Freeway Origin Destination matrices, not as simple as they seem

Satyanarayana Muthuswamy[1]

Gary A Davis[2]

David M Levinson[3]

Panos G Michalopoulos[4]

Satyanarayana Muthuswamy[1]
Traffic Engineer
KLD Associates Inc.
300 Broadway
Huntington Station, NY 11746
muth0012@tc.umn.edu
Phone: (612) 386-1458

Gary A Davis[2]
Associate Professor
University of Minnesota
500 Pillsbury Drive S.E
Minneapolis, MN 55414
drtrips@tc.umn.edu
Phone: (612) 625-2598
Fax: (612) 626-7750

David M Levinson[3]
Assistant Professor
University of Minnesota
500 Pillsbury Drive S.E
Minneapolis, MN 55414
levin031@tc.umn.edu
Phone: (612) 625-6354
Fax: (612) 626-7750

Panos G Michalpolous[4]
Professor
University of Minnesota
500 Pillsbury Drive S.E
Minneapolis, MN 55414
micha0001@tc.umn.edu
Phone: (612) 625-1509
Fax: (612) 626-7750

All communications to be sent to muth0012@tc.umn.edu

Submitted on 28 June 2002 to the Transportation Research Board for possible publication and presentation at the 2003 Annual Meeting

Word Count: 5061 + 9 (Tables and Figures) x 250 = 7311

ABSTRACT

Travel demand can be elegantly represented using an Origin-Destination (OD) matrix. The link counts observed on the network are produced by the underlying travel demand. One could use these counts to reconstruct the OD matrix. An offline approach to estimate a static OD matrix over the peak period for freeway sections using these counts is proposed in this research. Almost all the offline methods use linear models to approximate the relationship between the on-ramp and off-ramp counts. Previous work indicates that the use of a traffic flow model embedded in a search routine performs better than these linear models. In this research that approach is enhanced using a microscopic traffic simulator, AIMSUN, and a gradient based optimization routine, MINOS, interfaced to estimate an OD matrix. This approach is an application of the Prediction Error Minimization (PEM) method. The problem is non-linear and non-smooth, and the optimization routine finds multiple local minima, but cannot guarantee a global minima. However, with a number of starting "seed" matrices, an OD matrix with a good fit in terms of reproducing traffic counts can be estimated. The dominance of the mainline counts in the OD estimation and an identifiability issue is indicated from the experiments. The quality of the estimates improves as the specification error, introduced due to the discrepancy between the traffic flow model and the real world process that generates the on-ramp and off-ramp counts, reduces.

Keywords: travel demand, OD estimation, simulation, optimization

INTRODUCTION

Travel demand estimation is one of the most challenging and interesting procedures in transportation engineering. The process of demand estimation is an attempt to understand and predict the behavioral patterns of individuals, their choices on routes and trips. Over the years, many techniques have evolved to estimate travel demands in different forms. One of the most elegant forms of representing travel demand is an Origin Destination (OD) matrix, a table with the number of trips made between different points on a network. An OD matrix is an essential input to models that attempt to predict the impacts of transportation interventions like ramp metering, traveler information and capacity changes.

Theoretically, if the starting and ending points of all the trips made in a network are tracked, then the OD matrix can be observed. This is however, infeasible due to the resulting data management and privacy/surveillance issues. In addition, estimation of a sample of the OD matrix by tracking the start and ends of selected vehicles is cumbersome, time-consuming, and expensive and is ill-suited for tracking changes in travel demand needed for Advanced Traffic Management.

An attractive alternative to sample-based approaches is to estimate OD patterns from data routinely collected by surveillance and control systems, such as automatic traffic counts. A number of approaches to carrying out such estimations have been proposed over the past two decades, and most can be seen as variants of prediction error minimization (PEM) methods *(1)*. In PEM, one starts with a model, which generates predictions of observed quantities, as a function of a set of unknown parameters. One then chooses as parameter estimates those values that minimize the discrepancy between the predictions and observations. In particular, maximum likelihood methods *(2)* and *(3)*, Kalman filtering methods *(4)*, and neural network methods *(5)* can be seen as special cases of PEM.

In applying PEM to OD estimation, one encounters issues of identifiability, i.e. are the available data sufficient to estimate reliably the OD parameters. Identifiability problems arises when different OD patterns yield roughly equivalent predictions the observed data, so that it is difficult to determine which OD pattern actually generated the data. If the different OD patterns then lead to different predictions of the impacts of transportation policies, unresolved identifiability problems can compromise a policy analysis.

In this study, we revisit the problem of estimating freeway OD patterns from on-ramp and off-ramp counts, paying particular attention to identifiability issues. This estimation problem has recently become salient in Minnesota because of a need to predict the effects of ramp metering strategies, which in turn stems from increasing political opposition to ramp-metering, requiring that what metering is done be justified *(6)* and *(7)*.

The paper has been organized as follows. The first section is a discussion of the adopted method and describes the application of the PEM technique to OD estimation for freeway sections as a combination of simulation and optimization. The following two sections are descriptions of the two test sites and their related results and the last section has the conclusions of the study.

Muthuswamy, Satyanarayana, Gary A Davis, David M Levinson and Panos G Michalopoulos
Freeway Origin Destination Matrices: Not as Simple as They Seem
presented at Transportation Research Board 83rd Annual Meeting, January 10-14 2004, Washington DC.

METHODOLOGY

Traffic on a network can be modeled as a system where vehicles arrive randomly, traverse certain links of the network, and exit the system after a certain time-period. The arrivals follow a random process and the routes taken are a function of the driver's preferences. The travel times on the chosen route are a function of the traffic conditions.

The traffic conditions are a result of the different choices that drivers make and the actual driving characteristics (traffic flow model). The choice making process and the actual driving characteristics represent the travel behavior. The trip table represents the underlying traffic pattern for the network. In addition to this traffic pattern, if the travel behavior is known, the traffic conditions on the network can be reproduced as they are a direct result of the traffic pattern and travel behavior.

Extending this idea to freeway traffic, the inputs to the system would be a traffic flow model and an OD matrix. On a freeway, there is only one route for every possible trip, so there is no route selection process. The traffic conditions on a freeway are characterized by the on-ramp counts, off-ramp counts, and speeds. If we have an appropriate traffic flow model and the OD matrix, the traffic conditions can be reproduced.

If one of the inputs in this system is unknown, but the outputs and other inputs known, the unknown input can be estimated by matching a set of outputs corresponding to a set of inputs, to the actual conditions (observed outputs), in other words using the PEM method. The OD estimation problem is an example of such a case. The OD matrix is unknown, but the traffic conditions – the counts, speeds, density are known. If an appropriate traffic model is used, the OD matrix can be estimated by trying to reproduce the traffic conditions on the freeway. In other words using the traffic flow model, a search for the OD matrix is done in the feasible space of OD matrices and a particular matrix chosen based on its ability to reproduce the traffic conditions. Hence, the OD matrix estimation process can be defined as an optimization problem that searches for the optimal OD matrix that minimizes the deviations of the predicted and the actual traffic conditions.

The Minimization Problem

An OD matrix in its standard form is a trip table, where every cell entry 'T_{ij}' is the number of trips made from origin 'i' to destination 'j'. It can also be represented as a percentage matrix, where every cell 'b_{ij}' is the percentage of trips originating at origin 'i' that will end up at destination 'j'. The latter definition is chosen for reasons explained in the following section. The trip table is the product of the productions at the origins (on-ramp counts) and the percentage OD matrix. Now, using this trip table and a traffic flow model, the traffic conditions can be predicted.

On most freeways, the OD matrix is upper triangular as the downstream on-ramps cannot feed upstream off-ramps. In addition, the first origin will be the upstream mainline and the last destination will be the downstream mainline. The input for this system would the on-ramp counts and the percentage OD matrix and the traffic conditions that could be matched would be the off-ramp/mainline counts.

If it is assumed that this percentage OD matrix is constant over the peak period, the OD estimation problem can be defined as the search for that optimal matrix that minimizes the deviations from the actual off-ramp counts. The OD matrix when defined as a percentage matrix has to satisfy the constraints that the row sums add to 1. This implies that the sum of trips originating from an on-ramp have to match the on-ramp counts. Therefore, the OD estimation problem can be defined as a linearly constrained minimization problem.

In this setup, the upstream mainline and the downstream mainline are treated as the first on-ramp and the last off-ramp respectively. Hence, the OD estimation searches for that optimal OD matrix that best matches the off-ramp counts including the downstream mainline. The downstream mainline counts are typically an order or two higher in magnitude than the off-ramp counts. In order to avoid the minimization process from being dominated by the downstream mainline, the sum of the error terms are weighted based on their magnitude. Using the inverse of the standard deviations of the counts as the weights scales the variances equally and reduces the domination of the downstream mainline counts. The Non Linear Programming Problem (NLP) can be formally defined as

NLP: Minimize $$\sum_j \sum_t w_j (O_{tj} - \hat{O}_{tj})^2$$

Subject to $$\sum_i b_{ij} = 1.0$$

$$0.0 \le b_{ij} \le 1.0$$

Where,

O_{tj} - Actual off-ramp counts at ramp 'j' in time slice 't'

$\hat{O}_{tj}$ – Predicted off-ramp counts at ramp 'j' in time slice 't'

w_j – the weight for the ramp 'j' = inverse of standard deviation of O_{tj}

i – Origin index

j – Destination index

t – Time index

The solution to the above NLP is the estimate of the OD matrix 'B' that matches the actual off-ramp counts with the greatest accuracy.

The Time Invariant OD matrix

The trip table is constantly changing over every time slice because the inputs (on-ramp counts) are time varying. The justification of the assumption of a time invariant OD matrix needs to be addressed. Consider Figure 1, a conceptual model that relates the on-ramp counts and the off-ramp counts. At a very abstract level (Level 1), the whole process can be viewed as a Data generation mechanism that takes the on-ramp counts as inputs and gives the off-ramp counts as the outputs. This process can then be further broken down at Level 2 that involves the creation of the trip table and a Traffic Flow mechanism. At the lowest level, the traffic flow process can be broken down as a process that takes in the trip table and calculates the routes and the choice making process and then assigns the trips to the network and propagates the vehicles through the network. The OD estimation process involves calculation of the trip table from the observed on-ramp and off-ramp counts. For the best performance of the method, the process as defined in Level 3 must be replicated. The real world process cannot be exactly reproduced because of its complex nature and so a satisfactory approximation is required. The level of satisfaction is related to the need for the approximation and its simplicity. Therefore, in the OD estimation process, approximations to the above processes are used.

As shown in Figure 1 there are two components approximated in the data generation mechanism. The first is related to the creation of the trip table from the on-ramp counts and the second is the traffic flow model. The latter has been approximated with a microscopic traffic simulator. The first step that relates to the creation of trip tables from on-ramp counts is approximated by using a time-invariant percentage OD matrix and the on-ramp counts. The choice of the approximation is related to its simplicity and appeal to the intuitive sense of the process and is in line with the with Occam's Razor or the principle of parsimony.

It is known, that the OD estimation process in one time slice has an identifiability problem, as there are more unknowns than equations. Therefore, even over additional time slices, if it is assumed that the OD matrix is different for each time slice, it leads to the same problem. To solve the problem we need to assume that there is a time invariant OD matrix over some sub-set of the multiple time slices. Applying Occam's razor the simplest assumption of the OD matrix being constant over all the time slices is adopted in this research. The following discussion supports such an approximation.

Before proceeding with the discussion, some concepts on multinomial probability distribution are reviewed. The Multinomial distribution is the extension of the binomial distribution. A random experiment has multiple outcomes, say *m*. Each individual outcome X_i (i = 1, 2... m) has an associated probability p_i with it. In other words, if the same experiment were to be repeated '*N*' times, the '*N*' outcomes could be any combination of the 'm' possible outcomes and if, there are 'n_i' observations of outcome 'X_i' then the joint probability distribution would be as given in equation (1).

$$P(X_1, X_2 \dots X_m; p_1, p_2 \dots p_m; N) = \frac{N!}{n_1!n_2!\cdots n_m!} {p_1}^{n_1} {p_2}^{n2} \cdots {p_m}^{n_m} \tag{1}$$

This idea can be extended to an OD matrix. Consider one row, '*k*' in the OD matrix. The cell entries 'b_{kj}' that represent the percentage of trips from on-ramp '*k*' to off-ramp '*j*' can be interpreted as the probability that a trip originating at '*k*' ends up at '*j*'. When a vehicle enters a freeway at on-ramp '*k*' the OD matrix entry 'b_{kj}' corresponds to the probability that the vehicle will end up in destination 'j'. In this setup, there are as many outcomes as there are destinations and the associated probabilities for each of the outcomes are the OD matrix row entries 'b_{kj}'.

Consider the following experiment. At any given time interval, for every arriving vehicle at on-ramp '*k*', using the multinomial probabilities given by the OD matrix row entries 'b_{kj}', a destination is assigned. Based on this assignment, the trip table entries 'T_{kj}' are updated. The number of trials for this experiment is the on-ramp count 'Q_j'. The same experiment is repeated with all the rows of the OD matrix. As a result, the trip table for that time interval is generated.

The same set of experiments can be repeated over all the time slices using the same OD matrix and the time varying on-ramp counts. Since the on-ramp counts serve as the number of trials for each individual experiment, and they are time variant, the resulting trip table is also time variant.

Using the above argument, the time sliced trip tables can be visualized as outcomes corresponding to multiple experiments, using the OD matrix as the multinomial probabilities and the on-ramp counts as the number of trials. Therefore, the time varying trip table can be explained as the random outcome of an experiment using a set of fixed multinomial probabilities and time varying number of trials. Thus using the definition of

Muthuswamy, Satyanarayana, Gary A Davis, David M Levinson and Panos G Michalopoulos
Freeway Origin Destination Matrices: Not as Simple as They Seem
presented at Transportation Research Board 83rd Annual Meeting, January 10-14 2004, Washington DC.

the OD matrix as a percentage matrix, the idea of a time invariant OD matrix is posited. Since this research focuses on estimation of these percentage OD matrices, all further references to an OD matrix will correspond to this definition.

The Method

The methodology uses a traffic flow model and a search routine to match the off-ramp counts for a freeway section over the peak period in order to arrive at a time invariant OD matrix estimate, which corresponds to the underlying multinomial probabilities.

The off-ramp counts are an aggregate effect of choices and driving characteristics. If the traffic flow model can capture these effectively, the confidence in the OD estimate will be higher. The traffic flow model used in this research is the microscopic simulator, AIMSUN *(8)* and the search routine to minimize the objective function is MINOS *(9)*.

If the OD estimation results in a set of estimates, none having an exact match, the choice of the best solution could be made based on the closeness to the actual underlying OD matrix. In most cases, the OD matrix is unknown. This problem can be overcome by using a simulated data set. First, an OD matrix and on-ramp counts are assumed and the traffic is simulated in AIMSUN and the off-ramp counts are observed. Using these simulated off-ramp counts and the assumed on-ramp counts the OD matrix is estimated. If there are multiple estimates, the best solution can be chosen with reference to the assumed OD matrix.

The OD matrix is evaluated based on its ability to reproduce the off-ramp counts. However, the usefulness of the estimate is related to its ability to reproduce the traffic characteristics of the system as a whole. Typically, performance is evaluated using the system-wide Measure of Effectiveness (MOEs), like total travel, total travel time, average speed, total delay. Hence, the OD estimates will be evaluated not only by their ability to reproduce the off-ramp counts and the initial OD matrix, but also these system- wide MOEs. Multiple starting solutions (seeds) will be used to overcome the local minima problem, as MINOS is a gradient- based algorithm. A seed-generation process is also added to the OD estimation method, to provide the different starting solutions. The methodology is schematically shown in Figure 2. The method has been implemented as a FORTRAN77 program that interfaces the two components – AIMSUN and MINOS.

Starting solutions

The OD estimation problem has been posed as an optimization problem aimed at minimizing the weighted sum of squares of the deviation of the predicted off-ramp counts to the actual off-ramp counts. The reduced gradient algorithm as implemented in MINOS searches over the solution space. Typically, the search can start at any feasible point and proceed from there to finding the optimal point. Most search routines start at one of the bounds and search from there onwards. The efficiency of the search is a function of the nature of the problem, the algorithm, and the starting point.

Considering that the first two factors are held constant, a faster result can be expected if a start is made near the optimal solution. It is important to note that the algorithm is robust when a bad start results in the optimal solution. However, the speed of the convergence will be higher if we can start closer to the optimal solution. Hence, it becomes crucial to make a good guess of the solution and let the method search the space for the optimal solution from that point.

The starting solution (seed) is generally an 'educated guess' of the solution. To make that guess the available information on the problem must be used. In the OD estimation problem the detector data – on-ramp counts, off-ramp counts and mainline counts are readily available and could be used. In addition, the travel time and trip length details can be extracted from the geometry. Using this information, five different methods are used to estimate starting solutions for the minimization problem. The reason for the choice of these techniques is based on simplicity and ease of implementation.

The first method is a naïve equal-splitting percentage OD matrix. This is based on the assumption that all destinations are equally likely. The second method is one of the oldest estimates, proportional OD matrix. This OD matrix has percentages proportional to the off-ramp counts. The third method is the algorithm developed by *(10)* as implemented in *(11)*. It is the iterative trip table balancing method. The fourth seed was the gravity model proposed by Nancy Nihan as described in *(11)*. The impedance function is based on the gamma distribution. The last seed was generated using the turning percentages, assuming that the vehicles exit at an off-ramp based on the turning percentage independent of their origin.

TEST SITE – 1

The first test site was a small freeway section with one on-ramp and one-offramp. There are two origins and two destinations or four OD elements to be estimated. Simulated data was used to estimate the OD matrix. The data set was generated as shown in Figure 3. This data generation procedure removed the specification error in the OD estimation, because the model used for the data generation was the model used in the estimation. Using data corresponding to 1day's 3hour peak period composed of 5 min ramp counts, an OD matrix for this freeway section was estimated and the results are in Table 1. The results indicate that the final OD estimates are different, but have the same levels of performance with respect to reproducing the counts. As an additional experiment, data from 5 days having the same underlying OD matrix was used to estimate OD matrices. Table 1 has the related results. Interestingly, the results are similar – different estimates with same levels of performance with respect to reproducing the counts. The other important observation is that, the first row in the OD matrix (upstream mainline) is matched better, an indication of the dominance of the mainline. This warrants an investigation into the nature of the objective function.

This site has four OD elements to be estimated, but there are only two independent variables, as there are two constraints corresponding to each origin. Therefore, for a range of OD values, the objective function can be mapped and the 3-dimensional surface over which the optimal OD matrix is searched can be plotted. Figures 4 and 5 are the plots of the objective functions for 1-days and 5-days data respectively. The plots are related to the map as generated by the traffic flow model between the OD matrix and the ramp counts. The surface looks very spiky and the point at the bottom of both the figures is the optimal OD matrix. The plot indicates the existence of multiple local minima, a possible explanation for a set of OD matrices that are different but have the same levels of performance. In addition, the addition of more days has reduced the noise in the surface.

Using more days of data can be interpreted as adding more information into the system to facilitate the OD estimation. If the two days of data are the same, there is no additional information, so there is a need for difference in the data sets to help in the OD estimation. Hence using this observation and the mainline dominance, a data set was generated. This data set is called the *radical data set*. This represents a situation with one day having typical data and another day with on-ramp counts increased by one order of magnitude and held constant over the peak period, and the mainline counts halved. This is an attempt to make the data sets different and reduce the influence of the mainline, but have the same underlying OD matrix. Table 1 has the results from this experiment. The results show the same pattern, but seed 3 led to the actual underlying OD matrix. Since not all the seeds converged to the correct solution, there is not enough evidence to assure the finding of the actual solution. However, this experiment corroborates the existence of an identifiability issue related to OD estimation. In conclusion, the experiments with this test site helped check the implementation of the method and helped gain insight into the nature of the problem. The results indicate the existence of multiple solutions and an identifiability issue.

TEST SITE – 2

To test the performance of the method on a real site, a 6.5-mile northbound section of TH-169 was chosen between TH-55 and I-94. This section has 10 on-ramps and 11-off-ramps, an OD matrix with 11 origins and 12 destinations and 76 non-zero entries that need to be estimated. The data set used for the OD estimation was during the ramp-meter shutdown in the Twin Cities of Minneapolis and St. Paul. Five-minute counts for November 1, 2, and 3 2000 from the morning peak period (7am – 10am) were used for the OD estimation. To evaluate the performance of the method for a larger freeway section, a simulated data set was also generated as shown in Figure 3. The results from the simulated data set are discussed first, followed by the results from the real data set.

Simulated data set results

Table 2 has the performance of the OD estimates with respect to matching the off-ramp counts and system wide MOE's. Table 3 has the final OD estimates and the assumed OD matrix. The results indicate that the estimates match not only the counts but also reproduce the system statistics accurately. The final OD estimates from seed 3 and seed 5 are close to each other, but seed 2 is different. However, all the seeds perform well with respect to reproducing the counts.

Real data set results

Table 4 has the results for the OD estimation using the data from November 1, 2000. An additional experiment using a 15-minute warm-up time for the traffic simulator was also conducted, to account for the fact that when the simulation starts, the traffic model is empty, while in reality there are some vehicles in the system. The results are not as good as the simulated case. The mainline counts match best, and the final objective function values are higher than the simulated data set case. The only other common feature is that the estimate corresponding to seed 5 was the best estimate.

The relatively lower performance could be attributed to the lack of information in the system. In the simulated data set case, the on-ramp counts could be modified, but with real data, this cannot be done and so the specification of the problem needs to be changed to add the extra information. As a first modification, data from additional days were used.

In addition, two other modifications are proposed. The first adds the starting OD matrix in the objective function and the second is a 2-stage optimization process. The former, searches in the space around the starting solution under the assumption that the starting solution is a good estimate. The latter modification is based on the observation that the mainline counts are match best. Therefore, the OD estimation is broken down as a two-stage optimization where, in the first step, all the OD elements are estimated, and as a second-step the OD entries are estimated keeping the first row (upstream mainline, first origin) fixed. The related results are also in Table 4. The modifications improved the objective function value, but did not significantly improve the performance in terms of matching the ramp counts. This inability to improve the results could be a combination of the following issues: mismatch between the traffic flow model and real world, the identifiabilty issue, the nature of the map between the OD matrix and the off-ramp counts and bad data.

CONCLUSIONS

An offline method to estimate static OD proportions matrix for a freeway section over the peak period has been proposed. Most OD estimation methods have some form of a linear model to represent the relation between the on-ramp and the off-ramp counts and *(12)* showed that the approach having a traffic flow model embedded in the estimation process outperformed the linear models. This method enhances that approach. The appeal in the method is its simplicity. The problem has been defined as an optimization process with an embedded simulator that tries to find an optimal OD matrix that minimizes the weighted sum of the squared deviations of the off-ramp counts. The method is a combination of simulation and optimization, wherein the simulation component is the microscopic simulator AIMSUN and the optimization routine is MINOS.

The method does not need a prior estimate of the OD matrix, unlike most offline methods. As a part of the estimation process, using the time series of counts, estimates of the OD matrix are made using five different methods and starting solutions (seeds) are generated. These are used to start the search in the optimization process. This method can also be interpreted as an efficient updating scheme of the starting OD estimate.

Experiments were conducted on two test sites. The first test site was an imaginary test section of a freeway. The data set was simulated and the general observations on the results were the following. The starting solutions converged on different OD matrices that had comparable performance with respect to reproducing the counts, an indication of a many-to-one map between OD matrices and the objective function. The mainline proportions were matched best. The plots of the objective function over the space near the optimal point, gave very useful insight into the non-linearity and spiky nature of the map between the OD matrix and off-ramp counts as generated by AIMSUN. The experiments with the ‘radical’ day data set indicated the possibility of an Identifiability issue.

The second test site was TH-169. This had 76 non-zero entries that needed to be estimated. This is a sizeable increase in dimension over the network handled in *(12)*. Two sets of experiments were conducted on this site using a real and a simulated data set. The simulated data set resulted in very good estimates that matched the counts and the system statistics very well and the final solutions were close to each other and the true OD matrix. The real data set on the other hand did not produce equivalent results. New modifications to the method were proposed and the results did not improve significantly.

The inability to improve the estimates using real data to match the ramp counts as good as the simulated data set could be due to identifiability, bad nature of the map, inherent nature of the on-ramp demand patterns, and data discrepancy. The most important observation is the performance in the simulated data set. As discussed in the earlier sections, the process relating the on-ramp and off-ramp counts can be approximated as a data generation process. If a microscopic traffic simulator can approximate that process reasonably well, the OD matrix can be estimated accurately, which is evident from the simulated data set wherein the microscopic traffic simulator is the process that generated the data set. In addition, the mixed performance across the simulated and real data indicates that the ability to estimate the true OD matrix also depends on the on-ramp demand patterns. As a concluding note, the authors recommend against uncritical use of OD estimates and suggest that the ability of the estimation method to reliably estimate and OD matrix be verified before using in policy evaluation models.

Future work can enhance the performance of the microscopic traffic simulator with better calibration. The new modifications to the traditional objective function, the 2-step optimization and the new objective function need to be investigated. Finally, the identifiability issue of insufficient information in the off-ramp counts can be investigated by experimenting with alternate sites that have additional information, such as a small network with turning movement volumes.

REFERENCES

1. Ljung, L. (1987) *System Identification – Theory for the User*. Prentice-Hall, Englewood Cliffs, N.J, 1987.
2. Davis, G. (1993) Estimating the freeway demand patterns and impact of uncertainty on ramp controls. *Journal of Transportation Engineering,* Vol. 119 No. 4 July/August 1993, pp. 489-503.
3. He et. Al (2002) Estimation of time dependent O-D Demand and Route choice from Link Flows. *TRB 81st Annual Meeting* (Pre-print CD-ROM), Paper No. 02-2198 January 13-17 2002.
4. Ashok, K. (1996) *Estimation and prediction of time-dependent Origin-Destination flows,* Ph.D. Thesis, M.I.T
5. Ljung, L., Pflug, G., and Walk, H. (1992) *Stochastic Approximation and Optimization of Random Systems,* Birkhäuser, Berlin, 1992

6. Levinson, D et.al (2002) Ramp Meters on Trial: Evidence from the Twin Cities Ramp Meters Shut-off. *TRB 81st Annual Meeting* (Pre-print CD-ROM), Paper 02-2167 January 13-17 2002.
7. Michalopoulos, P and Hourdakis, J. (2002) Evaluation of Ramp Control Effectiveness in Two Twin Cities Freeways. *TRB 81st Annual Meeting* (Pre-print CD-ROM). Paper 02-2675 January 13-17 2002.
8. Transport Simulation Systems (2001) *Aimsun version 4.0 – Users Manual* April 2001.
9. Murtagh, B and Saunders, M (1983) *MINOS 5.5 Users Guide*.
10. Deming, W, and Stephan, F. (1940) On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *Annal of Mathematical Statistics 11* 1940, pp. 427-444.
11. May, D, and Willis, E, (1981) *Deriving Origin-Destination information from routinely collected traffic counts – Vol. 1,* Research Report, UCB-ITS-RR-81-8.

LIST OF TABLES AND FIGURES

Tables

Figures

Muthuswamy, Satyanarayana, Gary A Davis, David M Levinson and Panos G Michalopoulos
Freeway Origin Destination Matrices: Not as Simple as They Seem
presented at Transportation Research Board 83rd Annual Meeting, January 10-14 2004, Washington DC.

TABLE 1 Results for First Site

Assumed OD		d1	d2					
	o1	0.3250	0.6750					
	o2	0.2500	0.7500					
	Solution 1			Solution 2			Solution 3	
		d1	d2		d1	d2	d1	d2
1-days data	o1	0.3097	0.6904	o1	0.3235	0.6765	o1 0.3199	0.6801
	o2	0.5000	0.5000	o2	0.4051	0.5949	o2 0.4367	0.5633
R-squared		0.9535	0.9856		0.9617	0.9766	0.9565	0.9775
		f-value	2.435		f-value	2.338	f-value	2.389
		d1	d2		d1	d2	d1	d2
5-days data	o1	0.3152	0.6848	o1	0.3262	0.6738	o1 0.3183	0.6817
	o2	0.5000	0.5000	o2	0.3234	0.6766	o2 0.4211	0.5789
R-squared		0.9477	0.9749		0.9441	0.9792	0.9441	0.9762
		f-value	13.783		f-value	14.544	f-value	14.298
		d1	d2		d1	d2	d1	d2
Radical set	o1	0.3332	0.6669	o1	0.3135	0.6865	o1 0.3265	0.6735
	o2	0.1834	0.8166	o2	0.3241	0.6759	o2 0.2488	0.7512
R-squared		0.9851	0.9959		0.9890	0.9943	0.9849	0.9929
		f-value	4.902		f-value	4.446	f-value	3.446

TABLE 2 Results for Simulated Data set, Site-2

		Solution 5		Solution 3		Solution 2	
Fvalue		7.46		9.09		131.84	
Sq.dev		2209		2377		21740	
Ramp counts		% error	R-squared	% error	R-squared	% error	R-squared
D1		0.56	0.9927	1.94	0.9827	3.20	0.9345
D2		1.91	0.9742	1.92	0.9830	4.08	0.9477
D3		0.65	0.9794	0.96	0.9688	17.61	0.8815
D4		0.00	1.0000	1.07	0.9852	8.54	0.8058
D5		1.12	0.9883	2.07	0.9819	4.61	0.9202
D6		0.41	0.9941	1.18	0.9848	5.76	0.9297
D7		0.93	0.9932	1.43	0.9905	8.05	0.9298
D8		2.17	0.9721	1.45	0.9785	8.70	0.8152
D9		2.17	0.9328	1.78	0.9706	13.53	0.9207
D10		1.12	0.9768	2.46	0.9637	18.81	0.8970
D11		0.00	0.9921	1.25	0.9919	6.01	0.9802
System wide MOE's	Actual	Estimate	% error	Estimate	% error	Estimate	% error
Mean Flow (veh/hr)	6942	6938	-0.06	6949	0.10	6939	-0.04
Mean Speed (km/hr)	76.2	76	-0.26	74.3	-2.49	76.4	0.26
Total Travel (km)	112792.7	112778.7	-0.01	115292.6	2.22	112332.6	-0.41
Total Delay (hours)	63	65.8	4.44	110.7	75.71	55.5	-11.90

TABLE 3 OD estimates, Simulated Data set, Site-2

Actual

0.077	0.085	0.045	0.044	0.067	0.035	0.069	0.036	0.045	0.039	0.459
0.077	0.085	0.045	0.044	0.067	0.035	0.069	0.036	0.045	0.039	0.459
0	0.092	0.049	0.047	0.072	0.038	0.074	0.039	0.049	0.042	0.497
0	0	0.054	0.052	0.079	0.042	0.082	0.043	0.054	0.046	0.547
0	0	0	0.055	0.084	0.044	0.087	0.046	0.057	0.049	0.579
0	0	0	0	0.089	0.047	0.092	0.048	0.06	0.051	0.613
0	0	0	0	0	0.051	0.101	0.053	0.066	0.057	0.672
0	0	0	0	0	0	0.106	0.056	0.07	0.06	0.709
0	0	0	0	0	0	0	0.063	0.078	0.067	0.793
0	0	0	0	0	0	0	0	0.083	0.071	0.846
0	0	0	0	0	0	0	0	0	0.078	0.922
0	0	0	0	0	0	0	0	0	0	1

Solution 2

0.0798	0.0889	0.0336	0.0479	0.0815	0.0429	0.0762	0.0300	0.0299	0.0103	0.4791
0.0463	0.0486	0.0530	0.0375	0.0392	0.0014	0.0672	0.0510	0.0253	0.0526	0.5779
0	0.0551	0.0341	0.0173	0.0769	0.0135	0.0836	0.0534	0.0235	0.0505	0.5922
0	0	0.0518	0.0308	0.0717	0.0278	0.0646	0.0190	0.0681	0.0440	0.6222
0	0	0	0.0288	0.0704	0.0279	0.0444	0.0551	0.0841	0.0539	0.6353
0	0	0	0	0.0504	0.0535	0.0657	0.0530	0.0441	0.0675	0.6658
0	0	0	0	0	0.0399	0.0793	0.0485	0.0596	0.0720	0.7006
0	0	0	0	0	0	0.0977	0.0571	0.0494	0.0637	0.7321
0	0	0	0	0	0	0	0.0613	0.0722	0.0569	0.8096
0	0	0	0	0	0	0	0	0.0807	0.0774	0.8419
0	0	0	0	0	0	0	0	0	0.0749	0.9252
0	0	0	0	0	0	0	0	0	0	1

Solution 3

0.0780	0.0860	0.0447	0.0441	0.0678	0.0343	0.0698	0.0360	0.0452	0.0389	0.4552
0.0779	0.0857	0.0454	0.0445	0.0674	0.0353	0.0694	0.0362	0.0448	0.0388	0.4546
0	0.0930	0.0492	0.0482	0.0731	0.0383	0.0752	0.0393	0.0486	0.0420	0.4930
0	0	0.0543	0.0532	0.0806	0.0422	0.0829	0.0433	0.0536	0.0463	0.5436
0	0	0	0.0563	0.0852	0.0446	0.0877	0.0458	0.0567	0.0490	0.5748
0	0	0	0	0.0903	0.0473	0.0929	0.0485	0.0600	0.0519	0.6090
0	0	0	0	0	0.0520	0.1021	0.0533	0.0660	0.0571	0.6695
0	0	0	0	0	0	0.1077	0.0563	0.0696	0.0602	0.7062
0	0	0	0	0	0	0	0.0631	0.0780	0.0675	0.7915
0	0	0	0	0	0	0	0	0.0833	0.0721	0.8447
0	0	0	0	0	0	0	0	0	0.0786	0.9214
0	0	0	0	0	0	0	0	0	0	1

Solution 5

0.0766	0.0844	0.0448	0.0440	0.0670	0.0351	0.0692	0.0361	0.0450	0.0387	0.4589
0.0766	0.0843	0.0448	0.0440	0.0670	0.0351	0.0693	0.0361	0.0451	0.0388	0.4590
0	0.0913	0.0486	0.0476	0.0726	0.0381	0.0750	0.0391	0.0488	0.0420	0.4971
0	0	0.0534	0.0524	0.0799	0.0419	0.0825	0.0431	0.0537	0.0462	0.5470
0	0	0	0.0554	0.0844	0.0442	0.0872	0.0455	0.0567	0.0488	0.5778
0	0	0	0	0.0893	0.0468	0.0923	0.0482	0.0601	0.0517	0.6117
0	0	0	0	0	0.0514	0.1014	0.0529	0.0659	0.0567	0.6717
0	0	0	0	0	0	0.1068	0.0557	0.0695	0.0598	0.7081
0	0	0	0	0	0	0	0.0624	0.0778	0.0670	0.7928
0	0	0	0	0	0	0	0	0.0830	0.0714	0.8456
0	0	0	0	0	0	0	0	0	0.0779	0.9221
0	0	0	0	0	0	0	0	0	0	1

TABLE 4 Results, Real Data, Site-2

Solution 5	1-day	1-day warm up	3-days	New fobj	2-step
Fvalue	260.63	243.36	825.42	811.56	806.64
Sq.dev	21960	22626	75247	77924	74742
Ramp counts	% error	% error	% error	% error	% error
D1	14.58	15.41	17.25	19.21	17.45
D2	22.51	26.66	20.35	22.56	20.68
D3	15.81	15.65	20.28	21.27	19.12
D4	27.06	24.98	23.32	24.98	24.00
D5	24.43	19.06	32.29	34.38	32.42
D6	32.23	30.39	32.09	30.91	28.39
D7	18.91	24.00	17.85	19.07	17.95
D8	20.49	21.49	21.60	22.15	21.83
D9	29.44	36.86	27.93	28.06	28.12
D10	28.28	32.86	40.84	37.43	36.76
D11	6.03	6.65	7.20	7.24	7.19

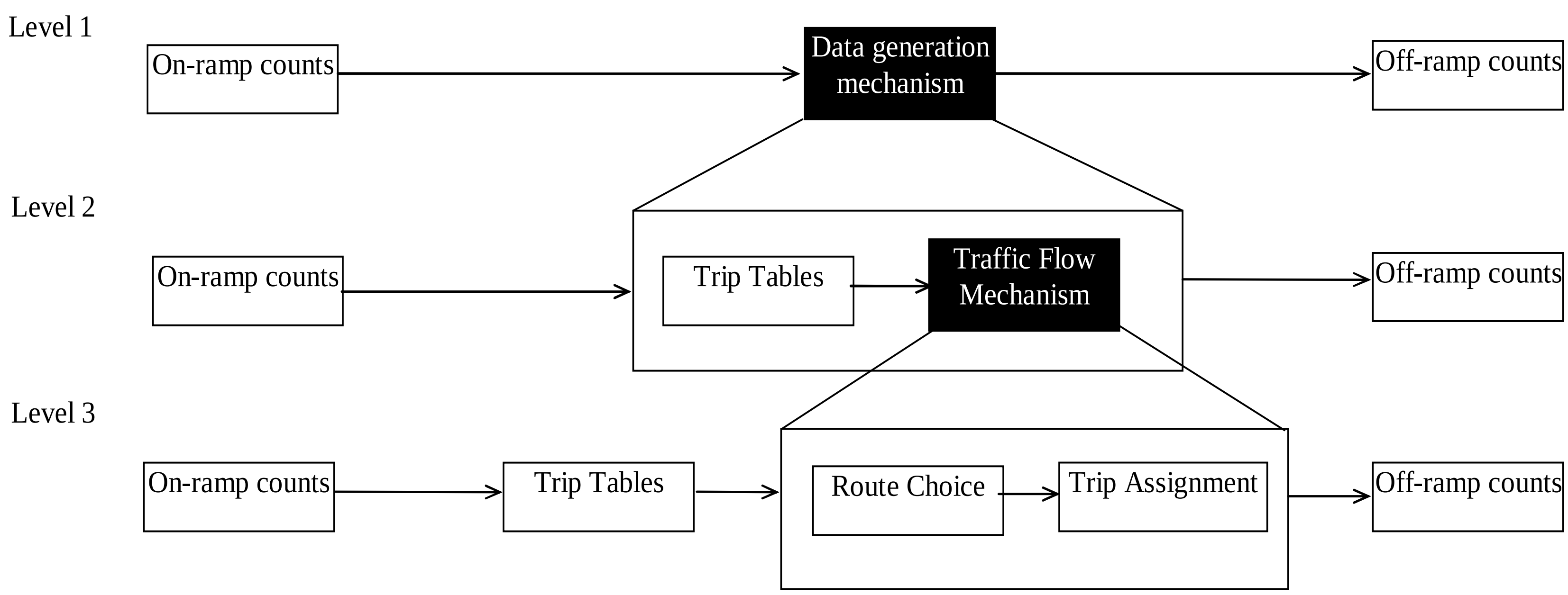

FIGURE 1 Data Mechanics

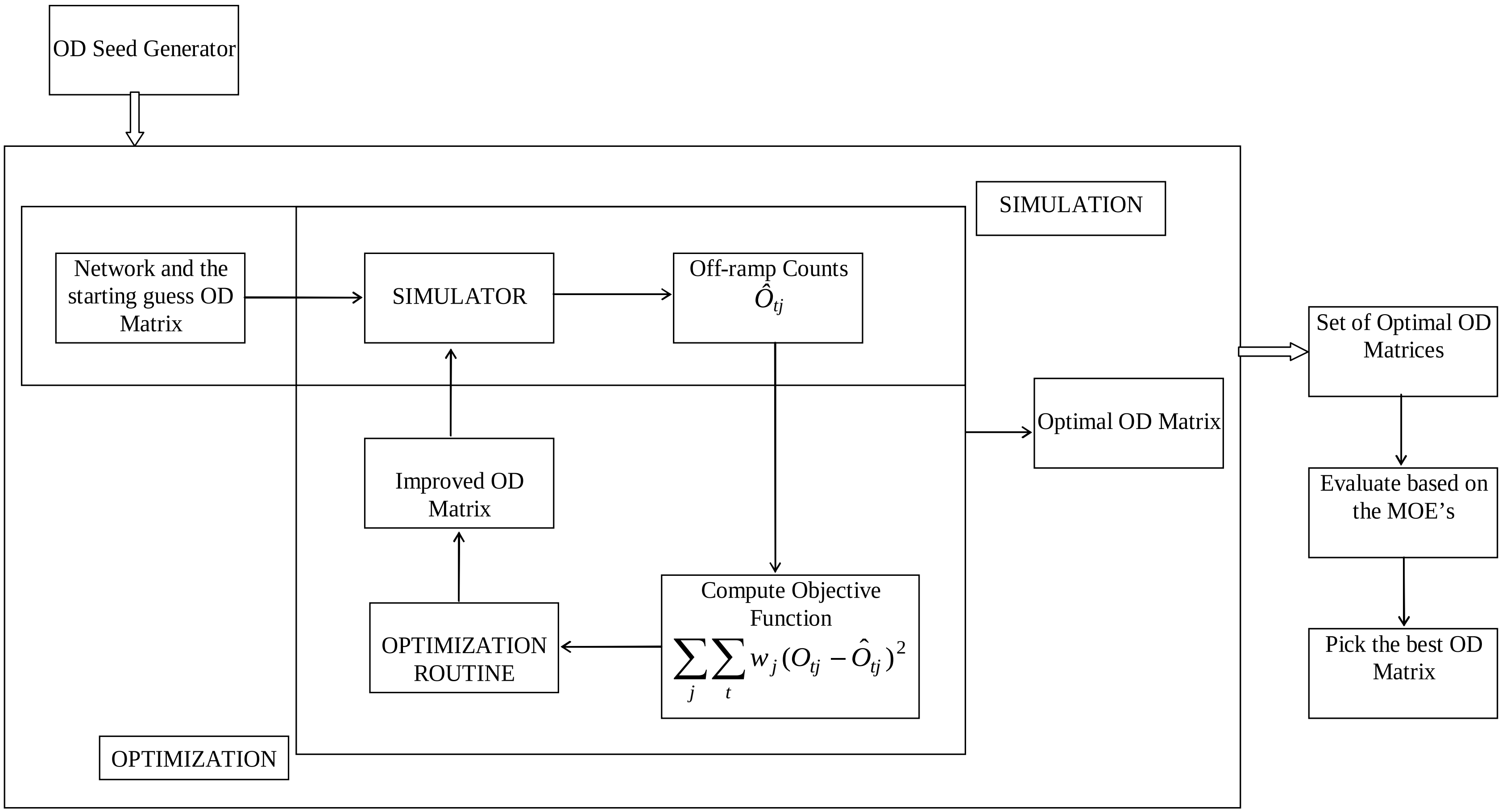

FIGURE 2 Methodology

Muthuswamy, Satyanarayana, Gary A Davis, David M Levinson and Panos G Michalopoulos
Freeway Origin Destination Matrices: Not as Simple as They Seem
presented at Transportation Research Board 83rd Annual Meeting, January 10-14 2004, Washington DC.

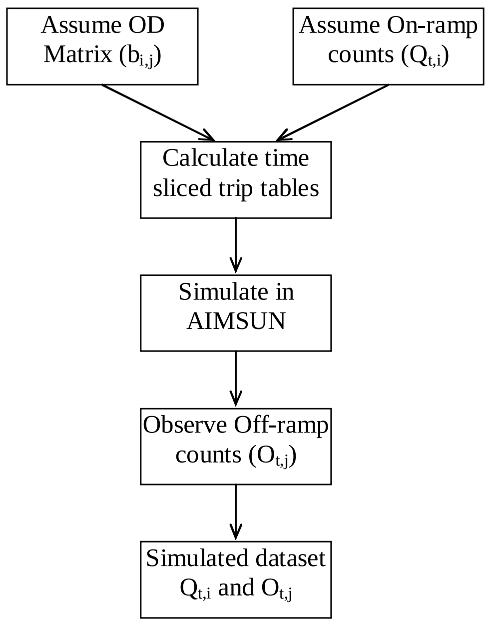

FIGURE 3 Data Generation Process

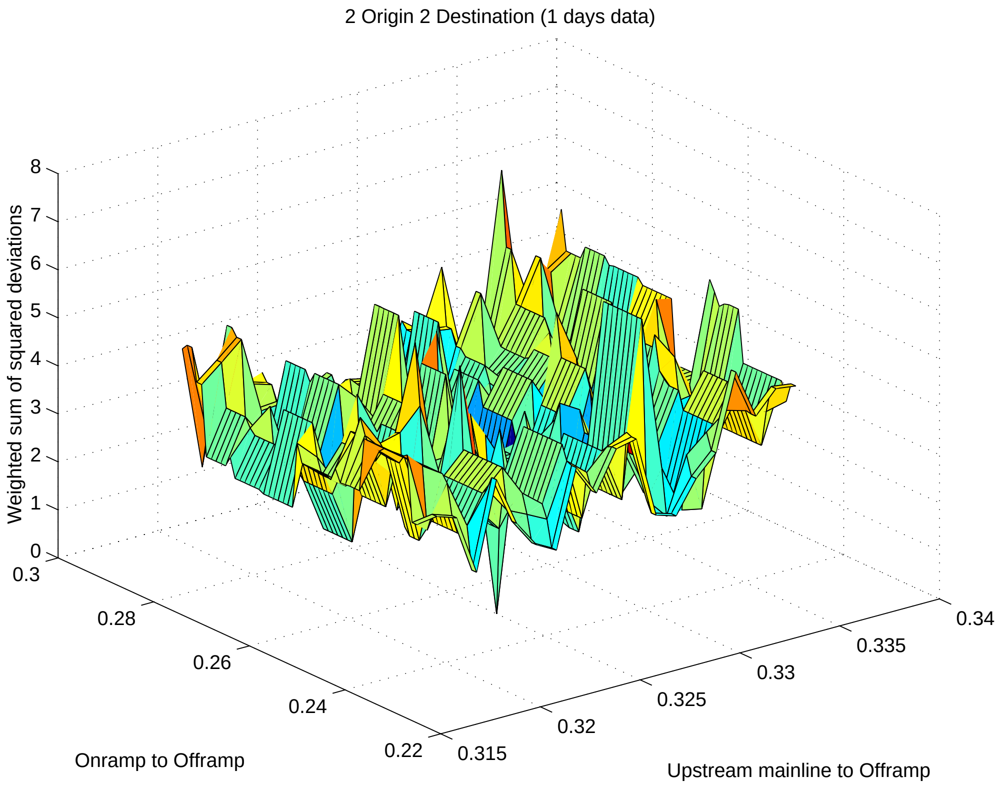

FIGURE 4 Objective Function for 1-Day's Data

Muthuswamy, Satyanarayana, Gary A Davis, David M Levinson and Panos G Michalopoulos
Freeway Origin Destination Matrices: Not as Simple as They Seem
presented at Transportation Research Board 83rd Annual Meeting, January 10-14 2004, Washington DC.

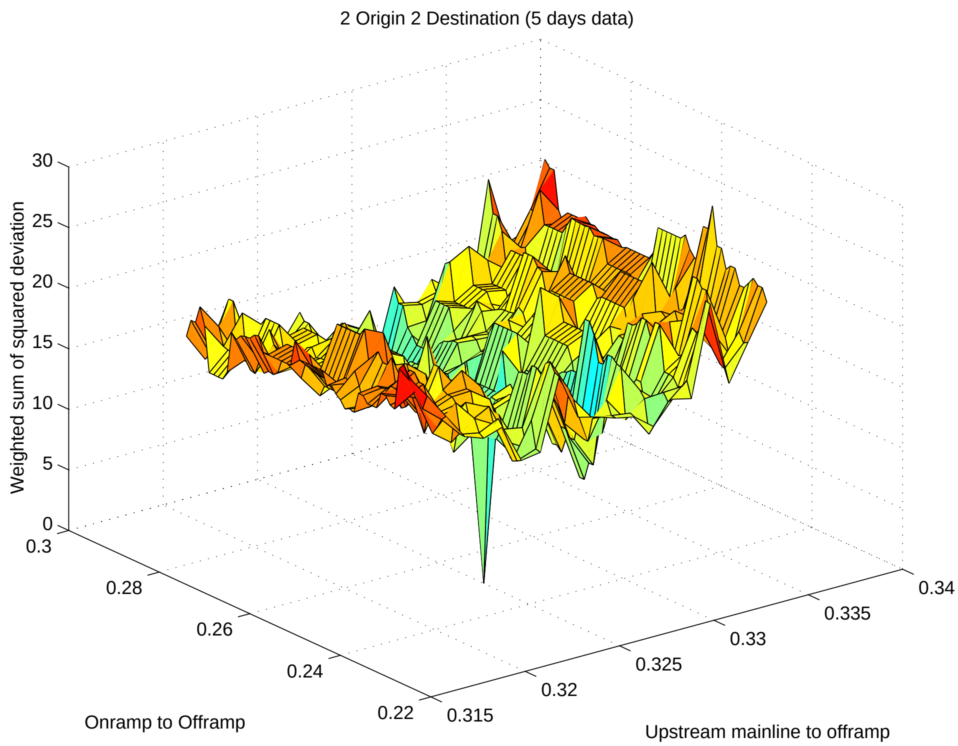

FIGURE 5 Objective Function for 5-Days Data