Improvements in Understanding, Calibrating, and Applying the Opportunity Model

EARL R. RUITER, Traffic Planning Analyst, Chicago Area Transportation Study, Chicago, Illinois

•THE Opportunity Model is the name given to the mathematical procedure developed by Morton Schneider to distribute, over all possible destinations, the actual destinations of all trips having a stated origin. The distinguishing feature of the model is its unique independent variable, intervening opportunities. Although new to the urban transportation field when proposed, this variable has been a feature of earlier models of human behavior in the fields of population migration and intercity travel (4).

Since the completion of its 1980 transportation plan for the Chicago area, the Chicago Area Transportation Study (CATS) has been involved in a number of projects in which detailed traffic forecasts are needed for relatively small areas within the large Chicago metropolitan region. Two of these areas, the Fox River Valley and Lake County, are shown in Figure 1 along with the original CATS region to indicate the types of areas with which CATS has been dealing. The application of the Opportunity Model of the trip distribution process to these areas was more difficult than the application to the entire metropolitan area. Calibration of the model to actual vehicle-miles of travel and to screenline counts was impossible with the two-parameter (two L values) model used in the earlier large area applications. These problems indicated a need for improved methods of applying the model. A study of the theoretical bases of the model and of calibration methods indicated that improved methods using multiple L values could be developed, two for each zone or group of zones, as parameters in the Opportunity Model.

Because of computer size limitations, it is felt that trip distribution applications to small areas within large metropolitan regions will become more important in the years to come, as such regions increase in size. These regions are becoming too large for complete inclusion in a single assignment run. An example is the Chicago region, in which the Chicago-Northwestern Indiana Standard Consolidated Area, as defined by the 1960 Census, includes eight counties (Fig. 1). By comparison, the 1956 CATS area includes parts of only four counties.

This paper discusses the improved understanding of the Opportunity Model which has resulted from the application of the model to small areas. It also explains the calibration methods developed and the results obtained with these methods. The Opportunity Model itself, rather than the CATS assignment system, which uses the model, is the major subject of the paper. Computer-oriented documentation of the assignment system in its entirety is provided in two CATS publications (8, 9).

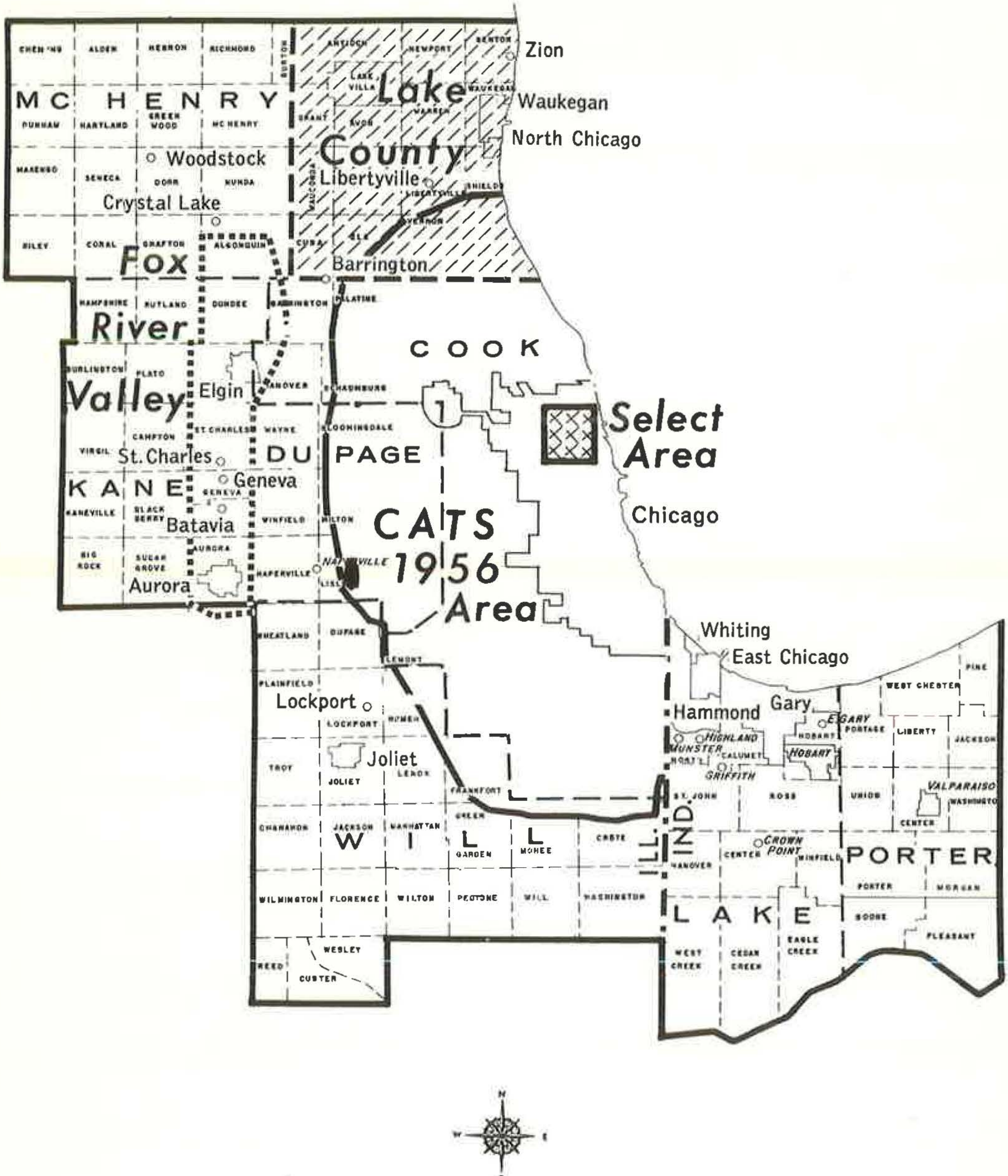
UNDERSTANDING THE OPPORTUNITY MODEL

Hypotheses and Mathematics

The hypotheses and mathematics underlying the Opportunity Model are given briefly as a starting point before the discussion of interpretations of the model and its parameter, and the presentation of relationships between the model and trip parameters (1).

The Opportunity Model is based on the hypotheses that (a) total travel time from a point is minimized, subject to the condition that every destination point has a stated probability of being accepted if it is considered; and (b) the probability of a

Paper sponsored by Committee on Origin and Destination.



Map Prepared By
CHICAGO AREA TRANSPORTATION STUDY

Figure 1. Assignment areas within Chicago-Northwestern Indiana standard consolidated area.

destination being accepted, if it is considered, is a constant, independent of the order in which destinations are considered.

The hypotheses lead to the following mathematical formulation, in terms of limitingly small quantities:

$$dP = L \left[1 - P(V) \right] dV \tag{1}$$

where

dP = probability that a trip will terminate when considering dV possible destinations;
P(V) = total probability that a trip will terminate by the time V possible destinations are considered;
V = possible destinations already considered, or subtended volume; and
L = constant probability of a possible destination being accepted if it is considered.

The solution of the differential Eq. 1 is

$$P(V) = 1 - e^{-LV} \tag{2}$$

The expected interchange from zone i to zone j (T_{ij}) is the volume of trip origins at zone i (O_i) multiplied by the probability of a trip terminating in j:

$$T_{ij} = O_i \left[P(V_{j+1}) - P(V_j) \right] \tag{3}$$

or

$$T_{ij} = O_i \left[e^{-LV_j} - e^{-LV_{j+1}} \right] \tag{4}$$

The subtended volumes (V's) are the sums of the possible destinations considered before reaching a given zone. As it can be assumed that a zone's trip origins equal its trip destinations over a 24-hr period, V_j canbe defined in terms of the trip origins reached before reaching zone j:

$$V_j = \sum_{k=1}^{j-1} O_k \tag{5}$$

where the O_k's are arranged in order of increasing travel time from zone i.

Although Eq. 5 could be substituted into Eq. 4 to express T_{ij} completely in terms of trip origins and the L value, it is more convenient to leave the equations as given in Eq. 4.

Initial applications of the Opportunity Model showed that it would be necessary to specify more than one value for L, because of the differing probabilities of acceptance associated with different types of trips. For example, people are more selective in choosing places to work than they are in choosing places to shop for groceries. Three trip subpopulations (short, long residential, and long nonresidential) with two L values (short and long) satisfactorily represented empirical trip data for large regions. A mathematical statement of the Opportunity Model, as used in the CATS assignment system, is, therefore:

$$T_{ij} = \sum_{k=1}^{3} O_{ik} \left[e^{-L_k V_{jk}} - e^{-L_k V_{j+1,k}} \right] \tag{6}$$

where k ranges over the three trip subpopulations.

The discussion of interpretations of the model and the L value which follows is based on just one of the trip subpopulations of Eq. 6. It is assumed that the model holds in the simple form of Eq. 3. When the time comes to speak of operational problems, the necessary trip subpopulations will be reintroduced.

Interpretations of Model

The Opportunity Model can be considered in its broadest sense as an explanation of human behavior, as stated in the two foregoing hypotheses. The fact that the model

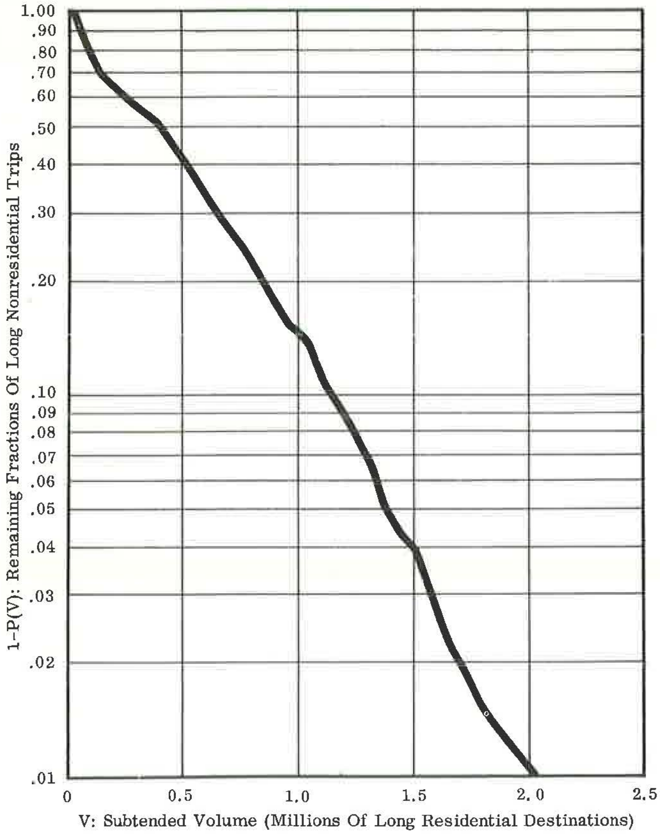


Figure 2. Cumulative distribution of long residential trips from CATS zone 001 according to number of opportunities.

has proved to be satisfactory for metropolitan regions indicates that, when averaged over a large area, people do behave as hypothesized in the model.

The model can be interpreted in a more limited sense by considering the mathematical expression of the model given in Eq. 2 as an equation which is to be fitted to empirical data by adjusting the parameter, L. Equation 2 can be changed to a linear form by rearranging and taking natural logarithms of both sides. This procedure results in the following:

$$-LV = \ln \left[1-P(V) \right] \tag{7}$$

Empirical values of V and 1-P(V) for a given zone can be plotted on semilog graph paper (Fig. 2). Theoretically, this plot will be a straight line for all trip types and all origin zones. The parameter, or L value, can be found by least squares regression. The procedure is very simple once values of V and 1-P(V) have been determined from survey data. These empirical values can be found once actual trip interchanges from a given zone are arranged in travel time order.

This interpretation of the Opportunity Model indicates that once the concepts of subtended volume (V) and probability of trip termination [P(V)] are understood, or simply accepted, the model can be thought of as a statement that a semilog relationship tends to exist between V and 1-P(V). The model expressed in a statement of this type may appeal to those who have difficulty visualizing the basic hypotheses of the model.

Interpretations of L Value

The curve-fitting approach previously discussed leads to a graphic interpretation of the L value. This parameter can be viewed simply as the slope of the straight line which best fits a set of empirical V and ln [1-P(V)] data.

Mathematically, the L value can be expressed:

$$L = \frac{-\ln [1-P(V)]}{V} \tag{8}$$

Two characteristics of the L value are evident in Eq. 8. The sign of L is always plus, because 1-P(V) is always less than one and its natural logarithm is always negative. The units of L are (1/opportunities), or (1/trip ends), as the numerator of Eq. 8 is unitless and the denominator has the unit opportunities, or trip ends. Experience with empirical data indicates that L is always very small, usually of the order of 10^{-5}, and always much less than one.

These three characteristics all support the interpretation of the L value as a modified probability quantity. Just as for other probability quantities, L lies between zero and one. Unlike more common probabilities, L is not unitless. It can be thought of as the probability, per individual opportunity or trip end, of destination acceptance. A reading of the hypotheses of the model shows that this interpretation of the L value is in agreement with the model's theoretical basis.

The interpretation of the L value leaves room for these parameters to vary from origin zone to origin zone. In fact, the interpretations given here are based on one origin zone and would be seriously limited if L values could not vary from zone to zone. The realization that multiple L values are desirable from an interpretive point of view was the first of two breakthroughs to CATS researchers attempting to apply the Opportunity Model in small areas. The second breakthrough was that the CATS assignment system was understood well enough so that it could be modified to accept multiple L values (9).

Relating L Values to Trip Parameters

The L value has been interpreted in terms of subtended volume and fraction of trips unsatisfied. For the extremely simplified situations in which trip end density is assumed to be constant, the L value can be expressed in terms of average trip length and average trip end density. As trip length and trip end densities are more common trip parameters than subtended volume, the expression obtained, although a simplification, provides insights into the nature of the L value.

In addition to the assumption of uniform trip end density, it is necessary to assume that the time ranking of possible destinations can be replaced by a distance ranking without loss of accuracy. This assumption would be true if the speed in all parts of the transportation system were constant, or nearly so.

Since the Opportunity Model is probabilistic in nature, the probabilistic concepts of mathematical expectation can be used to find desired averages. For example, average trip length ($\bar{r}$) may be found by performing the following integration:

$$\bar{r} = E(d) = \int_a^b f(V)\, dP(V) \tag{9}$$

where

- d = distance variable (mi);
- E(d) = expected value, or average, of distance variable (mi);
- f(V) = expression of distance in terms of variable V (subtended volume);
- dP(V) = density function of variable V; and
- a, b = lower and upper limits on V.

The density function needed can be obtained by differentiating Eq. 2:

$$dP(V) = Le^{-LV}\,dV \tag{10}$$

Because of the assumptions of constant density and distance ranking, subtended volume can be expressed in terms of distance as follows:

$$V = \rho\pi d^2 \tag{11}$$

where ρ = average trip end density (trip ends/mile²). Equation 11 can be solved for d to obtain f(V):

$$d = f(V) = \left[\frac{V}{\rho\pi}\right]^{1/2} \tag{12}$$

Equations 10 and 12, along with the limits on V (0 and ∞), can be substituted into Eq. 9:

$$\bar{r} = \int_0^\infty \left[\frac{V}{\rho\pi}\right]^{1/2} Le^{-LV}\,dV \tag{13}$$

Carrying out the integration and simplifying, the following expression for $\bar{r}$ is obtained:

$$\bar{r} = \frac{1}{2}\left[\frac{1}{\rho L}\right]^{1/2} \tag{14}$$

or, solving for L,

$$L = \frac{1}{4\rho\bar{r}^2} \tag{15}$$

Although it must be remembered that Eq. 15 is a gross approximation, it does indicate that L tends to be inversely proportional to trip end density and to the square of average trip length. Experience with CATS assignments has confirmed these tendencies.

A dimensional analysis of Eq. 15 indicates that, as ρ is expressed in trip ends per square mile and $\bar{r}$ is expressed in miles, the L value has the units (1/trip ends). Thus, Eq. 15 agrees dimensionally with the theoretical basis of the Opportunity Model.

CALIBRATION

The major benefit of the improved understanding of the Opportunity Model has been the enhanced ability to calibrate individual assignments. Calibration techniques are desired so that empirical data or predictions, such as total vehicle mileage and screenline traffic counts, can be duplicated by the assignment process without running a large number of expensive "trial and error" computer runs.

Criterion

The matching of actual or predicted average trip lengths by the assignment was the criterion which led to the most useful calibration techniques. Since the number of trips in an area must be specified before an assignment can be run, and since total vehicle-miles of travel is the product of total trips and their average trip length, the matching of average trip lengths means that actual or predicted vehicle-miles of travel will be matched by the assignment.

A case could conceivably be made for using the criterion of matching actual or predicted trip travel times which are more directly related to the Opportunity Model. The main reason for rejecting this criterion is that experience indicates travel time data obtained in travel surveys is much less reliable than travel distance data. Travel time

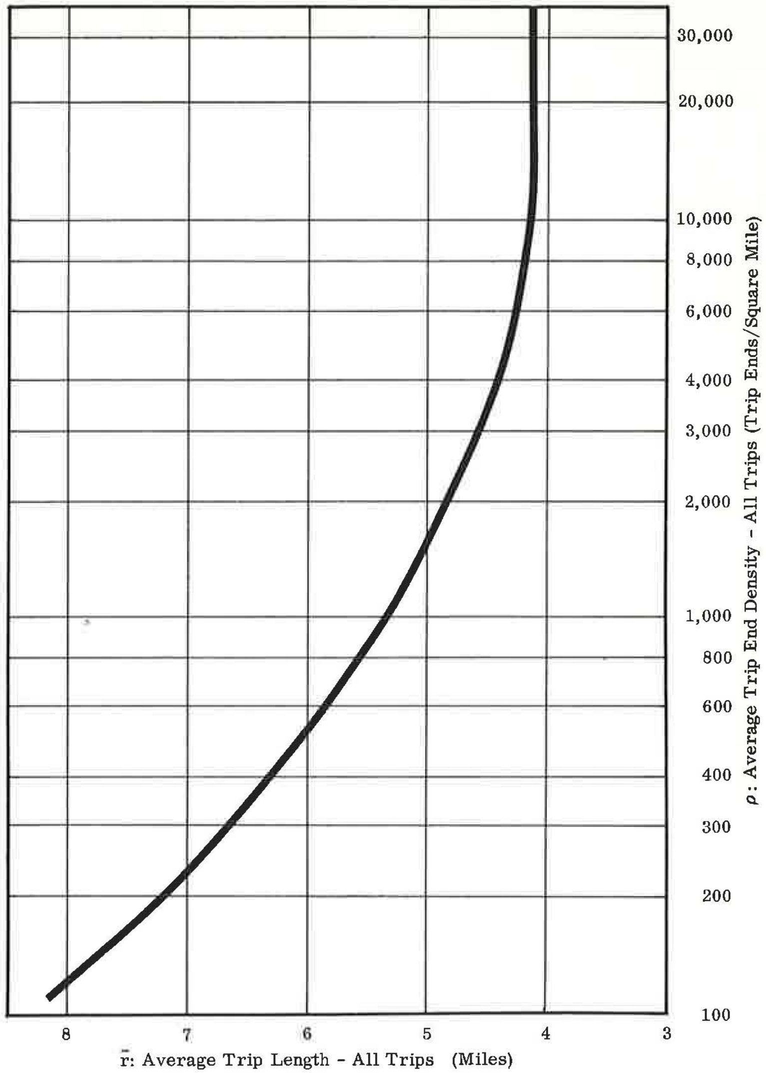


Figure 3. Relationship between total average trip length and total trip end density.

data must be estimated by the trip maker, and this estimation is often quite gross. Travel distances, on the other hand, can be calculated from origin and destination information, which typically is much more accurately reported in travel surveys.

The prediction of future average trip lengths must, of course, be accomplished when assignments serving as predictions of the future are to be calibrated. Although the problem of predicting future average trip lengths is far from completely solved, it is being dealt with. A National Cooperative Highway Research Project has as its goal the determination of trends in average trip lengths which can be observed over time on an area-wide basis (3, 7).

An investigation of zone-by-zone variations in average trip lengths, as they existed at the time of CATS 1956 travel surveys, has been conducted. One result of the investigation has been the discovery that an important source of zone-to-zone variation in average trip lengths is zonal trip end density. A hand-fitted plot of zonal average trip length vs zonal trip end density is shown in Figure 3. The fact that zones with few trips tend to have a long average trip length is apparent. This tendency is an affirmation of the hypothesis of the Opportunity Model which states that the satisfaction of trips and, therefore, average trip length, is affected by the number of available destinations.

As more sources of variation in zonal average trip lengths which can be applied to a future situation are discovered, it will be possible to predict zonal average trip lengths, once area-wide trends have been determined. For example, if it is assumed that the curve of Figure 3 will remain constant over time, future average trip lengths can be determined easily by reading values from the curve once future trips have been generated for each zone.

Single L Value Calibration Method

The first attempt to develop a calibration method was the application of Eq. 14 to the case in which only one L value is desired per trip population. This method has been attributed to Morton Schneider (5). Inasmuch as Eq. 14 is approximate, and the "constant" term is not exactly 0.25 in each case, the equation has been modified to a ratio form so that the results of an already completed assignment can be used as added information when planning a new assignment. The ratio form is

$$\frac{\bar{r}_1}{\bar{r}_2} = \frac{\sqrt{L_2 \rho_2}}{\sqrt{L_1 \rho_1}} \qquad (16)$$

where the average trip lengths ($\bar{r}_1$ and $\bar{r}_2$) and average trip end densities (ρ_1 and ρ_2) are obtained by considering the entire assignment area. The subscripts 1 and 2 refer to two particular times, places, or assignment runs.

The L value obtained by using Eq. 16 is only a first approximation. The results of the assignment run using this first L value must be investigated to determine the direction in which the second L value must differ from the first. This trial and error process must continue until the desired accuracy is achieved.

Multiple L Value Calibration Methods

The use of single L values for each trip population in an assignment run for a small area does not satisfy other criteria even when the total vehicle mileage criterion is met. For this reason, research at CATS has shifted to the calibration of multiple L value assignments. The remainder of the methods discussed in this paper, all multiple L value calibration methods, are labeled the empirical, statistical, and iterative calibration methods for convenience. Actually, all three are empirical, statistical, and iterative in some sense of the meaning of these words.

Empirical Calibration Method—Short Trips

The first multiple L value method developed at CATS for short trip L values is essentially an extension of the use of Eq. 16 to a number of groups of zones having similar average trip lengths and trip end densities. Individual L values then are obtained for each group of zones, termed a density class. The method is summarized briefly here; a complete description is given by Muranyi and Miller (6).

The method was developed before the assignment program had been modified to accept multiple L values automatically. Costly and error-prone manual stopping of the program and inserting of new L values was necessary; therefore, the number of density classes was held to five or six. The method is not limited to a small number of cases, however, and could be applied to each zone individually if desired.

One necessary revision of Eq. 16 was a change of the density variable. The use of an area-wide average would defeat the purpose of treating each density class individually. By using the average density within a three-mile radius of each origin zone, the area in which nearly all of the short trips from each zone would end was included. The densities used in the equation were, therefore, the average of these modified densities for all zones included in the density class.

The empirical nature of the method arises from the use of an initial, approximate assignment to a known situation to obtain the data corresponding to the subscript 1 in Eq. 16. In this respect, the method is similar to the single L value method. These data are used to obtain an empirical constant for each density class defined as follows:

$$C_i = \hat{r}_{1i} \sqrt{L_{1i} \hat{\rho}_i} \tag{17}$$

where

$\hat{r}_{1i}$ = average trip length for density class i, obtained from initial assignment;
L_{1i} = L value for density class i used in initial assignment; and
$\hat{\rho}_i$ = average trip end density for density class i.

Once values of C_i have been determined and values of $\bar{r}_i$ are obtained from given data or curves such as those in Figure 3, short L values for use in assignments to known situations can be found by using the following relationship:

$$L_i = \frac{C_i^2}{\hat{\rho}_i \bar{r}_i^2} \tag{18}$$

Future short L values can be obtained by assuming that the C_i vs $\hat{\rho}_i$ relationship found to hold for present assignments will continue to hold in the future. Eq. 18 then can be applied by using new densities and average trip lengths.

Empirical Calibration Method—Long Trips

It was not possible to define a meaningful average long trip end density corresponding to the density within a three-mile radius area used for short trips, so an adaptation of Eq. 14 could not be used for present assignments of long trips. Instead, Eq. 8 was used to insure that the correct number of trips from each density class would go farther than ten miles. The determination of the L value in each case is straightforward once V and P(V) have been determined from the available empirical data.

Future long L values can be determined by using Eq. 16, the ratio form of the approximate L value, the trip parameter relationship.

Statistical Calibration Method

The statistical calibration method was developed by Emilio Casetti at CATS when it became evident that Eq. 15 had serious deficiencies which limited its applicability to the prediction of zonal L values. The method uses multiple regression statistical techniques to determine relationships for test zones in a given area and with given trip ends, between arbitrarily chosen L values, resulting average trip lengths, and trip end densities determined within various cutoff points. Cutoff points indicate the truncation of the allocation of trips from a given origin at a given percent level. For example, a 60 percent cutoff point corresponds to the point at which 60 percent of the trips available in a given zone of origin have been allocated. The density within a given cutoff point varies depending on the L value used.

The equation to which a least squares fit is obtained is the following:

$$\log L = \log a_0 + a_1 \log \rho_1 + a_2 \log \rho_2 + \ldots + a_n \log \rho_n + a_{n+1} \log \bar{r} \tag{19}$$

where

ρ_i = trip end density within cutoff point i, using an arbitrarily selected L value;
a_i = coefficient obtained using multiple regression techniques; and
$\bar{r}$ = average trip length of all trips assigned with a given L value, using an arbitrarily selected cutoff point.

Because this equation is linear in the logarithmic transformations of the variables, standard multiple regression methods can be used to find the a_i coefficients. A more compact form of the equation is obtained by taking antilogarithms of both sides:

$$L = a_0 \cdot \rho_1^{a_1} \cdot \rho_2^{a_2} \cdots \rho_n^{a_n} \cdot \bar{r}^{-a_{n+1}} \qquad (20)$$

The family of curves represented by Eq. 20 was selected for use because Eq. 15 indicates a power relationship may be expected to exist between L, $\bar{r}$, and ρ. Eq. 20 represents the generalization of this power relationship to a large family of curves from which the best curve can be found.

The various ρ_i variables were introduced into the model to provide some measure of the density variations ignored in Eq. 15. Tests indicated that six density variables, corresponding to cutoff points ranging from 0.60 to 0.98 using an L value of 80×10^{-6}, resulted in a satisfactory predictive equation.

The cutoff point to be used in determining $\bar{r}$ is 0.80, as this value of $\bar{r}$ is most closely correlated with L.

The step-by-step procedure recommended for collecting the test data needed, and for using the resulting equation, is as follows. Inasmuch as a major feature of the method is its recognition of variable trip end density and the pattern of this variation is unique for each assignment area, it is not recommended that a_i coefficients found for one area be used to predict L values in another. Therefore, the test data must be collected for each new assignment area.

1. Select a workable number of test zones representing ranges of trip end densities and average trip lengths.
2. Select a reasonable number of test L values representing the probable range of actual L values.
3. Distribute trips using each L value from each test zone.
4. Calculate the trip end densities within each cutoff point for each test zone, using an L value of 80×10^{-6}.
5. Calculate the average trip length within the 0.80 cutoff point for each L value and for each test zone.
6. Use a standard multiple regression computer program to transform all data using a logarithmic transformation and to determine the regression coefficients, a_i.
7. For each zone or group of zones for which an L is to be determined, calculate the trip end densities within each cutoff point, using an L value of 80×10^{-6}. Also, determine the actual average trip length within the 0.80 cutoff point from survey data.
8. Use Eq. 19 or 20 to determine log L or L.

Obviously, the method is neither fast nor simple, but it is likely to be more accurate than the empirical method, because variable densities are recognized and included in the method. A computer program has been written to simplify the execution of steps 4 and 5. An indication of the accuracy of the method is that the multiple correlation coefficient obtained in the test case was 0.815.

Iterative Calibration Method

Both of the previously discussed multiple L value calibration methods were lacking in ease of application. The search was continued, therefore, for an easy and accurate calibration method which would use the fewest possible data not needed as input for an actual assignment run. A number of approximations on the order of Eq. 15 but using some sort of varying trip end densities were found no better than available methods.

It finally was necessary to use the same relationship between trip ends, average trip lengths, and L values as that existing in the assignment program. This relationship is simply the discrete version of Eq. 9:

$$\bar{r}_o = \frac{\sum_{j=1}^{n} d_{oj} \left[e^{-LV_j} - e^{-LV_{j+1}} \right]}{1 - e^{-LV_n}} \qquad (21)$$

where

$\bar{r}_o$ = average trip length for zone o,
d_{oj} = distance from zone o to zone j, and
n = total number of zones.

Although zones are ranked by time in the assignment system, it is necessary to assume for the calibration method that a distance ranking suffices. Otherwise, it would be necessary to use assignment output to determine assignment inputs.

The data needed to use Eq. 21 to find L values are $\bar{r}_o$'s, d_{oj}'s, and V_j's. However, the V_j's are just summations of O_i's (see Eq. 5), which are assignment inputs. The $\bar{r}_o$'s must be determined externally, just as with any of the calibration methods. The d_{oj}'s can be calculated if each zone is assigned X and Y grid coordinates. In summary, the data needed to solve Eq. 21 can be obtained if trip origins, X and Y coordinates, and average trip length are specified for each zone.

Equation 21 reveals that the L value cannot be isolated algebraically on one side of the equation. This is where iteration comes in. Iterative methods for solving non-linear equations such as Eq. 21 are presented in texts on numerical analysis, such as Hildebrand (2). A common iterative method is the modification of a nonlinear equation f(x) = O to the form x = F(x) and then by using the recurrence relation of the form:

$$x_{k+1} = F(x_k) \tag{22}$$

The procedure involves choosing an x_O as an initial approximation, finding $x_1 = F(x_0)$, and continuing until the difference between x_k and x_{k-1} is sufficiently small. Hildebrand points out that the method is guaranteed to converge only if:

$$\left| \frac{dF(a)}{da} \right| < 1 \tag{23}$$

where a is the true value of x.

In applying this method to Eq. 21, a function F(L) can be found by multiplying both sides of the equation by $L/\bar{r}_o$:

$$L = F(L) = \frac{L \sum_{j=1}^{n} d_{oj} \left[e^{-LV_j} - e^{-LV_{j+1}} \right]}{\bar{r}_o \left[1 - e^{-LV_n} \right]} \tag{24}$$

Since F(L) depends on a large number of parameters, it is difficult to check it for convergence in the general case. However, Eq. 14 is approximately true, so a test of the F(L) obtained by multiplying both sides of this simpler equation by $L/\bar{r}$ should indicate whether or not the F(L) of Eq. 24 will converge. Multiplication of Eq. 14 by $L/\bar{r}$ results in:

$$L = F(L) = \frac{1}{2\bar{r}} \left[\frac{L}{\rho} \right]^{1/2} \tag{25}$$

Differentiating,

$$\frac{dF(L)}{dL} = \frac{1}{4\bar{r}} \left[\frac{1}{\rho L} \right]^{1/2} \tag{26}$$

Once differentiation is complete, $\bar{r}$ can be replaced by its equivalent, as given in Eq. 14. The resulting value of the derivative is 0. 5, indicating that the condition expressed in Eq. 23 is met for the F(L) of Eq. 25 and, therefore, should be met for the F(L) of Eq. 15 with ρ replaced by O_o/A_o, where A_o is the area of zone o. This necessitates the addition of one more item to the list of zonal data needed, namely, the area of the zone.

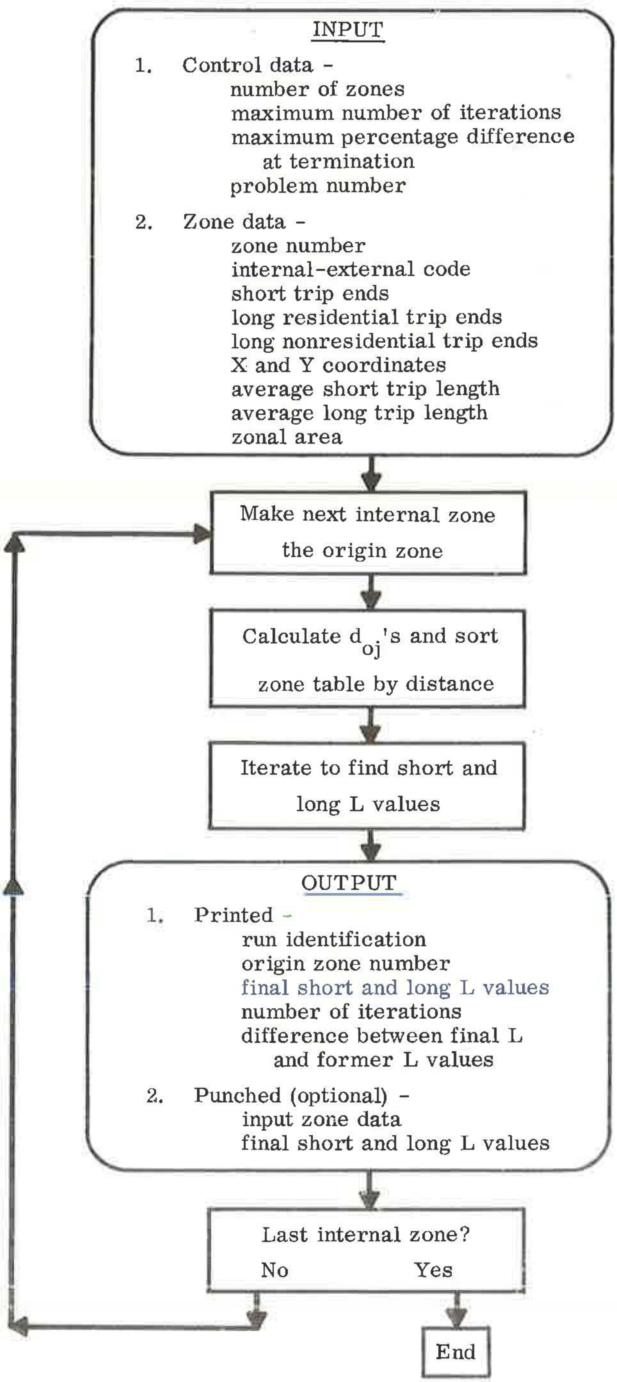


Figure 4. Generalized flow diagram of the L value calibration program.

TABLE 1

COMPARISON OF 1960 FOX RIVER VALLEY SURVEYED AND ASSIGNED TRIPS (Single L Value Model)[a]

Trip Type	Surveyed Trips	Assigned Trips	Ratio of Assigned to Surveyed
I-I[b]	42,700	44,200	1.04
I-B[c]	45,400	58,800	1.30
I-E[d]	49,700	62,800	1.26
Total	137,800	165,800	1.20

[a] Data derived from ref. 6.
[b] Trips with both origin and destination in the internal area.
[c] Trips with one end in the internal area and the other end in the buffer area.
[d] Trips with one end in the internal area and the other end in the external area.

TABLE 2

COMPARISON OF 1960 FOX RIVER VALLEY ACTUAL AND ASSIGNMENT (Multiple L Value Model)[a]

District of Origin	Actual[b] $\bar{r}_s$	Assigned $\bar{r}_s$	Ratio of Assigned to Actual
11-12	2.18	2.2	1.06
21-22	2.60	2.7	1.04
31-32	2.85	2.9	1.02
41-42	2.20	2.5	1.14
51-52	2.82	2.9	1.03

[a] Data derived from ref. 6.
[b] $\bar{r}_s$ = average short trip lengths.

TABLE 3

COMPARISON OF 1960 FOX RIVER VALLEY SURVEYED AND ASSIGNED TRIPS (Multiple L Value Model)[a]

Trip Type[b]	Surveyed Trips	Assigned Trips	Ratio of Assigned to Surveyed
I-I	42,700	45,900	1.07
I-B	45,400	51,400	1.13
I-E	49,700	54,300	1.09
Total	137,800	151,600	1.10

[a] Data derived from ref. 6.
[b] See explanation accompanying Table 1.

Although the foregoing iterative calibration method can be expected to be accurate and uses only easily determined zonal data, it would be far from easy to use if all calculations had to be performed by hand. It was to satisfy this requirement that a computer program was written to accept as input the zonal data and control information needed, to calculate iteratively both short and long L values for any or all zones in the assignment area, and to punch out these L values on cards which can be used directly as part of the assignment input. A generalized flow diagram of the program appears in Figure 4. The fact that L values are found only for zones coded "internal" means that any number of selected zones can be calibrated, or that all zones which will actually be used to send trips in the assignment program can be calibrated.

The program has been written in FORTRAN II and running time, when calibrating all zones, is slightly less than that for an assignment using the same computer. The number of iterations necessary to achieve an accuracy of 0.1 percent ranges from about 8 to 11.

RESULTS OBTAINED USING CALIBRATION PROCEDURES

Single L Value Method

The single L value calibration method was used in all assignments to the entire CATS area run before August 1965. These assignments have included not only the 1956 existing runs and the 1980 future runs, but also runs for a number of intervening years. In each case, a combination of trial and error methods and the use of the single L value calibration method have resulted, finally, in an acceptable assignment. The number of preliminary assignments has varied greatly, and in some cases has been reduced to one.

In two of the smaller areas within the Chicago metropolitan region, the Fox River Valley area and the Lake County area, the single L value calibration method and single L value assignments were tried a number of times, but never could be made to give acceptable results. An example of the problems involved is indicated in Table 1 which compares the final single L value run in the Fox River Valley with survey data. Although entirely internal trips have been quite accurately duplicated, trips between the internal area and the buffer and external areas are greatly overestimated. Results of this kind lead to the realization that multiple L values are necessary in the small area assignments.

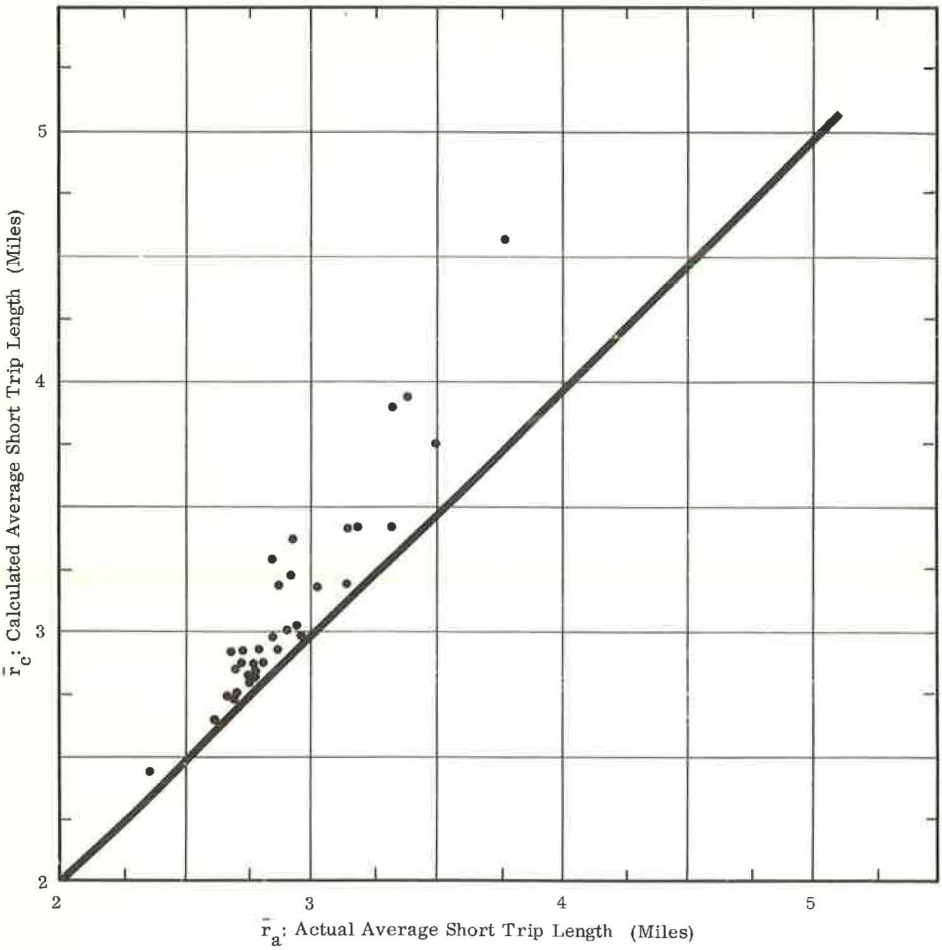


Figure 5. Comparison of calibration method input ($\bar{r}_a$) and select area assignment output ($\bar{r}_c$) average short trip lengths.

Multiple L Value Methods

Empirical Method. —The empirical method was used to calibrate multiple L values for present and future Fox River Valley and Lake County assignments. In both cases, the results indicate not only that the actual average trip lengths of the density classes are closely approximated by the assignments, but also that there is a general improvement in the quality of the assignments, as reflected in comparisons of trip survey and assignment origin and destination data. Table 2 gives actual and assignment average short trip lengths for final Fox River Valley assignments. Agreement is very good. Origin and destination comparisons in Table 3 indicate a general improvement, amounting to ten percent for all internal trips. Although trips with both origin and destination in the internal area are more poorly estimated when the multiple L value model is used, the total error is more uniformly distributed among the three groups of trips than it was when the single L value model was used.

Statistical Method. —No assignments have been calibrated by use of the statistical calibration method. The collection of the necessary test data was very time-consuming. And inasmuch as the results obtained through the empirical method for groups of zones in the Fox River Valley and Lake County assignments were considered sufficiently accurate, the more involved method was not attempted.

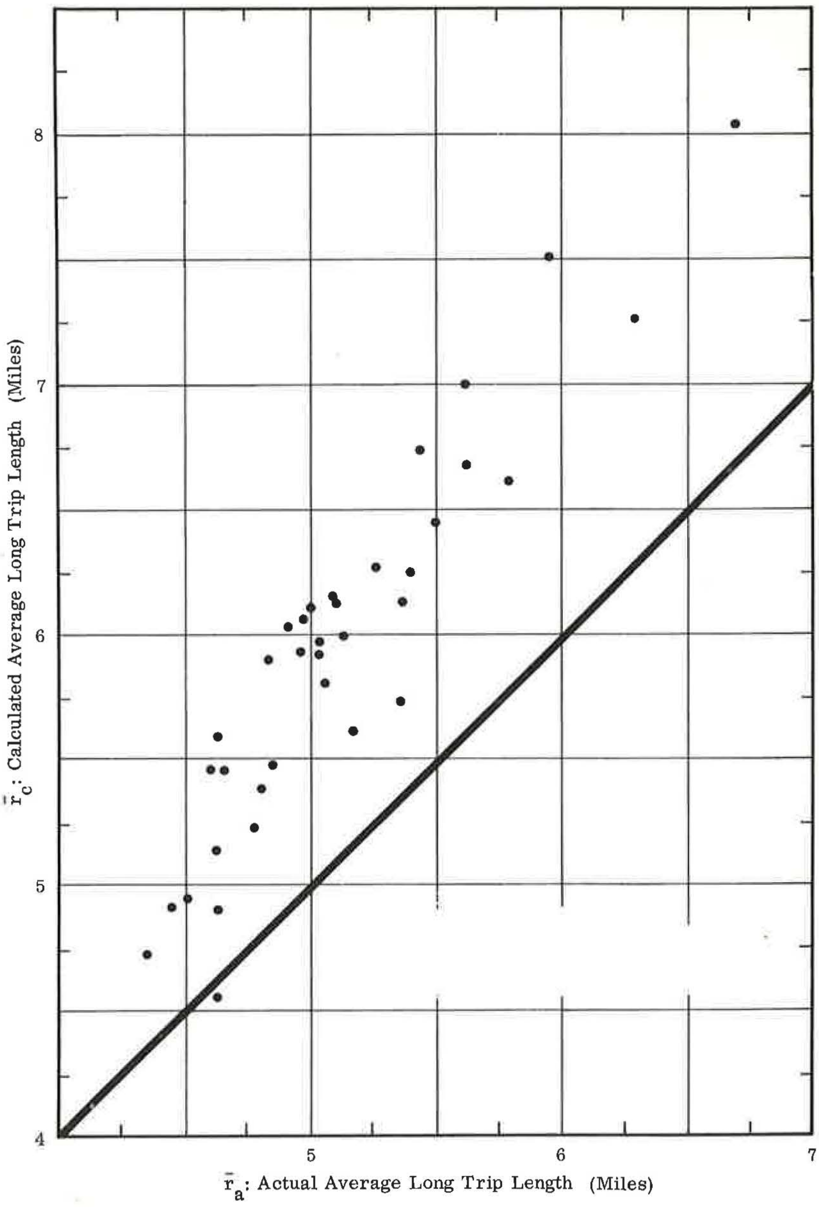


Figure 6. Comparison of iterative calibration method input ($\bar{r}_a$) and select area assignment output ($\bar{r}_c$) average long trip lengths.

Iterative Method. —The iterative method has not been used to calibrate any "production" assignment runs, but has been used for runs testing the ability of the Opportunity Model to distribute trips in a very small area. The area chosen is a 36-square mile section of Chicago lying between 1. 5 and 7. 5 miles north and west from the CBD. This area is identified as the "select area" (Fig. 1). The attempt to run a select area assignment was unique not only because of the small size of the area compared to the size of the metropolitan region, but also because the area cannot be considered to be even partially self-contained, as are the Fox River Valley and Lake County areas.

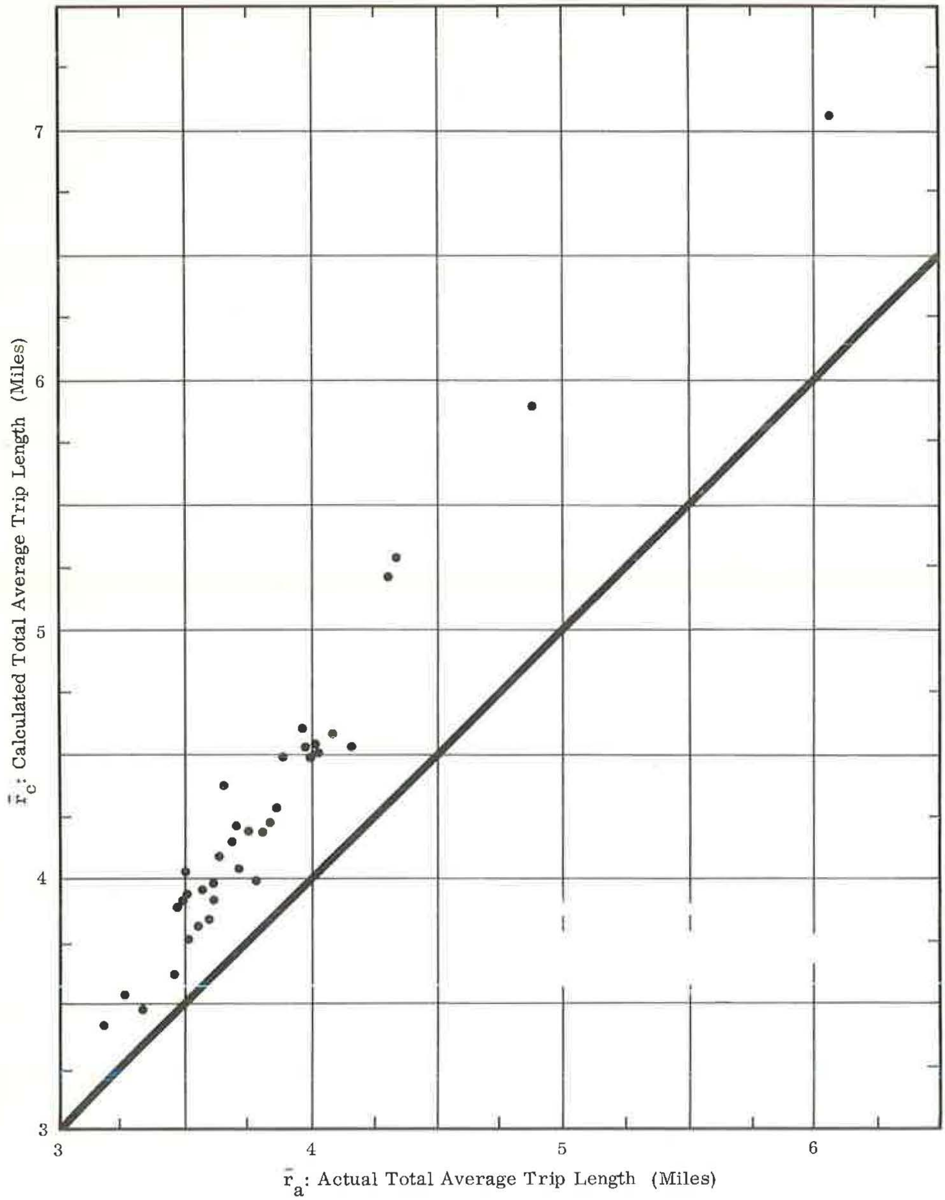


Figure 7. Comparison of iterative calibration method input ($\bar{r}_a$) and select area assignment output ($\bar{r}_c$) total average trip lengths.

The select area assignments involved the 1956 network and trip ends. Surveyed 1956 zonal average trip lengths were determined for both short and long trips. There were 36 one-square-mile zones for which trip end and average trip length information was available. Because it was desired to have as much detail as possible, the zones were divided to obtain 144 quarter-square-mile zones. The allocation of the three subpopulations of trip ends to the four smaller zones within each survey zone was based on the surveyed number of auto driver trip ends per quarter square mile, because these data were available and short and long trip ends were not available by quarter square mile. It was assumed that the average trip lengths for the square-mile zones would hold, also, for each of the four smaller zones.

TABLE 4

STATISTICAL MEASURES OF THE ACCURACY OF THE ITERATIVE CALIBRATION METHOD

Statistic[a]	Trip Population		
	Short	Long	All
$\bar{r}_c$ vs $\bar{r}_a$			
Mean of all $\bar{r}_a$ (mi)	3.00	5.12	3.83
$\bar{e}$ (mi)	0.211	0.828	0.458
MSE (sq mi)	0.1080	0.8063	0.2586
Due to e (sq mi)	0.0445	0.6853	0.2099
σ^2 (sq mi)	0.0635	0.1210	0.0487
RMSE (mi)	0.328	0.897	0.508
σ (mi)	0.252	0.348	0.221
Avg. range of variation of r_c's for quarter square miles within square mile zones (mi)	0.082		0.212

[a] $\bar{r}_c$ = output average trip length determined from select area assignments Nos. 9 and 10;

$\bar{r}_a$ = input average trip length, determined from CATS survey data;

$\bar{e}$ = average error in $\bar{r}_c$, equal to $\frac{\Sigma(\bar{r}_c - \bar{r}_a)}{n}$;

MSE = mean square error, equal to $\frac{\Sigma(\bar{r}_c - \bar{r}_a)^2}{n}$;

RMSE = root mean square error, equal to $\sqrt{MSE}$;

σ^2 = variance of errors, equal to MSE $- \bar{e}^2$; and

σ = standard deviation of errors, equal to $\sqrt{\sigma^2}$

The iterative calibration program was used to determine short and long L values for each of the 144 smaller zones. After running the assignment system using these calibrated L values, short, long, and total average trip lengths resulting from the assignment were calculated for each of the original 36 one-square-mile zones. Figures 5-7 show plots of actual average trip lengths ($\bar{r}_a$) and assignment-calculated average trip lengths ($\bar{r}_c$) are shown for short trips, long trips, and all trip averages.

A systematic error exists in this process of proceeding from $\bar{r}_a$ to L to $\bar{r}_c$. The output $\bar{r}_c$'s all are higher than they should be. It is believed that the major cause of this error is the distance ranking of destination zones used in the calibration program, whereas a time ranking is used in the assignment system. A distance ranking of destination zones minimizes average trip length subject to the L value. The time ranking must, therefore, result in at least as large and probably a larger average trip length than the distance ranking. The amount of divergence depends on the pattern of variations in speed on the various links of the network, as these variations are what cause the time and distance rankings to differ.

Various quantitative measures of the accuracy of the iterative calibration method are given in Table 4. The average errors found when $\bar{r}_c$ is compared with $\bar{r}_a$ for one-square-mile zones range from 0.211 miles to 0.828 miles, between seven and sixteen percent of the mean values. The dispersion of these errors is measured by their root mean square. The dispersions range from eleven to eighteen percent of the mean values. It is known now that systematic errors due to the distance ranking of destination zones exist, and it will be possible, in future applications, to adjust input averages downward by the amount of this systematic error. The standard deviation of the errors gives some indication of the accuracy of the method when this adjustment is used. Inasmuch

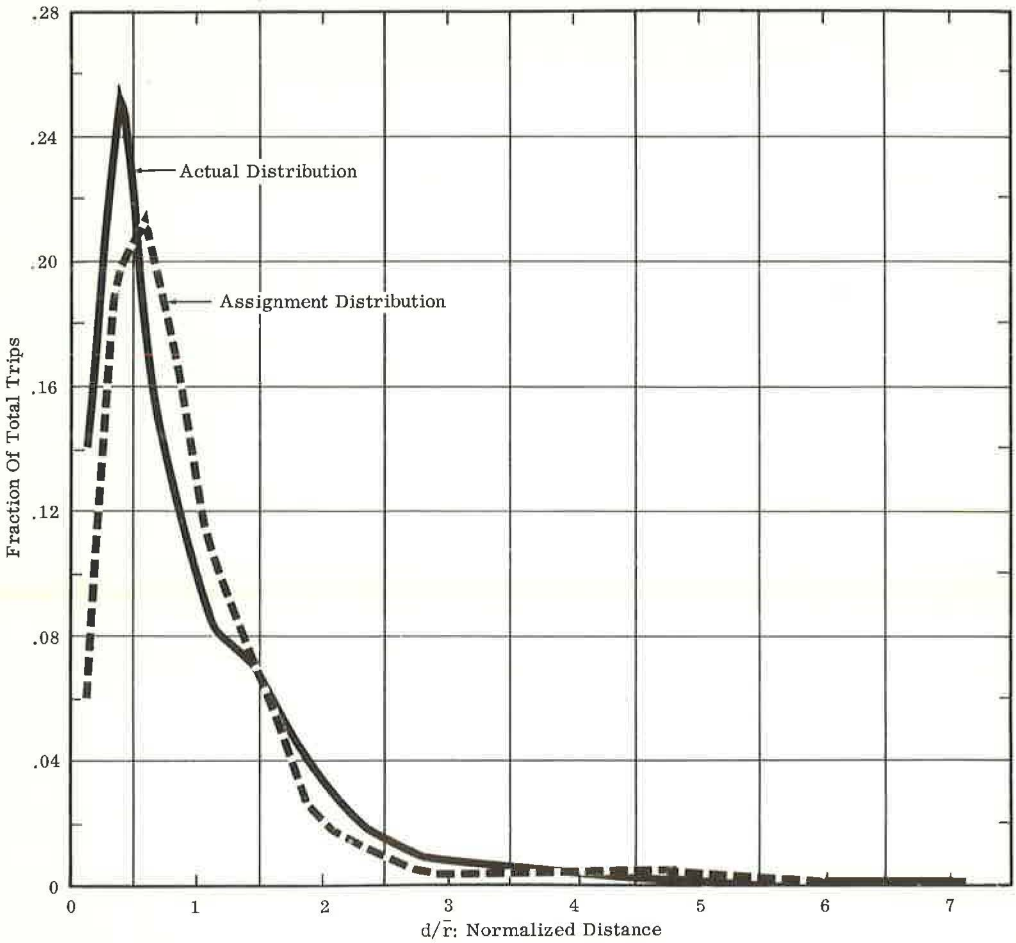


Figure 8. Comparison of actual and assignment output trip length distributions.

as these standard deviations all are less than ten percent, it appears that this calibration method will result in output average trip lengths which are within ten percent of the input averages two-thirds of the time. Furthermore, the average error will be very close to zero, so the total vehicle mileage of an assignment area will be very close to the observed amount.

To measure the ability of the calibration method to obtain the same average trip length in different zones, the range of variation of the output average trip lengths for each of the four zones within each square mile was determined. Table 4 gives the average range of variation for short trips and for all trips. These averages indicate that maximum zone-to-zone variations for equal input averages are only about three percent for short trips and six percent for all trips.

FUTURE CALIBRATION DIRECTIONS

Although the iterative calibration method does a good job of duplicating actual average trip lengths, further checks of the select area assignments indicate that the match of actual trip length distributions by assignment output distributions is poor (Fig. 8). The horizontal axis has been normalized in terms of the average trip length to indicate that the poor match is not the result of differences in the average trip lengths. Trips whose distances are from zero to about 0. 5 of the average trip length are underestimated, trips from 0. 5 to 1. 5 of the average are overestimated, and trips longer than 1. 5 times the average are underestimated.

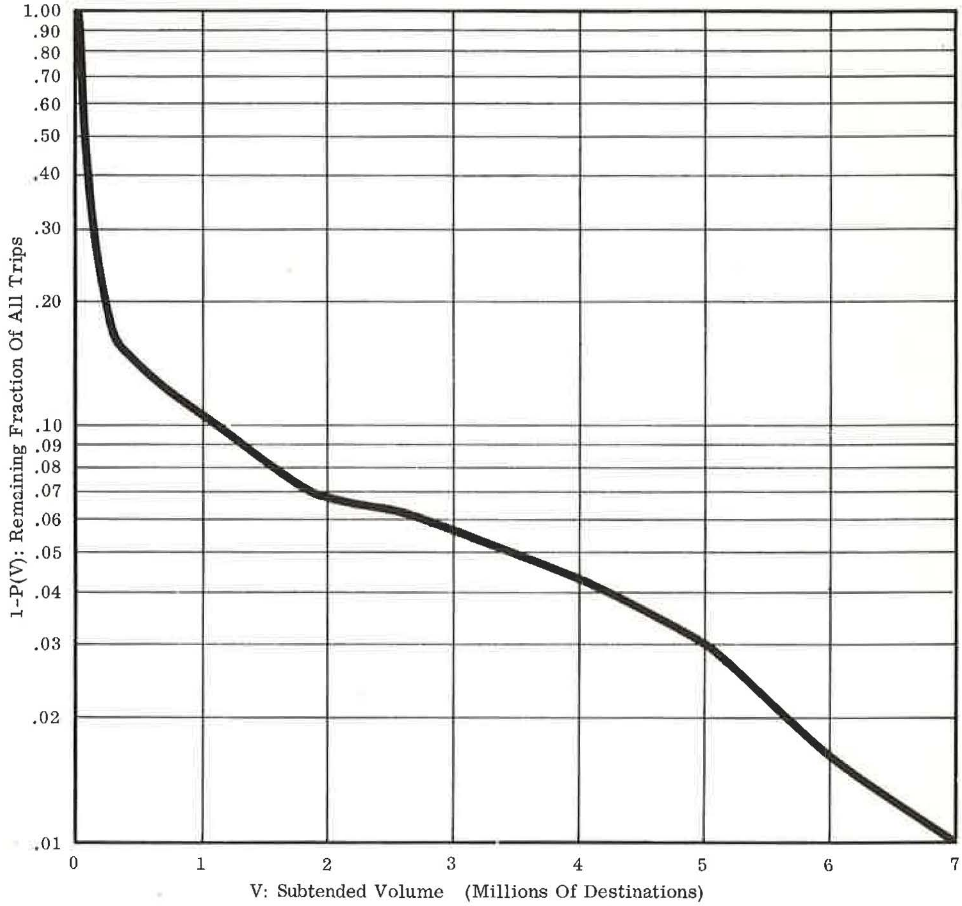


Figure 9. Cumulative distribution of all trips from CATS zone 487 according to number of opportunities.

These discrepancies can be interpreted in at least two ways and corrected in at least three ways. The first interpretation is that the model is all right, but that the trip split into long and short trips is faulty. This interpretation leads to two possible corrections: (a) keep the present trip split, but modify the observed average short and long trip lengths to obtain the correct total average trip length and the correct distribution; and (b) change the present trip split so that, using the short and long trip lengths corresponding to this split, the correct distribution is obtained.

A second possible interpretation of the trip distribution discrepancy is that the Opportunity Model's hypothesis of a constant probability of trip satisfaction (L value) is in error. Perhaps the probability of trip satisfaction is a function of V, the subtended volume. If this function could be found, trip splits would not be necessary, and the correct trip distribution would result when applying the modified model.

The three possible corrections mentioned are discussed next under the following headings: average trip length changes, trip split changes, and model changes.

Average Trip Length Changes

Investigation of the select area assignments indicates that if the average short trip length were set at about 0.6 of its actual value and if the average long trip length were modified upward so that the average total trip length remained the same, the two curves shown in Figure 8 would nearly coincide. Therefore, it is possible to change average trip lengths arbitrarily so that trip distributions will be matched. It is felt, however, that this type of correction is too arbitrary to be valid and should not be used unless acceptable methods of correction are unavailable.

Trip Split Changes

It is felt that a more acceptable correction of the Opportunity Model would be to change the definitions used in splitting trips into long and short subpopulations. Investigation has shown that some groups of trips presently classified "long" have shorter averages than groups classified "short." Also, experience with many assignments has indicated to CATS researchers that more "short" trips are needed so that trip distributions will be matched. Investigation presently is continuing to determine which long trips should be added to the short trip population.

Model Changes

The Opportunity Model implies a linear semilog relationship between 1-P(V) and V. However, this relationship can be demonstrated only for one of the trip subpopulations at a time. Figure 1, for example, shows the relationship for long nonresidential trips only. When the relationship is graphed for total trips, a curve like that in Figure 9 is obtained. A straight line would be a poor fit to this curve, but perhaps a relationship of the form $L = aV^b$ would provide a good fit. If so, trip splits would be unnecessary. All trips originating from a zone could be distributed by use of the following equation:

$$T_{ij} = O_i \left[e^{-aV_j^{b+1}} - e^{-aV_{j+1}^{b+1}} \right] \tag{27}$$

It would be necessary to change the second hypothesis of the Opportunity Model to allow for a variable L value instead of a constant. The second hypothesis could be changed to read:

> The probability of a destination being accepted, if it is considered, is a function of the number of destinations which already have been considered.

It is planned to investigate this approach to improving the Opportunity Model. The investigation will largely consist of curve-fitting, using data similar to those in Figure 9 and of determining methods of predicting the parameters needed to relate the variable L value to subtended volume.

SUMMARY AND CONCLUSIONS

The Opportunity Model has been analyzed by interpreting its hypotheses, its mathematical formulation, and its parameter, the L value. The L value also has been related to trip parameters. These analyses have served as the basis of a number of calibration methods which have been presented. The results obtained when these methods were applied to assignments at CATS have been given. Planned methods of improving both the calibration techniques and the Opportunity Model have been discussed.

The most promising calibration method developed so far is the iterative method, which provides a means of duplicating observed or predicted average trip lengths with a standard error of less than ten percent with one pass through a calibration program and one pass through the assignment system. Problems in matching observed trip length distributions indicate either that calibration methods must be concerned with more than matching averages, or that the Opportunity Model itself must be improved. CATS' future trip distribution research is expected to investigate both of these possibilities.

REFERENCES

1. Chicago Area Transportation Study. Final Rept., Vol. II, Chicago, 1960.
2. Hildebrand, F. B. Introduction to Numerical Analysis. McGraw-Hill, pp. 443, 1956.
3. Interim Report. Factors and Trends in Trip Lengths. Proj. 7-4, National Cooperative Highway Research Project, Feb. 1965.
4. Leonard, H. J., and Delano. A New Model for Distributing Short and Long Distance Travel. Presented at ASCE Transportation Engineering Conference, Minneapolis, May 1965.
5. Muranyi, T. C. Trip Distribution and Traffic Assignment. In Traffic Assignment Conference, pp. 21-22, Dec. 11, 1963.
6. Muranyi, T. C., and Miller, J. D. New Applications of the Opportunity Model to Traffic Planning and Design. Presented to Origin and Destination Committee, HRB Annual Meet., Jan. 1964.
7. Project Statement. Factors and Trends in Trip Lengths. Proj. 7-4, National Cooperative Highway Research Project, Feb. 1964.
8. Rathnau, A. S. Operations Manual—Traffic Assignment System. No. 34,320, CATS, Chicago, 1964.
9. Rathnau, A. S. Reference Manual—Traffic Assignment System. No. 34,330, CATS, Chicago, 1965.