Parametric Access Network Model

Antti Talvitie, University of California, Berkeley
Tom Leung, Department of Civil Engineering and Environmental Science, University of Oklahoma

Parametric models are calibrated for the access portions of rail and bus trips. The models are designed to predict average zonal travel times as a function of the transportation system, zone size, and volume-related characteristics of a zone. The calibrated models are access walking, driving, and bus-riding time for rail trips and walking time to a stop for bus trips. Corresponding models are developed for the within-zone variance of the access time. These models provide input to the existing travel demand forecasting process by systematizing the way in which the access times are currently obtained for network coding. The importance of these values for travel forecasting has been repeatedly demonstrated in the past. These models also enable the use of large zones to help simplify and speed up the transportation plan analysis and evaluation process. The predictive accuracy of the final models is evaluated in terms of standard indexes of forecasting accuracy. The results show that the coefficients of determination are high and that the coefficients of variation are low for all the models. Thus, the models should find an immediate use in transportation planning.

The demand for transportation depends, among other things, on the level of service, i.e., access time, in-vehicle time, and travel cost, provided by the transportation system. These variables both characterize the transportation system and serve as the basis for travel demands. Thus, the values of these variables are needed both for calibration of travel demand models and for forecasting purposes.

Currently, no satisfactory systematic methods exist for calculating the access-egress travel times even though some progress has been made in modeling the access travel times. Two studies (1, 2) were found to be pertinent starting points to the present research.

The research reported here is an extension of the work by Talvitie and Hilsen (2). It uses simulation to create the data on which the statistical calibration of access models can be based and, thus, does away with expensive data gathering. The models developed in the present research are also specific to a station and, if necessary, a bus line. In this way, the mixing of demand (which station to choose) and supply (what it takes to get to the station) sides is eliminated. Another new feature of the present access supply models is the explicit inclusion of intrazonal transportation system attributes in the models.

NEED FOR ACCESS SUPPLY MODEL DEVELOPMENT

Recent studies show that travel demand models must be policy sensitive and behavioral in order for them to be truly useful in transportation systems analysis. A traveler, confronted with questions of whether to make a trip, where to go, when to make the trip, which mode to take, and which route to choose, bases his decisions on the level of service provided by the system and the activity system around him. Research by Kraft and Wohl (3) pointed out that the level of service provided by the transportation system must be described for complete door-to-door trips.

Domencich, Kraft, and Valette (4) suggested that travelers react differently to different components of travel time and cost. It is desirable to segment the times and costs into their component parts so as to bring the effect of policy actions into much sharper focus.

The explicit modeling of the supply of access systems is also needed to obtain information on the access mode choice and access station choice. If the attributes of different access modes to different stations and lines are accurately represented, then the use of access-station selection models (5) becomes possible and the desired information on access mode and station can be obtained.

To model the mean (and variance) of the access times requires that the underlying system of transportation be defined; hence, it becomes possible to compute costs of such access systems and relate them to the performance (travel time) obtained by the system.

PROPOSED STUDY MODEL

In this study, two types of supply models for the access-egress portion of a trip are developed. The first type is inclusive models, which apply to all the people in the zone. The second type is restricted models, which apply only to those people who can choose the alternative whose time is being modeled.

The reason for developing both types of models is that it has not been clearly established yet which of the two types of models is the more appropriate counterpart for current (and future) travel demand models. Theoretical arguments tend to favor the restricted models (e.g., modeling choice is appropriate only if the choice exists); however, there are practical reasons for favoring the inclusive models (e.g., how do we know who in the zone does and does not have a choice?). Perhaps the consistency of the supply and demand models is a better yardstick; whichever supply model was used in developing the demand models should also be used in forecasting.

Three models that deal with a rail trip are calibrated. They are the access walking time to station, the access driving time to station, and access riding time in bus to station. The access driving time can also be applied to the access driving time to the ramp of the line-haul expressway for the automobile mode. One access model, walking time to a bus stop, is developed for the bus trip. However, bus walk time can also be considered a segment of a rail access trip if the traveler walks to a bus stop in order to take the bus to a rail station.

In addition to these models, which estimate the zonal mean access time, corresponding models for the within-zone variance or standard deviation of the access times are developed. These models are for the variance of the access time in a zone and not, of course, for the mean access time of the zone.

The supply models developed in this study have the following functional form:

$$\text{Access time}_j^m(L) = S(\text{zone size variable}_j, \text{transportation system variables}_j, \text{volume}_j) + e_j$$

where

- j = zone,
- m = access mode of travel,
- L = estimate of the access time,
- S = supply function, and
- e = error term.

Three types of variables are considered in the above model. The zone size variable describes the area of a zone. The volume variable is represented by trip density per square kilometer per day in a zone. Transportation system variables are separated into two groups: those characterizing the zone(e.g., area, spacing between arterials, and signalization) and those characterizing the transportation system serving the zone (e.g., bus stop frequency per kilometer, number of bus lines, distance of the station from the zone centroid). The parametricization of these variables allows us to develop statistical models that relate these variables to the mean access travel time of a zone.

DATA AND METHOD

In developing the supply models, a simulation approach is used. The values of the dependent variables are generated by the method of simulation; multiple regression analysis is then used to estimate the parameters of the model.

The input data set can be specified in the following three groups:

1. Characteristics of each zone—zone size, arterial spacing, traffic signals, and the number of lanes on the main arterials and intersections;
2. Characteristics of the public transit system—station coverage, frequency of the bus stations, bus route-kilometers, spacing between parallel bus lines; and
3. Volume characteristics—trip density.

In addition, the following assumptions were made.

1. Speed on local streets is 40 km/h (25 mph), and total delay is 10 s.
2. Capacity for each lane of the arterial is 1700 vehicles per hour of green.
3. Vehicles are evenly distributed among the lanes going in one direction.
4. If the signals are synchronized, the vehicle can either stop once or go through the system without delay. If unsynchronized, the vehicle has a 50 percent chance of stopping at each intersection it goes through.
5. Intersections are 0.8 km (0.5 mile) apart.
6. All study zones are squares.
7. Walking speed is 4.8 km/h (3 mph).
8. Buses stop at every station for 30 s.
9. Frequency of bus stops for each line within a zone is uniform.
10. Passengers can get on the bus anywhere along the line. The time spent to pick them up is negligible.
11. The volume on the arterial is a function of the arterial spacing and trip density (6).
12. The peak-hour volume on the arterials can be obtained as a percentage of the 24-hour traffic volumes. The percentage of the peak-hour volume in the morning peak direction was assumed to be 65 percent (7).
13. To approximate the mean delay at a signalized intersection, a typical volume relationship (8) was used. This was modified for the volume-capacity ratio to be on the x-axis.
14. The speed of the vehicle can be obtained as a function of the volume-capacity ratio (9).

These assumptions and the specific input data provide a framework for logical and mathematical relationships, analytical equations, and probabilities required to simulate the values of the explanatory variables. These simulation models are described below.

DRIVE MODELS

The input dat were used to derive the volume-capacity ratio for the zonal arterials. This volume-capacity ratio was used to obtain the zonal speed and the mean delay for each intersection. A randomly located individual was next generated inside the traffic zone, and his or her travel time on local streets and on the main arterials to reach the station-expressway ramp was calculated. For zones with unsynchronized signals, the number of intersections he or she had to go through in driving to the station was also recorded.

Similar models are developed for both inclusive and restricted cases. In the latter, no travelers were generated within 0.8 km (0.5 mile) of the station, because they are assumed to walk to the station only.

WALK MODELS

The walk models are very simple. Information on the location of the rail station or on the bus lines and the frequency of the bus stops in put, and a randomly located individual is generated. By assuming a walking speed of 4.8 km/h (3 mph), his or her walk time to the station or to the nearest bus line is easily obtained and recorded.

Again, similar models were developed for the restricted case, where people are generated only within the walking distance of 0.8 km (0.5 mile) of the rail station or 0.4 km (0.25 mile) of the bus line.

BUS RIDE MODELS

Speed and mean delay were derived for the bus ride models in the same way as in the drive models. Again, the travel time for a randomly located individual was obtained by recording the traveled distance, the number of stops the bus has to make, and intersections the bus has to go through to reach the station.

EXPLANATORY VARIABLES

From the simulation models, the mean and standard deviation of the zonal access time are obtained. They are the dependent variables for the supply models. The independent variables are the attributes of the zones and their intrazonal transportation system. The variables are defined below. (SI units are not given for the variables of this model inasmuch as its operation requires that they be in U.S. customary units.)

1. DISTI, distance from centroid, is the distance to the station (or to the expressway ramp) from the centroid of the zone in miles. If the station or ramp is 0.5 mile (0.8 km) outside of the zone, it is the straight-line distance from the centroid to the boundary of the zone.
2. AREA is the number of square miles in each zone.
3. Dummy is a variable to identify whether the station is inside the zone. Dummy equals 0 if the station is inside the zone and 1 if the station is outside.
4. COVER, coverage, is the ratio of the area of the circle of 0.5-mile (0.8-km) radius that is inside the zone to the area of the zone.
5. YI, Y_i or Y_o, is the smallest distance from the side of the zone to the nearest bus line in miles. It is positive if the line is inside the zone and negative if the line is outside.
6. BS, bus line spacing, is the distance between parallel lines. If there is only one bus line, spacing equals $2 \times (\mathrm{AREA}^{1/2} - \mathrm{YI})$.
7. COVER-B is the ratio that represents the portion of the area of the rectangle of area that is inside the zone.
8. FRE is the bus stop frequency in miles between stops.
9. DISTO is the straight-line distance from the boundary of the zone to the station or ramp in miles. If the station is inside the zone, DISTO = 0. Once the vehicle gets outside the zone, a synchronized signal system is assumed.
10. SIGN, signals, is a dummy variable for the synchronization of traffic lights in the zone. SIGN = 0 if synchronized, and SIGN = 1 × (average number of intersections) if unsynchronized.
11. TDPSQ, trip density, is the trip density of a zone per square mile per day. The value used in the expression is scaled down by dividing it by 1000.
12. LANE is a dummy variable for the number of lanes on the zonal arterials, and it equals 0 if two lanes and 1 if four lanes.
13. AS, arterial spacing, is the distance between parallel arterials.
14. DIST is the distance to the station from the boundary of the zone on a straight line joining the station and the centroid of the zone. It is positive if the station is outside the zone and negative if the station is inside the zone.

ESTIMATION OF MODEL COEFFICIENTS AND EVALUATION OF MODEL ACCURACY

The method of least squares was used to estimate the coefficients of the supply models.

The models developed were evaluated on the basis of standard indexes regularly used in econometric studies to measure predictive accuracy and goodness of fit. For this purpose the following measures are given in Table 1:

1. The coefficient of determination (r^2),
2. Coefficient of variation = (standard error of estimate)/(mean of the dependent variable), and
3. F-value of the model.

On the basis of these measures, the relative accuracy of the ordinary least squares models may be inferred. Another criterion for judging the performance of the models is the sign of the coefficients; the coefficients must have a proper sign if a model is to be useful.

ESTIMATION RESULTS AND EVALUATION OF MODELS

The forms for the models are given in Table 1. The estimated mean and standard deviation for the walk, drive, and ride models are given in Table 2. Table 2 also gives the model parameters and the statistical significance of the relationships measured by the parameters. The elasticities were calculated at the mean value of the variables.

Walk Models

Reasonably good results were obtained for the walk models. All the parameters have the correct sign, and most of them are highly significant.

In the walk, restricted models, coverage and its square are the important variables in determining the walk time to a rail station or to a bus stop. In addition, whether the bus line is within the zone and the distance adjacent bus stops appear to be relevant for the average zonal walk time to a bus stop.

For the walk, inclusive models, the size of the zone is an important variable. Logically, the larger the zone is, the longer the walk time is. Another obviously significant factor is where the station or bus lines are situated. In the walk to rail station model this is specified by the distance of the rail station from the centroid of the zone and whether the station is inside the zone. In the walk to bus stop, variables for number of lines and bus line spacing determine the number of lines serving the zone and their exact location. The elasticity with respect to these variables shows that walk time is very sensitive to the location of the bus lines and rail stations.

Examination of the t-statistics and the elasticity of the bus stop frequency variable shows that the bus stop frequency plays an important role in the walk, restricted model, although its presence is not significant in the walk, inclusive model. In the former, pedestrians walk to a bus line located closer than 0.4 km (0.25 mile) and a large part of the walk time consists of the distance between the bus stops. For the walk, inclusive model everybody inside the zone is allowed to walk to a bus stop. In this case, naturally the distance between the lines rather than the distance between the bus stops plays an important role in determining the walk time. Bus

Table 1. Form of models.

Model	Form	Elasticity About Mean	n	F-Value Mean	F-Value σ	CV Mean	CV σ	r^2 Mean	r^2 σ
1. Walk to rail station, inclusive	$T = a + bX$	$e_{\bar{x}} = b\bar{x}/\bar{T}$	39	352.45	175.22	0.08	0.11	0.97	0.94
2. Walk to rail station, restricted	$T = a + bx + cx^2 + dY$	$e_{\bar{x}} = (b + 2c\bar{x})(\bar{x}/\bar{T})$ $e_{\bar{y}} = d\bar{Y}/\bar{T}$	10	169.75	51.15	0.03	0.10	0.98	0.96
3. Walk to bus stop, inclusive	$T = a + bx$	$e_{\bar{x}} = b\bar{x}/\bar{T}$	66	166.44	120.68	0.19	0.26	0.89	0.85
4. Walk to bus stop, restricted	$T = a + bx + cx^2 + dy$	$e_{\bar{x}} = (b + 2c\bar{x})(\bar{x}/\bar{T})$ $e_{\bar{y}} = d\bar{y}/\bar{T}$	30	377.93	51.49	0.03	0.11	0.98	0.89
5. Drive to station or ramp, inclusive	$T = ae^{bx}$	$e_{\bar{x}} = b\bar{x}$	151	193.53	185.41	0.10	0.31	0.90	0.88
6. Drive to rail station, restricted	$T = ae^{bx}$	$e_{\bar{x}} = b\bar{x}$	151	257.64	176.47	0.10	0.40	0.91	0.88
7. Ride to rail station	$T = ae^{bx}$	$e_{\bar{x}} = b_{\bar{x}}$	54	103.60	49.78	0.06	0.12	0.92	0.92

Table 2. Means and standard deviations of model variables.

Model	Variable	Parameter Mean	Parameter σ	t-Value Mean	t-Value σ	Elasticity Mean	Elasticity σ
1	CONST	4.76	2.85				
	DISTI	12.3	4.3	14.42	10.37	0.45	0.42
	AREA	1.69	0.66	12.28	9.84	0.31	0.32
	Dummy	3.41	-1.12	3.40	2.30	0.04	-0.03
2	CONST	9.79	-0.46				
	COVER	-10.53	7.39	13.82	8.79	-0.19	0.68
	$COVER^2$	7.39	-4.46	10.67	6.05		
	Dummy		0.51		1.96		0.125
3	CONST	0.89	-0.25				
	AREA	0.18	0.11	2.16	1.65	0.10	0.11
	YI	1.91	0.98	3.42	2.19	0.10	0.08
	BS	3.84	2.71	17.80	15.62	0.72	0.91
4	CONST	4.92	0.23				
	COVER-B	-7.85	2.99	13.58	5.19	-0.29	0.28
	$COVER^2$-B	5.21	-2.08	12.12	4.89		
	FRE	5.37	1.66	24.16	7.55	0.32	0.29
	Dummy	0.32		2.66		0.03	
5	CONST	1.08	0.27				
	DISTO	0.35		7.68		0.35	
	SIGN	0.04	0.10	4.11	11.2	0.07	0.18
	AREA	0.04	0.05	6.86	8.75	0.19	0.24
	TDPSQ	0.06	0.08	13.9	17.8	0.73	0.98
	AS	0.76	1.04	13.3	17.4	0.57	0.78
	LANE	-0.35	-0.51	6.32	8.83	-0.26	-0.39
	DISTI	0.28	0.27	18.4	7.74	0.24	0.23
6	CONST	0.95	0.19				
	SIGN	0.07	0.10	9.79	10.04	0.12	0.18
	AREA	0.05	0.06	11.83	9.01	0.26	0.29
	TDPSQ	0.06	0.08	19.89	17.46	0.76	0.97
	AS	0.84	1.19	19.00	17.49	0.63	0.89
	LANE	-0.40	-0.61	9.59	9.35	-0.31	-0.46
	DISTI	0.28	0.30	11.31	7.73	0.25	0.25
7	CONST	1.52	0.34				
	$AREA^{1/2}$	0.07	0.09	7.34	7.48	0.39	0.84
	TDPSQ	0.02	0.05	3.00	5.88	0.26	0.65
	AS	0.47	0.69	5.15	6.22	0.41	0.62
	BS	0.45	0.34	6.13	3.60	0.45	0.31
	DISTO	0.53	0.14	16.73	3.66	0.53	0.14
	1/FRE	0.08	0.07	6.14	4.57	0.41	0.36
	YI		0.37		3.23		0.14

stop frequency did not appear to be significant for this model and was dropped off (a troubling result).

The standard deviation models were obtained by using the same or sometimes even fewer variables than were used in the mean models. Data in Table 1 indicate that the mean models are more accurate than the standard deviation models. The r^2 for the former is about 0.96 while it is 0.91 for the latter. Examination of the accuracy of the models with respect to their coefficients of variation shows that the standard error in the mean models ranges from 3 to 19 percent of the mean, while it ranges from 11 to 26 percent of the mean in the standard deviation models.

Drive and Ride Models

Semilog forms were used in these models. The linear model was rejected because it estimated some negative travel times, which are unrealistic. The semilog form was chosen because it ensured that the predicted travel time will always be positive and also because the relationship between travel time and volume resembles an exponential function. The model parameters all have the right sign and are highly significant.

The only parameter with a negative sign in the models is the variable for the number of lanes on the arterials. The addition of one lane to the zonal arterial will decrease the travel time as expected.

Zone size, trip density, spacing of arterials, and the distance to the station ramp from the boundary of the zone are important variables in determining both the drive and the ride time to a rail station or to a ramp of a line-haul expressway, while zone size, bus stop frequency, distance of the rail station from the boundary of the zone, and spacing between bus lines are the major contributing variables in the bus ride models. Also, the fact that all the supply elasticities are less than unity indicates that the supply is inelastic.

As with the walk models, the models for the mean travel time are better calibrated than the models for the standard deviation. r^2 for a mean model is about 0.91, and it is 0.89 for the standard deviation model. Examining the accuracy of the models with respect to their coefficients of variation shows that the standard error in the mean models ranges from 6 to 10 percent of the mean and from 12 to 40 percent of the mean in the standard deviation model. In summary, all the models appear to be quite good.

FUTURE RESEARCH

The research into network parametricization-network aggregation reported here will be continued. Specifically, three important and distinct areas of research have been identified for continuation of this research. First, because of the importance of access travel times on travel demand, the modeling of intrazonal transportation system will be continued by incorporating zonal (possibly trip end) density distributions in the model and redefining some of the variables to be clearer and more policy oriented. Around the stations and also along the guideway and bus lines, the development densities tend to be higher than elsewhere in the zone. The implication of this is that the access times and their variables, obtained from the present models, may be higher than in actuality. Also, including density distribution in the model enables the analyst to test the effect of zoning changes on travel demand and choice of mode.

Second, parametric models should be developed for line-haul facilities. Line-haul travel times can be modeled as a function of volume of travel, operating policy, and capacity and spacing of the line-haul facilities. This parametricization of the line-haul system (both between and within zones in order to keep the advantage of large zone sizes) allows the line-haul transportation system to be represented in the form of an equation (it can be envisioned as a link). As a first step, the networks will be parametricized for three modes: automobile, bus, and guideway.

As a result of these two research tasks, a parametric representation of the entire transportation system can be accomplished; that is, access-egress and line-haul can be represented by relatively simple equations instead of a large and involved network. These parametric equations should be extremely helpful in developing multimodal networks for detailed analysis by using current transportation model systems. By anticipating modal line-haul volumes and with the help of demand models perhaps, one can make initial estimates regarding the spacing, operating policy, and capacity of the line-haul system. Similarly, the sensitivity of these network components can be analyzed by using the model coefficients. In fact, the parametric network models, as proposed here, are the supply side analog of the developing behavioral travel demand models.

The objective of the third research effort is a more formal integration of the network supply and travel demand models for a powerful sketch planning tool. Even though a satisfying backward-seeking model would clearly be desirable, it may not be implementable for more than 10 to 20 zones at this time. The design testing approach (e.g., UTPS) is implementable for any number of zones. A range of multimodal alternatives can be quickly evaluated with respect to demand, capacity, and extent of line-haul and access facilities needed. This all can be done without coding networks because the networks are represented by equations. Sensitivity analysis of each plan can also be readily performed by using the coefficients of the demand and supply equations.

Another advantage of parametric network representation, coupled with large traffic zones, is the opportunity to humanly interpret and visualize the results. This contrasts with the often too detailed and unclear networks, too numerous zones, and thick computer printouts of line volumes, which, even when plotted, are only of marginal help in the initial stages of transportation system planning.

CONCLUSIONS

The development of parametric networks is possible. The equations given in this paper were developed quite quickly. Nevertheless, they appear to model the access system rather well and thus are immediately applicable. Given the relative importance of the access for travel demand and transportation planning, modeling of the access component should be included in the standard transportation planning model system such as UTPS. This should be a relatively easy task. Each zone is now characterized by its population, employment, and so forth; this description should be extended to include the few basic characteristics of the internal transportation system and the way it relates to the line-haul system.

Explicitly modeling the access supply and access mode-station choice provides another important advantage: the use of large zones. This in turn speeds up planning processes. Finally, large zones provide more reliable and quicker predictions of land use activities.

ACKNOWLEDGMENT

This report was produced as part of a program of research and training in urban transportation sponsored by the Urban Mass Transportation Administration, U.S. Department of Transportation. The results and views expressed are the independent products of university research and are not necessarily concurred with by the Urban Mass Transportation Administration.

REFERENCES

1. P. Rassam and R. Ellis. Access Characteristic Estimation System. Office of High Speed Ground Transportation, U.S. Department of Transportation, Northeast Corridor Transportation Project, Vol. 1, Dec. 1969.
2. A. Talvitie and H. Hilsen. An Aggregate Access Supply Model. Paper presented at 15th Annual Meeting, Transportation Research Forum, San Francisco, Oct. 1974.
3. G. Kraft and M. Wohl. New Directions for Passenger Demand Analysis and Forecasting. Transportation Research, Vol. 1, No. 3, 1967.
4. T. Domencich, G. Kraft, and P. Valette. Estimation of Urban Passenger Travel Behavior: An Economic Demand Model. HRB, Highway Research Record 238, 1968.
5. P. Liou and A. Talvitie. Disaggregate Access Mode and Station Choice Models for Rail Trips. TRB, Transportation Research Record 526, 1974, pp. 42-52.
6. Chicago Area Transportation Study. Vol. 4, Chicago, 1962.
7. Traffic Volume Characteristics. Nashville Metropolitan Area Transportation Study, 1959.
8. N. A. Irwin, N. Dodd, and H. Von Cube. Capacity Restraint in Assistant Problems. HRB, Highway Research Record 297, 1961.
9. Highway Capacity Manual—1965. HRB, Special Rept. 87, 1965.