Introduction

Although some might argue that the days of large-scale urban models and rational planning approaches are over, the literature in urban planning and geographical information systems demonstrates that many research groups continue to build mathematical models that allow planners and policymakers to assess the likely consequences of their decisions in terms of a variety of socioeconomic indicators and the behaviour of individuals. In fact, the use of such models in particular planning domains is more the rule than the exception, although it is fair to say that the integrative, large-scale models have been gradually replaced by dedicated models to predict, for example, locational choice behaviour of firms (see, for example, Wets et al, 1996a; 1996b), housing (see, for example, Schellekens and Timmermans, 1997; Timmermans et al, 1994; Timmermans and van Noortwijk, 1995; Van de Vyvere et al, 1998), shopping (see, for example, Baker, 1996; Fotheringham and Trew, 1993; Makin et al, 1997; Oppewal et al, 1997; Popkowski Leszczyc and Timmermans, 1996), or recreational/tourism choice behaviour (see, for example, Kim and Song, 1998; Louviere and Timmermans, 1990).

Large-scale models however have not disappeared (Sui, 1998). Land-use transport models have continued to receive attention (see, for example, Echenique, 1994; Hunt, 1994; Hunt and Simmonds, 1993; Pauley and Webster, 1991; Simmonds and Still, 1998; Still, 1998). New glasses have been used to serve the old wines. Conventional equilibrium-based models have been criticised for their rigorous theoretical underpinnings. It has been realised that urban systems are highly complex in the sense that they involve many different actors who interact in a multitude of different ways. Moreover, the dynamics of the systems are complex in the sense that significant probabilistic elements both with respect to random variations in exogenous factors and with respect to the stochastic nature of endogenous processes exist in the system which may lead urban systems to evolve along particular nonlinear trajectories over time. Conventional equilibrium models based on gravity and entropy have been criticised for their underlying fundamental difficulty of
estimating the likely future states of the system given the inherently complex nature of the underlying behavioural processes. As an alternative, cellular automata models (Clarke and Gaydos, 1998; Couclesis, 1985; Engelen et al, 1993; 1997; Phipps, 1989; Takeyama and Couclelis, 1997; White et al, 1997) and microsimulation models (Clarke et al, 1981a) have been suggested.

In fact, a renewed interest in large-scale models can be observed. It cannot be explained only by the emergence of alternative theoretical and modelling perspectives, but also by the rapid introduction of geographical information systems and the development of computer technology. An important factor hindering the development of large-scale models in the 1960s was the need to build and maintain large data files. The introduction of geographical information systems (GIS) has not only made this task of national, provincial, and municipal planning authorities much easier, but the mere existence of these systems more or less guaranteed a commitment to information management once these systems were introduced in these organisations. Moreover, many countries have experienced the emergence of an increasing number of commercial data providers (Batty, 1994; Wegener, 1994).

In line with this revival, the purpose of the research project reported in this article is to re-engineer a regional location model (Veldhuisen and Kapoen, 1977; 1978) using GIS-technology, data warehousing opportunities, and current views on human spatial behaviour. This system has been given the acronym RAMBLAS, which stands for a regional planning model based on the micro-simulation of daily activity patterns. It reflects the fact that daily activity patterns are used as a basis for predicting the spatial distribution of the demand for various services and traffic flows in the urban system, and that microsimulation is used as a tool to predict demand and traffic flows. Given the underlying philosophy, the main constraints underlying the development of the system were that only readily available information should be used to estimate the model, and that inexpensive GIS software should be used to graphically display the outcomes of the system.

This paper is organised as follows. First, we will discuss the central conceptual considerations underlying the system, which by and large are very similar to other activity-based traffic-forecasting models. Next, we will discuss to what extent RAMBLAS differs from existing or currently developed microsimulation models of activity patterns. Then, against this background, we will describe the architecture of the system, report data input, and discuss some particular modelling issues. We conclude with a discussion of research and development issues associated with the continuing development of the system.

Conceptual considerations

One of the key planning issues both in urban and transportation planning concerns the relationship between land use and transport demand. It is well known that land-use patterns have some, albeit not deterministic, impact on transport demand, and, vice versa, that traffic flows influence the distribution of particular land uses in an urban area. Key to this relation between land use and transportation are the activity patterns of individuals and households. It has been realised that transport patterns reflect the activities that individuals and households wish to pursue within their environment. It is the urban environment, including the transportation network, which defines the opportunities but also the constraints that individuals and households face when pursuing their preferred activities. It is this very notion that has led to the so-called activity-based approach (see, for example, Ettema and Timmermans, 1997), which aims at predicting which activities will be conducted when, where, for how long, the transport mode and route involved, and possibly with whom.
The current model shares the conceptual considerations which underlie most activity-based models. Following authors such as Hägerstrand (1970) and Chapin (1974), we view spatial structure as emerging from the activities of individuals and households. Activities result from individuals’ basic desires, which drive their propensity to engage in particular activities. The urban environment offers opportunities which are largely the result of the provision of services and facilities, but also of the quality of those services and facilities. However, the same environment also acts as constraints on behaviour. These constraints more or less dictate whether or not activities in fact can be performed in particular time-space settings. Hägerstrand’s capability, coupling, and authority constraints are relevant concepts in this regard. Individuals and households, however, experience different degrees of flexibility when conducting their activities, influenced by the nature of the activity, priority hierarchies, etc, which leads to adaptive behaviour and scheduling of activities over time and space (compare with Cullen and Godson, 1975).

Essential, then, for the development of RAMBLAS is the notion that the demand for facilities in urban environments is derived from people’s activity patterns. Activity patterns manifest themselves in terms of traffic flows as people have to travel in the conduct of their out-of-home activities, the intensity of which varies across the days of the week. As part of a regional planning model, the problem then becomes how to predict the activity patterns and related traffic flows.

Activity-based models

Over the last decade, several alternative modelling approaches have been suggested to predict activity patterns and related traffic flows, although hardly ever in an integrated fashion. In fact, this topic represents an actively researched topic as many different groups are currently working on the development of activity-based traffic-forecasting models. Most of these attempts can be seen as major efforts to improve the quality of land-use transport planning models by incorporating a sounder behavioural basis. Much of this work in the USA is currently being undertaken under the Travel Model Improvement Program (TMIP) initiative. Although there are substantial differences in approach (Ettema and Timmermans, 1997), there seems a general agreement that any advances should involve a greater focus on activities rather than the trip (Spear, 1996). To position our model in this rapidly growing field of research, we will first briefly discuss alternative modelling approaches.

Perhaps the best-known approach is that of the traditional models developed in time geography, which typically examine whether particular activity patterns can be realised within a specified time-space environment. An activity programme, which describes a set of activities of a certain duration to be performed within certain time windows is taken as input, and the purpose of the model is to examine the feasibility of the activity programme given the urban environment, the transportation system, and the institutional context. A combinatorial algorithm is typically used to generate all possible activity sequences, and to examine the feasibility of each sequence by (1) checking whether the interval between the end time of the previous activity and the start time of the next activity is sufficient to perform the activity plus the associated travel time; (2) testing whether the activity can start after the earliest possible start time and be finished before the latest possible end time; (3) checking that conditions about the sequencing of activities are not violated. The number of feasible activity schedules is often used as a measure of flexibility of the time-space environment. Examples are Lenntorp’s (1976) PESASP model and CARLA.

Although these models shed some interesting light on the impact of constraints on behaviour, they deny or at least do not incorporate people’s ability to reschedule their
activities. Therefore, other models attempt to predict actual behaviour. Dominant in
this respect have been the attempts to build utility-maximising models of activity
scheduling behaviour, which replace trip-based and tour-based models by activity-based
models. Early attempts stayed close to the gravity and entropy models (Kreibach, 1979;
Sparrmann, 1980; Swiderski, 1982; see also Axhausen and Herz, 1989), but these were
gradually replaced by nested logit models, derived from choice theoretical notions (see,
for example, Kitamura and Kermanshah, 1983; 1984; van der Hoorn, 1983). Perhaps
the most general model in this tradition is STARCHILD (Recker et al, 1986a; 1986b).
The model postulates that the generation and allocation of activities occurs at the house-
hold level. A household activity programme spawns individual activity programmes,
each implicitly reflecting decision rules and constraints at both the household and
individual levels. Activity participation is formulated as a constrained choice process
subject to the outcome of activity generation and allocation. The generation and
allocation of activities occur continuously over a multitude of time frames; however,
the execution phase is conveniently conceptualised as a daily pattern when the actual
participation and scheduling choices are implemented. The system consists of a series
of interlinked modules, which leads to a rather complex system. It may be for this
particular reason that researchers have looked for simpler operational versions.

The nested logit model, especially, has become rather popular in this regard. For
example, Kawakami and Isobe (1989) modelled the activity patterns of workers only.
They conceptualised the generation of activity patterns as a hierarchical choice pro-
cess. First, individuals are assumed to decide whether or not other activities will
be performed before or after work. Next, a pattern type (for example, visit a destination
after work on the way home, or go home first and then visit another destination) is
selected. Last, at the lowest level of the assumed hierarchical choice process, the exact
destination is chosen. These levels correspond with the trees of the nested logit model.
The probabilities of choosing a particular destination are assumed to be a function of
the travel time to and from that destination, the duration of the activity, and a set of
destination attributes, such as population size, suburban versus urban, and turnover in
the entertainment sector. The choice of pattern type is modelled in terms of the logsum
of the destination choice model (because a nested logit structure is used), total travel
time, total duration of out-of-home activities, and the amount of time spent at home.
Last, the choice probabilities of whether or not to perform additional activities were
expressed as a function of the logsum of the pattern type model, start and end times of
work, and age and gender. A more recent study in a similar vein is Ben-Akiva and
Bowman's (1995) model system. It is probably fair to say that this is currently seen as the
most advanced operational activity-based model of transport demand in this tradition.

Although the authors call it a computational process model, the Prism-Constrained
Activity Travel Simulator or PCATS (Kitamura and Fujii, 1997; Fujii et al, 1998) is
based also on principles of utility-maximising behaviour. It differs from the nested logit
models in that a different utility structure is assumed, and that several constraints
affecting travel choices are incorporated. In particular, Hägerstrand's concept of
time–space prism, travel-mode availability, and individuals' cognition of activity loca-
tions are incorporated in the model. PCATS assumes that individuals maximise the
utility associated within the open periods, subject to a set of constraints, where open
periods are defined as periods across the day during which an individual is free to travel
and become engaged in activities. They contrast with blocked periods, which are
characterised by commitments to conduct particular activities at particular locations.

A true computational process of activity-scheduling behaviour which is based on
decision heuristics as opposed to principles of utility-maximising behaviour has been
suggested by Gärling et al (1989). Their SCHEDULER model has however primarily
remained a conceptual model, although some numerical illustration has been given (Golledge et al., 1994). Computational process models are not based on algebraic equations, but rather use IF ... THEN rules to represent behaviour and underlying decisions. Most progress along these lines has been made in the context of the ALBATROSS system (Arentze et al., 1999), currently under development for the Dutch Ministry of Transportation. Decision tables are used to represent the rules individuals use to decide which activities to conduct where, when, for how long, and the transport mode involved.

Another model, which can be best considered as a hybrid model combining various utility-based and other modelling techniques, is AMOS, a dynamic microsimulator of household activities and travel over time and space (Pendyala et al., 1995). This model is part of SAMS, which also contains a socio-demographic simulator, an urban system simulator, a vehicle transactions simulator, a dynamic network simulator, and an emissions module. It is an activity-based model of travel decisions that simulates the scheduling and adaptation of schedules and resulting travel behaviour of individuals and households. Adaptation behaviour is treated as a learning process in which the individual gains knowledge about various aspects of the new travel environment as he or she attempts to adapt to it. Adaptation behaviour is viewed as a trial-and-error process in which the individual tries out alternative activity-travel options until a suitable option is found. The `satisficing' rule is used as a guiding principle. Starting from an initial set of activity-travel patterns, AMOS simulates each individual's adaptation process and finally determines how individuals and households will adapt to the new environment. Neural networks are used to determine which response options an individual may conceive as a result of changes in the travel environment.

Certain components of AMOS are very similar to yet another simulation model, SMASH, which is also based on notions of decision heuristics (Ettema et al., 1993). In particular, this model concentrates on the process of activity scheduling. The scheduling process is assumed to be a sequential process consisting of a number of consecutive steps. In every step the schedule, which is empty at the beginning of the process, can be adjusted by one of the following basic actions: (1) adding an activity from the agenda to the schedule; (2) deleting an activity from the schedule; (3) substituting an activity from the schedule with an activity from the agenda; (4) stopping the scheduling process. By repeatedly applying one of these basic actions, the schedule is constructed and adapted until a satisfactory schedule is created.

The aims and objectives of RAMBLAS differ in a number of important respects from the models just discussed. First, unlike alternative modelling approaches, the focus is not primarily on explaining or predicting household activity patterns. In contrast, our research effort is primarily concerned with the challenge of how (observed or predicted) activity patterns can be generalised to system-wide forecasts of transport demand and time-dependent traffic flows. To that effect, we avoided (at least until now) relatively sophisticated modelling tools to predict activity patterns, and insisted on relatively simple principles to predict activity patterns and related traffic flows. Our interest was to investigate the potential of such simplification, and we realised that, if necessary, further sophistication and detail could always be incorporated into the model. Second, although competing models have typically relied on specific local samples, our focus in this project is to explore how easily available existing national data can be used to predict regional traffic flows. It should be made explicit from the outset that this reliance on widely available data is made at the expense of the complexity of the model specification. For example, the current model excludes sensitivity to any price variables, and cannot capture any model substitution under transport policies.
It is because of this specific focus that microsimulation was used. As we will discuss in the next section, the principles used to predict the activity patterns do not lend themselves easily to be captured in terms of mathematical equations. If the problem would be expressed in terms of marginal and conditional probability distributions to generate, for example, trip matrices by time slice, some of the detailed multidimensional relationships would be lost. Moreover, feedback mechanisms could not be captured. It is for exactly these and similar reasons that microsimulation has been advocated long before in urban and transportation planning research (see, for example, Borgers and Timmermans, 1986; Clarke et al, 1981b; Goulias and Kitamura, 1992; Mackett, 1990; Wegener and Spiekermann, 1996), although it never has become a dominant modelling approach.

**RAMBLAS**

**Purpose**
The system is developed to estimate the intended and unintended consequences of planning decisions related to land use, building programmes, and road construction for households and firms. In this paper, we will not discuss the entire model system, but concentrate on the module that is concerned with the activity patterns of individuals and households. The main purpose of RAMBLAS is to predict the spatial distribution of individuals' activities and related traffic flows for a chosen time period given the forecasted spatial distribution of dwellings, the forecasted distribution of households over dwellings (Veldhuisen, 1985), and the transport network. Thus, the model allows planners to assess the likely effects of their land-use and transport plans on activity patterns and traffic flows.

**Data**
As explained above, the model was developed such that it is entirely based on readily and generally available statistical information. No dedicated data collection is involved: all data can be bought from official commercial and noncommercial sources. The model was developed for the Eindhoven region in the Netherlands, but the underlying principles can be directly applied to any other region in the Netherlands. The region consists of 33 municipalities and has approximately 680,000 inhabitants.

The spatial system underlying the model is given by the set of 400 zones defined by the Dutch Central Bureau of Statistics. These data are available in digital form. In particular, the borders of the zones are provided. This set of zones constitutes a standard for which many other data are collected and published. The model employs land use for various activities per zone and the number of different kinds of dwellings per zone. The spatial data are complemented by the road network, which is published in digital form by the Ministry of Transport. In addition to these spatial data, the model

<table>
<thead>
<tr>
<th>Table 1. Overview of the data used for the simulation.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Employment data per zone</td>
</tr>
<tr>
<td>Land use per zone</td>
</tr>
<tr>
<td>Population per zone</td>
</tr>
<tr>
<td>Road system</td>
</tr>
<tr>
<td>Dwellings by type</td>
</tr>
</tbody>
</table>

LISA, National Information System of Employment; ETIN, Economic-Technological Institute; WBR, Register of Neighbourhoods and Digital Boundaries; CBS, Central Bureau of Statistics; PTT, Post Office.
requires some nonspatial data. In particular, for each zone the number of households, classified into a series of segments, and the number of employees, again classified according to segments, are used. In table 1 we provide an overview of the data. These data are available for a 500 × 500 m grid network covering the Netherlands.

Input to the simulation of activity patterns and traffic flows
The input of the simulation model consists of the distribution of various types of households across the different kinds of dwellings per zone and the distribution of land uses and dwellings per zone. These variables are external to the simulation. Changes in these variables are externally monitored. Households are classified according to their size, and for each class the age and gender of household members is calculated. The spatial attributes of the area (that is, land use, dwelling stock, and road system) are treated as variables that can be manipulated by planning. The distribution of dwellings and employment over geographical space is partly dependent on land-use planning. The planning of the road system is also dependent on decisions of the various planning authorities. The spatial distribution of activities and trips are treated as dependent variables. Thus, the model enables us to predict the likely consequences of possible policy decisions on activity patterns and thus estimate the effectiveness of policy decisions. In particular, these decisions concern changes in land use, dwelling stock, and road construction.

Microsimulation of activity patterns
The aim of the microsimulation is to predict which activities will be conducted where, when, for how long, the transport mode involved, and which route is chosen to implement the activities. Remember that one of the conditions for model development was that no specific diary data needed to be collected. The idea was to use activity data from a national sample as input to the simulation. Now, there is some empirical evidence (Timmermans and van der Waerden, 1998) suggesting that activity patterns are correlated with socio-demographics. Therefore the basic idea underlying the simulation is to identify in the national sample various population segments, based on socio-demographic information, and derive segment-specific activity patterns. Moreover, for each out-of-home activity, the distribution of transport modes is derived from the national sample.

The first step in the microsimulation, then, involves for every individual in the Eindhoven region (1) identifying the corresponding population segment, and (2) drawing at random from the national distribution the activity agenda and transport mode. In principle at least two options are available to generate the activity agenda. First, one can decide to draw at random from the conditional distributions representing the conditional probability that an individual belonging to a particular population segment will be engaged in a particular activity. This option has the advantage that draws are made from larger numbers. It has the disadvantage, however, that the specific composition of activity agendas is unlikely to be generated. The second option, and the one chosen in the present study, is to sample for complete activity agendas, which implies that the specific composition of observed activity patterns is preserved.

In the present study, the INTOMART 1990 time-budget survey (TBO90) was used as input data to the simulation model. This survey consists of a nationwide sample of 3400 respondents who have given a detailed description of their activities for every quarter of an hour for a whole week. The total sample was divided into four subsamples, each of which reported their activities for one week in October 1990. The survey was repeated in 1995, and again just recently, but this information is not available at the time of writing. The 1990 data were used to build the simulation model, because it allows us to test the forecasting abilities of the model in figure research. The fact that this survey is
repeated allows the users to regularly update the model. In general, the data quality of the time-budget survey is believed to be good, although it has some associated problems typical for this kind of survey. In particular, the fact that 15-minute intervals are used implies that very short trips are underrepresented. There is also some evidence that the accuracy decreases towards the end of the week. A more detailed account of the reliability, validity, and data quality of these survey instruments is given in Arentze et al (1999).

Population segments were identified on the basis of gender, age, employment status, and educational achievement. In table 2 we list the sample sizes for the different segments. We show that the data density is acceptable-to-good for most segments, but also that only a few cases are available for other segments. In table 2 we show that 24 segments are identified based on gender, age, and employment status. To these 24 segments, two segments are added: (1) pupils going to elementary schools (boys and girls, age 4–12), and (2) pupils going to secondary schools (boys and girls, age 13–16). The national survey does not contain any information about these segments, and hence they needed to be added. It also meant that we had to assume the activity patterns of these groups.

The time-budget survey requests that respondents report their daily behaviour in terms of 183 activities. These activities were aggregated according to 7 activity classes: work, childcare, shopping, personal/medical care, school or study, social participation, and social contacts. For each of these out-of-home activities, the distribution of chosen transport mode was generated. These conditional distributions are listed in table 3, in which it is shown that, as expected, the distribution of transport mode varies by activity. For example, although a relatively large percentage uses the bicycle to travel to work, this means of transport is hardly used for shopping.

By using this data, the first step of the microsimulation results in an activity agenda for a simulated individual. The next step of the simulation addresses the problem of how this agenda is implemented in space and time. To that end, various additional operational definitions that drive the allocation of activities to particular

Table 2. Number of respondents in TBO90 according to gender, age, employment, and education.

<table>
<thead>
<tr>
<th></th>
<th>Male Total</th>
<th>Female Total</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>17 – 24</td>
<td>25 – 44</td>
<td>45 – 64</td>
</tr>
<tr>
<td>Employed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not employed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regular education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>89</td>
<td>1543</td>
<td>845</td>
</tr>
</tbody>
</table>

Table 3. Distribution of transport mode for out-of-home activities.

<table>
<thead>
<tr>
<th>Activity class</th>
<th>Car driver</th>
<th>Car passenger</th>
<th>Motor</th>
<th>Bicycle</th>
<th>Walk</th>
<th>Public transport</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Work</td>
<td>42.9</td>
<td>10.7</td>
<td>1.8</td>
<td>32.0</td>
<td>7.8</td>
<td>4.8</td>
<td>11620</td>
</tr>
<tr>
<td>Child care</td>
<td>22.5</td>
<td>5.5</td>
<td>0.3</td>
<td>42.7</td>
<td>28.0</td>
<td>1.0</td>
<td>7514</td>
</tr>
<tr>
<td>Shopping</td>
<td>31.2</td>
<td>16.2</td>
<td>0.0</td>
<td>0.9</td>
<td>32.9</td>
<td>18.8</td>
<td>8790</td>
</tr>
<tr>
<td>Personal/medical care</td>
<td>31.6</td>
<td>19.6</td>
<td>1.6</td>
<td>26.0</td>
<td>16.2</td>
<td>5.0</td>
<td>4423</td>
</tr>
<tr>
<td>School or study</td>
<td>12.7</td>
<td>6.5</td>
<td>4.3</td>
<td>60.1</td>
<td>8.1</td>
<td>8.3</td>
<td>4321</td>
</tr>
<tr>
<td>Social participation</td>
<td>30.2</td>
<td>17.6</td>
<td>1.0</td>
<td>29.9</td>
<td>18.6</td>
<td>2.6</td>
<td>2691</td>
</tr>
<tr>
<td>Social contacts</td>
<td>34.4</td>
<td>24.8</td>
<td>1.2</td>
<td>23.3</td>
<td>12.0</td>
<td>4.3</td>
<td>13381</td>
</tr>
</tbody>
</table>

Note: Motor, motorbike.
destinations were made. Although the TBO90 survey allows us to derive distance or travel-time distributions for each out-of-home activity, these distributions will not only be influenced by distance sensitivity of travellers, but also by the spatial structure (that is, the relative location of dwellings and facilities, employment, etc). Hence, the national distributions cannot be applied in any straightforward way to the study area. Assumptions about underlying mechanisms are required.

In the case of the work activity, we assumed that the travel time observed in the diary constitutes the time people are willing to travel to work, given the transport mode involved. In terms of the microsimulation, this means that a zone of employment is drawn at random from the total number of available jobs in the region, delimited by this maximum travel time. Job locations are drawn without replacement, hence the set of job locations is reduced during the simulation.

In the case of study or school, a different principle was employed. We assumed that parents of children going to elementary schools invariably choose the school nearest to their residence. Although this assumption is not perfect, it reflects the planning of the school districts in the Netherlands. For students going to secondary schools, we assumed that their action space is defined by an area of 45 minutes of bicycling time. Schools are drawn at random from this action space. The same principle is used for students of higher education, but now the distribution of employment in higher education is used as the distribution from which the school is sampled.

The latter principle is also used to determine the destination for shopping and services. The destination is drawn at random from the distribution of employment in the relevant services. There is some evidence that the relationship between employment in services and the number of visitors is fairly stable in shopping and medical care, but less so in sports and recreation (DTV, 1993). Hence, this operational decision may require some further thought in future research. As for the final activity classes (social participation and social contacts), the presence of other households, rather than employment, is used as the distribution from which the destination is sampled.

Having established these origin-destination pairs, the next step of the simulation involves the microsimulation of traffic flows. One of the problems in establishing the vehicle travel time from origin to destination is that the travel time is dependent on the traffic situation. The travel time on any time of the day is dependent on the intensity of traffic and the capacity of the links of the road system (see figure 1).

Figure 1. Road map of the study area.
Therefore, some method to calculate the speed of the network is required. In this present model, we decided to use the ‘speed-flow’ calculation method of the “Department of Transport in the UK” as described by de Dios Ortuzar and Willumsen (1994). Basically, this method involves the following steps and assumptions. The starting point is the set of observed (individual) arrival times at the destinations. Assuming that individuals will invariably use the shortest route, the travel times between the residential zones and the destination zones are calculated for all individuals in vehicle minutes on the basis of the original free-flow speeds. Next, the starting times are derived from the arrival and travel times by subtracting the travel times from the arrival times. For any chosen time interval, the vehicle trips are then simulated using the starting time. Given assumed speed-flow relationships, the speeds on the network are recalculated for every 10-minute interval. The new speeds are then used to recalculate the starting times and the flows on the network. Although in principle this can be repeated until the system reaches an equilibrium state, one iteration suffices to generate and calculate the traffic flows and speeds on the network. For every chosen interval, the traffic flows are displayed graphically on the computer screen (see figure 2).

**Figure 2.** Traffic density, Thursday morning.

**Implementation and illustration**

**Technical specification**

Technically, the simulation works as follows:

Step 1. Choose the day to be simulated and read in the TBO90 database for that day of the week. A total of 7 days a week are identified. Hence, the activity patterns are specific for the day of the week.

Step 2. Extract the geographical data of the zones and road system from MapInfo.-MIF and MID (ASCII) files.

Step 3. Calculate the shortest routes between the origins and destination using the DIAL algorithm, and for each origin order all destinations according to distance.
Step 4. Calculate the shortest routes between the origins and destinations in car-
minutes, based on the average speed on the links of the road system.

Step 5. Read the neighbourhood data concerning households and employment and
identify to which of the 26 population segments the individuals belong.

Step 6. Assume an average activity schedule for pupils and draw at random from
the TBO90 survey the activity agenda for all other individuals.

Step 7. From these agendas, determine the number of trips and, for each trip, the
travel mode, the travel time, and the begin time and the end time of the
activity.

Step 8. Dependent upon the activity class, determine the action space of the individ-
ual, given the mode of transport.

Step 9. Draw at random from the relevant database the destination where the activity
will be conducted.

Step 10. Calculate the shortest route and the speed between the residential zone of
the simulated individual and the destination.

The actual simulation can be best understood as a moving-time-window simulation.
For each minute, the model searches over all 640 000 people in the region to see
whether they make a trip and traces their moves along the transportation network.
In case an activity has been completed, the critical parameters of the next activity are
simulated.

Thus, for each individual at every minute of the simulation period the model
monitors whether he or she is at home, conducting an activity at another destination,
or travelling on the network. In the latter case, the shortest route information and the
simulated speed on the network is used to assign the individuals to a particular link of
the road network. Traffic flows are then updated on the computer screen.

RAMBLAS was programmed in Fortran for a PC system with a 200MMX processor.
The simulation of the work and school zones for the regional population is a `one-time'
operation and takes about 5 minutes. Reading the resulting file for a particular day
takes a few minutes. The simulation of the activities and traffic flows for a chosen
period of the day depends on the length of the period and takes from 30 seconds to a
few minutes.

Validation tests
As we have indicated, the simulation model is based on readily available national data.
This implies that the model may not be able to capture the regional idiosyncrasies as
the input is based on national relationships. Therefore, it is important to perform some
validation tests and compare the simulation results with the outcomes of other (local)
data sources. The results of some of these tests are described in this subsection. More
specifically, we conducted two validation tests: one compared the outcomes of the
simulation against another national survey that comprised provincial data, the second
one compared the outcomes of the simulation against older data for the same study
area.

A first test concerned the comparison with another national survey. The Central
Bureau of Statistic is responsible for the so-called Continuing Mobility Survey (OVG),
which collects data on various aspects of trips at the national level. Unfortunately,
published results are only available at the provincial level, but many regional planning
authorities will use these outcomes for their region. Moreover, the OVG data were
collected in 1995, hence there is a time difference between the two data sets. Last, the
OVG survey summarises the data over a complete year, whereas the TBO data cover
four weeks in October. When interpreting the results of the validation tests, these
differences should be kept in mind.
In table 4 we list the results of the comparison of the average number of trips by transport mode. Note that the motor and bicycle categories were combined. The data in table 4 suggest that the simulation model is reasonably consistent with the OVG data for the province of Noord-Brabant, although there is evidence of underprediction of the average number of trips for most transport modes. However, it should be remembered that the simulation model is based on data of activity participation recorded at 15-minute intervals. It has been widely reported in the literature [see Etterna et al (1998) for a review] that such diary data results in an underrecording of short trips. Moreover, the Ministry of Transport reported an 18% increase in car mobility over the period 1990–1995. Hence, when these considerations are kept in mind, it may be concluded that the simulation predicts the average number of trips by transport mode at an acceptable level.

Table 4. Comparison of average number of trips by transport mode.

<table>
<thead>
<tr>
<th>Transport mode</th>
<th>OVG</th>
<th>RAMBLAS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Netherlands</td>
<td>Noord-Brabant</td>
</tr>
<tr>
<td>Car driver</td>
<td>1.06</td>
<td>1.17</td>
</tr>
<tr>
<td>Car passenger</td>
<td>0.59</td>
<td>0.64</td>
</tr>
<tr>
<td>Bike</td>
<td>1.11</td>
<td>1.03</td>
</tr>
<tr>
<td>Walk</td>
<td>0.65</td>
<td>0.62</td>
</tr>
<tr>
<td>Public transport</td>
<td>0.17</td>
<td>0.09</td>
</tr>
<tr>
<td>Other</td>
<td>0.06</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Notes: OVG, Continuing Mobility Survey; bike, motorbike and bicycle.

The same data were also used to compare the person kilometres by activity class and transport mode. The results are listed in table 5. A problem that we encountered in this comparison was that the classification of activities differed between the two data sources. In table 5 we therefore list only the results of a subset of activity classes. An examination of this table suggests that the predictions of the model are satisfactory. If we examine the ‘total’ column the results indicate that the model consistently predicts higher person kilometres, but this is what we would expect as the TBO data on which the model is based is likely biased in the sense that short trips are underreported. Keeping this in mind, table 5 indicates that the predictions in many cases are very good, especially for the dominant transport modes of the various activity classes.

The OVG data were also used to compare the outcomes of the simulation with these data in terms of person kilometers by distance class and transport mode. The results are listed in table 6, which again provides evidence of the validity of the simulation model, especially for the noncar modes. Car use is slightly underpredicted, especially in the 10–20 and 20–50 km distance categories.

Table 5. Comparison of person kilometers by activity class and mode.

<table>
<thead>
<tr>
<th>Activity class</th>
<th>Car driver</th>
<th>Car passenger</th>
<th>Bicycle</th>
<th>Walk</th>
<th>Public transport</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OVG R</td>
<td>OVG R</td>
<td>OVG R</td>
<td>OVG R</td>
<td>OVG R</td>
<td>OVG R</td>
</tr>
<tr>
<td>Work</td>
<td>7.3</td>
<td>6.5</td>
<td>1.3</td>
<td>2.7</td>
<td>0.8</td>
<td>0.3</td>
</tr>
<tr>
<td>Shopping</td>
<td>1.8</td>
<td>1.9</td>
<td>1.2</td>
<td>1.1</td>
<td>0.6</td>
<td>0.7</td>
</tr>
<tr>
<td>Education</td>
<td>0.2</td>
<td>0.1</td>
<td>0.2</td>
<td>0.6</td>
<td>0.9</td>
<td>0.8</td>
</tr>
<tr>
<td>Recreation</td>
<td>1.5</td>
<td>2.4</td>
<td>2.0</td>
<td>2.4</td>
<td>0.6</td>
<td>0.4</td>
</tr>
<tr>
<td>Total</td>
<td>16.4</td>
<td>14.4</td>
<td>9.3</td>
<td>8.5</td>
<td>4.3</td>
<td>3.8</td>
</tr>
</tbody>
</table>

Notes: OVG, Continuing Mobility Survey; R, RAMBLAS.
A second validation test involved the comparison of the outcomes of the simulation model against data collected for exactly the same study area but in 1977. More specifically, these data allowed us to compare travel times in 1977 with travel times predicted by the model. During this time period, some important trends in the spatial structure of the study area and its transport system can be identified. First, the new residential areas were primarily built in the periphery of Eindhoven, the central city of the study area, and in the suburbs. Second, there has been a substantial increase in work locations along the highway system. Third, the commercial structure has largely remained the same, except of course the building of neighbourhood shopping centres in the new residential areas. Last, there has been a rather dramatic growth in car ownership and car use. Under these circumstances, one would expect a general increase in travel times, topping off at the maximum for the city region, which would be approximately 45 minutes. The results are portrayed in table 7, which shows that the results are consistent with this expectation. Predicted travel times are lower under 10 minutes travel time, higher than the observed travel times between 11 and 40 minutes, and there is hardly any difference for travel times higher than 40 minutes.

Table 6. Comparison of person kilometers by distance class and transport mode.

<table>
<thead>
<tr>
<th>Activity class</th>
<th>Car driver</th>
<th>Car passenger</th>
<th>Bicycle</th>
<th>Walk</th>
<th>Public transport</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OVG R</td>
<td>OVG R</td>
<td>OVG R</td>
<td>OVG R</td>
<td>OVG R</td>
<td>OVG R</td>
</tr>
<tr>
<td>0 – 5 km</td>
<td>1.0</td>
<td>1.0</td>
<td>0.6</td>
<td>0.4</td>
<td>1.5</td>
<td>3.9</td>
</tr>
<tr>
<td>5 – 10 km</td>
<td>1.6</td>
<td>2.8</td>
<td>0.9</td>
<td>1.0</td>
<td>0.8</td>
<td>3.6</td>
</tr>
<tr>
<td>10 – 20 km</td>
<td>3.2</td>
<td>2.4</td>
<td>1.8</td>
<td>1.1</td>
<td>0.7</td>
<td>6.2</td>
</tr>
<tr>
<td>20 – 50 km</td>
<td>4.2</td>
<td>1.6</td>
<td>2.2</td>
<td>0.9</td>
<td>0.3</td>
<td>7.7</td>
</tr>
<tr>
<td>&gt;50 km</td>
<td>6.3</td>
<td>6.6</td>
<td>3.9</td>
<td>5.1</td>
<td>0.1</td>
<td>12.5</td>
</tr>
<tr>
<td>Total</td>
<td>16.4</td>
<td>14.4</td>
<td>9.3</td>
<td>8.5</td>
<td>3.5</td>
<td>33.8</td>
</tr>
</tbody>
</table>

Notes: OVG, Continuing Mobility Survey; R, RAMBLAS.

Table 7. Comparison of travel times.

<table>
<thead>
<tr>
<th>Travel time</th>
<th>1977 survey</th>
<th>RAMBLAS</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 – 5 minutes</td>
<td>17.8</td>
<td>7.1</td>
<td>−10.7</td>
</tr>
<tr>
<td>6 – 10 minutes</td>
<td>23.9</td>
<td>21.6</td>
<td>−2.3</td>
</tr>
<tr>
<td>11 – 20 minutes</td>
<td>39.4</td>
<td>43.4</td>
<td>+4.0</td>
</tr>
<tr>
<td>21 – 30 minutes</td>
<td>13.7</td>
<td>18.1</td>
<td>+4.4</td>
</tr>
<tr>
<td>31 – 40 minutes</td>
<td>1.2</td>
<td>6.1</td>
<td>+4.9</td>
</tr>
<tr>
<td>41 – 45 minutes</td>
<td>1.8</td>
<td>1.1</td>
<td>−0.7</td>
</tr>
<tr>
<td>&gt;46 minutes</td>
<td>2.5</td>
<td>2.5</td>
<td>+0.0</td>
</tr>
</tbody>
</table>

Conclusions
The purpose of this paper was to report on the development of RAMBLAS. The core of the system consists of a microsimulation of daily activity patterns of households. The system was designed so as to be independent from expensive data collection, which is typically required for many current activity-based models. The guiding principle underlying the system was that the prediction of activity patterns and traffic flows should only require readily available and inexpensive data. To this effect, a national time-budget survey and GIS data on land use, population, and employment were used to predict regional activity patterns and traffic flows. It is in this very aspect that RAMBLAS differs substantially from competing activity-based models, which rely on principles of utility maximisation or decision heuristics. No doubt these models are
theoretically more appealing, but their estimation and application requires large amounts of detailed data, which often need to be collected specifically for the model at hand. Another major difference is that the prediction of traffic flows on the road network across the day is linked to the prediction of activity patterns, a feature which does not apply to the vast majority of activity-based models.

Although this guiding principle makes the model easy to apply in any region, it comes at the cost that the nationwide survey only allows some degree of breaking the sample down into segments. In our case, we decided to identify population segments reflecting the assumption that activity patterns are correlated with socio-demographic variables and allowing a link with regional population data. We did not make any attempt to break down the data by, for example, settlement type, although some will argue that activity patterns in suburban or rural areas are significantly different from those observed in urban areas. Implicitly, we assume that patterns are invariant across spatial context. If such contextual effects are to be incorporated in the model, alternative schemes of splitting up the national survey are required. It is our intention to apply data-mining or pattern-recognition techniques in future research to identify the segmentation that yields the most homogeneous and discriminating activity patterns.

Other operational decisions also require further thought. One key decision relates to the definition of action spaces. Observed data are used to define the maximum travel time, given the transport mode, that people are willing to travel in their conduct of a given activity. The validity of this assumption and its impact on the final simulation results requires further attention. Observed travel times typically reflect the distribution of available destinations, and hence may not be a valid indicator of willingness to travel. On the other hand, given our desire to base the entire simulation system on readily available data, it is not evident what alternative definition can be used.

The present version of the system is based on single-purpose trips only, which evidently is a rather strong assumption. Although single-purpose trips account for the vast majority of trips in the Netherlands, some trips are multipurpose. Moreover, some people believe that increasing time pressure and existing land-use policies will actually increase the share of multipurpose trips, especially for some population segments. Hence, we intend to incorporate the option of multipurpose trips in a future version of the model. Similar principles can be used in this effort, but the simulation will become more complex as one now has to simulate destination choice at successive stages of the tour and decide on route-choice behaviour in the case of complex trips. The simulation of the timing of activities also requires further attention and testing. The current version of the model is based on the national distribution of starting times of activities, and the departure time is simulated as a function of these starting times and the simulated speed along the network. Thus, the model does not allow for any other timing adjustments such as different starting times, shorter durations of activities, or rescheduling of activities. Although this seems a limiting assumption, again it is not readily evident how a more sophisticated system could be built from available data sources without attempts at generalising the relationships present in the data.

As we have argued before, the current version of the model is not sensitive to price variables and substitution. The inclusion of such options would make the model more similar to activity-based models of transport demand, derived from travel or activity diaries collected locally. It would be worthwhile, however, in future research to explore the possibility of conducting stated preference or interactive computer experiments to model change in activity patterns. The suggested microsimulation would still be relevant under such conditions.

A final research question that needs to be addressed in future research concerns the validation of the simulation model. This can be accomplished along the following lines.
First, the outcomes of the simulation can be compared with more detailed diary data collected in a particular region. Second, the simulated traffic intensities can be compared against traffic counts. In this effort, a systematic investigation of Monte Carlo errors is desirable. The estimates of model demand, traffic flows, etc., have distributions that are determined by the simulation process. A different set of random numbers will generate a different result. Our running of the model suggests that the standard errors of mean travel time and total demand, but also of flows by link, are small. Nevertheless, a more systematic investigation comparing Monte Carlo errors for various parameters as a function of operational decisions is desirable. The authors hope to report on such future research efforts in forthcoming publications.

Acknowledgements. The authors appreciate the constructive criticisms of three anonymous reviewers. The readability and positioning of the paper was substantially improved by their suggestions. Any errors remain the responsibility of the authors.

References

Arentze T A, Hofman E, Kalfs N, Timmermans H J P, 1999, “Data needs, data collection and data quality requirements of activity-based transport demand models”, keynote address, International Conference on Transport Survey Quality and Innovation, Grainau, Germany, 24 – 30 May; copy available from the authors


Ben-Akiva M E, Bowman J L, 1995, “Activity-based disaggregate travel demand model system with daily activity schedules”, paper presented at the Workshop Activity-Based Models, Eindhoven; copy available from the authors


Chapin F S, 1974 Human Activity Patterns in the City (John Wiley, New York)


Echenique M H, 1994, “Urban and regional studies at the Martin Centre: its origins, its present, its future” Environment and Planning B: Planning and Design 21 517 – 533


Ettema D F, Timmermans H J P, 1997 Activity-based Approaches to Travel Analysis (Pergamon, Oxford)