**THROUGH TRIP TABLES FOR SMALL URBAN AREAS:
A METHOD FOR QUICK RESPONSE TRAVEL FORECASTING**

Alan J. Horowitz
Civil Engineering and Mechanics Department
University of Wisconsin – Milwaukee
PO Box 784
Milwaukee, WI 53201-0784
e-mail: horowitz@csd.uwm.edu
Phone: 414-229-6685
Fax: 414-229-6958

Minnie H. Patel
Industrial and Manufacturing Engineering Department
University of Wisconsin -- Milwaukee

November 1, 1998

Through Trip Tables for Small Urban Areas: A Method for Quick Response Travel Forecasting

Alan J. Horowitz

Minnie H. Patel

Alan J. Horowitz, Civil Engineering and Mechanics Department, University of Wisconsin – Milwaukee, PO Box 784, Milwaukee, WI 53201-0784

Minnie H. Patel, Industrial and Manufacturing Engineering Department, University of Wisconsin – Milwaukee, PO Box 784, Milwaukee, WI 53201-0784

INTRODUCTION

An often overlooked, but essential, element of traffic forecasts for small urban areas is the preparation of through trip tables. Planning practice has been to build through trip tables by factoring an existing sampled through trip table or by judgment. Obtaining a set of sampled through trip tables is an expensive, error-prone and often incomplete undertaking for communities with many external stations in their networks. Creating a synthetic through trip table by use of conventional trip distribution theory (e.g., the gravity model) has not worked well, and there is rarely sufficient data to build good through trip tables purely from observations of ground counts. Smaller urban areas, in particular, have a need for serviceable through trip tables, because through trips constitute a sizable part of the traffic in a region.

Through trip table estimation is a conceptually difficult problem that has not been given much attention in the professional literature. The location of through traffic on a network is inherently a problem that should be solved with tools of traffic assignment but is traditionally modeled following the methods and philosophy of trip distribution. Almost two decades ago there were two attempts, one in Kentucky [1] and one in North Carolina [2], to create transferable empirical relationships for the structure of a through trip table. Although both studies selected variables for their empirical relationships on insights about the nature of through travel, neither had strong theoretical underpinnings or provided insights into important geographical effects.

Surprisingly, neither of the two aforementioned studies included information about the shape of the urban area, the location of the external stations, and barriers to travel and most information about network structure. This type of information is either already available or readily obtainable for every city. It stands to reason that the rudimentary geography of the urban area would have considerable influence over the values in a through trip table. For example, it is reasonable to hypothesize that two external stations in close proximity to each other would be less likely to exchange through trips than two external stations at opposite sides of the urban area.

The need for a good through trip table is greatest for small- to medium-sized communities, which have a comparatively larger amount of through traffic. Furthermore, the problem of obtaining good through trip tables has been compounded by a growing interest in truck movements in urban areas. For example, FHWA's recently published "Quick Response Freight Manual" [3] recommends that when doing a regional traffic forecast through trip tables should be created for each of three truck categories: four-tire vehicles, other single-unit vehicles, and combination vehicles. Each of these truck categories has considerably different behaviors on a network, and each differs from the behavior of passenger vehicles.

A complete construction of a through trip table also involves determining the fraction of trips at an external station that are through trips, as opposed to trips having one end internal to the study area (later referred to here as the "intended region"). This problem has been addressed by Modlin [2], but will not be considered in this paper.

The paper develops and tests a first-cut model for through traffic that relies almost entirely on (1) a gross description of the geography of the study area and its environs, and (2) estimates of through trip volumes at each external station. The model's data requirements are very modest, especially for a simplified form that applies to many small-city applications. This simplified form of the model is applied to Oshkosh, Wisconsin, and Tallahassee, Florida. Results of the model are compared with an actual, sampled through trip tables in both cases.

PREVIOUS MODELS OF THROUGH TRAFFIC

Pigman [1] and Modlin [2] developed similar empirical relationships for through trip tables. Modlin provided five regression equations found by analyzing through trips tables in 14 cities in North Carolina, one equation for each of five highway functional classes. For example the equation for principal arterials, a particularly important functional class, is:

$$Y = -7.40 + 0.55\ PTTDES + 24.68\ RTECON + 45.62\ ADT/CD$$

where

Y is the percentage distribution of through-trip ends form an origin station to a destination station,
PTTDES is the percentage of estimated through trip ends at a destination station,
RTECON is a route continuity dummy variable, and
ADT/CD is the ADT at the destination station divided by the ADT at all stations.

The values of Y that may be obtained from this equation generally do not satisfy the required structure of a trip table, so Modlin [2] suggests that the table be adjusted to an acceptable form by the Fratar method. None of the variables in this equation contain direct information about network topology or other aspects of urban area geography. Modlin reported goodness of fit in terms of R^2 values for each functional class. The R^2 values ranged from 69% for major collectors to 96% for interstate highways. Modlin's results were better than Pigman's.

The "Quick Response Freight Manual" (QRFM) [3] recommends that through trip tables be estimated by a set of doubly-constrained trip matrix equations similar in structure to a gravity model, but where friction factors are replaced by subjective weights. Since this method is the starting point for the current research, it will be explained in some detail. The through trip table may be approximated by:

$$t_{ij} = O_i D_j X_i Y_j w_{ij} \tag{1a}$$

$$X_i = 1 \Big/ \sum_j D_j Y_j w_{ij} \tag{1b}$$

and

$$Y_j = 1 \Big/ \sum_i O_i X_i w_{ij} \tag{1c}$$

where

t_{ij} is the number of trips between origin station i and destination station j;
O_i is the number of trips from origin station i;
D_j is the number of trips to destination station j;
X_i, Y_j are balancing factors so the trip table is consistent with all O_i's and D_j's; and
w_{ij} is an external weight associated with station pair i and j.

The external weights, w_{ij}, are determined subjectively with the guidance of a few rules. They should initially be set to 1.0, except for OD pairs that are unlikely to share trips for which w_{ij} should be set to zero. Examples of pairs of external stations not sharing trips include: (a) pairs of external stations on paths leading from the region to the same neighboring city; (b) pairs of external stations on two sides of the same divided highway; and (c) pairs of external stations where both are on paths leading to rural or suburban locations. After a trial computation of t_{ij}, it may be necessary to subjectively adjust the external weights. The use of a doubly-constrained set of trip matrix equations assures that the through trip table will have the proper structure, but the method for setting the external weights is not very precise or reproducible. Variations of this method have occasionally been used by MPO's in the United States for through trip table creation in urban areas and by states for through trip tables for their statewide models.

There has been considerable interest in recent years in using mathematical programming or statistical methods to estimate trip tables from ground counts (e.g., see [1, 2, 3]). Although promising at first glance, these methods are difficult to apply to through trip table creation, because through-trip ground counts are unavailable everywhere except at external stations. There do not yet exist any practical means of separating through traffic from internal traffic for most individual links. Should such means become available in the future, then there still would be a need for an approximate trip table to start the estimation process. An approximate trip table could be obtained for nearly any urban area from the method described in the next two sections.

ADDING GEOGRAPHY TO A THROUGH TRIP TABLE

It is important to identify the information that a planner might have readily available when attempting to build a synthetic through trip table:

A. Number of through trips for each direction at each external station;
B. The physical location of each external station;
C. The general geography of the intended region and external territories, such as the location of bodies of water, mountains, barriers to travel, and neighboring cities;
D. The traffic network internal to the intended region;
E. The functional class of each road at an external station;
F. The orientation of links that pass through boundaries of the region; and
G. Total volumes on ramps at key interchanges.

The number of through trips (item A) would need to be estimated from a small sample survey or by statistical relationships developed in the community or by relationships transferred from another city.

Since equation 1 could produce a perfect through trip table with perfect knowledge of each w_{ij}, the problem becomes one of determining a good set of external weights from the type of information listed above.

In order to make effective use of the data described above it is helpful to define a "catchment" area for each external station. A catchment area consists of all lands beyond the study area that are likely to contain the true (but unknown) origins or destinations of trips passing through a single external station.

An external weight, w_{ij}, can be interpreted in the same manner as a friction factor in a gravity model, that is, it is the propensity of an origin and a destination to share trips. As with friction factors, external weights do not have units. Only their relative sizes matter. In the next section, it is shown that one useful way of calculating the external weights is by setting them equal to the probability that a trip between i and j crosses the intended region.

The external weights should not be influenced by the amount of traffic at an external station, because this information is adequately captured by the O_i's and D_j's. For example, the size of a population center in the external territories beyond an external station would not be a critical factor in establishing the size of an external weight. By in large, an external weight measures the need to cross the study area to get between the catchment area of external station i and the catchment area of external station j. The following assumptions are used here to determine the external weights.

A. Each external station has a catchment area consisting of land outside the intended region; all through trips starting or ending within a catchment area must pass through the external station.
B. All points outside the intended region have identical characteristics and can belong to at most one catchment area.
C. The probability of a trip being made between two points declines with increasing distance between them.
D. Barriers to travel (lakes, rivers, mountains) in the external territories may force a high degree of interaction between external stations that would not occur otherwise.

Although not an assumption, it should be clear from this discussion that trips between many pairs of points would not pass through the intended region and would not be considered as through trips. It should also be evident that many points outside the intended region are irrelevant. They may be too far away from the intended region to matter or they may be in locations that do not generate any trips.

Network Design and Placement of External Stations

Different planning agencies use different rules for the placement of external stations on their networks. Some agencies arbitrarily place external stations along jurisdiction boundaries. A better approach adopted by other agencies is to try to minimize the amount of through traffic, while also trying to keep the network small. An underlying premise of this paper is the location of external stations should have been chosen to give the most accurate estimates of travel within the region. This paper also presupposes that the through trip table contains exclusively through trips. That is, the table excludes trips that first exit, then re-enter the region. These assumptions would suggest a region that is compact and convex, enclosing major interchanges and trip generators and either including or excluding beltways in their entirety.

EXTERNAL WEIGHT MATRIX FOR AN ARBITRARY REGION

This section develops the mathematics necessary for determining an external weight matrix that approximates the geography of the problem in terms of the shape of the urban area, the location of external stations, and barriers to travel. As mentioned previously, there are two regions of interest: an intended region and an external territory beyond the intended region. The external territory is divided into as many catchment areas as the number of external stations under consideration. Suppose that there are *n* external stations under consideration. The catchment areas are defined in such a manner that any trip originating or terminating in catchment area i uses external station i for making any trip between catchment areas i and the rest of the external territory. In addition, the potential for travel has symmetry. That is, the number of trip opportunities from catchment area i to catchment area j is equal to the number of trip opportunities from catchment area j to catchment area i.

Let $W = \lfloor w_{ij} \rfloor$ denote an external weight matrix of dimension *n*, the number of external stations under consideration, and whose $(i, j)^{th}$ element is equal to w_{ij}. Further define w_{ij} as being the probability that a trip between the catchment area of an external station i and the catchment area of an external station j crosses the intended region or crosses a barrier. Thus, w_{ij} may be expressed as:

$$w_{ij} = \int_{S_i}\int_{S_j} \frac{g\left(d\left(P_{ik}, P_{j\ell}\right)\right)}{\int_{S_i}\int_{S_j} g\left(d\left(P_{ik}, P_{j\ell}\right)\right)} \left(I\left(P_{ik}, P_{j\ell}\right) \vee B\left(P_{ik}, P_{j\ell}\right)\right) dP_{ik}\, dP_{j\ell} \qquad (2)$$

where,

P_{ik} denotes the Cartesian coordinates (x_{ik}, y_{ik}) of point *k* of catchment area i, for i = 1, 2, ..., n,

S_i is the set of all points of catchment area i, for i = 1, 2, ..., n,

$g\left(d\left(P_{ik}, P_{j\ell}\right)\right)$ is a distance decay function of the Euclidean distance between the points P_{ik} and $P_{j\ell}$,

$I\left(P_{ik,} P_{j\ell}\right)$ is an indicator function, equal to one if the line segment joining points P_{ik} and $P_{j\ell}$ passes through the intended region, and otherwise equal to zero,

$B\left(P_{ik,} P_{j\ell}\right)$ is an indicator function, equal to one if the line segment joining points P_{ik} and $P_{j\ell}$ crosses a barrier, and otherwise equal to zero,

$\vee$ is the Boolean 'or' operator representing $I \vee B$ to be equal to one if I and/or B is equal to 1, and $I \vee B$ to be equal to zero otherwise, and

dP_{ik} is the same as $dx_{ik} dy_{ik}$.

Here, a trip between any two points k and ℓ of catchment areas i and j (respectively) and passing through the intended region is determined by checking whether the line segment joining these two points passes through the intended region or whether the line segment joining these two points crosses a barrier. Thus, the Boolean function $I\left(P_{ik,} P_{j\ell}\right) \vee B\left(P_{ik,} P_{j\ell}\right)$ states that only trips passing through the intended region should be counted in the determination of the weights w_{ij}.

The distance decay function $g\left(d\left(P_{ik}, P_{j\ell}\right)\right)$ is a decreasing function of the Euclidean distance between the point k of catchment area i and point ℓ of catchment area j that adjusts for the declining likelihood of trips of increasing length. Also, the scaled distance decay function

$$\frac{g\left(d\left(P_{ik}, P_{j\ell}\right)\right)}{\int_{S_i}\int_{S_j} g\left(d\left(P_{ik}, P_{j\ell}\right)\right) dP_{ik}\, dP_{i\ell}}$$

represents the distance decay probability density function. Therefore, w_{ij} as defined by equation 2 gives the probability that a trip between the catchment area of an external station i and the catchment area of an external station j crosses the intended region. $W = \lfloor w_{ij} \rfloor$ is a symmetric matrix.

Often small- to medium-sized urban intended regions can be approximated by circular or rectangular or polygonal regions on a plane. The external weights defined by equation 2 can be evaluated numerically in these instances. A circular region in particular has broad applicability for smaller urban areas and will be treated in some depth.

EXTERNAL WEIGHT MATRIX FOR AN CIRCULAR REGION

In many practical applications intended (study) regions may be approximated by a circle on a plane. Although it is not necessary to have a full circular intended region (any pie segment will do), it is specified as circular for making equation 2 mathematically manageable. Figure 1 illustrates such a circular region. This method can be defined, roughly implementing the mathematics of the previous section, in four steps as follows:

Step 1 *Approximate the Intended Region Using a Circle*.
Let $E_1, E_2, \ldots\ldots, E_n$ denote the Cartesian coordinates of n external stations, respectively. Find the minimum covering circle by any acceptable method. Let *C* denote the center of the smallest radius circle that covers all *n* external stations, and r_1 be the radius of this circle. Let this circle represent or approximate the intended region.

Step 2 *Approximate the External Territory Using an Annular Region.*
Define a circle having the same center *C* and radius r_2 such that $r_2 > r_1$, where r_2 is selected judiciously to approximate the external territory of the intended region by the annular region thus defined (see figure 1).

Step 3 *Construct Catchment Areas S_i's*.
Under the absence of any travel barriers in the external territory, the number of catchment areas is equal to the number of external stations. Define catchment areas S_i, i =1, 2, ..., n, where S_i's are mutually exclusive and exhausts the external territory such that those trips that originate in catchment area S_i use its external station E_i, whenever the trips travel through the intended region. The external stations that do not lie on the boundary of the inner circle are forced to lie on the boundary of the inner circle by projecting them on to the boundary along the ray joining the center *C* and these external stations. For example, see the external station E_2 in figure 1.

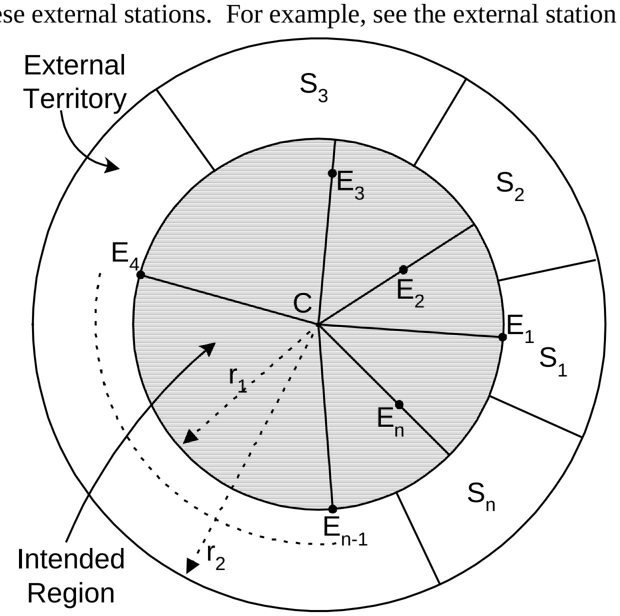


Figure 1. The Intended Region and the External Territory for a Circular Region

The external territory is divided into catchment areas by using the perpendicular bisectors of the line segments $E_1E_2, E_2E_3, ..., E_{n-1}E_n$, and E_nE_1 (see figure 1). Thus, catchment area S_i is defined by the perpendicular bisectors of the line segments $E_{i+1}E_i$ and E_iE_{i-1} for i = 2, ..., n-1. The catchment area S_1 is defined by the pair of perpendicular bisectors of the line segments E_nE_1 and E_1E_2, and the catchment area S_n is defined by the pair of perpendicular bisectors of the line segments E_nE_1 and E_nE_{n-1}, as shown in figure. 1.

Step 4 *Determine the Probability of a Trip between Each Pair of Catchment Areas that Pass through the Intended Region*

In this case, for each pair of catchment areas the trip between each pair of points, one point lying in one area and the other point lying in the other area will be treated as the line segment joining these two points. The numerical evaluation of equation 2 particularly can be exploited for the circular case as discussed below.

For this purpose, first transform equation 2 from Cartesian to polar coordinates. Then remove the uninhabited regions in the external territory by introducing a pair of fictitious external stations for each uninhabited region that is present. Finally, introduce a barrier function for each uncrossable region that is present in the external territory.

Evaluation of the Weight Matrix for the Circular Region

Consider w_{ij} as defined in equation 2 and transform Cartesian coordinates to polar coordinates. Then w_{ij} becomes

$$w_{ij} = \int_{r_1}^{r_2}\int_{\varphi_{1i}}^{\varphi_{2i}}\int_{r_1}^{r_2}\int_{\varphi_{1j}}^{\varphi_{2j}} g'\left(d\left(P_{ik}, P_{j\ell}\right)\right)\left(I\left(r_{ik}, r_{j\ell}, \varphi_{ik}, \varphi_{j\ell}\right) \vee B\left(r_{ik}, r_{j\ell}, \varphi_{ik}, \varphi_{j\ell}\right)\right) r_{ik} r_{j\ell} dr_{ik} d\varphi_{ik} dr_{j\ell} d\varphi_{j\ell} \tag{3}$$

where

$$g'\left(d\left(P_{ik}, P_{j\ell}\right)\right) = \frac{g\left(d\left(P_{ik}, P_{j\ell}\right)\right)}{\int_{S_i}\int_{S_j} g\left(d\left(P_{ik}, P_{j\ell}\right)\right) dP_{ik} d_{j\ell}},$$

r_{ik} is the Euclidean distance of point P_{ik} of catchment area S_i from the center *C*,

$r_{j\ell}$ is the Euclidean distance of point $P_{j\ell}$ of catchment area S_j from the center *C*,

φ_{1i} and φ_{2i} are the polar angles made by the boundaries of catchment area S_i $(\varphi_{1i} < \varphi_{2i})$,

φ_{1j} and φ_{2j} are the polar angles made by the boundaries of catchment area S_j $(\varphi_{1j} < \varphi_{2j})$,

$I\left(r_{ik}, r_{j\ell}, \varphi_{ik}, \varphi_{j\ell}\right)$ is an indicator function and is equal to one if the trip between the points (r_{ik}, φ_{ik}) of catchment area i and $(r_{j\ell}, \varphi_{j\ell})$ of catchment area j passes through the intended region, and is equal to zero otherwise,

$B\left(r_{ik}, r_{j\ell}, \varphi_{ik}, \varphi_{j\ell}\right)$ is an indicator function and is equal to one if the line segment joining two points (r_{ik}, φ_{ik}) of catchment area i and $(r_{j\ell}, \varphi_{j\ell})$ of catchment area j crosses a barrier, and is equal to zero otherwise, and

$\vee$ is the Boolean 'or' operator as defined earlier.

Removing Large Uninhabited Areas in External Territory

Suppose that there are large uninhabited areas such as bodies of water and mountains that must be incorporated in the mathematical formulation of the external weight given by equation 3. For each uninhabited area that is present in the external territory introduce two adjacent fictitious external stations such that the union of their catchment areas form the approximate area covered by the uninhabited area (see case study example). Once the external weights are obtained eliminate the rows and columns corresponding to the fictitious external stations from the weight matrix, W, since their purpose is served and they are no longer needed to derive the through trip table.

Description of a Barrier Function

A barrier function is used in equation 2 to allow a trip between a pair of points lying in two different catchment areas to go through the intended region whenever the line segment joining these two points crosses a barrier. A barrier is introduced for each uncrossable area that is present in the external territory (see for example, figure 3 depicting two barriers for the purpose of removing two lakes that are present in the external territory). As seen in figure 3, a barrier can be represented by a radial line passing through any arbitrary point within the uncrossable area.

Finding Minimum Covering Circle

Although the evaluation equation 2 looks overwhelming, the computations are numerically tractable. Using a spreadsheet type of application software such as Microsoft Excel, the center and the radius of the minimum covering circle can be found by solving the following convex programming problem:

Minimize $\quad d \quad$ (4a)

Such that $\quad (x_i - x_0)^2 + (y_i - y_0)^2 \le d \quad$ for i = 1, 2, ..., n (4b)

where,

(x_i, y_i) are the Cartesian Coordinates of external station i,

(x_0, y_0) are the Cartesian coordinates of the center *C* of the minimum covering circle of the external stations whose values need to be determined, and

d represents the squared distance of the maximum distant external station from the point (x_0, y_0).

Discussion of the Model

The circular model is not applicable to every city, although it should perform well in most isolated small cities and in many other larger cities, depending upon their geographical attributes. However, the circular model is unlikely to work well in communities where the external stations have been located haphazardly or arbitrarily. For example, the circular model will not give reasonable estimates for parts of nonconvex regions where truly internal trips may need to pass through two external stations, briefly exiting and then entering the region. Of course, the model assumes that all external-to-external trips behave as through trips, first entering and then exiting.

Usually concavity or other anomalies that affect the validity of a geographically-based through trip model can be easily determined by visually inspecting the locations of the external stations relative to the road network. Those external stations which violate the assumptions of the model should be removed from the analysis and handled as special cases.

One set of data not used by the model, but can be potentially helpful is the orientation of the link at an external station. The circular model assumes that links passing through external stations are radials. Although perfect radials are not essential for getting good results, those external stations that grossly violate the radial assumption should also get treatment as special cases.

CASE STUDY I: OSHKOSH, WI

The case study chosen to illustrate the major steps in finding external weights to calculate a through trip table is Oshkosh, WI. Oshkosh, a small city in east-central Wisconsin, is bounded on the east by Lake Winnebago and is bisected by the Wolf River, as illustrated in figure 2. Oshkosh's population is about 55,000. External travel is also substantially restricted by Lake Buttes des Morts and a string of other sizable lakes northwest of Oshkosh. The network, routinely used for local traffic forecasting in Oshkosh, was obtained from the Wisconsin Department of Transportation (WisDOT). The network was not modified in any way for this test, even though one external stations was located incorrectly and two other external stations could have been moved further away from the CBD to obtain better results. The network was used to ascertain the Cartesian coordinates of the external stations and to determine the shortest paths between external stations.

The network appears to be more rectangular then circular in shape. Nonetheless, the city was assumed to have a pie shape with a large slice removed for Lake Winnebago and a smaller slice removed for Lake Buttes des Morts.

Oshkosh had collected through trip tables for automobiles and trucks, separately. Because of the limited sample size for trucks, comparisons are made only to the total vehicle trip table.

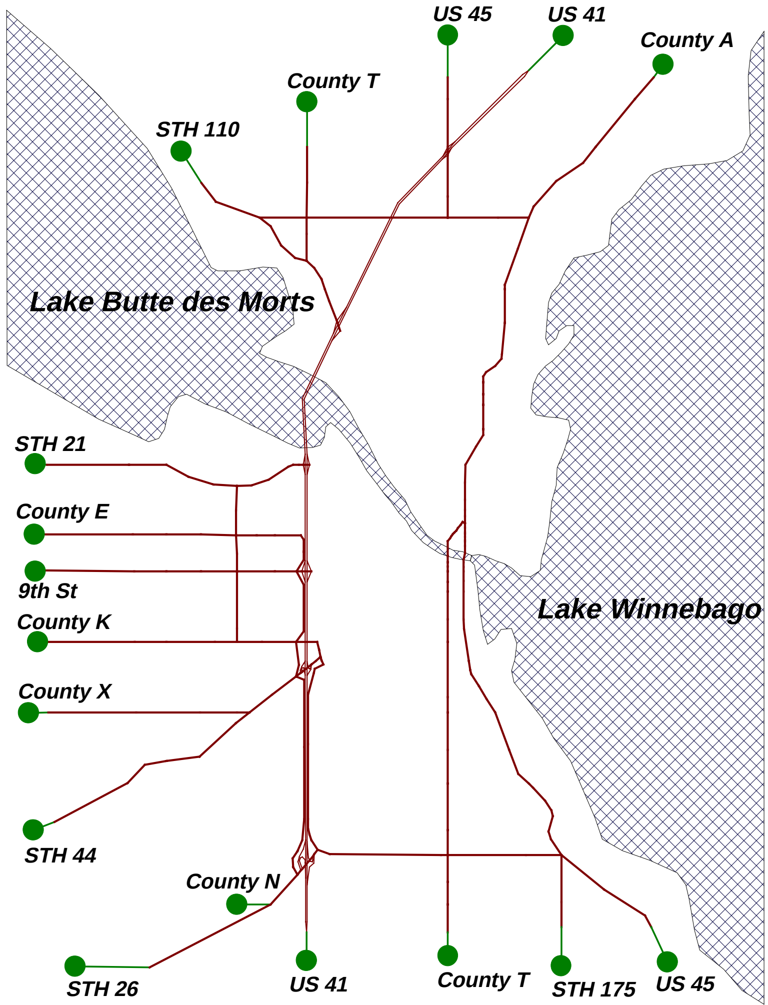


Figure 2. Oshkosh External Stations and Links on Shortest Paths between External Stations

The minimum covering circle for Oshkosh's external stations is illustrated in figure 3. In this case, the circle is defined by the two points lying farthest apart: STH 26 and County A. US 45 South lies close to the circle. Figure 3 also shows the locations of four dummy external stations that were required to remove the effect of the two neighboring lakes from the external weight matrix. Catchment areas for selected external stations are shown as shaded wedges. This figure shows the outer boundary to the catchment areas as having twice the radius of the covering circle; however, the radius of the outer boundary could be of any size. The analysis consisted of 21 catchment areas, 17 of which are associated with true external stations. Radials passing through two of the dummy external stations, one for each lake, are designated as barriers to travel.

The distance decay function arbitrarily selected to be similar to a gravity model friction factor function:

$$g\left(d\left(P_{ik}, P_{j\ell}\right)\right) = \beta e^{-\beta d(P_{ik}, P_{j\ell})} \qquad (5)$$

where d(·) is the Euclidean distance between two points in miles. Judging from gravity model calibrations in smaller cities, β should have a value of about 0.1 when d is measured in miles. The external weights were calculated by numerical integration as indicated previously for a circular region. The outside radius of the external territories was arbitrarily chosen as three times the value of the inside radius (22.92 miles and 7.64 miles, respectively). Each catchment area was divided into 400 cells -- 20 cells along a radial and 20 cells circumferentially.

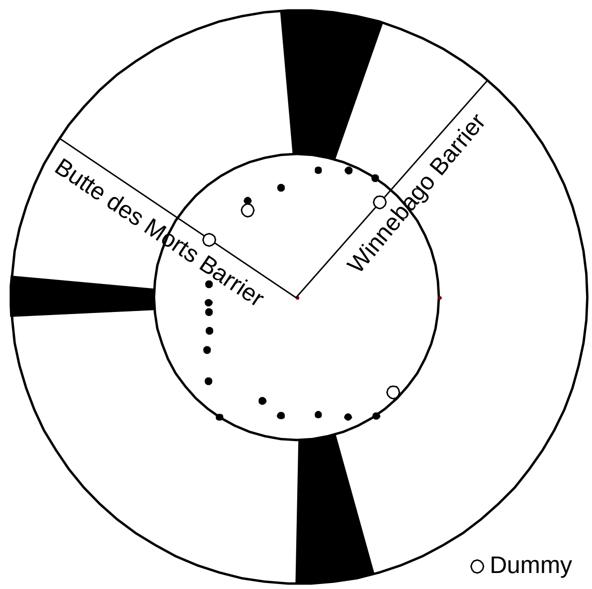


Figure 3. Minimum Covering Circle for Oshkosh, Barriers to Travel, Catchment Areas

CASE STUDY II: TALLAHASSEE, FLORIDA

Tallahassee, Florida, has two separate networks for travel forecasting purposes -- one used by the MPO and the state department of transportation and one used by the City of Tallahassee. They share many common elements, but differ somewhat in how they handle external stations. The MPO's network has 30 external stations, many of them quite minor, located along the border of the boot-shaped county of Leon. The only major population center in the county is the city of Tallahassee, which is located at the ankle of the boot. The City's network has just 13 external stations. The City's network was selected in order to eliminate all the external stations for which data was unavailable and to reduce the effect of arbitrarily locating external stations along the boundaries of an odd-shaped county. For the purposes of this analysis, one additional external station was eliminated from consideration because it contributed few trips and because it was located quite distant from all other external stations (in the toe of the boot). Including that external station would have almost doubled the area of the minimum covering circle. The remaining 12 external stations are shown in figure 4.

Tallahassee was selected to help validate the observations from Oshkosh and introduce additional considerations into the analysis. Although Tallahassee, like Oshkosh, is a radial city, it is free of strong barriers to travel, has a major interstate highway (I-10), and has a much larger network. All parameters were set identically to Oshkosh, except the radii of the outer and inner circles (46.8 and 15.6 miles, respectively). The outer circle clips a portion of Apalachee Bay on the Gulf of Mexico, possibly distorting the catchment areas for the four southern external stations. The outer circle also extends into Georgia, on the north. Furthermore, the inner circle, having a radius of 15.6 miles encloses an undesirably large amount of rural land.

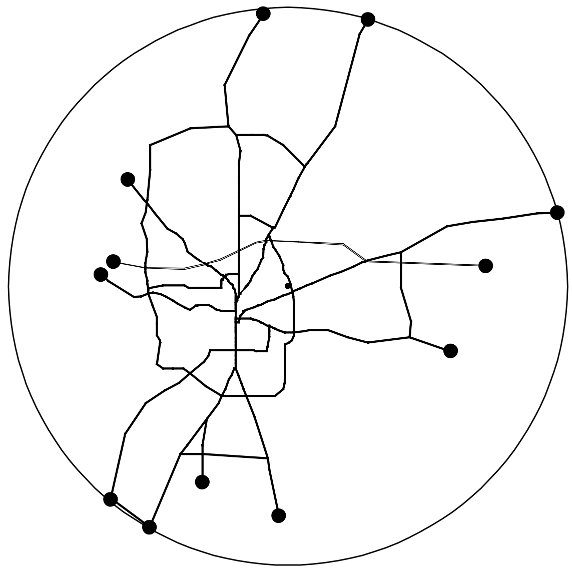


Figure 4. City of Tallahassee External Stations, Links on Shortest Paths Between External Stations, and Minimum Covering Circle

RESULTS

Computation of Through Trips

First the weight matrix W was computed using the procedure described earlier. For comparison purposes, three different cases of weight matrices were developed for Oshkosh

A. A weight matrix was developed from the theoretical model, with the coefficient β set equal to 0.1 in equation 5. The weights obtained using the proposed model were found to be insensitive to the values of β in the neighborhood of 0.1.

B. A weight matrix was prepared with all weights of 1's, except the diagonal elements that were set equal to zero (w_{ij} = 1 for i $\neq$ j, and = 0 otherwise).

C. A weight matrix was prepared with all weights set equal to gravity-model friction factors. Specifically, the following friction factors were used:

$$w_{ij} = 0.1e^{-0.1u_{ij}},$$

where u_{ij} is the over-the-network travel time in minutes between external stations i and j.

A typical value of the β coefficient for gravity models is about 0.1 for small urban areas, and this value was used to construct the weight matrix.

No subjective adjustments to the weights were made. The Furness procedure was run for fifteen iterations in all three Oshkosh cases to assure that the observed and estimated destination totals were reasonably equal for each external station. Of course, the estimated and observed origin totals are always equal with this procedure. Then for each case, the total sum of squares of errors were computed to obtain the percent of total variation (R^2) in observed trips about their mean value explained. The percent of total variation explained in each case is summarized in table 1. The percent variation explained was computed excluding the diagonal elements in cases A and B.

Table 1: Percent of Variation Explained by Three Synthetic Trip Tables, Oshkosh

Method of Computing Weights	Percent of Total Variation Explained
A. Probability from Catchment Areas	95.58%
B. All Unity	88.45%
C. Gravity Model Friction Factors	77.22%

It can be seen from table 1 that the proposed weight matrix explained is the highest percent of variation. This is not surprising as case A captures a large portion the underlying geographical information of the problem, while case B contains no information regarding geography, and case C misuses the geography. On balance, the results from the proposed procedure surpassed those of Modlin's model, although it is recognized that Modlin also tried to explain the ratio of through to non-through trips at external stations.

It is often difficult to interpret R^2 statistics for trip tables, so figure 5 shows desire lines between external stations for both actual and estimated trips (case A). Lines are drawn in proportion to the logarithm of the sum of volumes in both directions. OD pairs with fewer than 1 trip have been omitted. The logarithmic transformation tends to overemphasize the importance of low-volume OD pairs and visually exaggerates the error. Nonetheless, the desire line plots illustrate the overall agreement between the model and actual data.

In Tallahassee the model explained 99.9% of the variance in the trip table, after 50 iterations of the Furness procedure. Again, weights were not adjusted subjectively. The R^2 statistic was seriously inflated by the large amount of through traffic on I-10, although the overall fit was still judged to be good. The desire line plot, logarithmically transformed, for Tallahassee is shown in figure 6. Overall, the model only slightly outperformed the "all unity" case (case B in the previous section), which had 99.4% of the variance explained. Nonetheless, the model helped explain the largest proportion of the variance left unexplained by the "all unity" case.

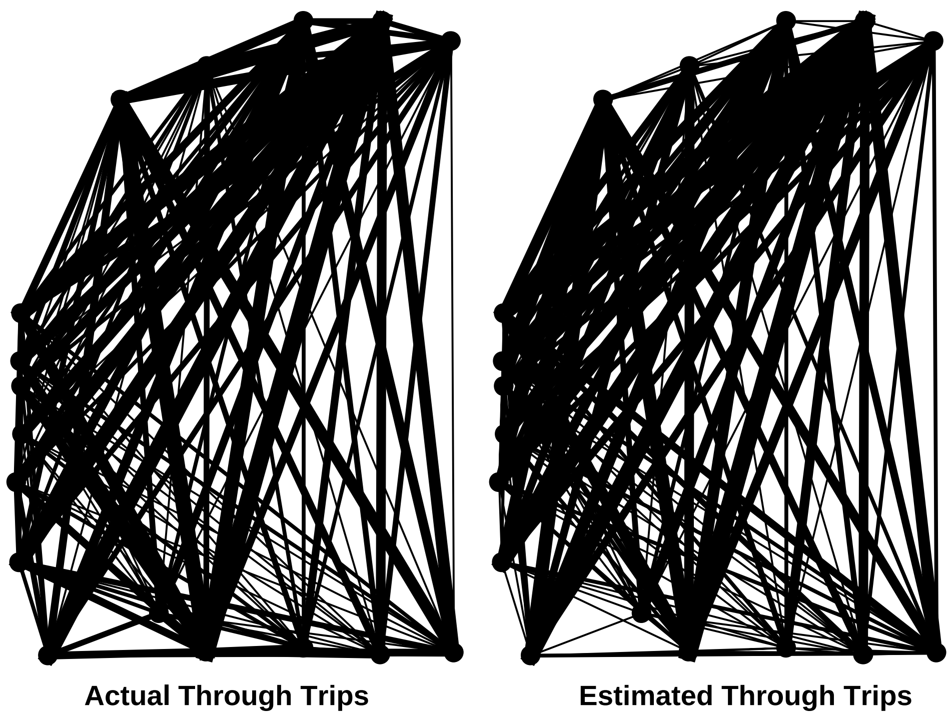
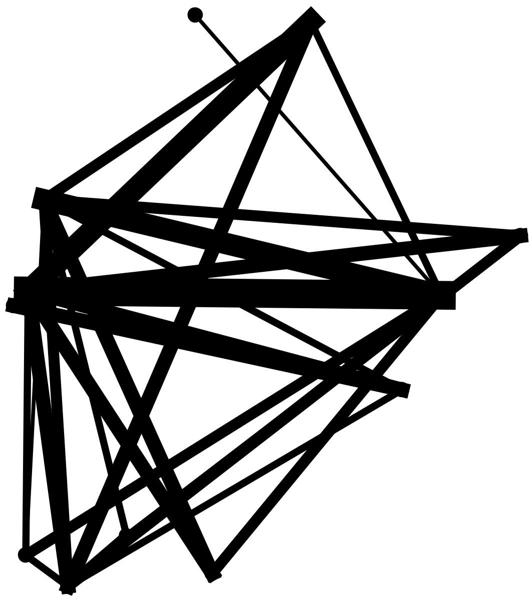
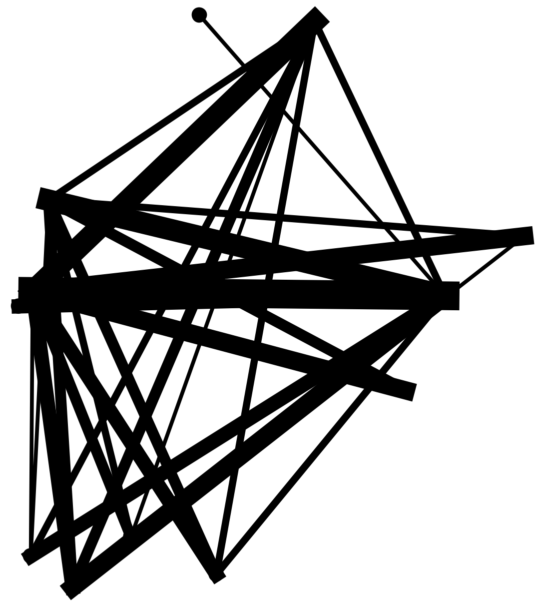


Figure 5. Comparison of Actual v. Estimated Desire Lines (Logarithmically Transformed, Widest Line is 17100 Vehicles) for Oshkosh





Actual Through Trips   Estimated Through Trips

Figure 6. Comparison of Actual v. Estimated Desire Lines (Logarithmically Transformed, Widest Line is 8800 Vehicles) for Tallahassee

Assignment of Through Trips

A more relevant comparison is the ability of the synthetic through trip table to produce reasonable assigned volumes. Although the actual through volumes are unknown, they can be approximated by assigning the actual through trip table to the traffic network. First, an all-or-nothing assignment was performed on an arbitrary through trip table (all elements positive) to identify those links that might be involved in carrying through trips. For Oshkosh the statistics in table 2 apply only to those links, as shown in figure 2, and for Tallahassee the comparison applies only to those links shown in figure 4.

Table 2 compares assigned volumes for Oshkosh from an actual through trip table to those of three synthetic trip tables: probability weights from catchment areas, all weights unity, and gravity model friction factors. The average through volume in each Oshkosh case was 1836 vehicles.

Setting all weights (except those on the diagonal) to unity can be interpreted as the absence of all spatial information. As can be seen in table 2, the RMS error from the method of this paper (probabilities from catchment area) produced an acceptable RMS error of only 10%. When all weights were set equal to one, the RMS error rose to 41%. The gravity model produced a remarkably bad RMS error of 179%.

In Tallahassee a similar comparison between actual and estimated link volumes gave an RMS error of 28 vehicles or 14%, roughly comparable to Oshkosh. The "all unity" case gave a slightly larger RMS error of 29 vehicles.

Table 2. RMS Errors of Assigned Volumes with Three Different Synthetic Trip Tables, Oshkosh

Method of Computing Weights	RMS Error
A. Probability from Catchment Areas	185.9
B. All Unity	753.8
C. Gravity Model Friction Factors	3294.6

DISCUSSION

The assigned link volumes, based upon the estimated trip table, must be checked against known ramp volumes at interchanges. If an assigned ramp volume exceeds the known volume, then adjustments to the estimated weight matrix are necessary. Such an adjustment requires identification of pairs of external stations making contributions to each violated ramp. The weights corresponding to these pairs of external stations must be scaled down appropriately. This step was not needed for either Oshkosh or Tallahassee.

In Oshkosh, the most noticeable errors in the trip table were observed to occur between state trunk highways (STH) and US highways. In many cases the model underestimated the actual number of trips between these two functional classes. This underestimate may be due to STH's serving as feeders to US highways for long distance travel. It is very difficult to incorporate this situation into the model without detailed, time-consuming study of the whole region. Feeder information, where available and appropriate, might be useful for manually fine-tuning of the estimated trip table.

Inspection of the errors in the trip tables, cell by cell, suggests that the model sometimes underestimates the number of trips between external stations on the same side of the city. It possible to anticipate some of this underprediction; the cause in Tallahassee and Oshkosh relates to travel between pairs of small population centers that are just outside of the region. The orientation of the links at the external stations, not used in the model presented here, is another clue to identifying problem OD pairs.

Although the model performed quite well in explaining Tallahassee's through trip table, it only slightly out-performed the "all unity" or zero information case. It was observed that a single large highway can distort standard measures of goodness of fit.

CONCLUSIONS

It is possible to estimate a through trip table from knowledge of through traffic volumes at each external station and the rudimentary geography of a region. The trip table is most easily calculated when the region is convex, and is even easier to calculate when the region can be approximated as a circle. However, the method illustrated here for through trip table estimation may perform poorly for regions that seriously violate one or more assumptions: severely nonconvex regions, regions where external stations have been arbitrarily located, regions where there are significant number of trips in the through trip table that are not true through trips, and regions where major highways make sharp turns or just clip a border.

Geographical data required for through trip table estimation include the physical locations of the external stations, the location of large tracts of uninhabited land, and the presence of barriers to travel. Through trip table estimation does not necessarily require information from a traffic network, although the orientation of links as the boundaries of the region may help in identifying problems and special cases.

A simplified procedure for through trip table estimation was applied without any calibration to Oshkosh and Tallahassee, achieving excellent results. The synthetic trip table explained almost 96% of the variation in the actual trip table in Oshkosh and almost 100% of the variation in Tallahassee. A better measure of the practical quality of the model, the RMS error in link volumes (from through trips only and for only links that are candidates for through trips), was approximately 10% and 14%, respectively. The quality of these results exceed those of earlier modeling efforts and greatly exceed those obtainable from a traditional gravity model.

The simplified version of the through trip table estimation procedure could be used in many small- to medium-sized urban areas, as is. However, geographic concepts of the model have practical significance in most cities when developing sampling schemes for purely empirical through trip tables, when extrapolating or plugging holes in empirical trip tables, when creating through trip tables by judgment, or when exploring the possibility of fitting a synthetic trip table to ground counts by statistical means.

REFERENCES

1. Pigman, Jerry G., Simulation of external-internal and through trips for small urban areas, Research Report 493, Division of Research, Kentucky Bureau of Highways, February 1978.
2. Modlin, David G., Jr., Synthesized through-trip table for small urban areas, <u>Transportation Research Record</u> 842, 1982, 16-21.
3. Cambridge Systematics, et al., Quick response freight manual, Federal Highway Administration, September 1996.
4. Fisk, C. S., On Combining maximum entropy trip matrix estimation with user optimal assignment, <u>Transportation Research B</u>, 24B, 1990, 363-373.
5. Bell, Michael G. H., The estimation of origin-destination matrices by constrained generalized least squares, <u>Transportation Research B</u>, 25, 1989, 13-22.

6. Lam, William H. K. and H. P. Lo, Estimation of origin-destination matrix from traffic counts: a comparison of entropy maximizing and information minimizing models, Transportation Planning and Technology, 16, 1991, 85-104.

ACKNOWLEDGEMENTS

Data for Oshkosh was supplied by the Wisconsin Department of Transportation. Data for Tallahassee was supplied by the Florida Department of Transportation and the City of Tallahassee.