28-Balakrishna

# ADVANCES IN ORIGIN-DESTINATION TRIP TABLE ESTIMATION FOR TRANSPORTATION PLANNING AND TRAFFIC SIMULATION

Ramachandran Balakrishna and Howard Slavin
Caliper Corporation

## Introduction

Travel demand matrices are crucial inputs to travel demand forecasting and traffic simulation applications. While it might be desirable to estimate these trip tables from survey data, this is not practical for large or complex networks due to the large sample sizes that would be required. Traffic counts, by comparison, are widely available and less costly to obtain making it highly desirable to build base year models that are consistent with measured counts. The problem of estimating trip matrices from counts has a long history in the transportation literature and a variety of methods are available for estimating the static matrices for a specific time period that are common to regional travel forecasting calibration efforts. While counts are often available for very short time intervals, this information is typically not used in travel demand modeling. However, for traffic simulation, dynamic counts are vital and estimation of dynamic trip tables that specify trips by origin, destination, and departure interval becomes an important, if not the most important problem in model development. Unfortunately, dynamic estimation of trip tables is vastly more complex and dynamic trip table estimation methods are best characterized as emergent, rather than proven.

## Static Trip Table Estimation

Planning studies have generally relied on *static* OD tables, containing total trip rates for extended time periods such as the AM and PM peaks or various off-peak periods. Early methods for estimating trip tables from link count data were based on simple trip distribution models of the gravity or entropy family and were estimated with linearized equations and least squares followed by non-linear least squares (McNeil and Hendrickson, 1985) and also by various maximum likelihood search processes (Van Zuylen and Willumsen, 1980).

The fact that answers are generated by these methods may tend to obscure the difficulties involved. For example, traffic counts are subject to measurement error, may be totally missing for some crucial locations, and may be grossly inconsistent with one another. More than one solution may therefore provide a reasonable fit to the observed counts. To eliminate this issue, some methods attempt to remain close to a prior or seed matrix. For example, the ME2 program (Willumsen, 1982) maximizes an "entropy" measure that keeps the estimated OD flows very close to user-specified target OD flows.

In the late 1980s and in the 1990s, a step forward was taken by the development of trip table estimation methods that were consistent with equilibrium traffic assignments. These have the virtue of seeking consistency with route choice behavior in the sense that an assigned matrix will be close to the measured counts. This is computationally much more intensive that the statistically estimated methods but has become practical due to vast improvements in computing power.

With respect to trip table estimation and equilibrium assignment, both single path and multiple path methods are encountered with the latter arguably quite preferable. Methods developed by Nielsen (1993, 1998) have been found to be very effective in empirical work. Nielsen's method has the advantage of treating counts as stochastic variables as well as working with any traffic assignment method. It therefore can be used with stochastic user equilibrium assignment as well as with user equilibrium assignment and also with transit assignment procedures. Caliper (2007) has extended Nielsen's method to include multiple user class matrix estimation, to incorporate turning counts as well as link counts, to constrain certain cells, and use statistical weights on the input counts. This method has been useful in estimating truck trip matrices for different size trucks. Spiess (1990) formulates a gradient-based approach to adjust a starting matrix so that it better reflects observed counts. The method is demonstrated using EMME/2 equilibrium assignments with a reasonably good starting solution.

Despite widespread use, some limitations of static O-D matrix estimation should be noted. First, within a fixed time period such as the AM peak period or peak hour, some trips have not been completed or may have started prior to the period. This leads to an inconsistency in accounting for tripmaking since the matrix is assumed to be assigned and the trips completed within the time period. Second, the quality of the estimation depends fundamentally upon the quality of the assignment model including its volume delay functions and convergence. It is well known that conventional assignment methods cannot take account of over-saturated conditions. Third, the path flows associated with equilibrium assignment are not unique. This suggests that the estimated O-D matrices are also not unique. Finally, since link volumes and speeds are dynamic in time, there is necessarily quite a bit of aggregation error in static trip table estimation processes.

**Dynamic Trip Table Estimation**

While static OD matrices may suffice for long-term planning, they do not capture the within-day temporal dynamics observed through peaking, queue formation and dissipation and spillbacks. Short-term planning studies therefore require *dynamic* OD tables in order to accurately measure the performance of the traffic network. Such tables represent trip departure rates during short time intervals such as 5 or 15 minutes.

Several approaches obtain dynamic OD profiles from one or more static OD tables. However, such methods are often not based on real traffic measurements; they use *ad hoc*, heuristic rules that are difficult to generalize beyond the examples used to derive them. For example, Boyles et al. (2006) generate a dynamic OD profile to match the total demand contained in a set of static tables. OD profiles generated by such methods do not attempt to match real-world data, need not reflect true traffic dynamics, and may even be unrealistic and counter-intuitive.

The widespread deployment of traffic surveillance sensors has made available a rich dataset of time-varying traffic measurements. Since these data are collected and archived continuously and automatically, they represent recent network conditions and contain useful indirect information about the underlying dynamic OD demand patterns. A logical approach is therefore to divide the analysis period into many short departure time intervals consistent with the data measurement intervals, and estimate an OD matrix for each interval so as to replicate the time-varying data..

Let the analysis period be divided into *H* departure time intervals, $h = 1,2,...,H$. Let the OD matrix for time interval *h* be denoted by $x_h$, and the measurements by $y_h$. Traffic measurements may be derived from various surveillance technologies. The most common measurements are vehicle counts obtained by loop detectors or roadside sensors. Other data include density, loop detector occupancy, speed, probe vehicle travel time and queue length.

The general dynamic OD estimation problem can thus be expressed as an optimization problem:

$$\begin{aligned} &\text{Minimize } z(x) = z_1\left(\hat{y} - y\right) + z_2(x - x^a) \\ &\text{subject to}: \hat{y} = Assign(x) \end{aligned} \tag{1}$$

where $z_1(\cdot)$ is a measure of fit between the measurements *y* and their modeled counterparts $\hat{y}$, the latter obtained by running an assignment model such as a dynamic traffic assignment (DTA), mesoscopic or microscopic simulator; $z_2(\cdot)$ is a measure of fit between the estimated OD flows and their seed values $x^a$. The subscript *h* has been dropped in the above notation for simplicity; the terms x, $x^a$, $y$ and $\hat{y}$ are assumed to consist of variables/data for all *H* intervals. Like in the static case, seed flows $x^a$ are required in order to eliminate the multiple solutions arising from sparse and inconsistent data coverage.

The most popular dynamic OD estimation method uses traffic count data and assignment matrices that are linear approximations of the function *Assign*($x$):

$$y_h = \sum_{p=h-p'}^{h} a_h^p x_p + v_h \tag{2}$$

where $a_h^p$ maps OD flows $x_p$ (departing in time interval $p$) to traffic counts $y_h$ in the current interval (Cascetta et al., 1993; Ashok, 1996); $v_h$ is an error term. This mapping accounts for the fact that the measurements (such as traffic counts) for interval $h$ will likely have contributions from vehicle departures in $h$ as well as prior intervals $h$-1, $h$-2, etc. Thus $p$' represents the longest trip length (in terms of time intervals). The assignment matrices may be computed from travel time estimates, like in Ashok (1996), or using the output of a mesoscopic or microscopic traffic simulator as in Tavana and Mahmassani (2000).

Since the assignment mapping is linear, the solution of the unconstrained least squares OD estimation problem using traffic counts is conceptually simple and involves a matrix inversion. However the simultaneous estimation of OD flows for all intervals requires the inversion of a massive block-diagonal assignment matrix (formed from individual $a_h^p$ matrices) that has been shown to be prohibitively expensive when compared with inverting just one assignment matrix (Cascetta and Russo, 1997; Toledo et al., 2003). The general approach thus has been to estimate the OD flows one interval at a time, or sequentially (rather than simultaneously), starting from $h$=1. The contribution of previous departure intervals to the current counts are now assumed to be fixed, and subtracted from $y_h$:

$$y_h - \sum_{p=h-p'}^{h-1} a_h^p \hat{x}_p = a_h^h x_h + v_h \tag{3}$$

where $\hat{x}_p$ denotes a prior estimate. The above heuristic assumes that a major share of $x_h$ reach the sensors within the same interval. This can be unrealistic on heavily congested networks, especially when the time intervals are short. The assignment matrix used in each optimization step can also be inconsistent with the assignment matrix obtained by re-assigning the new estimates, requiring potentially costly fixed-point iterations (Cascetta and Postorino, 2001).

The assignment matrix has other practical considerations for OD estimation. The addition of non-negativity constraints on the OD flow variables results in a more complex least squares problem. In addition, the calculation of the assignment matrix itself may not be easy: if the starting solution has very low flows, simulated assignment fractions may be unreliable since very few vehicles reach the sensors. If the starting flows are increased in an *ad hoc* way, artificial bottlenecks may be created that prevent vehicles for some OD pairs from reaching sensors. This latter aspect can potentially lead to a cycle of ever-increasing OD estimates. Finally, the assignment matrix provides an intuitive mapping between OD flows and link counts, but is harder to employ in the context of other data such as speeds or travel times. Some studies involving the DYNASMART DTA model have attempted OD estimation using section densities instead of counts (Sun and Porwal, 2000).

Gradient-based methods have recently been used to solve the dynamic OD estimation problem. For example, the Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm (Spall, 1998) has been employed to solve the original optimization problem (Equation 1) without imposing linear approximations or resorting to a sequential approach. The method provides a practical solution to simultaneously solve for the OD flows of several departure time intervals (Balakrishna, 2006; Balakrishna et al, 2007). It is a steepest descent-like method, in the sense that a new set of OD flows is obtained by moving along a search direction by a step size. However, the search direction is a stochastic approximation of the gradient at the current solution.

SPSA is a variant of the standard Finite Difference Stochastic Approximation (FDSA) approach in which the gradient is evaluated component-wise. Rather than perturb each variable in turn and re-run a simulation, SPSA perturbs all components simultaneously (though independently) and approximates a complete gradient vector from just two evaluations of the objective function $z(\cdot)$ irrespective of the number of variables, $n$. In comparison, FDSA requires at least $n$+1 function evaluations for a single gradient calculation (and 2$n$ if each variable is perturbed once on either side). This represents an $n$-fold improvement in per-iteration efficiency, which is significant when each (simulation) assignment run can take several hours.

The efficiency of SPSA has another dimension: the number of iterations to convergence. Spall (1998) has shown that SPSA and FDSA converge in a similar number of iterations, which preserves SPSA's computational superiority. SPSA possesses other advantages. The relationship between the data and the variables is directly captured through *Assign*($x$), and there are no linear approximations. Further, the assignment method i.e. the function *Assign*($x$) may be replaced with any model, of any fidelity. The objective function can also be supplemented with any available measurements, and not just counts.

*Recent results*

Two OD estimation methods (a sequential approach using the assignment matrix, and a non-linear approach solved with SPSA) were recently compared on the same dataset from downtown Los Angeles (see Balakrishna et al., 2007). The numerical results (sampled in Figure 1) illustrate the benefits of simultaneously estimating all departure time intervals at once, and moving away from linear approximations. In the Figure, "Reference" corresponds to the older approach, while "Calibrated" denotes the more accurate non-linear approach. "Observed" is the real-world measurement. On-going tests on a large example in California involving two classes of vehicles (single- and high-occupancy) and 281 zones also indicate the potential of the non-linear approach.

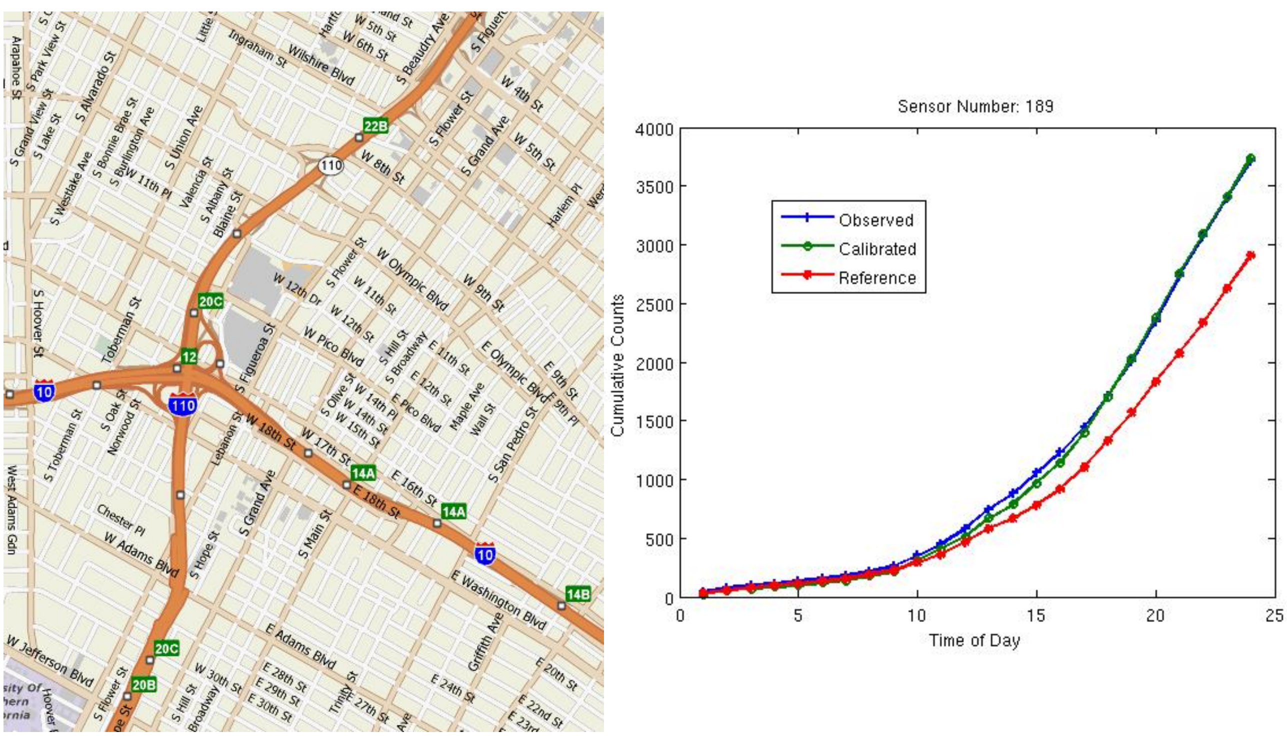


**Figure 1: Los Angeles Network and Cumulative Sensor Count Profiles**

## Conclusion

This paper motivates the need for estimating origin-destination (OD) trip tables from surveillance data, and provides references to popular static and dynamic OD estimation methods. The paper also discusses practical issues that significantly impact the accuracy of OD estimation, and presents some recent results from applying state-of-the-art estimators on a very large network in California. Despite the theoretical advances, dynamic OD estimation remains a challenging exercise owing to sparse sensor coverage, poor data quality and limited real-world applications. More tests on a wide range of networks is required before any method (existing or new) can be reliably adopted in practice.

## References


1. Ashok, K. (1996) Estimation and Prediction of Time-Dependent Origin-Destination Flows, *PhD thesis*, Massachusetts Institute of Technology.
2. Balakrishna, R. (2006) Off-line Calibration of Dynamic Traffic Assignment Models, *PhD thesis*, Massachusetts Institute of Technology.
3. Balakrishna, R., Ben-Akiva, M. and Koutsopoulos, H. N. (2007) Off-line Calibration of Dynamic Traffic Assignment: Simultaneous Demand-Supply Estimation, Transportation Research Record, No. 2003, pp. 50-58.
4. Boyles, S., Ukkusuri, S. V., Waller, S. T. and Kockelman, K. M. (2006) A Comparison of Static and Dynamic Traffic Assignment Under Tolls: A Study of the Dallas-Fort Worth Network, *TRB Innovations in Travel Demand Modeling*, Austin, TX.
5. Caliper (2007) Travel Demand Modeling with TransCAD Version 5.0 User's Guide, Caliper Corporation.
6. Cascetta, E., Inaudi, D. and Marquis, G. (1993) Dynamic Estimators of Origin-Destination Matrices using Traffic Counts, *Transp. Science* 27(4), pp. 363-373.
7. Cascetta, E. and Postorino, M. N. (2001) Fixed Point Approaches to the Estimation of O/D Matrices using Traffic Counts on Congested Networks, *Transp. Science* 29(4).
8. Cascetta, E. and Russo, F. (1997) Calibrating Aggregate Travel Demand Models with Traffic Counts: Estimators and Statistical Performance, *Transportation* 24, pp. 271-293.
9. McNeil, S. and Hendrickson, C. (1985) A Regression Formulation of the Matrix Estimation Problem, *Transp. Science* 19, pp. 278-292.
10. Nielsen, O. A. (1993) A New Method for Estimating Trip Matrices from Counts, *Institute of Roads, Traffic and Town Planning*, The Technical University of Denmark.
11. Nielsen, O. A. (1998) Two New Methods for Estimating Trip Matrices from Traffic Counts. *Travel Behaviour Research: Updating the State of Play*, Ortúzar, J., Hensher, D. and Jara-Diaz, S. (eds), pp. 221-250, Pergamon Press.
12. Spall, J. C. (1998) An Overview of the Simultaneous Perturbation Method for Efficient Optimization, *Johns Hopkins APL Technical Digest* 19(4), pp. 482-492.
13. Spiess, H. (1990) A Gradient Approach for the O-D Matrix Adjustment Problem, *Research report*, Centre de Recherche sur les Transports, University of Montreal.
14. Sun, C. and Porwal, H. (2000) Dynamic Origin/Destination Estimation using True Section Densities, *California PATH Program Report*, UCB-ITS-PRR-2000-5.
15. Tavana, H. and Mahmassani, H.S. (2000) Estimation of Dynamic Origin-Destination Flows from Sensor Data using Bi-level Optimization Method, 80th annual TRB meeting.
16. Toledo, T., Koutsopoulos, H.N., Davol, A., Ben-Akiva, M.E., Burghout, W., Andreasson, I., Johansson, T., and Lundin, C. (2003) Calibration and Validation of Microscopic Traffic Simulation Tools: Stockholm Case Study. *Transp. Res. Record* 1831, pp. 65–75.
17. Van Zuylen, H. and Willumsen, L.G. (1980) The Most Likely Trip Matrix Estimated from Traffic Counts, *Transp. Research* 14B, pp. 281-293.
18. Willumsen, L.G. (1982) Estimation of Trip Matrices from Volume Counts Validation of a Model Under Congested Conditions, PTRC 10th summer annual meeting, Transportation Analysis and Models seminar.