesting type of pedestrian therefore is the one in between these two extremes who has to decide what destinations, in what sequence and alone what route to choose, but, again, the modelling of their behaviour should be more elaborate than existing models suggest, because their behaviour might stem from various heuristics. Even for the simple case in which we assume pedestrian destination choice behaviour to be influenced by some distance variable only, there are still several decision heuristics that a pedestrian can apply to organize his journey. A
pedestrian might be involved in a sequential decision-making process. Because we have assumed some distance-driven process, this decision heuristic would amount to pedestrian behaviour that is not based on some *a priori* sequence of trip purposes, but rather on the actual configuration of land uses. A rational traveller would select that destination closest to the point where he/she entered the shopping area, which allows him/her to conduct one of the purposes. In the following segment of the journey, a second destination is sought closest to the one chosen at the previous stop. This process continues until all purposes have been conducted.

A little more complex situation occurs when a pedestrian is faced with a given set of destinations that has to be patronized. This set might be the result of a previous decision-making process in which an individual has decided which destinations to patronize, or might be the result of a purpose-sequence-driven process. This situation differs from the one discussed previously in that a choice constraint is imposed on the journey. Again, a pedestrian might apply various decision heuristics. As in the previous case, a pedestrian might apply a sequential decision-making process and minimize the distance travelled on each segment of the journey. This would imply that a pedestrian’s problem of deciding where to shop, in what sequence, and along what path through the network is broken down into a sequence of single choices. Since we have made the simplifying assumption that the nonlocational attributes are irrelevant, a pedestrian engaged in sequential decision-making attempts to minimize distance between all successive pairs of destinations that make up the entire journey. The application of this “local-distance-minimizing heuristic” implies that a pedestrian does not make any explicit simultaneous choice regarding the sequencing of destinations. The sequence of shop visits merely results from the distance-driven decision heuristics.

Alternatively, a pedestrian could choose destinations in such an order that the total distance travelled on the entire journey is minimized. This would imply that the pedestrian simultaneously decides about the sequencing of the destinations and his route choice behaviour. This “total-distance-minimizing heuristic” is an example of a simultaneous decision-making process in the sense that an individual attempts to minimize the total distance travelled. Mathematically, this heuristic is equivalent to the travelling salesman problem: in what order should a salesman located in a given city, who wishes to visit *N* cities once and only once, visit these cities as to minimize the total distance travelled?

A third and intermediate heuristic that might be called a “global-distance-minimizing heuristic” would be that pedestrians still attempt to sequence their visits in some optimal way, in the sense that the shortest route associated with every alternative sequence would result in longer distances travelled. Again, this would be an example of a simultaneous decision-making process.

So far, we have assumed that pedestrian sequencing, route and destination choice behaviour is regulated only by distance considerations. However, destinations differ in terms of their attractiveness variables and their relative location. The functional linkages of purposes and the spatial configuration of land uses influence the attractiveness/distance trade-off at each stop of a journey. Moreover, even if the sequence of destinations to be visited is known, the heuristics discussed above may be limited. A given sequence of destinations normally generates some total distance to be travelled. However, if these destinations are visited in a reverse sequence, the distance travelled remains the same. Hence, even in this simple case, a pedestrian will have to also decide on some spatial mechanisms underlying his sequencing behaviour.

Pedestrians may visit the destinations in order of proximity to entry point. This “nearest-destination-oriented heuristic” implies that a pedestrian will generally have to carry the things he/she bought along the route across longer distances. Alternatively, a pedestrian may decide to first visit the destination farthest from his/her entry point and then proceed back to the departure point. In general, such a “farthest-destination-oriented heuristic” will be associated with shorter carrying distances, ceteris paribus.

Evidently, it is very difficult to incorporate these heuristic in an algebraic model of pedestrian movement. Therefore, simulation seems the most appropriate modelling approach to develop a model of pedestrian destination, route choice and sequencing.
behaviour. It allows the researcher to simulate each pedestrian's behaviour from a set of rules that reflect the theoretical considerations discussed previously. Observed relative frequency distributions can be used to categorize sample respondents into behaviour types. Such a framework should allow the utility of alternatives to change, depending on both the kinds of shops that have been visited previously and in case of planned stops, the kinds of shops that will be visited later on during the trip. A model based on such principles is currently under development. The authors hope to report on the performance of this model in the near future.

Acknowledgements—This research was partly supported by the Dutch Organization for the Advancement of Scientific Research (NWO/SRO). The authors are grateful for the comments and suggestions of three referees, which helped them to clarify some of the ideas included in this article.

REFERENCES


