Traffic Forecasting in the Helsinki Metropolitan Area Transportation Study 1988

MATTI PURSULA AND HEIKKI KANNER

The models developed for traffic forecasting in the Helsinki Metropolitan Area, Finland, are presented. The basic traffic surveys for the Helsinki Metropolitan Area Transportation Study 1988 were done during 1987 and 1988. The main field studies were an origin-destination (OD) survey of automobile traffic, an OD survey of public transport, and a personal trip diary interview. Modeling of only internal trips made by inhabitants of the metropolitan area is described. The modeling was based on the trip diary interview. The model structure was basically a four-step model with feedback between the last three steps. Trip generation was calculated using production and attraction rates. Mode and destination choice were mostly modeled using nested logit models. In network loading, a standard multipath equilibrium assignment model (EMME/2 system) was used. Trips were divided into four categories according to trip purpose (home-based work trips, home-based school trips, other home-based trips, and non-home-based trips). Four alternative modes were included in the mode choice models. The population was divided into categories according to the different steps of the modeling. The most important categorization was the division according to a person's access to a car. Destination choice models included 117 alternative destinations. The coefficients of mode choice models were logical and the variables predictable. The first destination choice models had some theoretical deficiencies, which were partly abolished using constrained estimation. The model system with unconstrained models produced satisfactory forecasts. New models will be estimated and new forecasts will be made during 1992.

The study area, the Helsinki metropolitan area, consists of four cities: Helsinki (485,000 inhabitants), Espoo (165,000 inhabitants), Vantaa (150,000 inhabitants), and Kauniainen (8,000 inhabitants). The city center of Helsinki is located on a peninsula in the Gulf of Finland, and the metropolitan area forms a half circle around it with a radius of 25 to 30 km (total area 1031 km^2, land area 742 km^2). In the city center there are about 118,000 workplaces and 59,000 inhabitants. The number of jobs in the whole metropolitan area is about 442,000.

The car density in the area is about 320 cars per 1,000 inhabitants, and 60 percent of all households have at least one car. The public transport system of the area consists of bus and tram traffic, commuter and ordinary trains, and one subway line east of the city center. Of the 2 million internal daily trips of the inhabitants of the area, 46 percent are made by car, 32 percent by public transport, and 22 percent by bicycle or on foot.

M. Pursula, Laboratory of Transportation Engineering, Helsinki University of Technology, Rakentajanaukio 4 A, SF-02150, Espoo, Finland. H. Kanner, Viatek Ltd., Pohjantie 3, SF-02100, Espoo, Finland.

Buses dominate in public transport. Three-fourths of all public transport trips are made by bus. The share of public transport has been falling continuously during the last 20 years as a result of growing car density.

BASIC TRAFFIC SURVEYS

The basic traffic surveys for the Helsinki Metropolitan Area Transportation Study 1988 (LITU 88) were done during 1987 and 1988. The main field studies were an origin-destination (OD) survey of automobile traffic (*1*), an OD survey of public transport (*2*), and a personal trip diary interview (*3*).

The OD survey of automobile traffic was done at 122 sites on 10 cordon lines that divided the study area into 15 subareas. At each site a mail-back questionnaire was given to 10 to 20 percent of the drivers of the passing automobiles (except buses) between 6 a.m. and 8 p.m. The information asked included OD data, trip purpose, and number of passengers in the car. The drivers were also asked to draw the route of their trip on a map included in the questionnaire. The number of returned and approved questionnaires was about 86,000 (37 percent).

The OD survey of public transport was done by interviewing every fourth boarding passenger in every fourth bus, tram, and train (i.e., 6.25 percent of passengers). The questionnaire was short, and most of the passengers completed it during their trip and returned it directly to the interviewer. Only a small number were returned by post. The most important information in the questionnaire was OD data and information about the number of transfers needed during the trip in question. The number of accepted returned questionnaires was about 56,000, which is 5.6 percent of the 1 million boardings per day in the metropolitan area.

The trip diary interview was person based. The main reasons for the use of this technique were the easiness in sample formation and expansion and the good experience in some recent travel surveys in Finland (*4*).

The objective of the interview was to gather daily travel data plus socioeconomic and other background information from 7,000 inhabitants of the metropolitan area. Only persons 7 years of age or older were included in the original random sample, which was about 2.5 percent (18,000 persons) of the corresponding population. About 66 percent of these had a telephone.

The data were gathered by an informed telephone interview. This means that the questionnaire plus travel diary were sent in advance to people with a telephone (11,900), and after the survey date the person in question was phoned and the data were typed directly into computer memory. The number of accepted telephone interviews was about 6,100, and about 600 accepted answers were collected by mail to represent those persons who had no telephone or could not be reached by telephone.

The trip diary interview study formed the basis for the traffic modeling. In this work the other studies were used mainly for comparison and validation.

In all these surveys three basic zone systems were used. The data were gathered in a very detailed zone division with 282 zones. The models and forecasts were made in a division with 117 zones. For aggregation purposes a division with 19 zones was mainly used.

TRAFFIC MODEL SYSTEM

In this paper only the modeling of the internal trips of the inhabitants of the metropolitan area is described (*5*). These trips are about 90 percent of all person trips in the area, and the modeling is based on the trip diary interview. The modeling of external trips and commercial traffic was done with simple methods and is not discussed here.

The model structure is basically a four-step model with feedback between the last three steps. Trip generation is calculated using production and attraction rates. Mode and destination choice are mostly modeled using nested logit models. In network loading a standard multipath equilibrium assignment model (EMME/2) is used.

Trips are divided into four categories according to their purpose: home-based work trips (33 percent), home-based school trips (11 percent), other home-based trips (42 percent), and non-home-based trips (14 percent). In mode choice modeling there are four alternative modes: walk (and bicycle), car (driver or passenger), bus (and street car), and rail (train and subway). The access trips to rail are made on foot or by bus. Less than 2 percent of rail passengers use park-and-ride, and this mode is not included in the model.

The share of non-home-based trips in the data is low. There are two main reasons for this. The first is that, in the trip diary survey, short (less than 5 min travel time) non-home-based pedestrian and bicycle trips were purposely left out. The second is that non-home-based pedestrian and bicycle trips are not included in the models. The original share of non-home-based trips in the data is 19 percent. This is the same as in the corresponding survey in 1976 with no exclusion of short trips (*6*). In the latest nationwide study, the share of non-home-based trips is about 25 percent (*4*), and in a recent study in Oulu, a middle-sized Finnish city, it is about 33 percent (*7*).

The population is divided into different categories in different steps of modeling. The most important categorization is the division according to a person's access to a car. Persons with a driving license who, according to their own statements, practically always have access to a car for traveling belong to the category HAP (a Finnish acronym for persons who mainly use cars for traveling). Other people belong to the category EHAP (person who usually cannot use car for traveling). About 46 percent of people aged 18 years or more belong to the category HAP. This is clearly more than one person per car registered in the area.

Table 1 gives the main structure of the modeling system. The HAP/EHAP grouping is used in two trip-purpose categories: other home-based trips and non-home-based trips. Some minor categorizations of the population that are used during the modeling are not given in the table. The most important of these are the age grouping and the grouping of the people as working or nonworking (see, for example, Table 2).

TRIP GENERATION MODELS

The production and attraction rates are based on cross-classification analysis of the survey data. The division of the population into detailed categories is hard to forecast, so the trip production rates used in the traffic forecasts and given in Table 2 are based on very few categories.

The generation-attraction principle cannot be applied to non-home-based trips. The production rates given for this trip category in Table 2 were used only to check the total number of non-home-based trips in the metropolitan area. The actual trip locations for non-home-based trips were based on the number of inhabitants and jobs in the zones. The same principle was used for all the attraction rates calculated in the study.

The home-based school trips of persons aged 18 years or more are included in other home-based trips, and they are about 8.5 percent of all trips in this category. In this way the school trip category becomes homogeneous. For example, more than 75 percent of these trips are made on foot or by bicycle.

No trip matrix balancing was done in the study. The attraction rates were used mainly for validation and evaluation for the forecasts in a later phase of the process. They were also used in the calculation of the size variable in some of the destination choice models.

LOGIT MODELS FOR MODE AND DESTINATION CHOICE

Estimation of the Models

Nested logit models were used for mode and destination choice estimation for other home-based and non-home-based trips. The mode choice of home-based work trips was estimated with a logit model as well as the destination choice of the home-based school trips.

The destination choice of home-based work trips was modeled using a housing and workplace matrix from the population censuses. This matrix gives the location of the home and workplace of every working person in the area. The work trips of the present situation were simply distributed using this matrix. The matrix was transformed for future situations with growth factors based on a gravity model analogy. The approximation method is rough and is not discussed further here.

The mode choice of home-based school trips was based on the length of the trip. This could be done because walking

TABLE 1 Structure of the Modeling System of Internal Trips in LITU 88

Trip purpose	Trip generation rates: Production	Trip generation rates: Attraction	Mode choice	Destination choice
Home-based work trips	Trips/working person/day	Trips/employee/day	Logit model	Housing and working place matrix
Home-based school trips	Trips/school-aged person/day	Trips/inhabitant/day	Distance matrix	Logit model
Other home-based trips, HAP-persons*	Trips/HAP-person/day	Trips/inhabitant/day Trips/employee/day	Logit model	Nested logit model
Other home-based trips, EHAP-persons*	Trips/EHAP-person/day	Trips/inhabitant/day, Trips/employee/day	Logit model	Nested logit model
Non-home-based trips**, HAP-persons	Trips/HAP-person/day	Trips/inhabitant/day, Trips/employee/day	Logit model	Nested logit model
Non-home-based trips**, EHAP-persons	Trips/EHAP-person/day	Trips/inhabitant/day, Trips/employee/day	Logit model	Nested logit model

* HAP-person is a person that practically always has access to car for personal trips. Other persons are EHAP-persons.

** Walk (and bicycle) trips are excluded from this trip category.

and bicycling are so dominant in this trip category without direct access to a car. The mode and destination choice of school trips is not discussed further in the paper.

In other home-based and non-home-based trip categories, the nested logit models were estimated separately for HAP and EHAP persons. A stepwise estimation procedure was used (i.e., the mode choice model was estimated first, and the logsum term of that model was used as a variable in the estimation of the destination choice model). Simultaneous estimation was not possible in the original model work because of computer program restrictions.

The estimation of the destination choice models was done using all 117 zones of the metropolitan area as alternatives. At first, estimation using samples with 31 zones per alternative was tried just as in the corresponding study in 1976 (*6*). The results were poor, so the full choice set was used. The reason for the poor results with destination sampling was probably a bias in the methodology. The models were estimated according to the rules of random sampling even though the strategy used was basically a stratified importance sampling described, for example, by Ben-Akiva and Lerman (*8*). Random sampling strategy and the stability of the coefficients will be studied during 1992.

For each trip category where logit models were used for mode choice, two sets of models were estimated. The base models are detailed and include a larger number of variables

TABLE 2 Trip Production Rates (Trips/Person/Day) Used in Traffic Forecasts in LITU 88

Trip purpose	Population group	Trip production rate
Home-based work trips	Working persons, age 18-64 years	1.43
	Other population	0.03
Home-based school trips*	Age 7-17 years	1.77
	Other population	0.00
Other home-based trips**	Age 7-17 years	1.20
	EHAP-persons, age 18-	1.10
	HAP-persons	1.46
Nonhome-based trips***	Age 7-17 years	0.12
	EHAP-persons, age 18-	0.27
	HAP-persons	0.67

* The school trips of persons aged 18 years or more belong to the trip category other home-based trips.

** HAP-person is a person that practically always has access to a car for personal trips. Others are EHAP-persons.

*** Walk (and bicycle) trips are excluded from this trip category.

than the forecasting models that need less data and are used in the aggregate travel forecasts. The focus of this paper is in the forecasting models. Only one example of the base models is given here (Table 3).

Variables and Coefficients of Mode Choice Models

Table 3 gives the variables and coefficients of the mode choice models for home-based work trips. These models were estimated without categorization by access to a car (HAP/EHAP grouping).

The distance for walk trips is given in kilometers. This distance is for a one-way trip between home and work. The travel times, travel costs, and number of transfers in motorized traffic are calculated for a round-trip. Travel times are given in minutes, costs in FIM ($1 U.S. is about 4.3 FIM), and household income in thousands of FIM per month.

The total travel time includes walking, waiting, and in-vehicle times. These are calculated from the traffic networks with the EMME/2 program. For rail users a special procedure is used to give all these components.

Travel costs for public transport are based on the ticket type of the passenger. Travel costs for car (0.46 FIM/km) are out-of-pocket costs (*9*). For home-based work trips the travel costs of a car do not include parking costs. They are calculated in the parking index, which is a linear combination of parking costs and the logarithmic ratio of parking demand to parking capacity.

The base model in Table 3 has more variables than the forecasting model. For example, the possibility of a personally addressed parking place at the workplace and the possibility of a company car are included in the model. A special variable to describe the length of the access to rail stations is also used.

The model coefficients are logical, and the variables in the forecasting model are predictable. If an assumption of constant cost/income relationship is made, then no forecast is needed for these variables.

The value of travel time calculated from the forecasting model in Table 3 with a mean household income of 12,000 FIM/month is about 6.90 FIM/hr. This is about half of the price that was used in cost-benefit analyses by the Finnish National Road Administration in 1988 (*9*). From the base model the value of travel time for components of the work

TABLE 3 Variables and Coefficients of the Logit Models for Mode Choice of Home-Based Work Trips in LITU 88

Variable	Forecasting model Coefficient	t	Base model Coefficient	t	Modes
Ln(distance)	3.601	36.3			Walk
Total travel time	-0.02154	-9.0			Bus, car, rail
Trip cost/income	-2.236	-29.9			Bus, car, rail
Number of transfers	-0.5170	-14.7			Bus, rail
Cars/household	0.7896	5.1	-1.101	-5.6	Car/Bus,rail*
Parking index	-1.010	-4.9	-1.000	-5.7	Car
Walk-dummy	1.873	13.8	-0.7411	-3.9	Walk
Rail-dummy	0.3200	4.9	0.7837	4.7	Rail
Car-dummy	-1.420	-9.6	-3.300	-18.5	Car
Distance 0-10 km **			-0.8838	-28.2	Walk
Distance 10- km **			-0.1639	-4.8	Walk
Access walk time			-0.03476	-7.7	Bus, car, rail
In-vehicle time			-0.01338	-3.9	Bus, car, rail
Cost			-0.2468	-29.4	Bus, car, rail
Number of transfers bus			-0.6980	-12.9	Bus
Number of transfers rail			-0.4229	-9.9	Rail
Sex (female=0, male=1)			1.615	18.8	Car
Waiting time			-0.04733	-6.2	Bus, rail
Access time/road distance			-0.2781	-5.7	Rail
Reserved parking			0.9888	8.3	Car
Company car			1.293	11.4	Car
Number of observations = 4780	ρ^2 = 0.360		n = 4780	ρ^2 = 0.471	
Percent of correct predictions	66.6			72.8	

* The variable cars/household is used for car mode in the forecasting model and for bus and rail modes in the base model.

** If the distance <= 10 km then 'Distance 0-10' equals the distance and 'Distance 10 -' equals zero . If the distance > 10 km then 'Distance 0-10' equals 10 and 'Distance 10-' equals (distance-10).

trip can also be calculated: walk time 8.45 FIM/hr, waiting time 11.50 FIM/hr, and in-vehicle time 3.25 FIM/hr. The ratio of the component values is approximately 2.5:3.5:1. This is in reasonably good accordance with the international findings of traffic model studies (*10*).

Table 4 gives the models estimated for the HAP and EHAP populations for other home-based trips. In this trip category the model structure and most of the variables are the same as in the model for home-based work trips. The values of the variables are for round-trips except walking distance, which is given as a sum of one or two one-way distance variables in the same way as in the base model for home-based work trips. The parking costs are this time included in the cost/income variable. The parking conditions of the destination zone are

TABLE 4 Variables and Coefficients of the Logit Models for Mode Choice of Other Home-Based Trips in LITU 88

	HAP-population *		EHAP-population *		
Variable	Coefficient	t	Coefficient	t	Modes
Distance, 0-5 km **	-0.7838	-13.0			Walk
Distance, 5-10 km **	-0.5489	-7.6			Walk
Distance, 0-10 km ***			-0.7599	-29.6	Walk
Total travel time	-0.01943	-4.0	-0.01643	-6.2	Bus, car, rail
Cost/income	-0.7593	-8.4			Bus, car, rail
Number of transfers	-0.1875	-2.4	-0.4205	-10.0	Bus, rail
Ln(parking ratio-5)	-0.4376	-9.5	-0.3885	-9.2	Car
Acc. time/road dist.	-0.4418	-3.8			Rail
Walk-dummy	1.961	8.1	1.950	17.7	Walk
Rail-dummy	0.7967	2.8	-0.01898	-0.2	Rail
Car-dummy	1.660	9.8	-1.253	-12.7	Car
Number of observations = 2542		ρ^2 = 0.515	n = 4050	ρ^2 = 0.267	
Percent of correct predictions		77.0		60.8	

* HAP-person is a person that practically always has access to a car for personal trips. Others are EHAP-persons.

** If the distance <= 5 km then 'Distance 0-5' equals the distance and 'Distance 5-10' equals zero . If the distance > 5 km and <=10 km then 'Distance 0-5' equals 5 and 'Distance 5-10' equals (distance-5). If the distance > 10 km 'Distance 0-5' equals 5 and 'Distance 5-10' equals 5.

*** If the distance <= 10 km then 'Distance 0-10' equals the distance. If the distance > 10 km then 'Distance 0-10' equals 10.

described with the logarithmic parking ratio variable (demand/capacity).

The access time per road distance variable is calculated as the sum of the access times to and from rail stations divided by the direct road distance between origin and destination of the trip. The variable is a measure of difficulty in the use of a low-density rail network.

The value of travel time can be calculated only for main users of cars (the HAP population) because travel cost is not included in the EHAP models. The value of the time for 12,000 FIM/month income is 18.40 FIM/hr. This is clearly higher than the value for the whole population for home-based work trips.

Table 5 gives the coefficients of the mode choice models for non-home-based trips. Walk and bicycle trips are not included in this trip category, and one-way trips are used in the models. The parking demand variable, though, is a sum of the corresponding variables of both ends of the trip. In this way, parking costs and constraints of both ends of the trip affect the mode choice.

The value of travel time for non-home-based trips can be calculated for both population categories. The value for the HAP population (11.90 FIM/hr) is lower than that for the EHAP population (26.10 FIM/hr). The reason for this surprising result might be the exclusion of the walk trips from the estimation sample. Another reason might be in the differences of the purpose distribution of the non-home-based trips of the HAP and EHAP populations. HAP people clearly make more non-home-based trips than EHAP people, so it is possible that a bigger share of their trips are leisure trips (with a lower time value) than is the case for EHAP people.

The use of rail as a separate mode caused some problems in model estimation. In most models, the share of correctly predicted choices was lowest for rail. The problem was that

TABLE 5 Variables and Coefficients of the Logit Model for Mode Choice of Non-Home-Based Trips in LITU 88

	HAP-population *		EHAP-population *		
Variable	Coefficient	t	Coefficient	t	Modes
Total travel time	-0.05479	-3.2	-0.06058	-4.4	Bus, car, rail
Cost/income	-3.307	-8.0	-1.668	-5.1	Bus, car, rail
Number of transfers	-0.5439	-2.0	-0.2826	-1.4	Bus, rail
ΣLn(parking ratio-5)	-0.2102	-3.1	-0.2113	-3.3	Car
Rail-dummy	-0.3284	-1.4	-0.4938	-3.2	Rail
Car-dummy	1.767	5.1	-1.072	-4.0	Car
Number of observations = 1284		ρ^2 = 0.671	n = 652	ρ^2 = 0.199	
Percent correct predictions		89.1		40.9	

* HAP-person is a person that practically always has access to a car for personal trips. Others are EHAP-persons.

there were no good variables to differentiate between bus and rail. Some attempts were made with nested mode choice models, but the results were not any better. Because of the important role of some new rail construction proposals in the future, rail was kept as a separate mode in the models.

The ρ^2 values of the models are sometimes very high. This is especially the case in models for the HAP population. The reason for this is that the modal split in this population group is very one-sided: three-fourths or more of trips are made by car. A better indicator for the goodness-of-fit of the models would be a revised ρ^2 value that indicates the result in comparison with a model with alternative specific constants only. Unfortunately, the estimation program did not give these values, and they were not calculated afterwards.

Some estimates of model accuracy were done by sample enumeration. The results indicated that the estimates of car and bus shares are mostly quite good (average error mostly only a few percent). The results for rail are clearly worse, especially for HAP persons. This may be an implication of an IIA violation in the model structure for this population category. In practice the difference is not so serious, because HAP persons seldom use rail or any public transport alternative.

Variables and Coefficients of Destination Choice Models

Destination choice models for other home-based and non-home-based trips were estimated separately for HAP and EHAP populations. The models are nested logit models with feedback to mode choice models via the logsum variable. According to theory, the coefficient of the logsum variable should be greater than or equal to zero and less than or equal to one (*8*). If the coefficient is greater than one, some of the cross elasticities in the model can be illogical (*10*).

In some estimating programs it is possible to restrict the coefficient of the logsum variable to 1, but this was not in the program that was originally used during this work. Unfortunately, this led to coefficients that were clearly more than 1, as can be seen from the following tables where the model coefficients are given. After this occurred, a new program with the ability to restrict estimated parameters was used (*11*). In the following, results of unrestricted and restricted estimation are given, but so far no forecasts with the restricted models have been made.

In destination choice models, the attraction is described with a scale variable. Usually this variable contains a linear combination of the amount of different activities in the zone. The scale variable should be in logarithmic form, and its coefficient should be equal to 1 to give a model that is independent of the zone division (*8,10*).

The original estimation of the models of LITU 88 was done without restriction the coefficient of the scale variable to equal 1. Most coefficients turned out to be reasonable (i.e., they did not differ too much from one and were all less than 1). Trials of restricted estimation indicated only minor changes in other coefficients of the models, so these results are not given here.

The weighting of different activities inside the scale variable can also be problematic. The original estimation program that was used in this work could not estimate these weights simultaneously with other model coefficients. For this reason the weights were found partly by trial-and-error and partly by the use of usual attraction rates. With the new estimation program, these weights can be directly estimated. Some of the results are given here. The results indicate that the other coefficients of the model are not very sensitive to the weights.

With the new estimation program it is also possible to estimate the nested logit model for mode and destination choice simultaneously. This estimation makes better use of the data, and a new estimation of the models is going on. In this paper all models are based on sequential estimation of mode and destination choice.

Table 6 gives the results of model estimation for other home-based trips. The corresponding results for non-home-based trips are given in Table 7.

All the destination choice models given above have one to three alternative specific dummy variables. The city dummy is used if the destination of the trip is inside the city center area. Subcenters 1 and 2 refer to the main centers of Espoo and Vantaa. Subcenter 3 is the biggest subcenter inside the borders of Helsinki.

The models estimated with the new program without restrictions included the estimation of the weights in the scale variable, too. The exact results are not given here, but the values of the weights did not differ much from the corresponding values of the restricted estimation given in Tables 6 and 7. The model coefficients were also very near the coefficients of the free estimation of the models. This is natural because the coefficient of the scale variable in free estimation did not differ so much from 1.

The weighted values of the scale variable are clearly of different magnitude in the free trial-and-error estimation and in the new restricted estimation even though the coefficients of the logscale variables are very near each other. The explanation for this is in the nature of logarithmic variables. If a logarithmic variable in a logit model is common to all alternatives, a multiplication or division by constant inside the logarithm has no effect on the coefficient of this or any other variable in the model.

As was mentioned before, all the destination choice models were estimated using all 117 zones as possible choices. The great amount of possible choices results in low ρ^2 values. The basic test used for the goodness-of-fit of the destination choice models was their ability to replicate the trip length distributions of the travel data. An example is given in Figure 1.

The fit of the computed trip length distribution in Figure 1 is far from excellent (observed mean 7.78 km; forecast mean 6.24 km), and the model has to be developed further. It is possible that the simultaneous restricted estimation process that is going on at present will give better results.

TABLE 6 Variables and Coefficients of the Logit Model for Destination Choice of Other Home-Based Trips in LITU 88

Variable	HAP-population * Coefficient (free) **	t	Coefficient (restricted)	t	EHAP-population * Coefficient (free)	t	Coefficient (restricted)	t
Logsum	1.770	56.5	1.000	0.0	1.830	99.6	1.000	0.0
Logscale	0.7416	24.3	0.7992	21.7	0.9554	29.7	0.9158	34.3
Employment density	0.2629	8.3	0.1787	5.4	0.3022	22.3	0.1058	4.5
City-dummy	0.5908	6.5	0.2805	3.2				
Subcenters1&2-dummy	0.7699	9.0	0.9622	11.2				
Subcenter3-dummy	1.457	12.6	1.501	11.9	1.945	21.4	1.435	15.5
	ρ^2 = 0.189		ρ^2 = 0.180		ρ^2 = 0.338		ρ^2 = 0.288	
Variables in the logscale-variable								
Inhabitants	1.75	-	1.000	0.0	0.55	-	1.000	0.0
Retail employment	52.75	-	2.306	9.7	0.40	-	2.248	19.0
Service employment	24.00	-	0.8768	4.2	0.55	-	-	-
Industrial employment	1.00	-	-	-	0.10	-	-	-
Other jobs	15.50	-	0.2773	0.6	0.10	-	-	-

* HAP-person is a person that practically always has access to a car for personal trips. Others are EHAP-persons.

** Free : free estimation of the coefficient of the logsum variable

Restricted : the coefficient of the logsum variable was restricted equal 1.0

TABLE 7 Variables and Coefficients of the Logit Model for Destination Choice of Non-Home-Based Trips in LITU 88

	HAP-population *				EHAP-population *			
Variable	Coefficient (free) **	t	Coefficient (restricted)	t	Coefficient (free)	t	Coefficient (restricted)	t
Logsum	1.380	36.2	1.000	0.0	1.656	23.3	1.000	0.0
Logscale	0.7578	21.2	0.8593	16.3	0.7289	13.9	1.007	12.4
City-dummy	0.7717	8.8	0.5144	5.8	0.2716	2.5	0.1670	1.4
Subcenter1&2-dummy	0.4343	3.3	0.5144	3.8	0.7803	4.2	0.7307	3.9
Subcenter3-dummy	1.104	6.4	0.9700	5.4	1.328	5.4	1.374	5.2
	$\rho^2 = 0.216$		$\rho^2 = 0.208$		$\rho^2 = 0.256$		$\rho^2 = 0.246$	
Variables in the logscale-variable								
Inhabitants	1.00	-	1.000	0.0	1.00	-	1.000	0.0
Retail employment	57.00	-	3.356	8.2	43.00	-	2.554	6.4
Service employment	32.00	-	1.495	4.2	24.00	-	1.156	3.5
Industrial employment	8.50	-	1.630	5.3	5.00	-	-	-
Other jobs	20.50	-	1.317	2.5	18.00	-	0.5949	1.0

* HAP-person is a person that practically always has access to a car for personal trips. Others are EHAP-persons.

** Free : free estimation of the coefficient of the logsum variable

Restricted : the coefficient of the logsum variable was restricted equal 1.0

Fortunately, the case given in Figure 1 is the worst one. For example, the mean value of the observed trip lengths for home-based work trips is 9.96 km, and the forecast value is 9.62 km. For other home-based trips the values are 7.02 km (observed) and 7.28 km (forecast). For non-home-based trips the values are 7.78 and 6.24 km. This indicates that the models for non-home-based trips have to be developed further.

OTHER MODELS

The forecasting process used in LITU 88 needs some additional traffic models. The most interesting of these are the car ownership models and the model for HAP/EHAP division of the population. For both purposes logit models were used.

Table 8 gives an example of car ownership models. These models are, unlike the other logit models of the study, household based. According to the model, the size of the household has a strong impact on car ownership as well as the type of housing.

In Table 9 there is an example of the HAP/EHAP model. A better model can be estimated if the sex of the person is included. However, the model in Table 9 was used for forecasting because it was assumed that the difference in access to cars between males and females will diminish in the future.

DISCUSSION OF RESULTS

The model system was used to produce basic traffic forecasts for the present situation (as a part of the model validation) and for two future situations. The basic forecasts give the daily trip matrixes divided by trip purpose and mode. The forecasting was done separately for morning and evening peak periods and for the rest of the day. This way, the differences in trip purpose and destination choice during different times of day were taken into account.

The forecasting was based on aggregate data. The zonal means of the variables included in the models were used for all individuals of the zone. This of course is a source of aggregation error, and this must be kept in mind when total results are referred to in the following discussion.

Comparison of the forecast for the present situation with the results of the traffic studies is the final test of the modeling system. The model estimation cannot be based solely on the

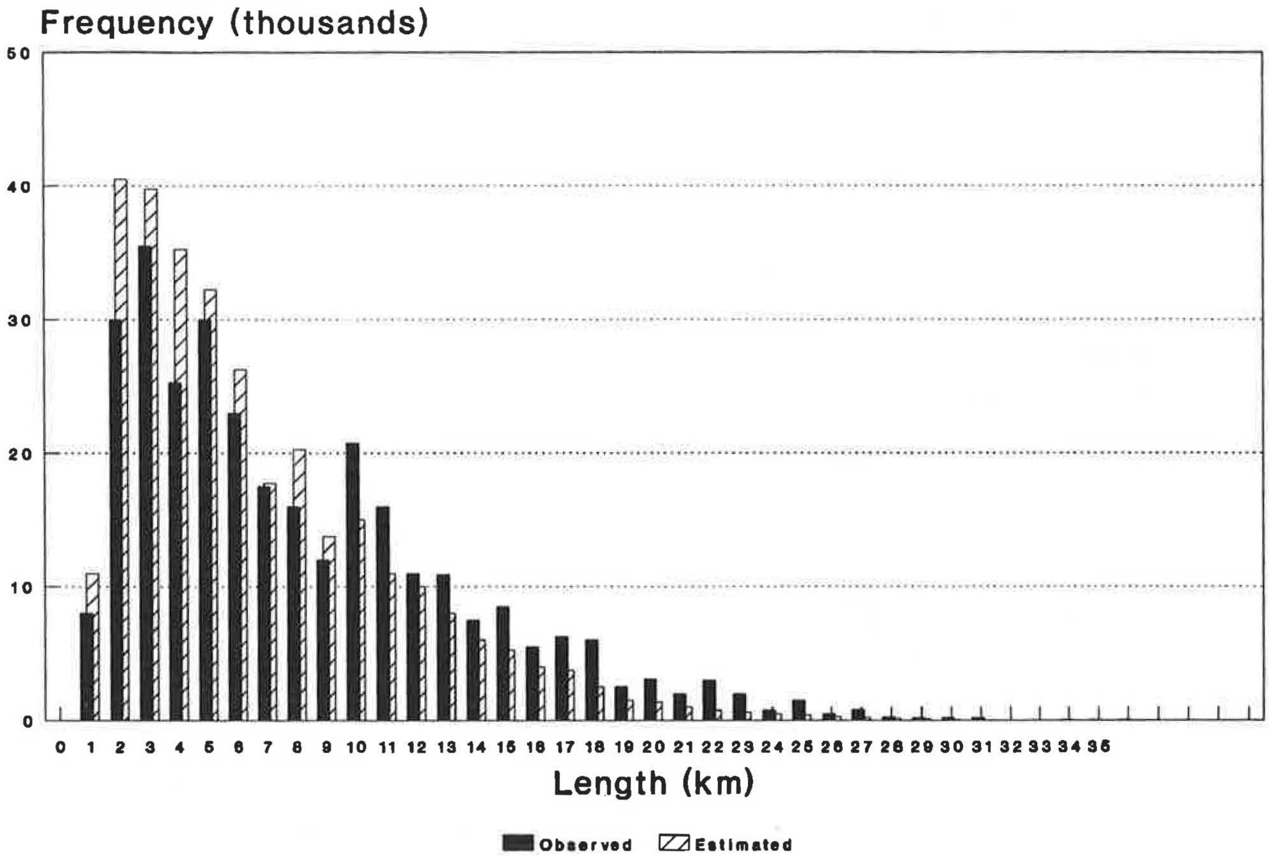

FIGURE 1 Observed and estimated trip length distributions for non-home-based trips in LITU 88.

TABLE 8 Logit Model for Household Car Ownership in LITU 88

Variable	Coefficient	Alternative
Income	0.2014	Car(s) in household
Household size	0.5881	Car(s) in household
Flat-dummy *	1.216	No car in household
City-dummy	1.418	No car in household
No car -dummy	2.302	No car in household

$\rho^2 = 0.468$

*Flat-dummy equals one if the household lives in an apartment house outside the city center and zero otherwise.

TABLE 9 Logit Model for the Division of the Population into HAP/EHAP* Categories in LITU 88

Variable	Coefficient	t	Alternative
Car in household (0/1)	3.093	27.9	HAP
Household income	0.06382	6.4	HAP
Number of workers in household	0.6438	10.5	EHAP
HAP-dummy	-2.658	-21.9	HAP

$\rho^2 = 0.347$

* HAP-person is a person that practically always has access to a car for personal trips. Others are EHAP-persons.

statistical indicators like ρ^2. One must also consider the performance of the models in forecasting.

During these comparisons, the alternative specific dummy variables of the original mode choice models were corrected by an iterative method, presented by Talvitie (*12*), to give a better replication of the present situation.

Some practical problems of the modeling work are also worth mentioning here. First, the travel times for buses were estimated with the EMME/2 system. The assignment procedure of the program tries to minimize the sum of the travel time of all passengers between each origin and destination pair (*8*). The procedure gives fairly stable total travel time, but its components (waiting, walking, and in-vehicle time) are very sensitive to the way the lines and network are coded. The assignment procedure also tends to give more transfers than are actually made during the trips.

The travel times for trains were calculated with a separate procedure where EMME/2 and some tailor-made programs were used. Here too the way the EMME/2 network was coded caused some problems in the estimation in station choice.

Second, the correct way of estimating the intrazonal distances and travel times, both in private and public transport, turned out to be problematic. The estimation was done in advance in network coding without direct connection to modeling, and this was probably one reason for the difficulties in replicating the correct trip length distributions. The intrazonal distances and travel times were originally coded in the 282-zone system. The aggregation of the zones to a 117-zone system has perhaps even strengthened the deficiencies of the values.

In spite of these difficulties, the forecasts of the model system described here turned out to be satisfactory in most cases. The forecasts so far were made with models with some theoretical deficiencies (for example, in the coefficient of the logsum variable), and many problems still exist. The estimation of new models is going on, and new forecasts will be made during spring 1992.

The final method of forecasting is still under discussion, too. There are three possible alternatives. The first is to use disaggregate models with zonal means as described here. The second is to make the forecasts using sample enumeration. This method has not been used in Finland earlier, and many practical problems have to be solved before full-scale applications. Sample enumeration will anyway be used to make further checks on model performance, especially in destination choice.

The third way to make the forecast is to use either the first or the second method to calculate growth factors that can be used to modify the present car and public transport trip matrixes. The basic idea of this method is to make better use of the trip matrixes that were explored with the big OD field surveys mentioned in this paper.

ACKNOWLEDGMENT

This work was done as a part of a consulting contract between Viatek Ltd. and the Helsinki Metropolitan Area Council. Many persons from both organizations have been involved in this 2-year-long effort.

REFERENCES

1. *Origin-Destination Survey of Automobile Traffic in the Helsinki Metropolitan Area in 1987–1988* (in Finnish). Publication B 1990:3. Helsinki Metropolitan Area Council, Helsinki, Finland, 1990.
2. *Origin-Destination Survey of Public Transport in the Helsinki Metropolitan Area in 1988* (in Finnish). Publication B 1990:5. Helsinki Metropolitan Area Council, Helsinki, Finland, 1990.
3. *Travel Characteristics of the Helsinki Metropolitan Area Population in 1988* (in Finnish). Publication B 1990:2. Helsinki Metropolitan Area Council, Helsinki, Finland, 1990.
4. *Nationwide Travel Survey 1986* (in Finnish). Publication TVH 713422. Finnish National Road Administration, Helsinki, Finland, 1988.
5. *Traffic Models for the Helsinki Metropolitan Area* (in Finnish). Publication B 1990:15. Helsinki Metropolitan Area Council, Helsinki, Finland, 1992.
6. H. Rintamäki. *Helsinki Metropolitan Area Transportation Study 1976* (in Finnish). Publication 50. Helsinki University of Technology, Otaniemi, 1980.
7. *The Home Interview Survey of Oulu Region Transportation Study* (in Finnish). Report 6, Oulu Region Transportation Study. Oulu, Finland, 1991.
8. M. Ben-Akiva and S. Lerman. *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press, Cambridge, Mass., 1985.
9. *Road Traffic Driving Costs in 1988* (in Finnish). Publication 11. Research Center, Finnish National Road Administration, Helsinki, Finland, 1988.
10. S. Algers, J. Colliander, and S. Widlert. *Logitmodellen. Användbarhet och generaliserbarhet*. Byggforskningsrådet. Rapport R30:1987. Stockholm, Sweden, 1987.
11. A. Daly. *Current Status of Alogit 3.1*. The Hague Consulting Group, The Hague, the Netherlands, 1990.
12. A. Talvitie. *Refinement and Application of Individual Choice Models in Travel Demand Forecasting. A Guide to the Development and Application of Disaggregate Choice Models*. State University of New York at Buffalo, Buffalo, 1981.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.