ORIGIN DESTINATION DISAGGREGATION USING FRATAR BIPROPORTIONAL LEAST SQUARES ESTIMATION FOR TRUCK FORECASTING

University of Wisconsin – Milwaukee
Paper No. 09-1

National Center for Freight & Infrastructure Research & Education
College of Engineering
Department of Civil and Environmental Engineering
University of Wisconsin, Madison

Author: Alan J. Horowitz
Center for Urban Transportation Studies
University of Wisconsin – Milwaukee

Principal Investigator: Alan J. Horowitz
Professor, Civil Engineering and Mechanics Department, University of Wisconsin – Milwaukee

February 13, 2009

Origin Destination Disaggregation Using Fratar Biproportional Least Squares Estimation for Truck Forecasting

Abstract

This working paper describes a group of techniques for disaggregating origin-destination tables for truck forecasting that makes explicit use of observed traffic on a network. Six models within the group are presented, each of which uses nonlinear least-squares estimation to obtain row and column factors for splitting trip totals from and to larger geographical areas into smaller ones. The techniques are philosophically similar to Fratar factoring, although the solution method is quite different. The techniques are tested on a full-sized network for Northfield, MN and are found to found to work effectively.

Introduction and Mathematical Underpinnings

It is often desirable to obtain a highly detailed origin-destination table for vehicles or commodities, when only a much more aggregated table is available. These situations typically arise when survey data are organized into fairly large districts (zip codes, cities, counties or states) in order to preserve confidentially or simply to provide meaningful flow comparisons when the number of data samples is limited. Commercial vehicle and freight data, in particular, are prone to this type of spatial aggregation.

For the purposes of this discussion, the aggregated OD table will be said to contain trip data between "districts", while the disaggregated OD table will be said to contain trip data between "zones". Traditional practice has been to disaggregate a district-level origin-destination table by factoring it along its rows and columns, simultaneously. That is:

$$T_{ij} = A_i B_j \tau_{kl} \tag{1}$$

where:

i = an origin (row) in the disaggregrated (zonal) table and where i is an element in the set of zones I;

j = a destination (column) in the disaggregated (zonal) table and where j is also an element in the set of zones I

k = an origin (row) in the aggregated (district-level) table and k is an element in the set of districts K;

l = a destination (column) in the aggregated (district-level) table and l is also an element in the set of districts K;

T_{ij} = the disaggregated origin-destination table, zone-to-zone;

A_i = a row split factor for each zone table origin;

B_j = a column split factor for each zone table destination;

τ_{jl} = the aggregated origin-destination table, district-to-district.

The sets of splits, A_i and B_j, have the effect of spreading a large number of trips between an origin and a destination into smaller numbers of trips between, perhaps, many origins and many destinations. Each i and each j is associated with one and only one k or l, respectively. So for notation purposes, it is necessary to define two further groups of sets: L_k^A and L_l^B, which keep track of the structural relationship between the two tables. That is,

L_k^A = the set of *i* rows that are associated with row *k* in the aggregated district table;
L_l^B = the set of *j* columns that are associated with column *l* in the aggregated district table.

Zones nest into districts and no zone may occur in multiple districts. Each district table index can be computed as a function of a given zonal table index. This is, when *i* is known, then *k* can be found by referencing the set, L_k^A.

It should be recognized that this traditional practice ignores the possibility that there are special zone-to-zone interactions that are hidden in the aggregation at the district-to-district level. For example, a large factory might ship to a large warehouse, creating a particularly large OD flow between two zones that might not be apparent by just looking at the flow between the two respective districts.

Origin-destination tables are often thought to be symmetric over a 24-hour period for passenger travel; however, commodity flow tables cannot be assumed to be symmetric and vehicle flow tables, both passenger and freight, are rarely symmetric for periods of time shorter than a day.

Although the term "origin-to-destination" is used in this discussion, the procedures developed herein are equally applicable to "production-to-attraction" flows for passenger travel and "production-to-consumption" flows for commodities.

The amount of data available to determine A_i and B_j varies considerably depending upon the planning problem. Very often planners will calculate the splits from socioeconomic data or by applying trip generation equations, as they might have been prepared for a travel forecasting model. Another common method is to determine the splits by observing the amount of travel in each disaggregated zone, such as the zone's VMT (vehicle miles of travel).

Another possible data source is traffic counts on individual links. Individual traffic counts are difficult to use directly for determining the splits because any one count is not usually associated with any specific zone. Indeed, the relationship between a traffic count and the number of trips that are generated in a nearby zone, as will be seen, is quite complex.

Recent work on estimation of origin-destination tables from traffic counts has direct implications for the OD table disaggregation problem. One technique in particular, Fratar biproportional least-squares estimation, can be suitably modified to create needed row and column splits. In particular, Fratar biproportional estimation seeks the solution of this nonlinear, least-squares minimization problem (Horowitz, 2005) to obtain sets of row and column factors to refine a rough (or "seed") table at the same level of aggregation:

$$\min P = \sum_{\forall a \in A} w^a \left(C^a - s \sum_{\forall i \in I} \sum_{\forall j \in I} p_{ij}^a x_i y_j T_{ij}^* \right)^2 + z \sum_{\forall i \in I} \sum_{\forall j \in I} T_{ij}^{*2} \left(1 - x_i y_j\right)^2 \qquad (2)$$

where

x_i = row (origin) factor for zone *i*;
y_j = column (destination) factor for zone *j*;

C^a = ground count for link direction a, with each direction on two-way links tabulated separately, and a is an element in the set of all counted directions A;
T_{ij} = number of trips between origin i and destination j to be estimated;
T_{ij}^* = seed trip table;
p_{ij}^a = estimated proportion of trips between zones i and j that use link direction a (as determined by an equilibrium traffic assignment);
I is the set of zones, i = 1 to N or j = 1 to N;
A is the set of link directions;
w^a = link weight for link direction a;
z = the trip table weight; and
s = a scale factor that is either set to 1 or selected automatically to scale the trip table to produce the correct average traffic count before optimization.

For example, this equation might be useful for approximating a peak-hour origin-destination table, zone-to-zone, from peak-hour traffic counts and from a 24-hour origin-destination table, also zone-to-zone. Seed tables are often built from survey data, behavioral travel theory or expert judgment. The estimation finds the best compromise set of origin and destination factors that gives good agreement with traffic counts and does not deviate hugely from the seed table. It is also mathematically necessary to constrain the factors to be greater than zero, and it is quite desirable in most circumstances to keep them within reasonable bounds.

If the aggregated, district-level OD table is perfect, then the following relationship must hold:

$$\tau_{kl} = \sum_{\forall i \in L_k^A} \sum_{\forall i \in L_l^B} T_{ij} \tag{3}$$

However, it is entirely possible that the district-level OD table is less than perfect, because it too is subject to various data collection errors or inadequacies in theory. In such cases, it may be appropriate to avoid using Equation 3 as a strict constraint.

It is highly likely that a planner can find some data to suggest how the district-level OD table might be disaggregated, justifying the use of splits, A_i and B_j, at least tentatively. However, traffic counts might suggest that different splits are better for the purpose. Therefore, Equation 1 should be modified to include the information coming from all sources:

$$T_{ij} = s x_i A_i y_j B_j \tau_{kl} \tag{4}$$

where x_i and y_j are empirical modifiers, somehow derived from traffic counts, of A_i and B_j. Thus, s, x_i and y_j have similar purposes to the same variables in Equation 2.

There are limits as to how many x's and y's can be estimated, given the amount of data available to the problem from the district-level OD table and from the traffic counts. It is also entirely possible that a given zone's traffic (origins or destinations) might not travel on any of the counted links; in such cases the x's and y's must default to 1.0, with all the factoring carried by the predetermined s, A's and B's.

Beyond Equation 3, a zonal OD table that perfectly conserves trips must require these relationships to be satisfied:

$$\sum_{\forall i \in L_k^A} x_i A_i = 1, \quad \text{for all district rows, } k$$

$$\sum_{\forall j \in L_l^B} y_j B_j = 1, \quad \text{for all district columns, } l$$

However, a perfectly-conserving zonal OD table may not be desired because of data collection errors, so these relationships could be considered just approximate for many situations.

The strength of the solution depends upon how much traffic data are available and the sizes of the two OD tables. There are 2*N* variables, where *N* is the number of zones in the zonal OD table. For example, if the district-level OD table has 12 districts and if there are 400 traffic counts, then there are 544 data items in the estimation (12^2+ 400). This means that the estimation can reliably expand this table to at most 272 zones, i.e., half the number of data items, but probably a lot less. The locations of the 400 traffic counts matter. More splits can be estimated when the traffic counts are spread evenly throughout the region and where the counting stations are located on major roads.

There are a number of different ways to formulate the estimation methodology, and six of these ways are discussed in this paper, along with some extensions:

Model I: District-level OD table is perfect
Model II: District-level OD table is approximate
Model III: District-level OD table is perfect, OD's are affected by trip utility
Model IV: District-level OD table is approximate, OD's are affected by trip utility
Model V: District-level OD table is approximate, link-to-link flows are available
Model VI: District-level OD table is approximate, some zone-to-zone flows are special

Model I: District-Level OD Table Is Perfect

The least squares estimation for this model tries to match ground counts while fitting the district-level OD table exactly.

$$\min P = \sum_{\forall a \in A} w^a \left(C^a - s \sum_{\forall i \in I} \sum_{\forall j \in I} p_{ij}^a x_i A_i y_j B_j \tau_{kl} \right)^2 \tag{5a}$$

where *k* and *l* are taken to be functions of *i* and *j*, respectively. These constraints must hold.

$$\tau_{kl} = s \sum_{\forall i \in L_k^A} \sum_{\forall i \in L_l^B} x_i A_i y_j B_j \tau_{kl}, \ \forall k \in K, \forall l \in K \tag{5b}$$

$$x_i \geq 0, \ \forall i \in I$$

$$y_j \geq 0, \ \forall j \in I$$

This method can be implemented within the mathematics of Model II, if a suitably large value of *z* (see next section) is selected to assure that the first constraint is satisfied.

As a practical matter, all of the models must account for the presence of external stations on the network. There can be no "intrazonal" trips within external stations. This is perhaps a minor detail, but it can be handled by introducing another factor G_{ij} (near where A_i and B_j appear in all expressions) which applies to all zones within the set E of external stations, where $E \subset I$. Thus,

$$G_{ij} = 0 \text{, if } i = j \text{ and } i \in E \tag{6a}$$

$$G_{ij} = 1 \text{, otherwise} \tag{6b}$$

This same variable should be introduced in the implementation of any of the models described here.

The scale factor, s, has been retained from Equation 2. A scale factor can help eliminate systematic errors in data collection or adjust for different sets of units. For example, it is conceivable that an aggregated freight OD table can be given in units of tons of while the link volumes can be given in units of trucks. At the surface, s appears to be entirely redundant. However, upper and lower bound constraints placed on x's and y's can sometimes make it desirable to keep them close to 1, and a value of $s \neq 1$ allows this to happen more readily.

Model II: District-Level OD Table Is Approximate

Model II is more consistent with past practice in OD table estimation from ground counts, where the seed table is considered to be, at best, a rough approximation of reality. Assuming that the district-level OD table is approximate provides some flexibility to the estimation and recognizes that there may be serious inconsistencies between the ground count data and the survey data that were used to build the district-level OD table.

$$\min P = \sum_{\forall a \in A} w^a \left(C^a - s \sum_{\forall i \in I} \sum_{\forall j \in I} p_{ij}^a x_i A_i y_j B_j \tau_{kl} \right)^2 + z \sum_{\forall k \in K} \sum_{\forall l \in K} \left(\tau_{kl} - s \sum_{\forall i \in I} \sum_{\forall j \in I} x_i A_i y_j B_j \tau_{kl} \right)^2 \tag{7}$$

$x_i \geq 0, \quad \forall i \in I$

$y_j \geq 0, \quad \forall j \in I$

Again, k and l are functions of i and j. Judgment as to which are most accurate, either ground counts or aggregated OD flows, is expressed by the set of link-direction weights, w^a, and by the sole table weight, z. The actual effects of these weights are not obvious and the effects are best evaluated after the optimization has been completed.

Model III and IV: OD's Are Affected by Trip Utility

Model III (perfect district-level table) can be treated as a special case of Model IV (approximate district-level table). Both models introduce the idea from a gravity model of trip distribution or a logit model of destination choice that there is less likelihood of a trip between a pair of zones if there is considerable spatial separation between them. Spatial separation is measured by a traveler's "utility", which is almost always increasingly more negative or less positive as trip distance increases. Most travel forecasting models calculate a value of utility primarily from the travel time between the two zones. Define:

U_{ij} = utility of travel from zone i to zone j; and
V_{kl} = utility of travel from district k to district l;

Utility in these cases are deterministic and can be obtained directly from the traffic network. The district-to-district utility of travel may be found by taking a weighted average of all zone-to-zone utilities. Thus,

$$V_{kl} = \frac{\sum_{\forall i \in L_k^A} \sum_{\forall j \in L_l^B} x_i A_i y_j B_j U_{ij}}{\sum_{\forall i \in L_k^A} \sum_{\forall j \in L_l^B} x_i A_i y_j B_j} \tag{8}$$

which would need to be recomputed repeatedly as the set of x's and y's become better known. If the A's and B's are known fairly well, then initially, $x_i = 1$ and $y_j = 1$.

Assuming a logit or maximum-entropy relationship for destination choice, then the following correction, F_{ij}, might be needed for zone pairs from particularly large districts:

$$F_{ij} = e^{U_{ij}} \big/ e^{V_{kl}} , \tag{9}$$

and k and l are functions of i and j.

Within the framework of a typical travel forecasting model, utility would likely be calculated from the travel time between two zones. That is,

$$U_{ij} = \phi t_{ij} \tag{10}$$

where

t_{ij} = the travel time between zones i and j, and
ϕ = a negative constant.

Model IV's objective function can be obtained by slightly enhancing Model II:

$$\min P = \sum_{\forall a \in A} w^a \left(C^a - s \sum_{\forall i \in I} \sum_{\forall j \in I} p_{ij}^a x_i A_i y_j B_j F_{ij} \tau_{kl} \right)^2 + z \sum_{\forall k \in K} \sum_{\forall l \in K} \left(\tau_{kl} - s \sum_{\forall i \in I} \sum_{\forall j \in I} x_i A_i y_j B_j F_{ij} \tau_{kl} \right)^2 \tag{11}$$

ϕ could be adopted from a trip distribution model for the region, or it could be obtained directly through the optimization process.

Model III could be implemented by setting z to a large number.

Model V: District-level OD Table Is Approximate, Link-to-Link Flows are Available

Model III could be further expanded to include information coming from toll road transponders. Readers could be placed throughout the community, and it would then be possible to gain a better understanding of the movement of traffic. Reader locations are neither origins nor destinations, but are points in between. Data collected from many readers could greatly enhance

OD table disaggregation beyond what could be accomplished through ground counts alone, because these data embody spatially precise flows. Since not every vehicle has a transponder, any reader location also needs to be a count location, so reader locations are a subset of links in the set of counted links, $R \subset A$. Reader locations would be ordered by time of day, so:

g = the first reader on a trip, closest to the origin; and
h = the second reader on a trip, closest to the destination.

It is possible for a vehicle to traverse many readers along a single trip, but all reader data need to be organized into pairs of locations to be compatible with a method known as "select link analysis" within a travel forecasting model. So further define:

p_{ij}^{gh} = estimated proportion of trips passing between link directions *g* and *h* that have their origin in *i* and their destination in *j* (as determined by an equilibrium traffic assignment); and
Q_{gh} = expanded, sampled number of trips between link directions *g* and *h*, as ascertained by reader and count data.

$$\min P = \sum_{\forall a \in A} w^a \left(C^a - s \sum_{\forall i \in I} \sum_{\forall j \in I} p_{ij}^a x_i A_i y_j B_j \tau_{kl} \right)^2 + z \sum_{\forall k \in K} \sum_{\forall l \in K} \left(\tau_{kl} - s \sum_{\forall i \in I} \sum_{\forall j \in I} x_i A_i y_j B_j \tau_{kl} \right)^2 + r \sum_{\forall g \in R} \sum_{\forall h \in R} \left(Q_{gh} - s \sum_{\forall i \in I} \sum_{\forall j \in I} p_{ij}^{gh} x_i A_i y_j B_j \tau_{kl} \right)^2 \quad (12)$$

where *r* is a weight on the transponder component.

Model VI: District-Level OD Table Is Approximate, Some Zone-to-Zone Flows Are Special

In some circumstances it might be necessary to account for known high interactions between specific pairs of zones. The number of such zone pairs must be kept to just a few so as not to overwhelm the estimation process. Therefore, it is assumed that these zone pairs can be identified in advance, even if the actual level of interaction is unknown. Such interactions can be incorporated into the model by defining a mask H_{ij} for special zone pairs:

$H_{ij} = 1$, if the interaction between *i* and *j* is special, and (13a)
$H_{ij} = 0$, otherwise. (13b)

The additive adjustment for special trips between an OD pair, M_{ij}, can then be inserted into any of the previous models. For example, for Model II the objective function becomes:

$$\min P = \sum_{\forall a \in A} w^a \left(C^a - s \sum_{\forall i \in I} \sum_{\forall j \in I} p_{ij}^a A_i B_j \left(x_i y_j + H_{ij} M_{ij} \right) \tau_{kl} \right)^2 + z \sum_{\forall k \in K} \sum_{\forall l \in K} \left(\tau_{kl} - s \sum_{\forall i \in I} \sum_{\forall j \in I} A_i B_j \left(x_i y_j + H_{ij} M_{ij} \right) \tau_{kl} \right)^2 \quad (14)$$

and $M_{ij} \geq 0$,

if it is assumed that this variable is used only to increase the number of trips for a zone pair, which would be typical. But more generally,

$$x_i y_j + H_{ij} M_{ij} \geq 0,$$

which allows M_{ij} to be either positive or negative.

Further Variations

Constraints on Zonal Factors. Additional constraints on x_i and y_j could be imposed to keep the results within reasonable bounds throughout the optimization and in the results. These constraints override the nonnegativity constraints. Thus,

$$x_{\min} \leq x_i \leq x_{\max}, \quad \forall i \in I \tag{15}$$
$$y_{\min} \leq y_i \leq y_{\max}, \quad \forall i \in I$$

where both x_{min} and y_{min} are greater than zero.

Single Station OD Survey. Depending upon the availability of data, further additions could be made to the methodology. For example, it might be possible to include single-station origin-destination (SSOD) survey data into the objective function. An SSOD survey is typically conducted on a rural road, perhaps at a weigh station or at a rest area. Each driver is asked to give his/her origin and destination among other information. SSOD surveys are comparatively inexpensive. Unfortunately, SSOD data are heavily biased toward long trips and toward trips that just happen to be passing by that particular weigh station or rest area. However, the select link analysis feature in travel forecasting models can emulate the results from an SSOD survey with all its issues, so SSOD survey data can be included in the objective function without distortion.

Multiclass Traffic Assignments for Trucks. Truck routing in urban networks would likely be influenced by congestion on streets that carry both freight and passenger traffic, with the largest number of vehicle serving passengers. To properly handle this case, it is necessary to implement the estimation methodology within a multiclass equilibrium traffic assignment. Ground counts would be provided only for trucks, but the p_{ij}^a array would be influenced by the presence of passenger vehicles that exacerbate delays for trucks along streets and at intersections.

Bilevel Solution Algorithm

The OD disaggregation problem, any model, is solved by embedding it within a travel forecasting framework, as illustrated in Figure 1. Two separate input OD tables must be provided: the district-to-district OD table and a seed zone-to-zone OD table. The reason for the seed zone-to-zone table is to obtain a traffic assignment that can be used to compute the p_{ij}^a array and to obtain an initial set of delays on links and at intersections. A good source of a seed zone-to-zone OD table is Equation 1. The seed zone-to-zone OD table does not directly contribute to the creation of origin and destination factors, but the seed table is retained in the MSA averaging process and can slightly influence assigned volumes.

Bilevel algorithms similar to Figure 1 must be used when the OD table has not been fully determined at the point of initial traffic assignment and congestion is present on the network. An accurate traffic assignment requires accurate link delays that require the correct loadings, which can only be found for congested networks after a sizable number of MSA iterations.

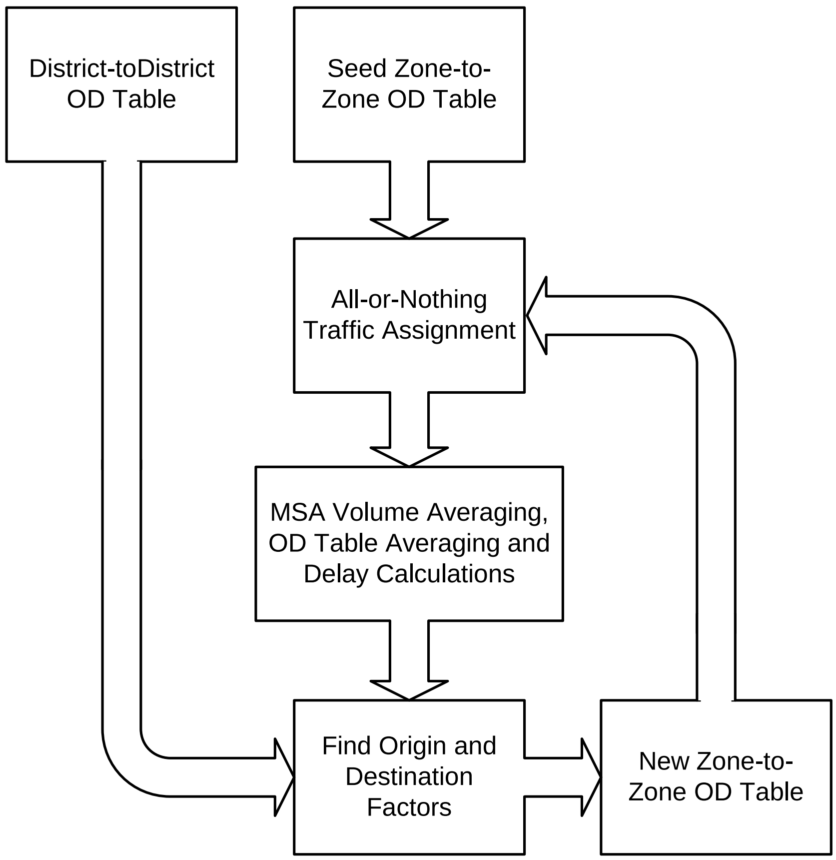


FIGURE 1. Bilevel Algorithm for Solving the OD Table Disaggregation Problem

Computational Considerations

Models I through V are constrained nonlinear minimization problems in 2N dimensions. Beyond dimensionality, computation time and memory usage are also heavily influenced by the number of link directions on the network, which has direct impact on the number of nonzero elements in the p_{ij}^a array. The size of the aggregated (district-level) OD table has little effect on computational efficiency.

Memory usage is largely dictated by the p_{ij}^a array, even when stored in a manner that squeezes out all zero elements. Computation tests presented in this paper were done within a 32-bit Windows environment, which has a practical memory limit of 3 GB. For example, a moderate-sized metropolitan network with 2000 zones and an average of 250 counted links used between each zone pair in a multipath traffic assignment would have a nonzero p_{ij}^a array size of 10^9 cells, which would more than fill any 32-bit Windows computer.

Computation time can also be an issue, as large problems could take days or even weeks of computation time on a standard desktop computer. In order to reduce computation time, the experimental software has been written to simultaneously use all processors in a multiprocessor computer. Substantial parallel processing power is now available in affordable desktop computers that come equipped with four, eight and even 16 processors.

The algorithm for solution is the gradient projection method with PARTAN. Searches in the gradient projected direction are stopped when the step size, η, decreases beyond:

$$\eta < \theta\sqrt{2N} \tag{16}$$

where θ is a suitably small number and 2N is the number of variables. The optimization is terminated when the relative change in the objective function between PARTAN steps is smaller than another small arbitrary number, determined through a trial and error process.

Computational Tests, All Vehicle OD Table Estimation in Northfield

The Northfield, MN network was selected for testing the computational properties of OD table dissaggregation. These tests involved passenger, commercial and freight vehicles in a single-class traffic assignment. The Northfield network is shown in Figure 1. It has 29 zones and 12 external stations, which were organized into 11 districts. External stations were treated similarly to zones in the tests. Because all streets were included in the network, there were 819 links, but just 60 link directions had traffic counts. The number of ground counts is undesirably less than the number of variables. In Figure 1, zone boundaries are included as thin black lines. The spider-like structures show the relationship between zones and districts.

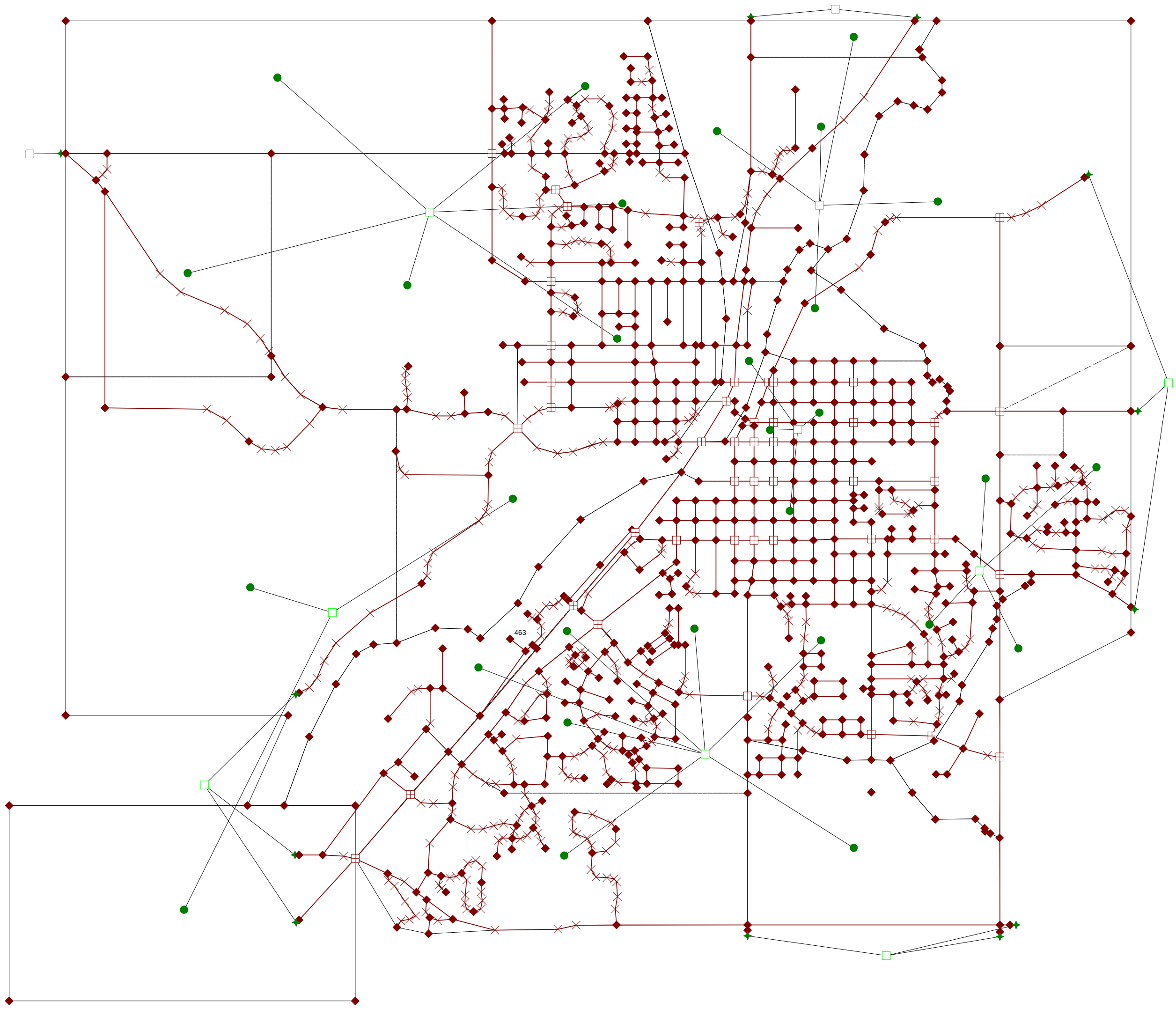


FIGURE 2. Northfield Test Network

The aggregated OD table was created by a gravity model at the zone-to-zone level with home-based-work, home-based-nonwork, and nonhome-based trip purposes for passengers, then aggregated to the district level. The zone-to-zone OD table was retained for comparison purposes. All parameters were taken from NCHRP Report #365 for passenger travel. Because the aggregated OD table omitted any consideration of freight or commercial vehicles, substantial disagreement with the ground counts was anticipated. So not only would the model be expected to disaggregate the OD table, but it would also be expected to correct for errors inherent in the aggregated OD table caused by omitting many trucks.

Zonal characteristics were available that could have permitted the creation of fairly good sets of zone splits, A_i and B_j, but there was a particular interest in seeing what a cruder set of zone splits would accomplish. So for these tests all A_i's and B_j's were set to the reciprocal of the number of zones in their respective districts.

Models I through IV apply to the Northfield case. Optimization parameters were set as follows:

- All link weights, w^a, were set to 1;

- The OD table weight, z, was set to 100 for Models II and IV and 10,000 for Models I and III;
- There was prior scaling of the OD table, i.e., $s \neq 1$, to at least account for the omission of trucks from the district-to-district OD table;
- All x's and y's were constrained to be between 0.2 and 5.

All simulations and optimizations used "area spread" equilibrium traffic assignment, which loads traffic at almost all intersections and dispenses with centroid connectors, which are common devices in travel forecasting networks. This assignment method is able to assign the vast majority of intrazonal trips to the network and is highly multipath. Equilibrium was achieved by running 40 iterations of the method of successive averages (MSA), which is more than in most travel forecasting applications, but not sufficient to reduce convergence error to a negligible amount. The time period of the simulations was a full 24-hours.

A statistical summary of the four models (I through IV) and an ordinary simulation are shown on Table 1. Data are given for the 60 link directions with ground counts.

TABLE 1. Summary of Computational Tests, Inexact Input Data

Model	Average Ground Count	Average Assigned Link Volume	RMS Difference in Volumes	RMS Difference in Aggregate OD Table
I	3840	3781	1152	0.3
II	3840	3814	1008	6.9
III	3840	3851	1385	1.6
IV	3840	3862	1013	9.2
Simulation	3840	2594	2238	0.0

The RMS difference in the OD table for the simulation was zero because the zone-to-zone OD was not changed during the simulation and the district OD table was built by aggregating the zone-to-zone OD table. The average aggregated OD flow was 554 vehicles.

The simulation performed very poorly in matching ground counts in relation to Models I to IV, even though it had the advantage of being provided a zone-to-zone OD table (41 by 41) instead of a district-to-district (11 by 11) OD table. A major contributor to the error of the simulation was an average of approximately 1200-vehicle systematic underestimate of all ground counts; presumably many of these were trucks. Another possibility for the underestimate is that the simulation does not have enough congestion to create diversion due to equilibrium effects and traffic is unrealistically being kept on routes with slight advantages in free travel time and were not counted.

Models III and IV did slightly worse than Models I and II, which was unexpected given that the district-to-district OD table was created with a gravity model and Model IV differs from Model II by making gravity-type adjustments. Table 1 shows that Model I preserves the district-to-district OD tables (to within one-half of a trip), with only a slight increase in the RMS difference between the forecast and the ground counts over Model II. Even the 9.2 trip error in the aggregated OD table for Model 4 is not large.

As an example, Figure 3 shows a map of the computed destination factors from Model II. These destination factors have been multiplied by the scale factor, s. Darker red hatching indicates zones that have destination factors between 0.2 and 0.4 which the darker blue hatching has destination factors between 3.2 and 4. The spider-like structures have been emphasized to show the relationship between zones and districts. A zone could have a large destination factor because it has more activity, overall, than its companion zones or because it has activities that generate a disproportionate amount of travel not accounted for in the NCHRP Report #365 parameters, such as truck travel. The map confirms intuition by having about as many blue zones as red zones in each district. The results for origin factors and for Model II were similar. It is difficult to further interpret Figure 3 without considerable local knowledge.

The scale factor, s, was selected by the algorithm to be between 1.12 and 1.17, depending upon the model. The 1996 Quick Response Freight Manual (Cambridge Systematics, 1996) states that commercial vehicles make up 10.5% of traffic on urban principal arterials, so these scale factors are somewhat greater than what would be expected if they were only accounting for the absence of trucks in the district-to-district OD table.

All of zones changed from their original values of x_i and y_j of 1.0, and all of the factors fell easily within the constraints of Equation 15. These initial tests of Models I to IV demonstrate that they can give plausible results, but these tests do not demonstrate that the results are accurate.

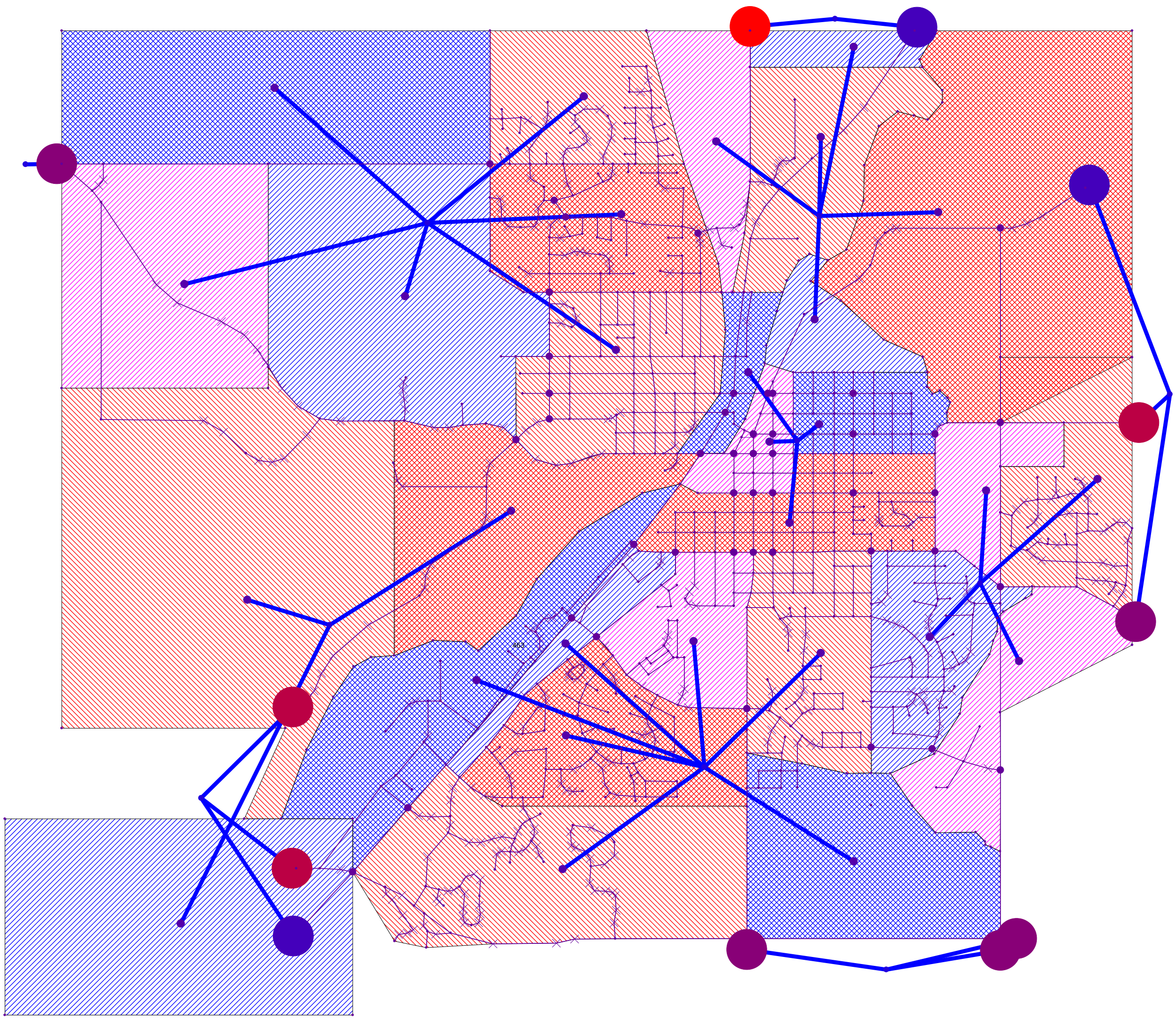


FIGURE 3. Destination Factors for Model IV at Zones (Shaded Areas) and at External Stations (Dots) (Darker Blue Indicates Larger Factors, Darker Red Indicates Smaller Factors and Magenta is Neutral)

To better gauge the accuracy of the models in reconstructing an underlying zone-to-zone OD table, these steps were performed:

1. Create a reasonable zone-to-zone OD table for Northfield, in this case by adopting the output table from the last iteration of Model II of the previous tests. Retain the origin and destination factors.
2. Build a district OD table from the zone-to-zone OD table.
3. Assign the zone-to-zone OD table to the network and obtain link volumes.
4. Set the "ground count" on all arterial links equal to the computer link volumes.
5. Using the district-to-district OD table and the "ground counts", compute a new zone-to-zone OD table.
6. Compare the original and new origin and destination factors.

This procedure resulted in 527 “ground counts” on the network. For these tests, the OD table weight, z, was left at from 100, even though there were a greater number of ground counts than the earlier tests. It should be noted that the input zone-to-zone table was computed from Equation 4, so the relationship between the zonal and district-level OD tables perfectly adheres to the theory, and the computed “ground counts” are consistent with both OD tables. Table 2 shows that the optimization, as expected, is finding a solution that is very close to an exact fit to both the district-to-district OD table and the “ground counts”. The average OD flow in the district-to-district table was 623 vehicles, so the errors in matching the aggregated OD table is just 1% and the error in matching ground counts is less than 3%. The remaining small differences between assigned volumes and “ground counts” in the OD table are attributed mainly to convergence error of the equilibrium traffic assignment algorithm. The test was repeated by eliminating the bounds on x’s and y’s (Equation 15) to determine of these bounds were inhibiting the estimation process. The results in Table 2 for the unconstrained optimization, although slightly improved, indicate that the constraints were not a serious issue in finding origin and destination factors.

TABLE 2. Summary of Computational Tests, Artificial Input Data

Model	Average Ground Count	Average Assigned Link Volume	RMS Difference in Volumes	RMS Difference in Aggregate OD Table
II Constrained	2138	2103	59	6.5
II Unconstrained	2138	2104	53	5.7

The most interesting outputs of these tests are shown in the scatter charts of Figures 4 and 5, which compare the results of the optimization with the known origin and destination factors. The original and computed sets of factors compare very well to each other.

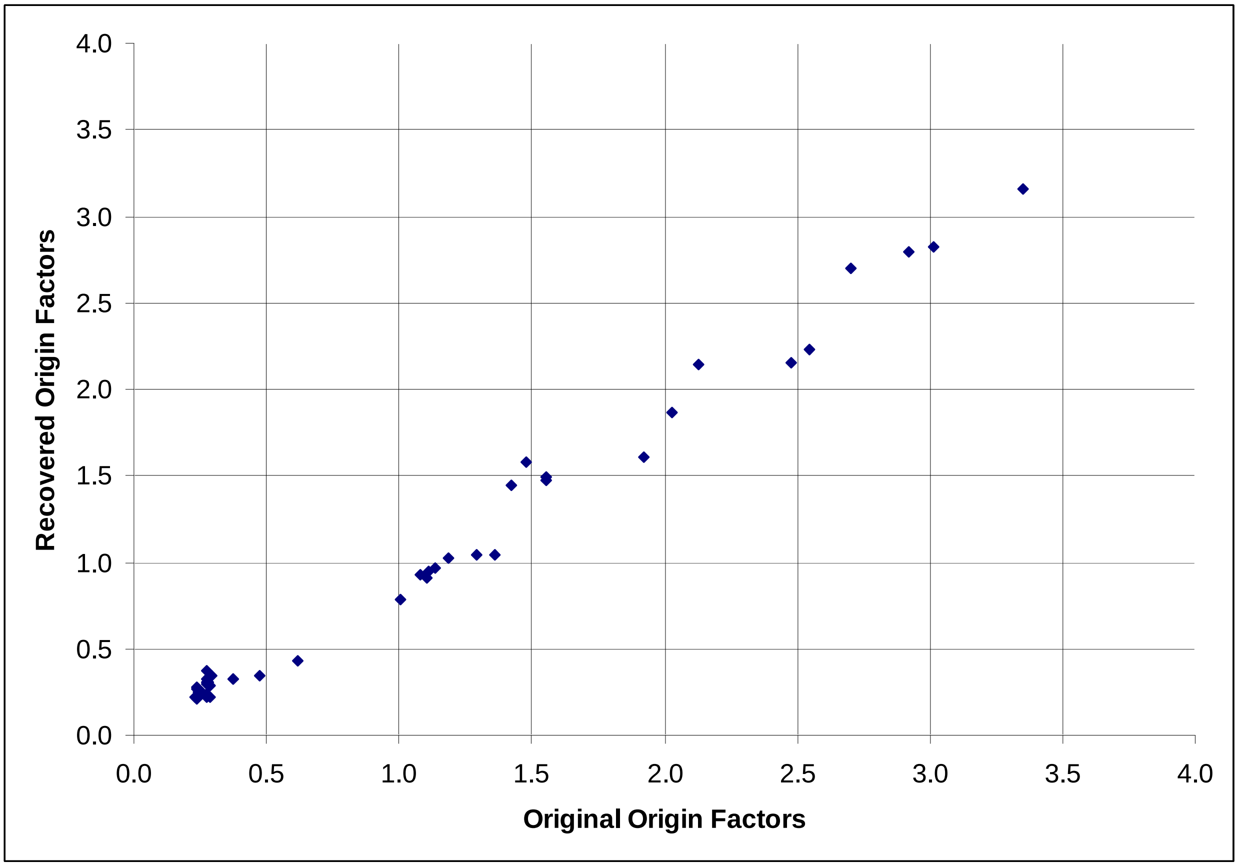


FIGURE 4. Scatter Chart Showing the Computed Origin Factors Against the Known Origin Factors

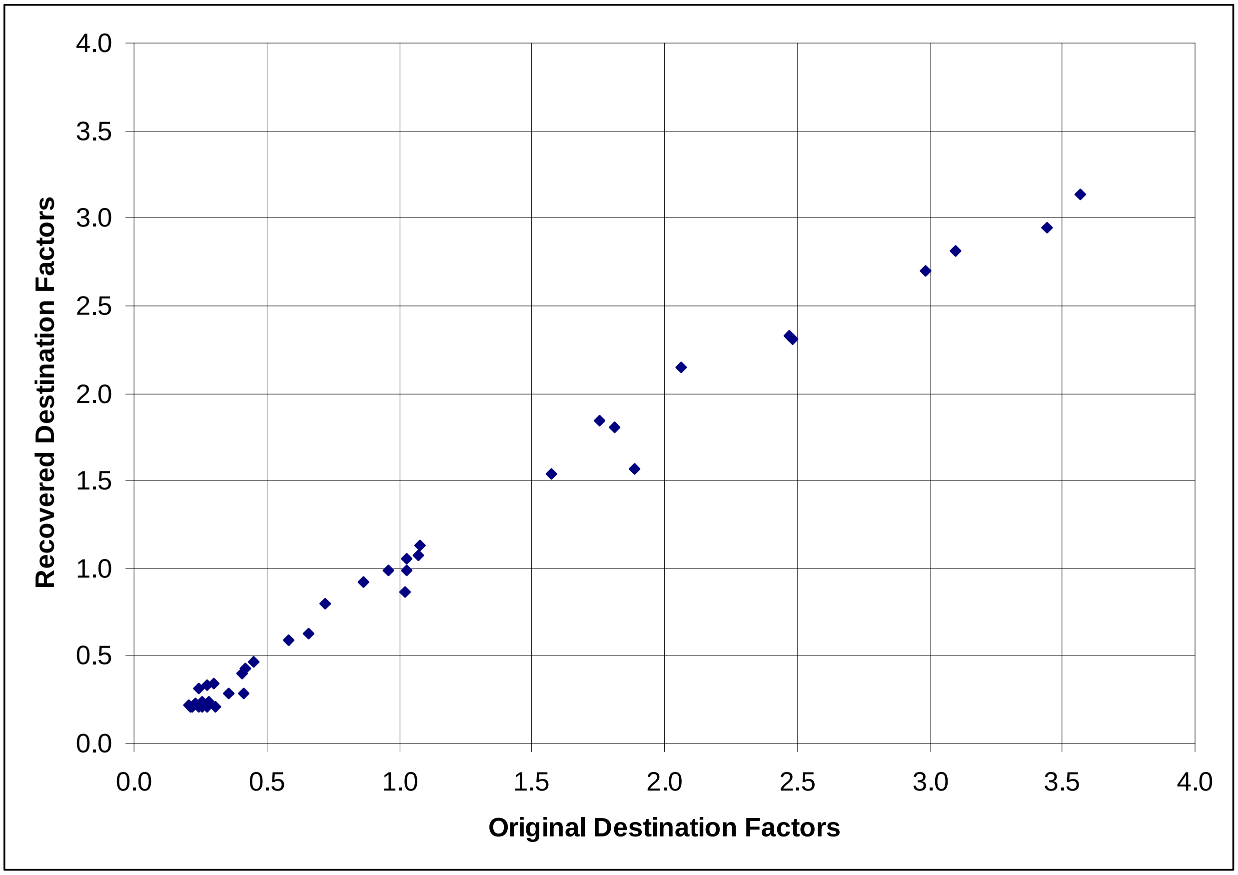


FIGURE 5. Scatter Chart Showing the Computed Destination Factors Against the Known Destination Factors

The tests of Table 2 and Figures 4 and 5 were idealized so that they could be readily interpreted. They do not represent the sternest test of the models and solution algorithm. However, these tests indicate that if the data fit the theory of the model, if there is consistency between the aggregated OD table and ground counts and if there are sufficient number of ground counts, a least squares optimization can do very well in estimating origin and destination factors.

Conclusions and Recommendations

This paper outlines several optimization models that can disaggregate origin destination tables by using information from ground counts. Four of the optimization models were tested on real data from Northfield, MN and were found to work effectively.

One of these optimizations models was tested on realistic, but artificial data, on the same Northfield network and was found to be able to accurately reproduce known underlying origin and destination factors in ground counts.

These methods developed in this paper are intended for commercial vehicle or freight forecasts because origin-destination data for these flows are often aggregated to a level where they are no longer useful for planning purposes.

Additional tests are needed on full-scale freight networks.

Acknowledgements

This study was funded by the Center for Infrastructure Research and Education, a national university research center of the US Department of Transportation.

References

Alan J. Horowitz, “Tests of a Family of Trip Table Refinements for Quick Response Travel Forecasting”, *Transportation Research Record Journal*, #1921, 2005, pp. 19-26.

Cambridge Systematics, et. al., “The Quick Response Freight Manual”, Federal Highway Administration, DOT-T-97-10, September 1996.

Appendix I: Setting Up and Running an OD Table Disaggregation

The OD table disaggregation procedures (Models I, II, III, IV and VI) are implemented within an experimental version of the Quick Response System II (QRS II) software. Networks are prepared with the General Network Editor (GNE).

Networks must be built with the QRSDynamicEx.dta application schema. Networks built using the QRSDynamic.dta can be upgraded to the correct schema using File Append within GNE.

Networks are prepared in the usual way, with ground counts entered for the "ground count" attributes for one-way and two-way street links. Links without ground counts should have the "ground count" attributes left at zero. Centroids and external stations must have unique names. Districts are established in the usual way, with each zone belonging to a single district. Each district tag must have a unique name. Initial origin and destination splits (*A*'s and *B*'s) are entered on centroids and external stations in the District Share: Origin and the District Share: Destination attributes, respectively.

There should be no demographic data on any of the centroids and no productions and attractions at external stations. However, if the traffic assignment has a time period of greater than 1 hour, then QRS II needs a way to determine the time-of-day of travel. To assure QRS II makes reasonably good time-of-day assumptions, it is necessary to place a very small number of dwelling units, retail employees, and nonretail employees on one and only one centroid (e.g., du = 0.1; re = 0.01; nre = 0.04). These values will create fractional intrazonal trips at this centroid, but gives QRS II enough information to compute a time-of-day distribution of traffic.

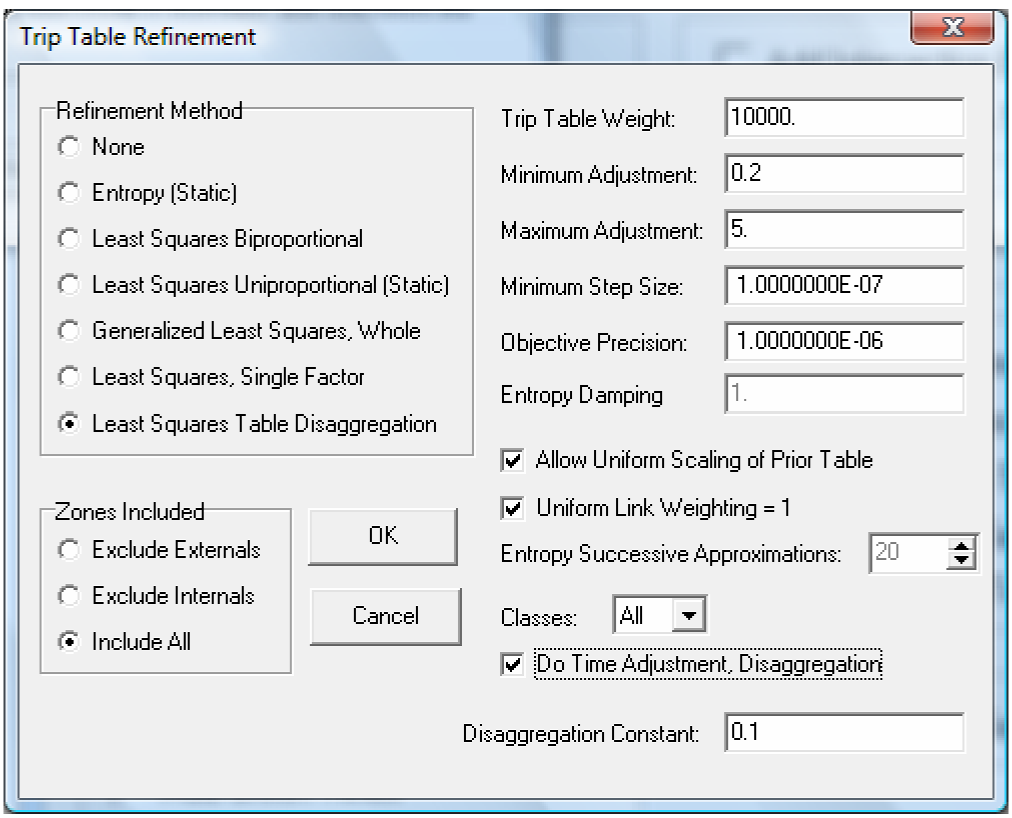
One "add" file is required, AddDTrips.txt. This file is given in the same format as AddVTrips.txt, as described in the QRS II Reference Manual. AddDTrips.txt contains the district-to-district OD table. Row and column names are the names of the district tags within the network. This file must be placed in the "project" folder for the run, i.e., the same location as Param.txt. The project folder is displayed on the QRS II main window. An example AddDTrips.txt file, organized as a single table, is shown below.

```
Externals 2
Externals 3
District1
District 2
Externals 1
District 3
District 4
District 6
District 5
Externals 4
Externals 5
END OF ROWS
Externals 2
Externals 3
District1
District 2
Externals 1
District 3
District 4
District 6
District 5
Externals 4
Externals 5
END OF COLUMNS
   0.0     325.0       0.2       0.3     216.0       0.2       0.0     0.2       0.1    2982.0     714.0
```

```
 343.0       0.0       0.2       0.1     934.0       0.1       0.0     0.1       0.0    1466.0    1535.0
   0.2       0.2    3519.9    1281.0       0.2    3000.1     278.6  1216.3    1640.5       0.2       0.1
   0.3       0.1    1287.2     932.3       0.3    1796.3     279.2   393.0    1072.1       0.2       0.2
 204.0     838.0       0.2       0.3       0.0       0.4       0.0     0.2       0.2     745.0      80.0
   0.2       0.1    3015.0    1801.7       0.4    3916.1     904.6   753.0    2975.6       0.2       0.3
   0.0       0.0     277.1     276.9       0.0     898.5     144.7   215.6     635.5       0.0       0.0
   0.2       0.1    1230.0     396.3       0.3     755.4     218.2   411.8    1047.0       0.5       0.2
   0.1       0.0    1644.2    1070.5       0.2    2966.8     639.1  1041.5    4031.2       0.3       0.3
3673.0    1770.0       0.2       0.2     858.0       0.2       0.0     0.4       0.3       7.0      21.0
 761.0    1548.0       0.1       0.2      77.0       0.3       0.0     0.2       0.3      21.0       0.0
END OF TABLES
```

An optional “add” file is AddODMask.txt. This file implements Model VI by telling QRS II which specific OD pairs should be given special treatment. Row and columns consist of centroid or external station names. Values are either 1 (special treatment) or 0 (no special treatment). OD pairs not shown are not given special treatment. AddODMask.txt is in the same format as AddVTrips.txt. QRS II needs to be alerted to the presence of AddODMask.txt on the Add Files dialog box. AddODMask.txt is placed in the project folder.

There are a variety of parameter settings for QRS II, depending on what must be accomplished. The parameter settings shown below on the Refinement dialog box are for a static equilibrium traffic assignment, all vehicle classes at once. Parameters not shown should be left at QRS II’s default or at the appropriate values for travel forecast in that particular network.



Least Squares Table Disaggregation must be selected. Minimum and maximum adjustments bound the values of x’s and y’s. The trip table weight is the value of z. This value must be determined experimentally. Minimum Step Size and Objective Precision are convergence criteria. Smaller numbers for these criteria make an optimization run longer and make the results more precise. Allow Uniform Scaling of Prior Table tells QRS II to find a value for the scale factor, *s*, before the optimization commences. Otherwise, *s* will be set equal to 1. Do Time Adjustment, Disaggregation invokes Models III and IV where the negative of the Disaggregation Constant is used to calculate the adjustments with Equation 15.

QRS II produces several reports of direct interest. Among these are:

ODFactors.txt	Contains the origin factors and destination factors (*x*'s and *y*'s) for each zone. The first value in the file in the first row is the scale factor, *s*. Rows 2 to N+1 contain the zonal factors. These rows are ordered as in Nodelabl.txt. Masked OD adjustments may appear at the end of this file.
AddATrips.txt	Contains the equilibrium averaged zone-to-zone OD table. As such, it combines the results of many optimizations.
AddRTrips.txt	Contains the last iteration zone-to-zone OD table. This table is computed by applying the factors in ODFactors.txt directly to the data in AddDTrips.txt.
Output.dta	Contains the assigned, equilibrium averaged volumes on the network (also found in LinkVols.txt).
History.txt	Contains optimization statistics and computation time.

It is possible to use the methods in this paper to disaggregate an OD table for a single vehicle class, but a multiclass assignment is necessary. These changes must be made to the above procedure to get a disaggregated table for a chosen class:

- A separate class other than the base class must be defined on the Multiclass dialog box by giving it a class letter (e.g., "T").
- The class must be chosen on the Refinement dialog box where it says "Classes".
- Data for other classes may be given within in the usual way, such as AddVTrips files, AddPTrips.txt, or through demographics variables on the network.

No changes are needed for AddDTrips.txt, except that AddDTrips.txt is specific to the chosen class. For a class to have different paths, its inputs to path building must differ in some important way from the base class (i.e., class specific extra times in an AddETime file).