AN OVERVIEW OF PCATS/DEBNETS MICRO-SIMULATION SYSTEM: ITS DEVELOPMENT, EXTENSION, AND APPLICATION TO DEMAND FORECASTING

Ryuichi Kitamura
Department of Urban Management, Kyoto University, Sakyo-ku, Kyoto 606-8501, JAPAN
Department of Civil and Environmental Engineering, University of California, Davis, USA
rkitamura@termws.kuciv.kyoto-u.ac.jp

Akira Kikuchi
Department of Urban Management, Kyoto University, Sakyo-ku, Kyoto 606-8501, JAPAN
kikuchi@term.kuciv.kyoto-u.ac.jp

Satoshi Fujii
Department of Civil Engineering, Tokyo Institute of Technology, JAPAN
Toshiyuki Yamamoto
Department of Geotechnical and Environmental Engineering, Nagoya University, JAPAN

ABSTRACT

The micro-simulator of individuals' daily travel, PCATS, and a dynamic network simulator, DEBNetS, are integrated to form a simulation system for urban passenger travel. The components of the simulation system are briefly

described, and three areas of on-going system improvement are described, i.e., (i) introduction of stochastic frontier models of prism vertex location, (ii) adoption of a fine grid system for quasi-continuous representation of space, and (iii) use of MCMC algorithms to handle colossal choice sets. Application case studies demonstrate that micro-simulation is a practical approach for demand forecasting and policy analysis, especially in the area of demand management.

INTRODUCTION

Traffic flow on road networks is dynamic. The flow rate on a roadway section varies substantially in the course of the day; the direction of the dominant traffic flow changes over time; and queues at bottlenecks grow, then dissipate. Simulating traffic flow on networks along a time axis is an effective approach to represent network dynamics. In fact a number of network simulators have been developed and their effectiveness in traffic control and transportation planning is being evaluated.

Inputs to dynamic network simulation may take on the form of a set of origin-to-destination (O-D) trip matrices by time of day, each of which represents trip interchange over a certain time interval within a day. Since such intervals are artificial and real traffic flow varies continuously over time irrespective of the set of time intervals adopted by the researcher, it would be more desirable to adopt shorter intervals for improved accuracy. At the limit where the length of the interval diminishes, inputs are represented by individual trips to be loaded onto the network along a continuous time axis.

The conventional, trip-based passenger travel demand forecasting methods, such as the four-step procedures, have been unable to provide inputs to dynamic network simulation. This is simply because these forecasting methods have not incorporated a time axis into their analytical frameworks. The spatial dimension has been captured in these methods by subdividing the study area into traffic zones. The time dimension, on the other hand, has never been represented properly in any analytical manner. This is quite curious in light of the fact that congestion, the long-time preoccupation of transportation planners, is a consequence of the concentration of demand in time, as well as in space. The O-D trip matrices by time interval that many transportation planning agencies have developed are by and large based on daily O-D traffic; heuristic procedures and rules of thumb are applied to the daily traffic to generate trip matrices by time interval.

One approach to generating inputs to dynamic network assignment is to simulate individuals' travel behavior and generate trips over a time axis; car trips generated in the simulation can be loaded onto the road network as they are generated. This approach has many advantages over the use of trip matrices by time of day. First, the latter method is incapable of capturing the effect of various planning measures, e.g., congestion pricing or parking surcharges by time of day. Second, the latter method is valid only when demand can be viewed as inelastic. The foundation of the former approach, on the other hand, is the recognition that travel demand is elastic. It is also based on the recognition that individuals may modify their behaviors when faced with changes in their travel environment. Consequently this approach is capable of accounting for behavioral changes caused by planning measures.

An increase in travel time due to congestion may cause the individual to change the route, departure time, destination, travel mode, or any combination of these. It may also be the case that the traveler chose not to make the trip altogether. These, and possibly other reactions of individuals, can be represented while simulating individuals' behaviors when reactions to changes in the travel environment are properly captured in the simulation procedure. Furthermore, when the simulation of individuals' behaviors and the simulation of network flow are interactively performed repeatedly, it may become possible to simulate individuals' learning and adaptation to changes in the travel environment, thus simulating the process in which individuals' behaviors collectively approach "equilibrium."

Dynamic network simulation also has many advantages. It has many applications in real-time traffic control as well as the prediction of network performance. More precise estimation of the magnitude and duration of congestion will become possible by dynamic network simulation. This will in turn make the prediction of the cost of congestion more precise. The analysis of transient phenomena, such as non-recurrent congestion due to a traffic accident, can be performed more appropriately. In addition, the environmental impacts of automotive traffic can be assessed more accurately because of the explicit inclusion of the time dimension. For example, the CO_2 concentration at an intersection can be estimated more adequately using simulation outputs. Roles of automotive traffic in the "heat island" phenomenon now found in large metropolitan areas can be better examined. The formation of photochemical smog, which depends on the clock time when NO_x and other contributing substances are emitted, can also be estimated more adequately.

Returning to the simulation of individuals' behaviors discussed above, it is worthy to note in this context that it offers additional merit of being capable of addressing how various planning measures may influence the patterns of trip chaining and the frequency of cold starts. Conventional methods of demand forecasting, which are trip-based, do not adequately capture how patterns of trip chaining may be influenced by policy measures, and how these changes may in turn affect traffic flow and pollutant emissions, as well as the frequency of cold starts. The combination of the simulation of individuals' travel behaviors and dynamic simulation of network flow can thus be a very powerful tool for traffic management and planning.

As a more exact and versatile tool for traffic management, demand forecasting and policy analysis, a micro-simulation model system of individuals' daily travel behaviors and network flow has been developed at Kyoto University. The main body of the system comprises two simulators, PCATS and DEBNetS. PCATS stands for Prism Constrained Activity-Travel Simulator, and, as its name indicates, simulates an individual's daily activity and travel while explicitly incorporating Hägerstrand's space-time prism (Hägerstrand, 1970). DEBNetS, which stands for Dynamic Event-Based Network Simulator, is a micro-meso scale simulator of network flow. These two simulators are integrated to perform, in the nomenclature of the conventional four-step procedures, trip generation, trip distribution, modal split and network assignment. The critical difference from the conventional procedures, however, is the inclusion of a time axis; trips are generated and network flow is simulated along time in the PCATS/DEBNetS system. Another important difference is that PCATS simulates not individual trips but the series of trips generated by an individual over a course of the day, while maintaining the spatial continuity of the trips and coherently representing various constraints on trip making.

In this paper, the PCATS/DEBNetS system is first outlined, with emphasis on PCATS. Also discussed is a synthetic household generator and demographic simulator called HAGS. HAGS is used to generate data for long-term forecasting with PCATS/DEBNetS. Ongoing efforts to extend the capabilities of the simulation system are then described. These are:

- incorporating models of prism vertex locations,
- adoption of a fine (10m × 10m) grid system instead of a system of large traffic zones, and
- use of a Markov Chain Monte Carlo (MCMC) method to handle colossal choice sets.

Following this, results of empirical applications are summarized. Problems that can be addressed with these enhancements are then discussed.

OUTLINE OF THE PCATS/DEBNETS SYSTEM

Three simulators, PCATS, DEBNetS and HAGS, are outlined in this section, with more emphasis placed on PCATS because it has several unique features that are noteworthy. HAGS stands for Household Attribute Generating System and generates disaggregate data of household and person attributes.

There are several important differences between the conventional travel demand forecasting procedures and the proposed micro-simulation system that are expected to facilitate more authentic representation of travel demand. They are:

- Clock time is the key dimension of the simulation system; activities and trips are generated for each person, then properly expanded and loaded on the network along a continuous time axis;
- The series of activities and trips made by an individual are simulated for the period of one day; thus spatial and temporal continuity of activities and travel is maintained and interdependencies among trip attributes are represented; and
- Hägerstrand's prisms are evaluated for each individual in the simulation along with coupling constraints associated with private modes of travel and operating hours of public transit; activities and trips are generated within the confines of these constraints (Kitamura, Fujii et al., 2000).

Incorporating space-time prisms and other constraints into PCATS implies that the travel environment is depicted more realistically for each individual. No individual and household attributes are aggregated by zone to eliminate loss of information and resulting statistical inefficiency and prediction error; in fact, as discussed later in this paper, parts of on-going effort are dedicated to the use of a fine grid system instead of traffic zones. The PCATS-DEBNetS system thus takes full advantage of the advent of fast computing and data management capabilities that are now available and eliminate many of the constraints under which the four-step procedures were constructed in the 50's and 60's.

PCATS[1]

Behind the development of PCATS lies the recognition that various constraints imposed on individuals' activity and travel are not well

[1] The discussions of this section are based on Kitamura, Fujii et al. (2000).

represented in conventional models of travel behavior.[2] Emphasized in PCATS, therefore, are the constraints imposed on the individual's movement in geographical space along time. Because the speed of travel is finite while the time available for travel and activity is limited, the individual's trajectory in time and space is necessarily confined within a certain region. This region is called "Hägerstrand's prism." PCATS first identifies the set of prisms that govern an individual's behavior, then generates activities and trips within each prism while observing coupling constraints involving private travel modes and operating hours of public transit.[3]

PCATS first determines for each individual the periods in which he is committed to engage in a certain activity, or a bundle of activities, at a predetermined location.[4] These periods are called "blocked periods." The complement of the set of blocked periods for an individual comprises the set of "open periods." A Hägerstrand's prism is established for each open period, i.e., given the mode of travel being used, it is determined for each zone whether the zone can be reached within the open period and, if so, how much time can be spent in the zone.

Blocked periods for workers are typically determined by work schedules. As in Damm (1982), then, a worker's day may be assumed to include three prisms: one before work, one during the lunch break, and one after work. The beginning time of the first prism before work and the ending point of the last prism after work are not observed. As discussed later in this paper, stochastic frontier models have been developed to determine unobserved locations of prism vertices (Kitamura, Yamamoto et al., 2000; Pendyala et al., 2002; Yamamoto et al., 2002), and are being incorporated into PCATS.[5]

In PCATS, the probability associated with a daily activity-travel pattern is decomposed into a series of conditional probabilities, each associated with an activity bundle and the trip to reach the location where it will be pursued,

---

[2] For details, see Fujii et al. (1997) and Kitamura & Fujii (1998). The latter reference contains validation results of PCATS.

[3] In the current version of PCATS a zone system is used to construct a prism, which defines the universe of alternatives that exist for each activity-travel decision (activity type, duration, location and travel mode).

[4] In this study the set of activities pursued at a location is called an activity bundle, and the activity bundle is treated as the unit of analysis. Therefore there always is a trip between two successive activity bundles. It is in general unknown from standard travel survey data if an activity bundle is fixed in time and location. In the exercises presented in this paper, work, work-related business and school activities are treated as fixed activities.

[5] The results reported in this paper are based on versions of PCATS that had not incorporated the stochastic frontier models.

given current conditions and the past history of activity engagement. In this sense PCATS has a sequential structure. Models of activity engagement are now being developed to adopt a two-tier structure in which the decision of activity engagement in the respective prisms is first made, then attributes of activities and trips are determined (Kitamura et al., 2001), and will be incorporated into PCATS in the future.

The conditional probability of pursuing an activity bundle is further decomposed to yield the following three sets of model components: 1) activity type choice models, 2) destination and mode choice models, and 3) activity duration models. The activity type choice models are two-tier nested logit models. The upper tier comprises three categories of activity bundles: (A) in-home activities, (B) activities at or near the location of the next fixed activity, and (C) general out-of-home activities. Nested under the first category are two lower-level alternatives: (A-1) engage in out-of-home activities subsequently in the prism, and (A-2) do not engage in out-of-home activities within the prism. The alternatives nested under (C) include the following activity types: meal, social, grocery shopping, comparison shopping, hobbies and entertainment, and sports and recreation. These are defined in terms of the trip purposes as adopted in the travel diary data used for the implementation of the PCATS components. The destination and mode choice models are also nested logit models. Alternative destinations constitute the upper-level alternatives, and available travel modes are nested under each destination alternative.

The type of the first activity bundle in an open period is determined using an activity type choice model. The models are specified such that the probability that a given activity type will be selected decreases as the time available in the prism becomes shorter relative to the distribution of activity durations for that activity type. Given the activity type, a destination-mode pair is next determined using a destination and mode choice model.

As noted earlier, geographical zones are used in the versions of PCATS that are used to produce results reported in this paper. The extension of a prism is evaluated for each travel mode, and destination-mode pairs are excluded from the choice set if they do not fall in the prism for the mode. Again, the amount of time available at the destination is one of the determinants of the choice probability along with the attributes of the destination zone and the trips by respective travel modes to the destination. In the application examples presented later in this paper, travel modes are classified into auto, public transit, walk, and bicycle in the case study for Kyoto, and into auto, public transit, and "others" in the study for Osaka.

Following these, the duration of the activity at the destination is determined using the activity duration model corresponding to the activity type. PCATS initially adopted standard hazard-based duration models with Weibull distributions (Fujii et al., 1997; Kitamura, Fujii et al., 1998, 2000). Standard hazard models, however, do not appear adequate when accounting for the presence of prism constraints because the hazard is presumably influenced by the presence of constraints. For example, it is unlikely that an individual would terminate an activity just a few minutes before the time when a constraint so dictates. Instead, he would extend the activity by a few minutes so that there will be no window of time which is too short to engage in another activity and therefore will be wasted. This situation, then, can be best represented by the distribution of activity durations that has a spike (probability mass) at the time point when the constraint dictates that the activity be terminated. It was suspected that this is one of the reasons why activities tended to be over-generated in an earlier version of PCATS (Iida et al., 2000). To resolve this problem, split population survival models (Schmidt & Witte, 1989) are adopted in PCATS. The split population survival model assumes a two-stage structure. In the first stage, it is determined whether the entire time available in the prism is allocated to the activity (i.e., the activity is terminated when the constraint dictates so). Given that the entire amount of available time is not allocated to the activity, then the time allocated to the activity is determined in the second stage. The version of PCATS applied in Osaka (see "APPLICATION CASES") incorporates split survival models.

Once the attributes of an activity bundle are all determined, the procedure is repeated for the next activity bundle in the same prism. Activity and travel in each open period is thus simulated by recursively applying these model components, while considering the history of past activity engagement. The procedure is repeated until each open period is filled with activities. Note that activity starting and ending times are determined based on the simulated activity durations and, in case of auto trips, travel times obtained from DEBNetS.

DEBNetS[6]

DEBNetS takes as its input auto trips produced by PCATS and replicates the dynamics in network traffic flow along the time axis. DEBNetS determines each vehicle's speed based on simplifying rules, adopts event scanning, and thereby reduces computational requirements. The behaviors of individual vehicles are represented by applying speed (u)-density (k) relationships to

[6] This section is based on Kitamura, Fujii et al. (1998). The development of DEBNetS is presented in Fujii et al. (1998).

roadway segments, with the assumption that each segment is internally homogeneous (the u-k relation proposed by the Bureau of Public Roads is used in the simulation). A network link is therefore divided into a number of homogeneous segments. As a vehicle enters a segment in the simulation, the travel speed of this vehicle is determined based on the number of vehicles in the segment at the time of its entrance, using the u-k relationship specified for the segment. The time the vehicle exits the segment is determined based on the travel speed thus evaluated.

Because of this representation of vehicular movement on a network, computation is required of each vehicle only when a vehicle exits a segment to enter a new segment. If the vehicle is located at a branching point, then a segment to enter is selected based on shortest path information (updated every 15 minutes in the application cases reported later in this paper), and the travel time to traverse that segment is determined. If the vehicle is not at a branching point, then only the travel time on the new segment needs to be calculated. Computation is thus required infrequently for each vehicle, only as often as there are segments on the route taken from the origin to the destination. Because of this, event scanning is adopted in DEBNetS. Trips produced by PCATS are loaded onto the network exactly at the clock time when they are supposed to start.

In the case studies, the CO_2 emissions as a result of a vehicle trip, say trip i, is evaluated as $C_i = \sum_{j \in \Omega_i} L_j f_c(u_{ij}, \delta_i)$, where C_i denotes the total CO_2 emissions by vehicle trip i in the study area, Ω_i the set of link segments on the route of trip i, L_j the length of segment j, u_{ij} the mean travel speed to traverse segment j in trip i, δ_i the type of the vehicle used for trip i, and $f_c(u_{ij}, \delta_i)$ the emissions factor per unit distance for vehicle type δ_i at speed u_{ij}. Vehicles are classified into: small vehicles, which include passenger cars and light duty trucks, and others. The emissions factor, f_c, was developed based on the mean fuel consumption rate by vehicle type and speed class available from the City of Kyoto, and the mean automotive CO_2 emissions per liter of fuel available from the Japanese Ministry of Environment. The emissions by individual trips thus estimated are aggregated in the simulation to produce total emissions in the study area by time of day.

Many decisions simulated in PCATS, e.g., travel mode choice and destination choice, are in part based on auto travel time information, which is output from DEBNetS. Auto trips that are fed to DEBNetS, on the other hand, are output from PCATS. It is in general the case that travel times used in PCATS are not identical to those in DEBNetS. For this reason, these two simulators are applied iteratively until convergence in travel times is obtained.

DEBNetS Enhancement and Validation[7]

Earlier applications of the PCATS-DEBNetS system have indicated that the upstream propagation of congestion must be appropriately represented, and that computational time must be reduced for practical application of the simulation package. For the former, procedures are introduced to regulate the rate of the outflow from the last (most downstream) segment of a link to better reflect the capacity of the traffic lights as well as congestion in the downstream link.

Because of the way traffic flow is simulated in DEBNetS, computational requirements can be effectively reduced by event scanning. Event scanning, however, calls for the task of sorting events according to the time of their occurrences. The time required for this sorting increases as the number of vehicles in simulation increases, and can be a substantial portion of the total computation time. A scanning scheme, which is a combination of event scanning and periodic scanning, is therefore adopted to reduce this sorting time; traffic flow is simulated by event scanning, while events are sorted and vehicle data are updated by periodic scanning. This reduced computational time quite substantially to about one-seventh.[8]

Using the empirical morning peak benchmark traffic data made public for the purpose of validating simulation models by the Road Traffic Simulation Systems Clearing House,[9] a validation run was performed with DEBNetS. With a value of 0.95, the correlation coefficient between simulated hourly link traffic volumes and observed volumes indicates good fit (Fig. 1). Correlation coefficients of 10-minute traffic volumes have an average of 0.830 between 8:00 A.M. and 10:00 A.M (Fig. 2).[10] Link traffic speeds observed with

---

[7] This section is based on Kikuchi et al. (2002)

[8] The application to the City of Osaka (Kikuchi, Fujii et al., 2000; Kitamura et al., 2001) involved a network with 2,994 links, 1,050 nodes, 292 centroids and a total of 990,575 trips. A DEBNetS run for a 24 hour period took 17 hours and 31 minutes using a personal computer with two Pentium III, 800 MHz CPUs and Sun OS 5.8, with shortest path trees updated every 15 simulation minutes. With the improved scanning scheme, the same run took 2 hours and 25 minutes with shortest path trees updated very 5 simulation minutes.

[9] http://trans1.ce.it-chiba.ac.jp/ClearingHouse/

[10] Correlation coefficients between 7:50 A.M. and 7:59 A.M. are excluded as they were heavily influenced by the initial condition of the simulation, which started with an empty network. Correlation coefficients have an average of 0.859 between 8:00 A.M. and 8:29 A.M., 0.842 between 8:30 A.M. and 8:59 A.M., 0.822 between 9:00 A.M. and 9:29 A.M., and 0.798 between 9:30 A.M. and 9:59 A.M.

10-min. intervals, on the other hand, are not well represented with an overall correlation coefficient of 0.32, suggesting needs for further improvement of the simulator.

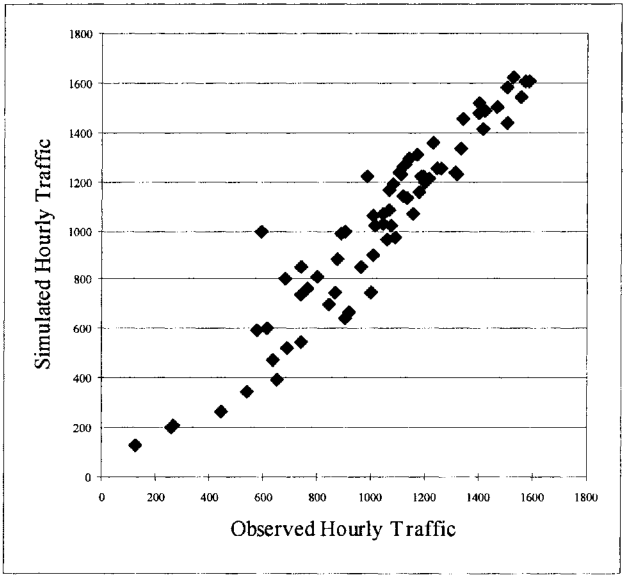


Figure 1. Correlation between Observed and Simulated Hourly Link Traffic Volumes.

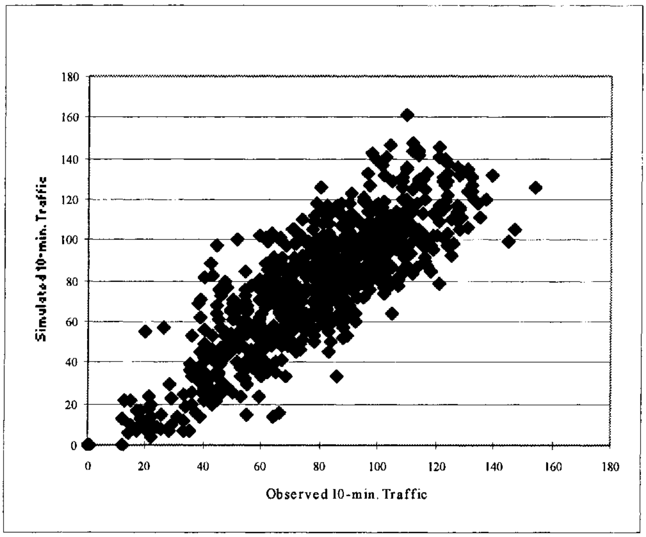


Figure 2. Correlation between Observed and Simulated 10-min. Link Traffic Volumes.

HAGS[11]

The case study for Osaka described in "APPLICATION CASES" of this paper is an application of the micro-simulation model system to long-range forecasting with a horizon year of 2020. This calls for the generation of future household and person attributes at the household or person level by zone, so that the disaggregate model components of the micro-simulation system can be applied. This is done by HAGS, which comprises: Household Distributor, Household Ager, and Fixed Activity Generator.[12]

The Household Distributor determines the distribution of attributes of households in the respective zones based on data from the census, travel surveys, and other sources. An iterative proportional fitting (IPF) method (Beckman et al., 1996) is applied to base-year marginal distributions of pertinent household and person attributes, obtained for each zone from the census and other data, along with their area-wide joint distributions obtained from base-year travel survey data, to yield a frequency distribution of households by their attributes for each zone. Each zone is then populated by cloning households from the travel survey data according to the distribution obtained for the zone.

The Household Ager simulates the aging of the base-year households generated by the Distributor through the horizon year. Simulated events include: birth, death, marriage, divorce, employment, and residential relocation. The probabilities of these events are obtained from the census and other available statistics while accounting for the trends of decreasing household size, increasing single-person households, and increasing labor force participation by women. Household vehicle ownership is endogenously determined using models estimated on the base-year travel survey data. Future driver's license holding by individuals is estimated using a simple cohort model. New households generated through these events are assumed to stay in the study area with those probabilities that replicate observed social changes in the regional population.

The Fixed Activity Generator determines the beginning and ending times of fixed activities for each household member from the Household Ager. Since only work (including work-related business) and school activities are considered as fixed activities in the versions of PCATS so far developed, work/school starting and ending times are generated probabilistically for each

[11] This section is based on Kitamura, Fujii et al. (1998).

[12] Details are reported in Nishida et al. (2000)

worker or student based on their distributions observed in the base-year travel survey data. The household and person attributes thus generated are input to PCATS in the micro-simulation.

EXTENSION

This section presents some of the improvements being made to the PCATS-DEBNetS system. These improvements are expected to make the simulation system more realistic and versatile.

Incorporating Models of Prism Vertex Locations

One of the important features of PCATS is its representation of various constraints imposed on individuals' activity and travel. In particular, Hägerstrand's prisms are evaluated for each individual to represent the spatial and temporal constraints imposed on him. The vertex of a prism (its beginning or ending point), however, is not always observable.[13] In earlier versions of PCATS, a fixed time of day is arbitrarily selected for unobserved prism vertex. This, of course, is not a desirable approach because there is no theoretical or empirical basis for the selection of the particular time of day, and also because there is no reason to believe that every individual's prism vertex is located at the same time point.

Stochastic frontier models have been developed to estimate the location on a time axis of an unobserved prism vertex (Kitamura, Yamamoto et al., 2000; Pendyala et al., 2002; Yamamoto et al., 2002). A stochastic frontier model depicts the value of an unobserved variable based on the relation that the unobserved variable is greater (or smaller) than some observed variable. In this case, the relations hold that the location on a time axis of an unobserved origin vertex is earlier (smaller) than the beginning time of the first trip observed in the prism, and that the location of an unobserved terminal vertex is later (greater) than the ending time of the last trip in the prism.

Results of model estimation have indicated that an hour of commute pushes back the terminal vertex of a worker's evening prism by 52.4 minutes; a female member of a household with children has a terminal vertex that is about 100 minutes earlier than that of her male counterpart; and the terminal

[13] The terminal vertex of a worker's morning prism before work, is set in PCATS at the work beginning time, while its origin vertex is unobservable. For a worker's evening prism after work, on the other hand, the origin vertex can be set at the work ending time, while the terminal vertex is not observable.

vertex of a younger single person tends to be located later. It is also shown that a worker's terminal vertex is located 96 minutes earlier on a day when he is not working. It has also been shown there exists a mean difference of 1.46 hours between the observed trip ending time and the estimated location of the terminal vertex of a worker's evening prism (Kitamura, Yamamoto et al., 2000).[14]

In the studies referenced above, the location of a prism vertex was estimated by the stochastic frontier models using person and household attributes and commute characteristics as explanatory variables. The causal relation assumed in these stochastic frontier models, where commute characteristics affect the location of the prism vertex, is reversal of that assumed in PCATS; in PCATS, prism constraints influence commute characteristics. Models are therefore re-estimated for PCATS implementation without commute characteristics as explanatory variables, except for commute distance which is treated as exogenous.

Earlier versions of PCATS produced too many trips in later hours of the day, which can be attributed at least in part to the arbitrary value of 3:00 A.M. used as the location of the terminal vertex of an evening prism. This problem can be resolved by introducing the modified stochastic frontier models into PCATS and adopting more realistic values for prism vertices.

Adoption of a Fine Grid System: From Zones toward Coordinates

As earlier discussions of this paper have alluded to, the primary reason for the use of traffic zones in conventional demand forecasting models, including the four-step procedures, is the computational limitations that existed in the 50's

---

[14] There remains the question of whether the vertex location as depicted by the stochastic frontier models in fact represents the prism constraint in the strict sense of Hägerstrand. One could argue that what the models depict may represent a threshold which an individual subjectively holds as the earliest possible starting time or the latest possible ending time for a trip, which may not necessarily coincide with actual constraints that are governing. For example, a commuter may believe that he cannot possibly leave home before a certain time in the morning, but he may do so for a business trip. Models of prism vertices are estimated in this study with empirical data without any information on the individual's beliefs or perceptions of prism constraints. Yet, observed travel behavior is governed by subjective beliefs and perceptions, e.g., "I must return home by midnight" or "I cannot possibly leave home before 6:30 A.M." Thus some ambiguity does exist about the nature of what the models depict. It is nonetheless considered reasonable to assume here that they offer a useful measure for the practical purpose of determining the earliest possible departure time or latest possible arrival time for a trip (Kitamura, Yamamoto et al., 2000).

and 60's. Information collected at household and individual levels through travel surveys has been aggregated to zonal values (e.g., means or medians) and some models are estimated using zonal statistics. It has been pointed out that this leads to gross statistical inefficiency because most information is lost in the process of aggregation.

Less attention has been directed to the problem of error and low resolution in spatial representation that result from the use of traffic zones. A traffic zone typically covers quite a large area, introducing inaccuracies in representing level-of-service (LOS) attributes of a trip and creating difficulties in handling short trips typically made on foot or by bicycle.

The most typical case for the former may be found in the representation of walking in a transit trip. Since all LOS attributes are zone-based, one representative value is assigned to the walking distance to a transit stop for all transit trips generating from a zone. Walking distance, however, varies greatly from location to location within a zone. Using a representative value may lead to serious errors in analysis. This applies to non-linear models, including discrete choice models, even when the representative value is an accurate zonal average.

Short trips are difficult to represent with a zone system because they tend to be internal to a zone, or, "intra-zonal." Again, the same, "representative," zonal values are assigned to all intra-zonal trips. All intra-zonal trips, then, would have the same probability of being made by auto or on foot. This would make the analysis unrealistic, especially when zones are large as is the case in the fringe of a metropolitan area. In fact mode choice models used to exclude non-motorized modes of travel in most planning regions. Presumably this is at least in part due to the difficulty in dealing with short trips within the framework of conventional zone systems. One approach to overcome these shortcomings and extend the applicability of the model, is to adopt finer zones, or, a coordinates system to represent spatial location.

Attempts have been made to develop a location reference system and to construct models on it as components of PCATS. In Kikuchi, Kobata et al. (2000), a fine grid system is defined and the study area is subdivided into 10 m × 10 m parcels, and the location of each parcel is referenced using a coordinates system. The study area is a rectangular area (13 km east to west, 11 km south to north) that centers around the central business district of Kyoto, Japan. The grid system produced about 140 million 10 m × 10 m parcels. After eliminating parcels on which no opportunities for activities exist (e.g., river water, forests, railroad tracks, roadways), approximately 74 million parcels qualify as potential trip destinations (Kikuchi, Fujii et al., 2001).

The location of railroad stations, bus stops and other transportation facilities are input to the database using geographical information system (GIS) software.[15] Parcel-to-parcel LOS attributes are determined first by simulating auto and bus traffic on a network of major roadways for each hour of the day.[16] This simulation produced the travel time, number of transfers, and transit fare between each pair of nodes on the network. LOS information between 10 m × 10 m parcels is obtained by systematically inter/extrapolating the LOS data obtained from the simulation. The duration of a walk or bicycle trip is evaluated by applying a constant to the parcel-to-parcel straight-line distance obtained by the GIS software.

The land use data developed in the study is based on information compiled for 3,635 neighborhood units in the City of Kyoto. Each neighborhood unit typically comprises of housing units on the two block faces that share a street segment. Since no information is available on how land uses are distributed within each neighborhood unit, they are uniformly distributed to parcels that lie within a neighborhood unit. Obviously this is an approximation. Ideally the land use database should be developed based on information on each plot of land, as is done in Portland, Oregon.

Models of destination and mode choice are developed based on the grid system and the land use database thus developed (Kikuchi, Kobata et al., 2000), and applied to evaluate selected TDM measures (Kikuchi, Fujii et al., 2001). Such applications call for the development of methodologies to efficiently handle the huge number of alternatives that are involved in choice models defined on the grid system. The following subsection is concerned with such methodologies.

Application of Markov Chain Monte Carlo (MCMC) Methods to Handle Colossal Choice Sets

Even with the advent of fast computers, simulating discrete choices requires a substantial amount of time when the choice set is large, because evaluating choice probabilities for all alternatives in the choice set requires a substantial amount of computation time. For example, consider the multinomial logit model, where the probability that alternative j will be chosen by individual i from the choice set, Ω_i, is given as $PD_i(j) = \exp(\beta' X_{ij}) \big/ \sum_{l \in \Omega_i} \exp(\beta' X_{il})$. To simulate a choice according this choice model, one must evaluate $PD_i(j)$

[15] The GIS software used in this study is SIS (Spatial Information System) V5.2, Informatix, Inc.

[16] At this point, this simulation is independent of the network traffic simulation of DEBNetS.

for all alternatives in Ω_i, which requires that the denominator of $PD_i(j)$ be evaluated. This, however, involves computing J_i exponential functions, where J_i is the number of alternatives in Ω_i. Although this wouldn't be a problem when zones are used as the alternatives of destination choice, it imposes serious computational problems when J_i is as large as hundreds of thousands.

When destination choice is formulated as a multinomial logit model, the Markov Chain Monte Carlo (MCMC) algorithm applies quite well to simulate choices. The algorithm is shown schematically in Fig. 3. In the figure, r_i refers to an alternative. Once the procedure is repeated large enough a number of times and the influence of initial condition has diminished, r_i's can be drawn with large intervals to form a sample of alternatives that are drawn according to the choice probabilities as indicated by the model.[17]

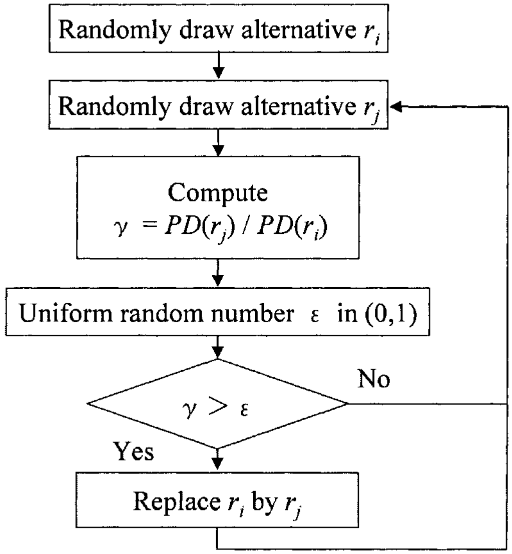
Most critical in this application is the fact that the algorithm only requires the ratio of two choice probabilities, but not choice probabilities themselves. The ratio,

$$\gamma = PD_i(r_j) \big/ PD_i(r_k) = \exp(\beta' X_{ir_j}) \big/ \exp(\beta' X_{ir_k}),$$

does not involve the denominator of the logit choice probability, and thus can be evaluated very easily. This substantially reduces the computational requirements for choice simulation. The accuracy of the algorithm was tested by simulating destination choice in an abstract uniform circular city where the distribution of destination locations can be theoretically determined. The result indicated that the MCMC algorithm produced a distribution of destination locations that is statistically not different from the theoretical distribution (Kikuchi, Yamamoto et al., 2001).

An example of simulation results is presented in Fig. 4. The figure shows a set of destination locations that are chosen under the existing condition, and another set of locations that are chosen when the service level of public transit is improved. In this simulation, the individual who is located at "S," whose next fixed activity must take place at "N," is choosing a destination for a discretionary activity. It can be seen that the distribution of destination locations expands with the improvement in service level. In fact the sum of travel times from the current location (S) to the destination, then to the next

[17] Previous applications of MCMC algorithms in the transportation field can be found in Hazelton et al. (1996) and Yamamoto et al. (2001). Also see Hajivassiliou et al. (1996) and Chiang et al. (1999).

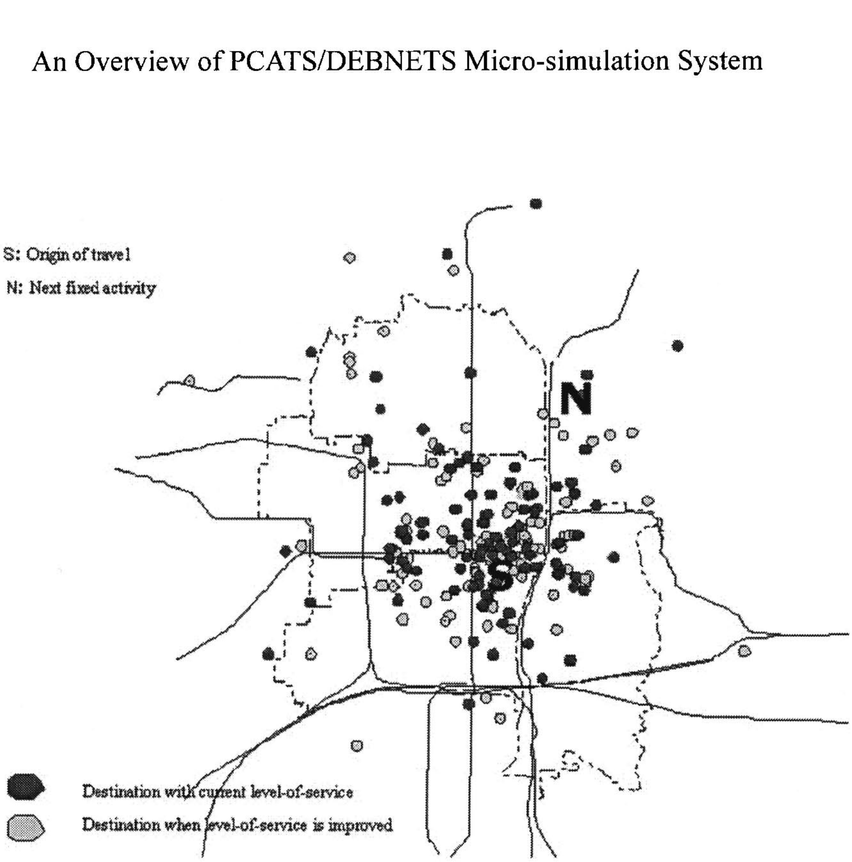


Source: Kikuchi, Yamamoto et al. (2001)

Figure 3. MCMC Algorithm.

fixed activity (N) increases from an average of 4,396 m to an average of 4,533 m with the improvement. The variance in travel distance increases from 2.1×10^6 to 3.6×10^6.

Another improvement being made is concerned with destination choice. The destination choice models of PCATS are being modified to better reflect recent findings on destination choice behavior. Using data from the Los Angeles metropolitan area, Kitamura, Chen et al. (1998) show that the travel time from a potential destination to the home base is as influential a factor of destination choice as the travel time from the origin to the potential destination. It has also been shown that closer locations tend to be chosen toward the end of the day, and distance to a destination is positively correlated with the time spent there. The study, however, does not incorporate the space-time prism into its analytical framework. The destination choice models in PCATS are now being reformulated to reflect these findings and improve their predictive capability. In addition, as noted earlier, the two-tier model of activity engagement and activity attributes is being developed for implementation in PCATS.



Source: Kikuchi, Yamamoto et al. (2001)

Figure 4. Distribution of Simulated Destination Locations.

APPLICATION CASES

The PCATS-DEBNetS system was first applied to evaluate the effectiveness of several transportation planning measures in reducing CO_2 emissions in the City of Kyoto, Japan (Fujii et al., 2000; Kitamura, Fujii et al., 1998, 2000). It was then applied in Osaka (Iida et al., 2000; Kawata et al, 1999; Kitamura, Fujii et al., 2000), Toyota (Kikuchi et al., 1999) and Ashiya, Japan.[18] It is currently being implemented in Tampa, Florida. The Osaka application deployed the demographic simulator, HAGS, to facilitate long-term forecasting. Drawing from Kawata et al. (1999), Iida et al. (2000) and Kitamura, Fujii et al. (2000), results from the Osaka application are presented in this section. This case study is based on a zone system.

[18] Also see Arentze et al. (2001)

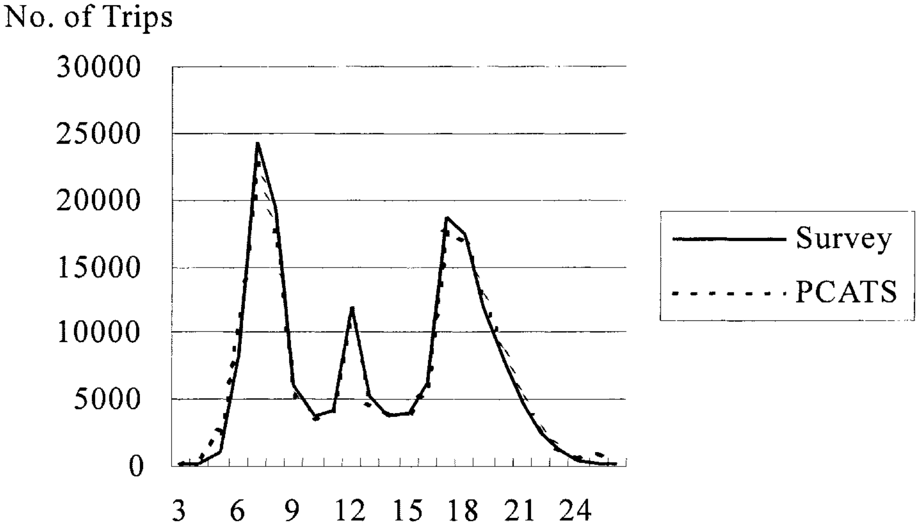
The City of Osaka is the largest of the three major cities in the Kei-Han-Shin (Kyoto-Osaka- Kobe) metropolitan area of Japan. It has a population of 2.6 millions and 1.1 million households (as of October, 1995), and an area of 221.27 km^2. The city is served by ten lines of the Japan Rail's networks and 15 rail lines operated by other private rail companies. The City of Osaka operates eight subway lines and a people mover system. According to the 1990 household travel survey, the share of auto trips is about 17%, while public transit accounts for about 34% of all trips. The population of the city has been declining since 1990.

In this application, level-of-service (LOS) variables are evaluated for public transit systematically using public transit operation schedules, network connectivity data, and fare schedule data. For each origin-destination pair and with 10-minute intervals throughout transit operating hours, the total travel time, number of transfers, and transit fare are evaluated for the fastest route. The results are aggregated into 2-hour intervals and used as transit LOS variables in PCATS. At this stage of model development, transit LOS variables are all static, and the effect of road congestion, which does influences bus operation, is not represented (bus, however, is a minor mode of public transit in Osaka where subway networks are well developed.)

The study area is defined by the ring road that surrounds the City of Osaka and areas in its periphery. Residents in the study area, and those who reside outside the area but worked or studied in the area, are included in the simulation. The records of a total of 103,462 individuals were drawn from the results of the 1990 household travel survey. The home and work locations, other person and household attributes, and the fixed (work and school) activities of these individuals are retained, but all other activities and trips are deleted from the records. The resulting records, supplemented with land use data and network data, constitute the base of the simulation. A PCATS run on this database took approximately 6 minutes on a Pentium II (300 MHz) Linux machine. Outputs of PCATS are quite similar to trip records in a household travel survey data set, and indicate the purpose, origin, destination, mode, beginning and ending times of each trip, from which the location, beginning time and duration of each out-of-home activity can be inferred.

PCATS was first run to examine how well it replicates the 1990 data. The resulting mean number of trips per person per day is 2.87 for workers and 2.49 for non-workers. These compare with the average trip rates of 2.75 and 2.63, respectively, obtained from the survey data. The error in the PCATS prediction is 4.7% for workers and –5.3% for non-workers. Although these results still represent over- or under-prediction of quite a few trips, they at the same time represent a substantial improvement in the model system's accuracy compared with the earlier results from the Kyoto application.

The number of trips generated is shown by time of day in Figs. 5 and 6 for workers and non-workers, respectively. Overall trip generation is well replicated along the time axis by PCATS. Comparing the two figures indicates that trip generation by workers is better replicated than that by non-workers. This is presumably because non-workers tend to have larger unblocked periods and more degrees of freedom in their activities, making prediction more difficult.

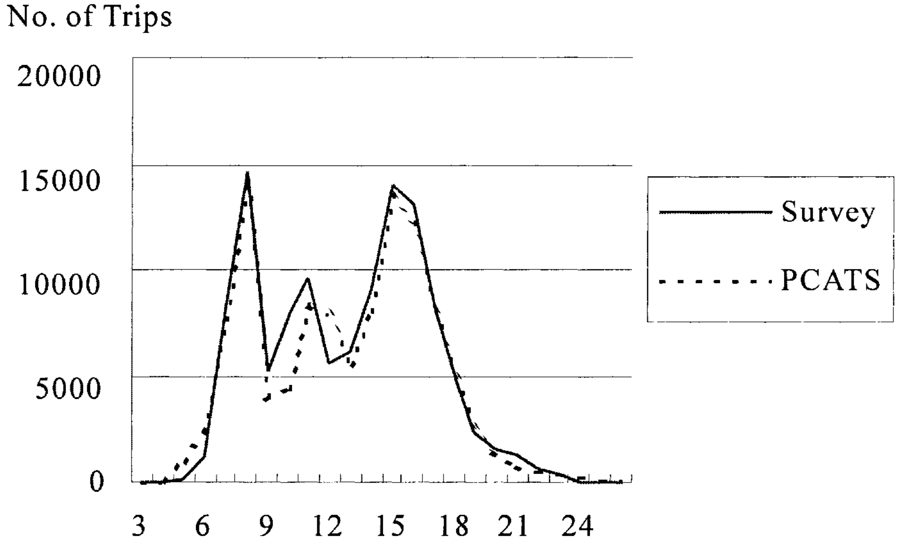


Source: Kawata et al. (1999)
Figure 5. The number of trips generated by time of day: Workers.

The network adopted for the Osaka case study had 3,057 links, 1,098 nodes and 289 centroids, of which 36 are for external traffic. A DEBNetS run took approximately 30 minutes using one processing unit of Fujitsu VPP-500.[19] A preliminary comparison of travel times indicated that DEBNetS overestimated travel speeds on toll roads (an average of 36.8 km/h was estimated while the observed 1994 average was 28.5 km/h), while reasonable estimates were obtained for surface roads (19.7 vs. 18.8 km/h). Presumably this was because the same assignment method as adopted by the regional planning agencies was used in this study, and the expressway tolls were converted to equivalent

[19] VPP-500 consists of 15 processing units, each having a capability of 1.6G flops.

travel times and added to the actual link travel times in the assignment. As a short-term solution to reduce the discrepancies, link constants were calibrated using a simple algorithm, and then added to the link travel time. This yielded an average estimated expressway travel speed of 29.7 km/h (Kikuchi, Fujii et al., 2000). In the long run, more behavioral route choice models will be incorporated into DEBNetS. The average absolute prediction error for daily traffic volume was 11.4% on 7 major links on the network, and a prediction error of 8.7% was obtained for screen-line traffic on 8 major screen lines.



Source: Kawata et al. (1999)

Figure 6. The number of trips generated by time of day: Non-workers.

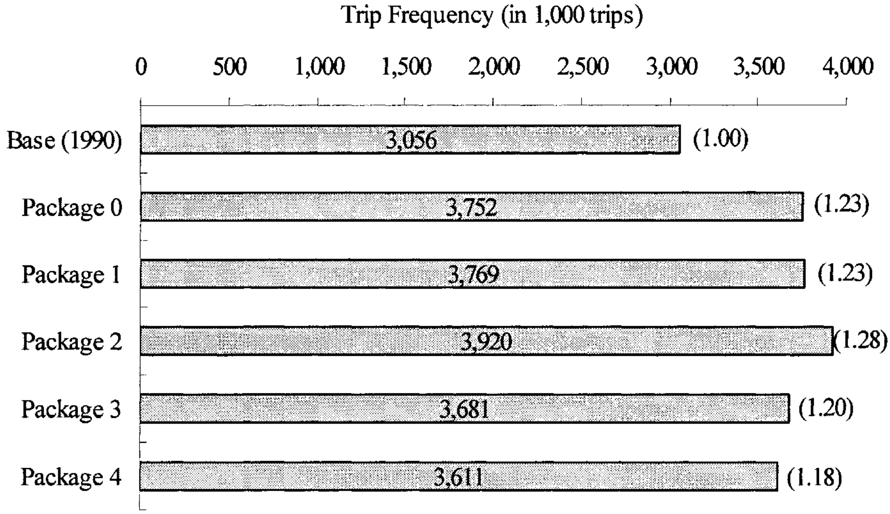
In the long-range forecasting for 2020, the population of the City of Osaka is assumed to decrease from the current 2.60 millions to 2.51 millions. Zonal population, employment and land use characteristics are adjusted for 2020 first assuming the completion of planned housing and other development projects. Balances of population and employment are then distributed to zones. Weights for individuals are determined based on the population age distributions by municipality for 2020, which were obtained using a cohort method. These weights are applied to household members aged by HAGS. As noted earlier, HAGS reflects trends towards smaller households, later marriage, and increasing labor-force participation by women in Japan.

The following policy packages are examined in the study:

Package 0. Do nothing
Package 1. Execute planned infrastructure development projects
Package 2. Package 1 plus circumferential and radial arterials and other facilities to disperse through traffic
Package 3. Package 1 plus new rail lines, LRT, and other measures to reduce auto use
Package 4. Package 1 plus minimal infrastructure projects (a circumferential roadway and LRT) and introduction of congestion pricing and transit malls in the central city.

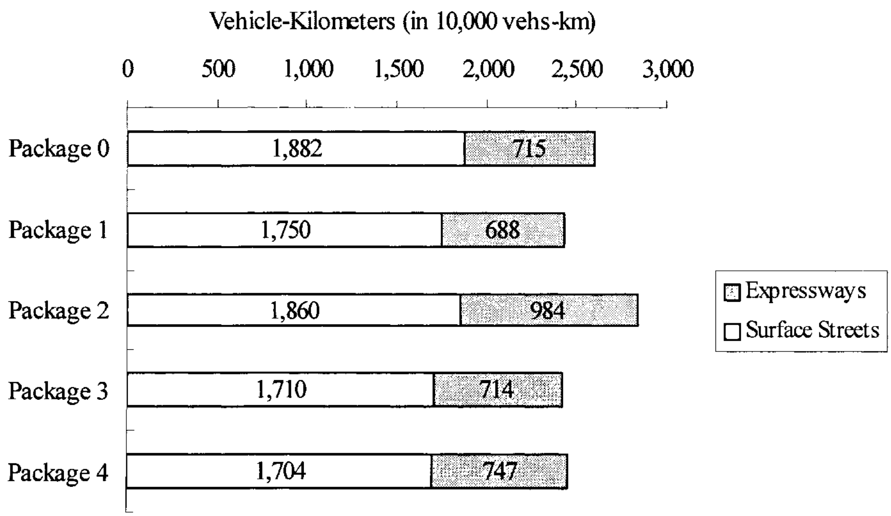
Despite the decline in the total population in 2020, the number of auto trips increases substantially (Fig. 7). This is due to increases in driver's license holding among women and older individuals. Another factor is the large-scale residential development projects planned in the waterfront area, which will lead to increases in population in areas where public transit service levels are low. Vehicle-kilometers traveled within the City of Osaka, however, do not increase proportionally with the number of auto trips (Fig. 8). In fact, only Package 2, which is auto-oriented, yields a vehicle-kilometer total that is larger than that of Package 0 (do-nothing alternative). Reasons for this are difficult to pinpoint with the results so far tabulated, but it may be the case that dispersed residential locations tend to produce either shorter trips, or more trips that are made outside the city boundaries and therefore are not included in the tabulation here. It is also conceivable that the various measures implemented in the respective packages tend to shorten the length of auto trips.

Fig. 9 presents estimated CO_2 emissions within the City of Osaka for the respective policy packages. From Figs. 8 and 9, it can be seen that vehicle travel can be reduced by implementing the planned infrastructure projects (Package 1) and by developing public transit and adopting measures to suppress auto use (Packages 3 and 4). Only auto-oriented Package 2 produces more CO_2 emissions than does Package 0. The results make evident that investing in road facilities promotes more auto use. Interesting is the result that Package 4, which involves the development of a circumferential road, has more vehicle-kilometers traveled than Package 3, which involves only transit development; yet the former package results in less CO_2 emissions than the latter because of the implementation of congestion pricing. The effects of auto restriction measures in Package 4 can also be seen in Fig. 10 which shows vehicle-kilometers traveled within the CBD area; Package 4 has CBD vehicle-kilometers that are about 20% less than those of Package 0, and about 14% less than those of Package 3.



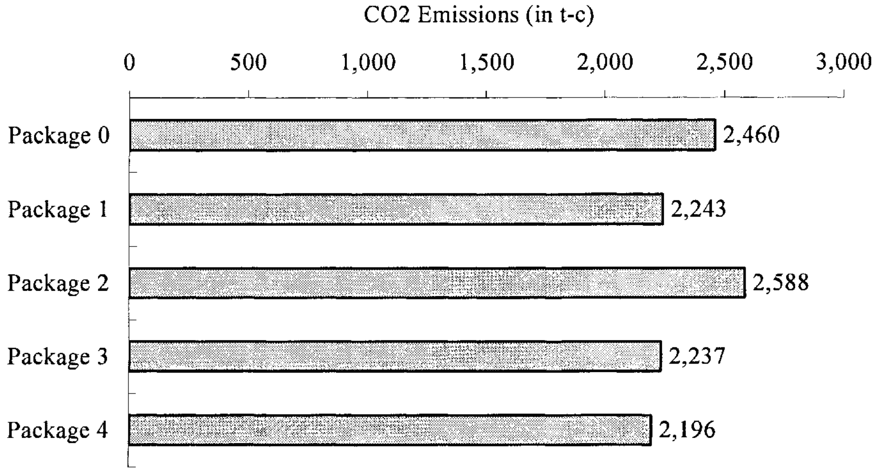
Source: Kawata et al. (1999)

Figure 7. Auto Trips Generated in the City of Osaka.



Source: Kawata et al. (1999)

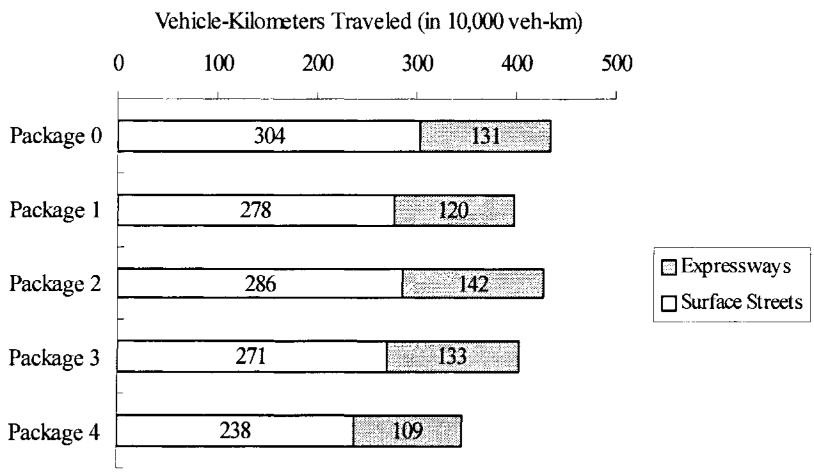
Figure 8. Vehicle-Kilometers Traveled within the City of Osaka.



Source: Kawata et al. (1999)

Figure 9. CO_2 Emissions by Vehicles within the City of Osaka.

The infrastructure investment and transportation control measures in these packages tend to improve traffic flow. The mean auto travel speed increases by about 11% with Package 1 and Package 3, and by 21% with Package 4 which includes auto restriction measures in the CBD. The most auto-oriented Package 2 produces the smallest improvement in travel speed of 9%.



Source: Kawata et al. (1999)

Figure 10. Vehicle-Kilometers Traveled in Osaka CBD.

CONCLUSION

Application results of PCATS and DEBNetS, along with HAGS, are starting to demonstrate that micro-simulation is a practical approach for demand forecasting and policy analysis, especially in the area of demand management. The micro-simulation model system is by no means data hungry; it is based on data that are typically maintained by planning agencies and have been used in conventional travel demand forecasting. Because the time axis is explicitly incorporated into the model system and because it represents individuals' behavior in space and time over a one-day period, the system simulates travel demand in a more coherent manner and applies to a wide range of planning measures. HAGS makes this tool applicable to long-range forecasting using disaggregate models. The rich data the simulation system offers as output can be analyzed to produce information on travel demand, traffic condition, or pollutant emissions.

The application examples of the micro-simulation system presented in this paper are one of the earliest attempts in which all steps of the four-step procedures are performed by micro-simulation. As such, the model system is yet to be refined. This paper has described some of the refinements being undertaken to improve the PCATS/DEBNetS system. It is hoped these efforts will make the micro-simulation system a more practical, accurate and versatile tool for urban travel demand forecasting and policy analysis. In addition, incorporating behavioral models of route choice with real-time information is being planned to make DEBNetS a versatile tool for network traffic management.

REFERENCES

Arentze, T., A. Borgers, F. Hofman, S. Fujii, C. Joh, A. Kikuchi, R. Kitamura, H. Timmermans and P. van der Waerden (2001). Rule-based versus utility-maximizing models of activity-travel patterns: A comparison of empirical performance, In D. Hensher (ed.) *Travel Behaviour Research: The Leading Edge*, Elsevier Science, Oxford, pp.569-583.

Beckman, R.J., K.A. Baggerly and M.D. McKay (1996). Creating synthetic baseline populations, *Transportation Research A*, **30A**, 415-429.

Chiang, J., S. Chib and C. Narasimhan (1999). Markov Chain Monte Carlo and models of consideration set and parameter heterogeneity, *Journal of Econometrics*, **89**, 223-248.

Damm. D. (1982). Parameters of activity behavior for use in travel analysis, *Transportation Research*, **16A**(2), 135-148.

Hägerstrand, T. (1970). What about people in regional science?, *Papers of the Regional Science Association*, **24**, 7-21.

Hajivassiliou, V., D. McFadden and P. Ruud (1996). Simulation of multivariate normal rectangle probabilities and their derivatives: theoretical and computational results, *Journal of Econometrics*, **72**, 85-134.

Hazelton, M.L., S. Lee and J.W. Polak (1996). Stationary states in stochastic process models of traffic assignment: a Markov Chain Monte Carlo approach, In J.-B. Lesort (ed.) *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, Pergamon, Oxford, pp. 341-357.

Fujii, S., A. Kikuchi and R. Kitamura (2000). A micro-simulation analysis of the effects of transportation control measures to reduce CO_2 emissions: a case study in Kyoto City, *Traffic Engineering*, **35**(4), 11-18 (in Japanese).

Fujii, S., M. Okushima, A. Kikuchi and R. Kitamura (1998). Development of a network flow simulator and evaluation of travel time, *In the Proceedings of the 17th Annual Meeting of the Japanese Society of Traffic Engineers*, pp. 694-695 (in Japanese).

Fujii, S., Y. Otsuka, R. Kitamura and T. Monma (1997). A micro-simulation model system of individuals' daily activity behavior that incorporates spatial, temporal and coupling constraints, *Infrastructure Planning Review*, **14**, 643-652 (in Japanese).

Iida, Y., M. Iwabe, A. Kikuchi, R. Kitamura, K. Sakai, Y. Shiromizu, D. Nakagawa, M. Hatoko, S. Fujii, T. Morikawa and T. Yamamoto (2000). Micro-simulation based travel demand forecasting system for urban transportation planning, *Infrastructure Planning Review*, **17**, 841-847 (in Japanese).

Kawata, H., Y. Iida and Y. Shiromizu (1999). Case study of evaluation for comprehensive transportation policy, *The Proceedings of the Infrastructure Planning Review Annual Meeting*, **22**(1), 511-514 (in Japanese).

Kikuchi, A., S. Fujii and R. Kitamura (2001). Evaluation of transportation policies by micro-simulation of individuals' behaviors on a coordinates system, *City Planning Review*, **36**, 577-582 (in Japanese).

Kikuchi, A., S. Fujii, Y. Shiromizu and R. Kitamura (2000). Calibration of DEBNetS on a large-scale network, In *the proceedings of the 20th Annual Meeting of the Japanese Society of Traffic Engineers*, Tokyo, pp. 49-52 (in Japanese).

Kikuchi, A, Y. Kato, T. Macuchi, S. Fujii and R. Kitamura (2002).

Improvement and verification of dynamic traffic flow simulator "DEBNetS", *Infrastructure Planning Review*, **19** (in press, in Japanese).

Kikuchi, A., A. Kobata, S. Fujii and R. Kitamura (2000). A mode and destination choice model on a GIS database: from zone-based toward coordinates-based methodologies of spatial representation, *Infrastructure Planning Review*, **17**, 841-847 (in Japanese).

Kikuchi, A., T. Yamamoto, K. Ashikawa and R. Kitamura (2001). Computation of destination choice probabilities under huge choice sets: application of Markov Chain Monte Carlo algorithms, *Infrastructure Planning Review*, **18**(4), 503-508 (in Japanese).

Kikuchi, A., R. Kitamura, S. Kurauchi, K. Sasaki, T. Hanai, T. Morikawa,S. Fujii and T. Yamamoto (1999). Effect Analysis of Transportation Policies using Micro-Simulation Method - A Case Study of Toyota City -, *The Proceedings of the Infrastructure Planning Review Annual Meeting*, **22**(1), 817-820 (in Japanese).

Kitamura, R., C. Chen and R. Narayanan (1998). The effects of time of day, activity duration and home location on travelers' destination choice behavior, *Transportation Research Record*, **1645**, 76-81.

Kitamura, R. and S. Fujii (1998). Two computational process models of activity-travel behavior, In T. Gärling, T. Laitila and K. Westin (eds.) *Theoretical Foundations of Travel Choice Modelling*, Pergamon Press, Oxford, pp. 251-279.

Kitamura, R., S. Fujii, A. Kikuchi and T. Yamamoto (1998). Can TDM make urban transportation "sustainable"?: A micro-simulation study, Paper presented at *International Symposium on Travel Demand Management*, Newcastle, UK.

Kitamura, R., S. Fujii, T. Yamamoto and A. Kikuchi (2000). Application of PCATS/DEBNetS to regional planning and policy analysis: Micro-simulation studies for the Cities of Osaka and Kyoto, Japan, In *the Proceedings of Seminar F, European Transport Conference 2000*, pp. 199-210.

Kitamura, R., T. Yamamoto, K. Kishizawa and R.M. Pendyala (2000). Stochastic frontier models of prism vertices, *Transportation Research Record*, **1718**, 18-26.

Kitamura, R., T. Yamamoto, K. Kishizawa and R.M. Pendyala (2001). Prism-based accessibility measures and activity engagement, Paper presented at the *80th Annual Meeting of the Transportation Research Board*, Washington, D.C., January.

Nishida, S., T. Yamamoto, S. Fujii and R. Kitamura (2000). A household attributes generation system for long-range travel demand forecasting with disaggregate models, *Infrastructure Planning Review*, **17**, 779-787 (in Japanese).

Pendyala, R.M., T. Yamamoto and R. Kitamura (2002). On the formation of time-space prisms to model constraints on personal activity-travel engagement, *Transportation*, **29**(1), 73-94.

Schmidt, P. and A. Witte (1989). Predicting criminal recidivism using split population survival time models, *Journal of Econometrics*, **40**, 141-159.

Yamamoto, Y., R. Kitamura and R.M. Pendyala (2002). Comparative analysis of time-space prism vertices for out-of-home activity engagement on working days and non-working days, Submitted to *Geographical Analysis*.

Yamamoto, T., R. Kitamura and K. Kishizawa (2001). Sampling alternatives from a colossal choice set: an application of the MCMC algorithm, *Transportation Research Record*, **1752**, 53-61.