Australasian Transport Research Forum 2013 Proceedings
2 - 4 October 2013, Brisbane, Australia
Publication website: http://www.patrec.org/atrf.aspx

A study on the generalized TFlowFuzzy O-D estimation

Mohammad Yousefikia [a,*], Amir Reza Mamdoohi [b], Sara Moridpour [c], Mohammad Hossein Noruzoliaee [d] and Alireza Mahpour [e]

[a] PhD student, School of Civil, Environmental & Chemical Engineering, RMIT University, Melbourne, Australia
[b] Assistant professor, Faculty of Civil & Environmental Engineering, Tarbiat Modares University, Tehran, Iran
[c] Lecturer, School of Civil, Environmental & Chemical Engineering, RMIT University, Melbourne, Australia
[d] PhD student, Department of Civil and Materials Engineering, University of Illinois at Chicago, Chicago, United States
[e] PhD student, Faculty of Civil & Environmental Engineering, Tarbiat Modares University, Tehran, Iran

Abstract

It is an essential prerequisite to know the number of trips made between origin destinations (O-D) pair of a network. Typically, O-D matrices can be estimated by applying a series of travel demand models, based on expensive and time consuming widespread field studies. From 1970s, some methods have been proposed and widely used for directly estimating O-D matrices from observed traffic counts and an initial O-D matrix. These methods could be considered as the inverse of the traffic assignment problem in which an O-D matrix is found that reproduces the observed traffic counts when assigned to the network.

TFlowFuzzy (TFF), as one of the most practical and realistic O-D estimation methods models observed traffic counts as imprecise values based on fuzzy sets theory. In this method, a key element in the estimation of a trip matrix is the “route choice proportions” for each link. In the TFF method, route choice proportions are computed only once and held constant for all iterations in the matrix estimation process which is a very simplifying assumption. Recently a generalized version of TFF is proposed in which the route choice proportions are updated successively at each iteration. This modification considers the route choice proportions endogenously. This allows for a more precise estimation of the O-D matrix. In this research the proposed modified TFF is implemented on Tehran provincial intercity network. Results indicate a considerable improvement in the goodness of fit of the modified TFF in comparison to the traditional algorithm. The goodness of fit is increased by 12 percent. Furthermore, the modified TFF represents better results in descriptive statistics in terms of less error values and intervals. This shows the effectiveness of the proposed algorithm as it can replicate the real data more precisely.

Keywords: O-D matrix estimation, modified TFlowFuzzy, route choice proportions

1. Introduction

Many aspects of transportation planning rely on knowing the number of trips made between origin and destination pairs, also known as the origin-destination (O-D) matrix, in a transportation network. Typically, O-D matrices can be obtained by applying a system of models that compute the approximate number of trips generated and distributed with different modes of transportation in the network. These methods rely on the calculation of O-D matrix based on field studies such as household, roadside, and terminal surveys (Ortuzar and Willumsen, 2011). As a matter of fact, these methods have three unpleasant characteristics: They are expensive, time-consuming to conduct, and the results of such studies tend to lose their validity fairly rapidly (Friedrich et al., 2000).

[*] Corresponding author: Email address: s3419605@student.rmit.edu.au

From 1970s some methods have been proposed and widely used for estimating O-D matrices directly from the current data of observed traffic counts collected on a set of links and prior information expressed as a target/old O-D matrix, without first estimating a series of travel demand models such as trip generation, modal split, and trip distribution (Sheffi, 1984). This approach could be considered as the inverse of the traffic assignment problem: finds an O-D matrix that, when assigned to the network, reproduces the observed traffic counts. Estimating O-D matrix from traffic counts, due to less costs and the ability to keep demand matrices up-to-date, are increasingly gaining interest (Friedrich et al., 2000).

The formulation of the matrix estimation problem based on traffic counts has one serious weakness: traffic counts are assumed to be known input values without any uncertainty. As a result, traffic counts obtain an inappropriate weight since any count only provides figures that represent a snapshot situation which are subject to considerable sampling error. For this reason an algorithm named TFlowFuzzy (TFF) was developed based on the Rosinowski's research (Rosinowski, 1994) that models the counts as imprecise values based on fuzzy sets theory. In the TFF algorithm, the route choice proportions for each link are computed only once and held constant during the whole calculation process. Yousefikia et al. (2013) proposed a modified algorithm for TFF, as a generalized version, in which the route choice proportions were updated successively at each iteration and allows for a more precise estimation of the O-D matrix.

This paper builds upon the research of Yousefikia et al. (2013) and addresses the implementation of their proposed methodology on the network of Tehran provincial roads and study the effectiveness and precision of the proposed methodology in estimating the O-D demand matrix on an uncongested network. In This research the effectiveness of the proposed modified TFF algorithm to replicate the real data on an uncongested network will be studied.

The remainder of the paper is organized as follows: Section 2 gives a brief review of the related literature in the field of O-D matrix estimation. The basic mathematical framework for the proposed methodology is presented and discussed in detail in section 3. Finally, the methodology is implemented for a Tehran provincial road network in order to evaluate the effectiveness of the proposed modified TFF algorithm.

2. Literature review

Different approaches to estimate O-D matrices from traffic counts have been proposed which could be classified into two broad categories: the first category is based on traffic modeling concepts. These approaches include Maximum Entropy (ME) model and combined models for traffic planning. An estimate of the O-D matrix is obtained by direct solution (entropy measure) or by estimating the parameters of the combined model. The second category, statistical inference approach, includes the Maximum Likelihood (ML), Generalized Least Squares (GLS), and Bayesian inference models. Here, the traffic volumes and the target O-D matrix are assumed to be generated by some probability distributions. The target O-D matrix in traffic modeling approach is normally obtained as an outdated O-D matrix while statistical approach rely on a target O-D matrix obtained from sample surveys. The joint use of (full or partial) traffic count and target matrix information has resulted in a number of other trip matrix estimation methods, such as the Bayesian inference approach (Perrakis et al., 2012; Maher, 1983), least squares approach (Cascetta, 1984; Bell, 1991; Yang et al. 1992), and maximum likelihood approach (Spiess, 1987; Lo et al., 1996). Despite various implicit parameter assumptions and optimization principles, all seek a matrix that represents some form of trade-offs between a target trip matrix and observed traffic counts. These trade-offs appear in the model constraints or objective functions.

Maximum entropy theory was first used by Van Zuylen and Willumsen (Willumsen, 1981; Zuylen and Willumsen, 1980) for the most-likely-trip-matrix estimation problem based on traffic counts. By assuming that the underlying traffic flows follow a known proportional routing pattern, these models resort to a simple iterative balancing method for solutions (Xie et al., 2010)

Some authors included congestion effects in the estimation problem in which the dependence of link costs, path choices, and assignment fractions on link flows are considered. Equilibrium assignment models are particularly adopted for such cases. Nguyen first introduced the equilibrium-based approach to estimate O-D matrices through a mathematical program (refer to Leblanc and Farhangian, 1982). Fisk imposed the User Equilibrium (UE) routing principle to a similar matrix estimation problem, resulting in a ME model subject to a variation inequality constraint (Fisk, 1988). Also, the trip matrix estimation problem including congestion effects by considering different traffic assignment models was solved by some other researchers (Yang et al., 1994; Cascetta and Posterino, 2001; Yang et al., 2001).

Bi-level programming approach has been used for the problem of O-D matrix estimation in case of congested networks. In this approach, the upper-level problem is the trip matrix estimation and the lower-level problem represents a network equilibrium assignment problem. Yang extended the bi-level programming problem by including link flow interaction and developed a model with heuristic algorithms which was later discussed by Kim (Yang, 1995; Kim et al., 2001). Lundgren and Peterson (2008) developed a heuristic bi-level problem solving it by a descent algorithm and demonstrated the algorithm using a large size network for the city of Stockholm.

Many nonlinear real-life problems in the field of transportation planning and traffic control have been solved by fuzzy logic systems. Its application can be found in O-D matrix estimation problem, too. Fuzzy logic is used to model situations in which user decisions are so complex that it is very hard to develop a mathematical model. Xu and Chan (1993) estimated O-D matrix with fuzzy weights. Reddy and Chakroborty (1998) proposed a bi-level optimization approach on a multipath, fuzzy inference based flow dependent assignment algorithm for generating the route choice proportions which along with the observed link flows are then used in the O-D matrix estimation (refer to Bera and Rao, 2011). Shafahi and Faturechi proposed a new fuzzy O-D matrix estimation method which considered link data as a fuzzy value that vary within a certain bandwidth (Shafahi and Faturechi, 2009).

Given the fact that traffic counts only provide figures that represent a snapshot situations which are subject to considerable sampling error, based on research by Van Zuylen and Willumsen (1978, 1980), Boserhoff (1985) and Rosinowski (1994), the TFF O-D matrix estimation employs a fuzzy approach and represent a technique for computer-aided processing of demand matrix more realistically. This method has been widely used and implemented on several networks (Yousofikia, 2012; Ren and Aziz, 2009). Evaluation of the results of these researches indicates that the precision of results can be evolved by generalizing model mathematical background.

In all the pre-mentioned models of O-D estimation using maximum entropy optimization problem, route choice proportions are assumed to be constant and flow independent. Yousefikia et al. (2013) proposed a generalized O-D estimation method in which route choice proportions are updated successively. Evaluation of the results of these researches indicates that the precision of results can be evolved by generalizing model mathematical background. In this research the effect of this generalization will be studied on the precision of O-D estimation of the rural network of Tehran province roads as an uncongested network.

3. Methodology

In this section, the optimization algorithm for the O-D estimation problem, proposed by Yousefikia et al. (2013) is presented. The steps of the algorithm are as follows.

Step 1- Data input

Starting point of the O-D estimation is observing the travel demand for O-D pairs which describes the trip demand patterns of an earlier state. The trips of any O-D pair contribute a certain share to each traffic count. The counted volumes correspond to the sum of all O-D trips travelling on this specific link. The observed link flows and the initial O-D matrix are the input data of the algorithm. Obviously, the first step of iterations is done based on existing information of link flows and O-D matrices. The notations of variables used in this step are presented as follow.

ε: Maximum permissible change of O-D which is estimated at each successive iteration

n: iteration counter

$T_{rs}^{n=0}$: Number of trips from origin *r* to destination *s* in the primary O-D matrix

P_{rs}^{n}: share of each link from the calculated number of trips from origin *r* to destination *s* at iteration *n*

Step 2- Assign O-D matrix to the network

At every iteration (*n*), the target matrix related to that iteration (T_{rs}^{n}) assigns to the transportation network by the UE traffic assignment model, first proposed by Wardrop (Sheffi, 1984), by means of Visum software package. Having known the calculated flows on network links by implementing UE model, route choice proportions (P_{rs}^{n}), are then computed. The UE model of traffic assignment is based on the fact that a behavior on the individual level creates equilibrium at the system level. Flows on links (whose travel times are assumed to vary with flow) are said to be in equilibrium when no trip-maker can improve his/her travel time by unilaterally shifting to another route.

Step 3- Estimate O-D matrix

The mathematical formulation for O-D estimation problem used in this research is as follow.

$$Max - \left(\sum_{rs\epsilon OD} T_{rs}^{n} \, ln\left(\frac{T_{rs}^{n}}{\hat{t}_{rs}}\right) - T_{rs}^{n}\right) \tag{1}$$

S.t.

$$\sum_{rs\epsilon OD} T_{rs}^{n} \times P_{rs}^{n} = \tilde{v}_a \;;\; \forall a \tag{2}$$

$$T_{rs}^{n} \geq 0 \tag{3}$$

Where;

ε permissible change in O-D estimated in successive iterations

T_{rs}^{n} number of trips from origin r to destination s at iteration n

$\hat{t}_{rs}$ number of trips from origin r to destination s in the base year (initial target matrix)

$\tilde{v}_a$ observed traffic count on link a with varying bandwidth

$P_{rs,a}^{n}$ route choice proportions for link a at iteration n (elements of assignment matrix)

n Iteration number

Step 4- Convergence criterion

A convergence criterion based on the permissible difference between the estimated matrix and the target matrix is checked for process termination. If it is not met, the procedure continues with redoing the aforementioned steps, but with the underlying fact that the route choice proportions are updated by assigning the new estimated matrix instead of the initial target matrix.

$$E = \sum_{rs} (T_{rs}^{n} - T_{rs}^{n-1})^2 \tag{4}$$

If $E < \varepsilon$ stop, otherwise$T_{rs}^{n} = \frac{T_{rs}^{n} + T_{rs}^{n-1}}{2}$ → go to step 1

The flow of the overall process is illustrated in Figure 1.

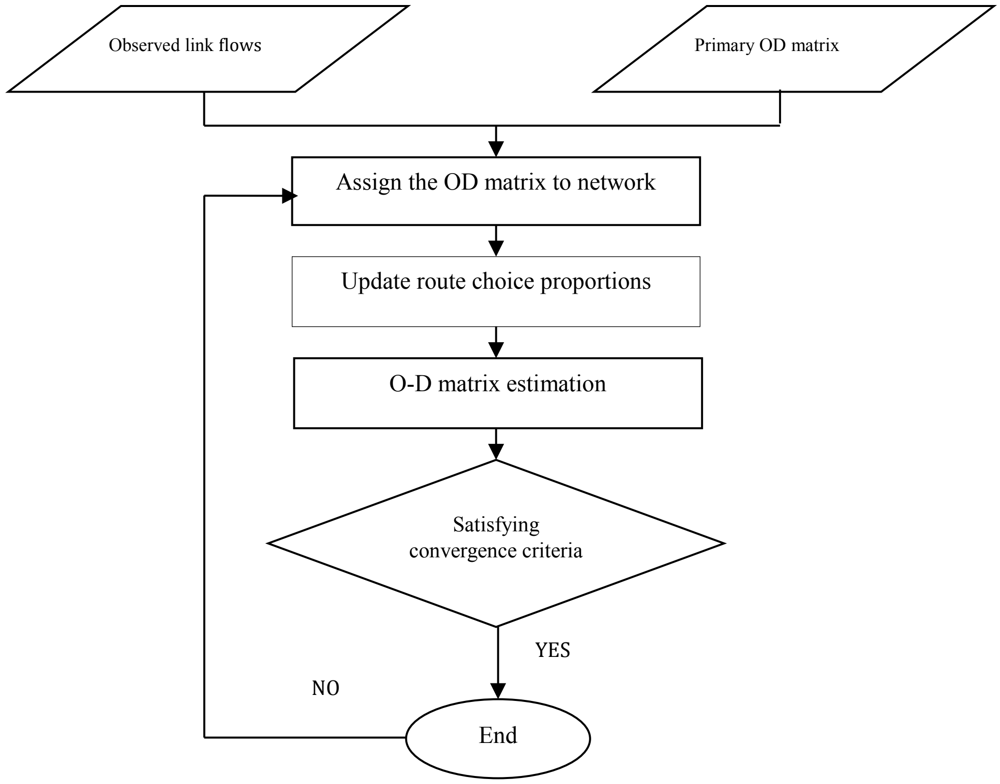

Figure 1. Overall analogy of the proposed methodology

4. Case study

The proposed methodology for the modified TFF is implemented on the network of Tehran province intercity roads. This study area consists of 35 TAZs and the corresponding network includes 750 nodes and 626 links with a total length of 4049 Km (Figure 2). Traffic counts on 67 links observed during a day in 2012 the initial target matrix which is extracted from a four-step model are considered as the main inputs.

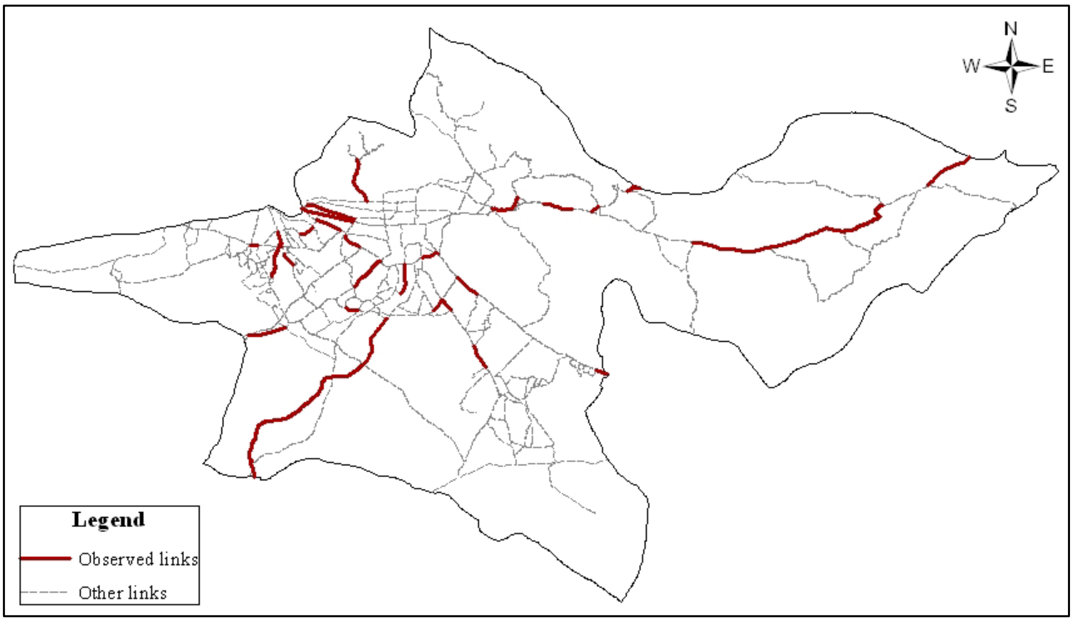

Figure 2. Intercity transportation network of Tehran province

Scatter plots of the calculated traffic flows (estimation) versus the observed traffic counts (observation) for the traditional TFF algorithm implemented on Tehran network is depicted in Figure 3. The calculated traffic flows on network links (model outputs) are obtained from assigning the estimated O-D matrix to the network. As shown in Figure 3, the traditional TFF can replicate the real traffic counts with a coefficient of determination (R2) of 0.85.

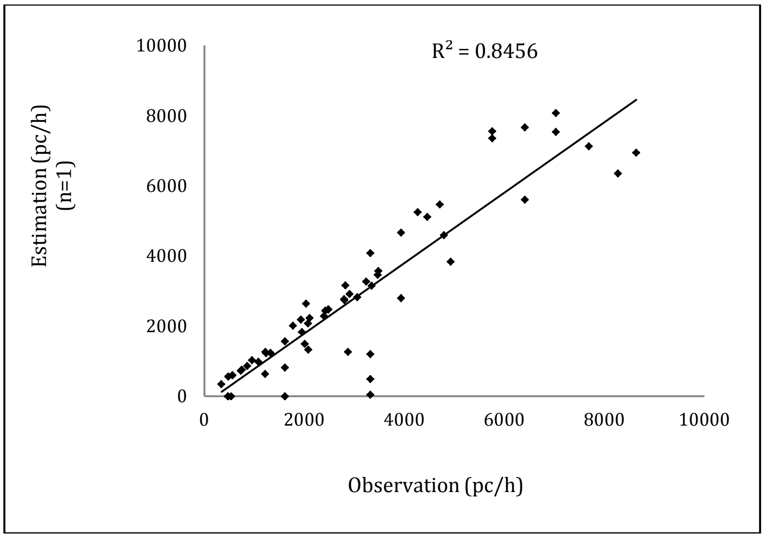

Figure 3. Scatter plot of estimated vs. observed link (traditional TFF)

In order to evaluate the effect of route choice proportions endogenization, the iterative procedure of the modified algorithm is implemented by running a Visual Basic Application (VBA) script using Visum Component Object Model (COM) interface. The convergence criterion for the modified TFF of Tehran province roads is satisfied at iteration 4. As seen in Figure 4, the modified TFF algorithm reproduces the real traffic counts with an R^2 value of 0.95.

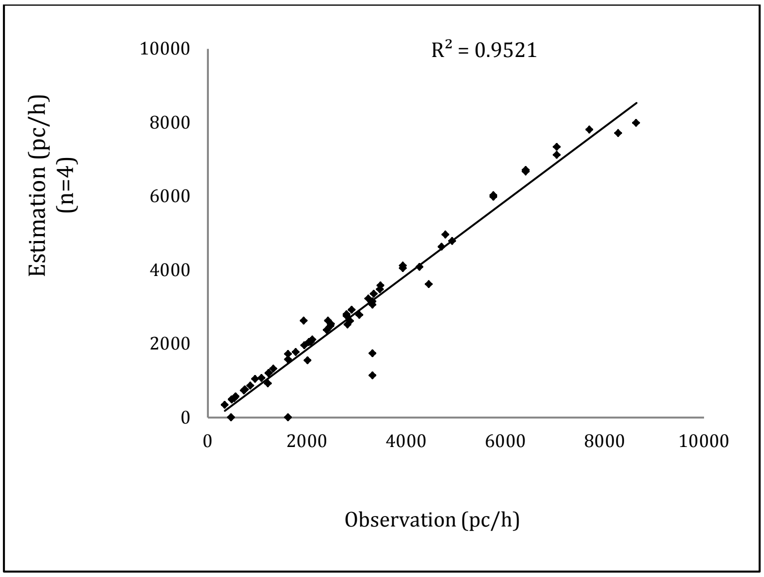

Figure 4. Scatter plot of estimated vs. observed link volume (modified TFF)

Results indicate a considerable improvement in the goodness of fit of the modified TFF in comparison to the traditional algorithm as it is increased by 0.10 (nearly 12 percent). This proves the effectiveness of the proposed modified TFF algorithm as it can replicate the real data more precisely.

Furthermore, descriptive statistics of relative errors (difference between calculated and observed flow of each link) are reported in Table 1. As seen, the modified TFF represents better results in all statistics (measures of central tendency and dispersion of the error) in terms of less error values and intervals.

Table 1. Descriptive statistics of the error term

Statistic	Tehran intercity network	
	Traditional TFF	Modified TFF
Minimum	0.1	0.0
Maximum	100.0	100.0
Mean	23.6	12.9
Mode	100.0	100.0
Std. dev.	30.9	27.2
1st quartile	3.3	0.4
Median	12.6	2.7
3rd quartile	26.7	7.9

5. Conclusions

Origin-Destination (O-D) matrix is one of the fundamental elements of transportation planning in the demand side. Obtaining this matrix by conventional methods needs considerable amount of time, cost, and manpower. For 30 years, some demand matrix estimation methods have been proposed and widely used for estimating O-D matrices using the current data of observed traffic counts. In these methods, by using traffic counts in some links and an initial O-D matrix (target matrix), a demand matrix is estimated which can replicate observed link flows and stay close to the initial matrix.

TFlowFuzzy (TFF), as one of the practical and more realistic O-D estimation methods, models the counts as imprecise values based on fuzzy sets theory. In this method, a key element in the estimation of a trip matrix from traffic counts is the route choice proportions for each link. In the TFF method, rout choice proportions are computed only once and held constant in all iterations in the matrix estimation process which is a very simplifying assumption. Recently a generalized version of TFF is proposed in which the proportions are updated successively at each iteration. This modification, considering the route choice proportions endogenously, allows for a more precise estimate of the O-D matrix.

In this research the proposed modified TFF is implemented on Tehran province intercity roads as an uncongested network. Results indicate a considerable improvement in the goodness of fit of the modified TFF in comparison to the traditional algorithm as it is increased by 0.10 (nearly 12 percent) for Tehran provincial network. This proves the effectiveness of the proposed modified TFF algorithm as it can replicate the real data more precisely.

Applications of the modified TFF are suggested to be implemented on other transportation networks. Further research can focus on the effect of other behavioral assumptions (rather than UE assignment methods) such as Dynamic User Equilibrium (DUE) on the precision of this modification

6. Acknowledgment

The authors would hereby like to thank Tarahan Parseh Transportation Research Institute and Tehran Department Of Transportation (DOT for making available most of the data needed for this research.

7. References

Ashok, K., 1996. Estimation and prediction of time-dependent origin-destination flows.*PhD thesis in transportation systems, Massachusetts Institute of Technology, USA*.

Bell, M., 1991. The estimation of origin-destination matrices by constrained generalized least squares. *Transportation Research Part B: Methodological*25, 13-22.

Bera, S. and Rao, K. V., 2011. Estimation of origin-destination matrix from traffic counts: the state of the art. *Journal of European Transport*49, 3-23.

Bosserhoff, D., 1985.StatischeVerfahrenzurErmittlung von Quell-Ziel-Matrizenim öffentlichenPersonennahverkehr. *Ein Vergleich, Dissertation, University of Karlsruhe, Deutschland*.

Cascetta, E., 1984. Estimation of trip matrices from traffic counts and survey data: a generalized least squares estimator. *Transportation Research Part B: Methodological*18, 289-299.

Cascetta, E. and Postorino, M. N., 2001. Fixed point approaches to the estimation of O/D matrices using traffic counts on congested networks. *Journal of Transportation Science*35, 134-147.

Fisk, C. S., 1988. On combining maximum entropy trip matrix estimation with user optimal assignment. *Transportation Research Part B: Methodological*22, 69-73.

Freidrich, M., Nokel, K. and Mott, P., 2000. Keeping passenger surveys up-to-date; A fuzzy approach. *Transportation Research Record: Journal of Transportation Research Board, Washington DC*1735, 35-42.

Kim, H., Beak, S. and Lim, Y., 2001. Origin-destination matrices estimated with a genetic algorithmfrom link traffic counts. *Transportation Research Record: Journal of Transportation Research Board, Washington DC*1771, 156-163.

LeBlanc, L. J. and Farhangian, K., 1982. Selection of a trip table which reproduces observed link flows. *Transportation Research Part B: Methodological*16, 83-88.

Lo, H., Zhang, N. and Lam, W., 1996. Estimation of an origin-destination matrix with random link choice proportions: a statistical approach. *Transportation Research Part B: Methodological*30, 309-324.

Lundgren, J. T. and Peterson, A., 2008. A heuristic for the bilevel origin-destination matrix estimation problem. *Transportation Research Part B: Methodological*42, 339-354.

Maher, M. J., 1983. Inferences on trip matrices from observations on link volumes: a bayesian statistical approach. *Transportation Research Part B: Methodological*17, 435-447.

Ortúzar, J. D. and Willumsen, L. G., 2011.*Modeling Transport*. Wiley, Chic ester, USA.

Perrakis, K., Karlis, D., Cools, M.et al., 2012. A Bayesian approach for modeling origin-destination matrices. *Journal of Transportation Research Part A: Policy and practice*46, 200-212.

Reddy, K. H. and Chakroborty, P., 1998. A fuzzy inference based assignment algorithm to estimate O-D matrix from link volume counts. *Journal of Computers, Environment and Urban Systems*22(5), 409-423.

Ren, J. and Rahman, A., 2009. Automatically balancing intersection volumes in a highway network.*12th TRB Transportation planning application conference*.

Rosinowski, J., 1994. Entwicklung und ImplementierungeinesÖPNV-MatrixkorrekturverfahrensmitHilfevonMethoden der TheorieunscharferMengen (Fuzzy-Sets-Theorie). *Master thesis, University of Karlsruhe, Deutschland*.

Shafahi, Y. and Faturechi, R., 2009. A Practical OD matrix estimation model based on Fuzzy set theory for large cities. *23rd European Conference on Modelling and Simulation*. Madrid, Spain.

Sheffi, Y., 1984.*Urban transportation networks: Equilibrium analysis with mathematical programming methods*. Prentice-Hall, New Jersey, USA.

Spiess, H., 1987. A maximum-likelihood model for estimating origin-destination matrices. *Transportation Research Part B: Methodological*21, 395-412.

Van Zuylen, J. H., 1978. Theinformation minimizing method: validity and applicability to transport planning. *New Developments in Modelling Travel Demand and Urban Systems* (edited by GRM Jansen et al.).

Van Zuylen, J. H. and Willumsen, L. G., 1980. The most likely trip matrix estimated from traffic counts. *Transportation Research Part B: Methodological*14, 281-293.

Willumsen, L. G., 1981. Simplified transport models based on traffic counts. *Journal of Transportation*10, 257-278.

Xie, C., Kockelman, K. M., Waller, S. T., 2010. A maximum entropy Method for origin-destination trip matrix estimation. *Transportation Research Record: Journal of Transportation Research Board, Washington DC*.

Xu, W. and Chan, Y., 1993. Estimating an origin-destination matrix with fuzzy weights. Part 1: Methodology, *Transportation Planning and Technology*17, 127-144.

Yang, H., Sasaki, T., Iida, Y. and Asakura, Y., 1992. Estimation of origin-destination matrices from link traffic counts on congested networks. *Transportation Research Part B: Methodological*26, 417-433.

Yang, H., Sasaki, T. and Iida, Y., 1994. The Equilibrium-based origin-destination matrix estimation problem. *Transportation Research Part B: Methodological*28, 23-33.

Yang, H., 1995. Heuristic algorithms for the bi-level origin-destination matrix estimation problem. *Transportation Research Part B: Methodological*29, 231-242.

Yang, H., Meng, Q. and Bell, M. G. H., 2001. Simultaneous estimation of the origin-destination matrices and travel-cost coefficient for congested networks in stochastic user equilibrium. *Journal of Transportation Science*35, 107-123.

Yousefikia, M., 2012. Estimation of origin-destination matrix from link volume counts using TFlowFuzzy method: implemented for Mashad city. *Master of Science thesis, Tarbiat Modares University, Iran*.

Yousefikia, M., Mamdoohi, A. R. and Noruzoliaee, M. H., 2013. A modified OD Estimation; updating route choice proportions. *Journal of the proceedings of the institution of civil engineering (ICE)-Transport,* in press.