Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 23C (2017) 887–900

Transportation Research Procedia

www.elsevier.com/locate/procedia

22nd International Symposium on Transportation and Traffic Theory

Modeling Origin-Destination Uncertainty Using Network Sensor and Survey Data and New Approaches to Robust Control

Lee K. Jones [a],Nathan H. Gartner [b,*], Michael Shubov [a], Chronis Stamatiadis [b], David Einstein[a]

[a] *Department of Mathematical Sciences University of Massachusetts Lowell, Lowell, MA, 01854, USA*
[b] *Department of Civil and Environmental Engineering, University of Massachusetts Lowell, Lowell, MA, 01854, USA*

Abstract

This study develops new methods for network assessment and control by taking explicit account of demand variability and uncertainty using partial sensor and survey data while imposing equilibrium conditions during the data collection phase. The methods consist of rules for generating possible origin-destination (OD) matrices and the calculation of average and quantile network costs. The assessment methodology leads to improved decision-making in transport planning and operations and is used to develop management and control strategies that result in more robust network performance. Specific contributions in this work consist of: (a) Characterization of OD demand variability, specifically with or without equilibrium assumptions during data collection; (b) Exhibiting the highly disconnected nature of OD space demonstrating that many current approaches to the problem of optimal control may be computationally intractable (c) Development of feasible Monte Carlo procedures for the generation of possible OD matrices used in an assessment of network performance; and (d) Calculation of robust network controls, with state-of-the-art cost estimation, for the following strategies: *Bayes, p-quantile and NBNQ* (*near-Bayes near-Quantile*). All strategies involve the simultaneous calculation of controls and equilibrium conditions.

Peer-review under responsibility of the scientific committee of the 22nd International Symposium on Transportation and Traffic Theory.

Keywords: OD uncertainty; robust optimization; signal control; near-Bayes near-Quantile strategy; disconnected OD space

* Corresponding author. Tel.: +1-978-934-2289; fax: +1-978-934-3052
E-mail address: Nathan_Gartner@uml.edu

Peer-review under responsibility of the scientific committee of the 22nd International Symposium on Transportation and Traffic Theory.
10.1016/j.trpro.2017.05.049

1. Introduction

State of the art in transport network modeling is based on the estimation of travel demands between origins and destinations in the form of a trip table by various techniques. Once a trip table has been estimated by any of these methods, analysts typically proceed to evaluate the performance of various planning, design or operational alternatives for the transportation network assuming the demands established in the table are unaffected by those alternatives. The analysis culminates with a network assignment procedure that generates the volumes and costs associated with those volumes on the paths and links of the network. This classical approach is realized by network equilibrium models and algorithms as described in detail by Patriksson (1994), Florian and Hearn (1995), and more recently by Marcotte and Patriksson (2007). Since the input data are, generally, noisy or incomplete the resulting demand data used in the analysis involve considerable uncertainty. Furthermore, the assumption of immutability of the demands in the face of changing network conditions is quite unrealistic.

This study develops robust optimization (RO) methods for network assessment and control that take explicit account of demand variability and uncertainty while observing equilibrium conditions on the observation space. The methods consist of rules for generating possible origin-destination (OD) matrices and the calculation of quantile network costs. This approach can lead to improved decision-making in transport planning and operations and can be used to develop management and control strategies that result in more robust network performance. As pointed out by Bertsimas et al. (2011), RO is an approach to optimization under uncertainty, in which the uncertainty model is not stochastic, but rather deterministic and set-based. Instead of seeking to immunize the solution in some probabilistic sense to stochastic uncertainty, this approach seeks a solution that is feasible for *any* realization of the uncertainty in a given set. Whereas, in general, a robust solution is not optimal for all realizations of the uncertain data, this approach performs well even under severe case scenarios. A number of strategies that are developed in this paper are a manifestation of this approach. The motivation for this approach is twofold. First, the model of set-based uncertainty is a suitable application of parameter uncertainty. Second, computational tractability is an important motivation and goal. These characteristics have has been responsible for the considerable success of RO in many application areas.

Several approaches exist for network analysis with OD uncertainty. For example, Xie et al. (2010 and 2011), derive a maximum entropy estimate of the OD matrix using available link flow measurements. For this estimate, equilibrium assignments and total costs can be determined as a function of the signal controls and then minimized. Maximum likelihood and generalized least squares methods have been used by other authors (van Zuylen and Willumsen, 1980; Bell, 1991). Numerous other approaches have been proposed in the literature, see Bell and Iida (1997) and Yang (1994 and 1995). Bierlaire (2002) proposes a measure of quality for estimated OD tables, called the *total demand scale*. It measures the uncertainty due to the network topology and the route choice assumptions and is complementary with other measures of quality. However, in many cases, when existence of a compatible equilibrium is imposed as a constraint on the uncertainty space and some criterion is optimized to choose a single OD matrix for analysis, the problem may become (as shown below) infeasible.

In Jones et al. (2013), the unknown OD matrices were characterized by the space of all non-negative matrices with fixed row and column sums, corresponding to the given total out- and in- flows in vehicles per hour as specified by sensors at sources and sinks. Probability distributions on this space were then proposed based on additional data to represent the relative uncertainty of actual OD values, and Bayesian and minimax strategies were developed for calculating optimal traffic signal controls. For instance, the Bayes optimal controls are signal settings which minimize the expected system cost at equilibrium flows, or system-optimal flows, with respect to the probability distribution.[†]

Other authors have applied RO techniques to address uncertainty in special cases of traffic control. Yin (2008) investigates methods of signal optimization for pre-timed signal control under demand fluctuations. For the case of an isolated intersection, he develops a timing plan whose performance is near optimal in an average sense, and is fairly stable under any realization of uncertain traffic flows. Ukkusuri et al. (2010) present a dynamic system-optimal signal control model with a stochastic representation of the OD demands within an RO framework. Three probability distributions are considered for each OD demand from which a set of discrete scenarios is generated. Traffic flow

[†] Jones et al (2013) labeled the calculation *ECO* (for equilibrium constrained optimization); i.e., calculating signal controls that minimize network travel costs while observing the equilibrium constraints.

and performance is represented by means of an embedded cell transmission model. This approach is applied in a fairly simple network with a single multi-phase signalized intersection offering limited route choices.

The interdependency between signal timings and user route choices has been described extensively in the literature. Comprehensive discussions can be found in Lee and Machemel (2005).and in Mitsakis et al. (2011). Simonelli et al. (2012) consider the optimal network sensor location problem accounting for the variability of the OD matrix estimate. Castillo et al. (2015) provide an integrated approach, based on mathematical optimization, for the sensor location, flow observability, estimation and prediction problems in traffic networks.

The presentation in this paper proceeds as follows: Section 2 initially defines the uncertainty space by characterizing the space of OD matrices (termed *feasible*) in a network which are consistent with observations expressed as a set of path flow constraints. Assuming there is an equilibrium assignment that is also consistent with the observations, the uncertainty space is further narrowed to a space of *admissible* OD matrices. Section 3 provides a demonstration that the latter, more restricted OD space, may be highly disconnected. This has significant implications on the development of algorithms for calculating robust controls. Section 4 presents Monte Carlo methods for generating OD matrices that (a) satisfy observed constraints, and (b) also provide equilibrium solutions satisfying the constraints. Section 5 applies the methodology in a moderately sized network and calculates robust controls for two uncertainty sets: a Monte Carlo sample size of 100 feasible matrices consistent with the measured data and an independently generated Monte Carlo sample size of 100 admissible matrices having equilibrium flows that are consistent with the data. Controls and associated network costs are calculated for the following strategies: *Bayes, p-quantile and NBNQ* (*near-Bayes near-Quantile*). All calculations are based on the state-of-the-art cost function given in HCM2010. Furthermore, the Appendix provides a proof that the total cost function using the HCM2010 formula is separately, but not jointly, convex in the controls and in the flows. This function serves as a basis for all the subsequent cost calculations.

2. Characterization of the space of admissible OD matrices in a network

For a traffic network $\mathfrak{R}$, or sub-network, usually only partial information concerning the matrix of the flows (in veh/hr or veh/min) from each source to each sink (the OD matrix) is available. Such information can be obtained from sensor and/or survey data. Some sources may be also acting as sinks but there are no loops from a source to itself, hence the matrix has zeros at precisely these locations. Let m be the number of possible locations that may have positive entries. It is generally desirable to have a full OD matrix of flows from each source to each sink so that either equilibrium or system optimal routes for travelers can be determined. Total cost can then be determined for existing or planned networks or sub-networks. Knowledge of the OD matrix would allow the minimization of the cost by optimizing over signal settings and route volumes. Unfortunately, there is almost always considerable uncertainty in the origin-destination demands and one can only restrict the OD matrix to lie within a certain space depending on the information available. We consider cases where this information consists of linear constraints satisfied by the OD entries (perhaps obtained from a survey sample of users), or linear constraints on link flow measurements (perhaps measured by sensors) which are feasible for the OD matrix. Since linear constraints in the OD values can be expressed as linear constraints in the path flows, and linear constraints on link flows may be rewritten as linear constraints on path flows, we can thus consider a general case where all constraints are expressed in terms of path flows.

The uniform distribution on the set of possible OD tables plays an important role in the analysis: uniformly generated tables from the space may be used to estimate expected cost if the proposed distribution on the OD space is uniform. Or, one may require cost estimates assuming a conditional Gaussian (or any other) distribution on the OD space. Uniform generation of tables from the space can be used via the accept-reject or importance sampling method to generate the conditional Gaussian and hence the cost estimate for use in finding optimal controls.

Alternatively, we may consider the uniform distribution as least informative on the space and calculate, for fixed controls and fixed traveler routing, the p-quantile cost for this distribution (that cost which exceeds exactly 100 p% of the costs under the uniform assumption). This criterion can then be optimized appropriately.

In general, the incomplete information, in the form of linear equality or inequality constraints on certain link and path flows, may or may not include total source/sink out/in flows. In this paper we consider mathematical properties of the set of OD tables that satisfy these general constraints resulting from the observed data; i.e., there are flow assignments for these OD tables that satisfy the constraints. We call these *feasible OD tables* $\mathfrak{M}_0$.

Some transportation planners may observe their data in transient states before equilibrium develops. $\mathfrak{M}_0$ is then the appropriate space for designing optimal controls. Others may want to assume that data was collected from a system in equilibrium. Hence, we further consider properties of the class $\mathfrak{M}$ of such matrices subject to the additional condition of equilibrium in the observation phase. We call the subspace $\mathfrak{M}$ the set of *admissible OD tables*. In particular, when we consider the set of OD matrices as a continuum, we show that the subspace $\mathfrak{M}$ of OD matrices such that there are equilibrium flows that satisfy the constraints, may not be convex or even not connected. This makes the estimation (and use in optimal control strategy development) of a single (so-called, best) OD matrix in $\mathfrak{M}$ more challenging as optimization is required over a domain which may be highly separated. Even with this possibility our Monte Carlo approach allows us to accurately estimate expected costs over $\mathfrak{M}$ (with relative accuracy proportional to the inverse square root of the sample size) and solve for robust control strategies, allowing for cost control under uncertainty. We show how to generate OD tables for use in this general situation. We give analysis and estimate p-quantile control performance here for a representative moderate sized network.

Since traffic planners may want to consider either $\mathfrak{M}_0$ or $\mathfrak{M}$, depending on the conditions during observation, we consider both $\mathfrak{M}$ and $\mathfrak{M}_0$ in our example. We generate OD tables in $\mathfrak{M}$ by first generating them in $\mathfrak{M}_0$ and then checking whether an equilibrium assignment satisfies the constraints. For our moderate sized network example, this procedure provides a sufficient sample for 2% relative accuracy. For very large networks, more efficient methods have been developed in Jones and O'Neil (2002). We intend to implement these in future work. The problem of generating such matrices is also important in statistical analysis for testing independence of row-column effects after sampling two properties of individuals in a population, see Diaconis and Efron (1985) and Holmes and Jones (1996).

We assume flows take on non-negative real values. More general constraints on path flows (other than just total in- and out- flow at sinks and sources) must be considered. Total in/out flows may or may not be included for some or all sources/sinks. If we assume only fixed OD matrix row and column sums as constraints on our candidate OD matrix, then there always exists equilibrium travel assignment flows satisfying the OD demands given by this matrix.[‡] However, in general, there may be OD matrices for which flows exist satisfying the OD demand and satisfying a set of linear constraints on path flows, but for which there are no equilibrium assignments satisfying the flow constraints.

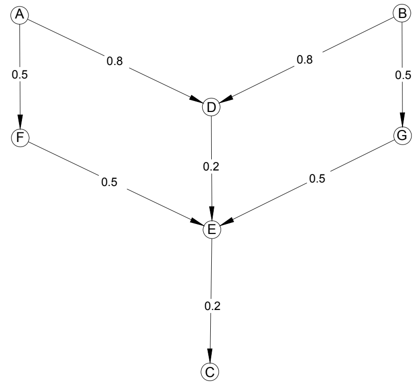
It is straightforward to give examples of OD demand for which there exist flows satisfying both the demands and the given path flow constraints, but for which no such flows are equilibrium. Simply make the constraints force flows which are non-equilibrium. Hence, a general problem presents itself: For a given class of path flow constraints (including constraints on the OD matrices), characterize the set $\mathfrak{M}$ of OD matrices which admit equilibrium flows that satisfy both the OD demands given by the matrix and the given path flow constraints; i.e., characterize the set of admissible OD tables. The equilibrium requirement introduces non-convex constraints. In fact, as we will see in the next section, $\mathfrak{M}$ may not be connected. We will also show that $\mathfrak{M}$ could be the union of a number of separated domains (connected components) which is exponential in the number of sinks. Hence the optimization of some criterion over the OD space $\mathfrak{M}$ leading to an optimal choice for "the OD matrix" may well be computationally intractable as well as not having the robustness of our approach, which keeps the OD space variable. We propose to generate (uniform) random admissible OD tables so that we may characterize network performance with given signal settings and obtain robust network controls. We do this by accurately estimating expected costs over the uncertainty space using the tables generated. This avoids the potential problem of finding one of many separated sub-domains where an OD criterion is optimized. We believe this approach can be feasible for very large networks using a variety of new Markov Chain methods from Jones and O'Neil (2002).

To demonstrate the possible disconnected nature of the admissible OD space, we first present an example of a small disconnected $\mathfrak{M}$ in the next section. Then we combine numerous copies as sub-networks forming a network of k sinks for an OD space with 2^k connected components.

[‡] This is because, in the proof of existence of equilibrium, only total flows from each source to each sink are constrained and thus the row and column sums are automatically fixed.

3. On the Connectedness of the Admissible Origin-Destination Space

Consider the network in Fig.1. There are only 2 sources A and B and only one sink C. The links are the directed arcs AD, AF, BD, BG, FE, GE, DE and EC. As shown in Fig.1, these have lengths .8, .5, .8, .5, .5, .5, .2 and .2 respectively. Assume the costs are a constant $C^* \cdot (length\ of\ a)$ per unit of flow (e.g., per vehicle) for all links a, plus a total delay cost $D_a(V_a)$ per unit flow (vehicle) due to traffic controls at the head of arc a, for all arcs except EC, and where V_a is the flow on a. We use the "state of the art" delay formula in HCM2010:



Flow constraints:

$$V_{DE} = 1/2$$

$$V_{EC} = 1 + \varepsilon$$

Figure 1: A simple network with a disconnected admissible O-D space.

$$D_a(V_a) = d_{3,a} + \frac{0.5 \cdot c \cdot (1 - x_a)^2}{1 - min\{x_a, s^{-1} \cdot V_a\}} + 900 \cdot T \cdot \left[\frac{V_a}{s \cdot x_a} - 1 + \sqrt{\left(\frac{V_a}{s \cdot x_a} - 1\right)^2 + \frac{K}{(s \cdot x_a)^2} \cdot V_a} \right]$$

The control vector x has components x_a representing green time ratio (green time divided by the positive cycle time c) for (intersection at head of) arc a. See a detailed description in the Appendix.

We assume all degrees of saturation X_a (i.e. the quantities $(s \cdot x_a)^{-1} \cdot V_a$) are less than one; then the term $(1 - \min\{x_a, s^{-1} \cdot V_a\})$ above simplifies to $(1 - s^{-1} \cdot V_a)$. We take all green time fractions x_a equal to 1/3 (this might be the case with no pedestrian traffic at E, moderate pedestrian traffic at D and high pedestrian traffic at F and G). Hence, we assume that all 7 delay functions (for the 7 links connecting to a signal) are the same function $D(V_a)$. It is shown in the Appendix that the first and second derivatives of each $D(V_a)$ *wrt.* V_a are positive if $K < 4sx_a$, which we assume. Our proof works for any function $D(V_a)$ as long as the first two derivatives are positive. Hence our example is valid for very general networks.

The constraints are: the total flows on DE and EC are $1/2$ and (1+ε) respectively. We take ε such that $D(\varepsilon) = .5\,D(0) + .5\,D(1/2)$. Since $D(V_a)$ is strictly increasing, this equation has a unique solution for ε in (0, .5). (Note $V_a \leq 1 + \varepsilon < 1.5$ so s, K can be chosen as assumed.)

Consider the following OD matrices:

$$M_1 = \begin{bmatrix} 1 \\ \varepsilon \end{bmatrix} \quad \text{and} \quad M_2 = \begin{bmatrix} \varepsilon \\ 1 \end{bmatrix}$$

where the entries are the total flows from A to C and B to C. For M_1 we describe an equilibrium flow assignment which satisfies the two constraints: first put flow ½ (out of source A) on path AFEC and ½ on path ADEC. Then put flow ε (out of source B) on path BGEC and 0 on path BDEC. Now a vehicle considering traveling from B along path BGEC would have a cost of $[D(\varepsilon) + D(\varepsilon) + .5\,C^* + .5\,C^* + .2\,C^*]$ while one traveling from B on path BDEC would have the same cost, namely $[D(0) + D(½) + .8\,C^* + .2\,C^* + .2\,C^*]$. Similarly, a vehicle considering traveling from A along path AFEC would have a cost of $[2D(½) + .5\,C^* + .5\,C^* + .2\,C^*]$, while one traveling from A on path ADEC would have the same cost, namely $[2D(½) + .8\,C^* + .2\,C^* + .2\,C^*]$. Hence M_1 is admissible. By symmetry, M_2 is also admissible.

Next note, from the constraint on EC, that the total flow into C is (1+ε). Now, if $\mathfrak{M}$ is connected, there exists a continuous real function p(s), so that

$$M(s) = \begin{bmatrix} 1 + \varepsilon - p(s) \\ \\ p(s) \end{bmatrix}$$

is admissible for $0 \leq s \leq 1$ with $p(0) = \varepsilon$ and $p(1) = 1$ (i.e., $M(0) = M_1$ and $M(1) = M_2$). For any small δ ($0 < \delta < 0.1$) we can find an s in (0,1) with $p(s) = \varepsilon + \delta$. Assuming M(s) is admissible, when changing from M(0) to M(s) we are increasing the total flow out of node B by δ, sending $\alpha\delta$ of the increase along path BGEC and $(1-\alpha)\delta$ along path BDEC, where $\alpha = \alpha(\delta)$ is a real number depending on δ. To maintain the constraints the flow out of A along path ADEC must decrease by $(1-\alpha)\delta$ and decrease by $\alpha\delta$ along path AFEC. We now write the equilibrium conditions for potential flow out of nodes A and B, respectively:

$$\begin{aligned} .5C^* + D(1/2 - \alpha\delta) + .5C^* + D(1/2 - \alpha\delta) + .2C^* \\ = .8C^* + D(1/2 - (1-\alpha)\delta) + .2C^* + D(1/2) + .2C^* \end{aligned} \tag{1}$$

$$.5C^* + D(\varepsilon + \alpha\delta) + .5C^* + D(\varepsilon + \alpha\delta) + .2C^* = .8C^* + D\big((1-\alpha)\delta\big) + .2C^* + D(1/2) + .2C^* \tag{2}$$

Now we show that $\alpha = \alpha(\delta)$ is bounded by 1 above and 0 below: After cancelling constant terms in (1), we note that α cannot exceed 1, since then $1/2 - \alpha\delta$ would be less than both $1/2 - (1-\alpha)\delta$ and $1/2$, while D is strictly increasing. By a similar argument, α cannot be less than 0, since $1/2 - \alpha\delta$ would be greater than both $1/2 - (1-\alpha)\delta$ and $1/2$. In (2) we see that α cannot exceed 1 as flows must be nonnegative. After cancelling constant terms in (2) , we see that α cannot be negative since then, because D is strictly increasing, we have

$$D(\varepsilon + \alpha\delta) + D(\varepsilon + \alpha\delta) < 2D(\varepsilon) = D(0) + D(1/2) < D\big((1-\alpha)\delta\big) + D(1/2)$$

which contradicts (2).

Finally we rewrite (1) and (2) as first order Taylor series in δ obtaining, after cancellation of constant terms and noting that $o(\alpha\delta) = o(\delta) = o\big((1-\alpha)\delta\big)$ since α is bounded,

$$2D'(1/2)(-\alpha\delta) = D'(1/2)(-(1-\alpha)\delta) + o(\delta) \tag{1'}$$

$$2D'(\varepsilon)(\alpha\delta) = D'(0)\big((1-\alpha)\delta\big) + o(\delta) \tag{2'}$$

Since δ may be arbitrarily small and positive we can equate coefficients of δ as it approaches 0 solving for a limiting value for α. For (1′) we get the limiting $\alpha = 1/3$ while for (2′) we get the limiting $\alpha = D'(0)/\big(2D'(\varepsilon) + D'(0)\big)$ which is less than 1/3 since $D'(\varepsilon) > D'(0)$. Hence the connectedness assumption is not valid.

The disconnectedness of the OD space in a larger network may result in an extremely large number of dis-connected sub-regions, making the optimization of an objective function over such a space prohibitively time consuming. Consider a simple extension of the previous example, shown in Fig. 2, consisting of k subnetworks, each of the form of the network in Fig. 1, with an additional source S and links from S to each of the sink nodes $C_1, C_2, \ldots C_k$. For each of the k subnetworks the same link lengths and link flow constraints are used as in Fig. 1, while the links from node

S to C_1, C_2, … C_k are all of equal length, one unit. An additional constraint is that the summation of the flows on links SC_i is equal to the outflow at S of 100:

$$\sum_i V_{SC_i} = 100$$

It can be easily proven that for such a network there will be at least 2^k disconnected regions in $\mathfrak{M}$.

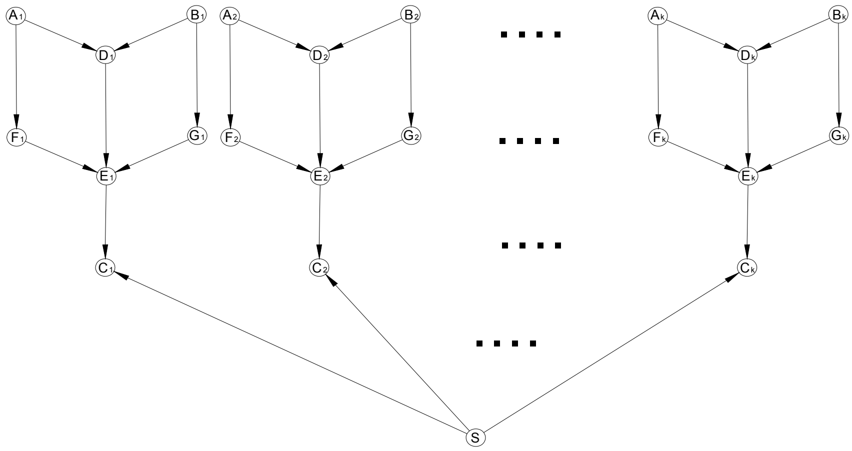


Fig. 2: Network with a highly disconnected admissible OD space $\mathfrak{M}$.

3. Algorithm for uniformly generating admissible OD tables in a general network

Since $\mathfrak{M}$ is not necessarily connected, rejection algorithms can be used; namely, first, uniformly generate an OD table from $\mathfrak{M}_0$ (consisting of those for which assignments that satisfy the flow constraints exist) and then check if it admits an equilibrium solution which also satisfies the constraints, rejecting it if not and accepting otherwise. For uniform generation of the OD tables having assignments satisfying the constraints, our approach is to combine the "hit-and-run algorithm" by Lovasz and Vempala (2004) applied to the convex set $\mathfrak{M}_0$,which iteratively chooses a random point along a random line segment from one boundary of $\mathfrak{M}_0$ to another, with a set constructed in the flow space which projects onto the line segment. The details are as follows:

Consider an OD matrix M as an m dimensional vector which corresponds to a feasible (for M) n-dimensional vector π in the path flow space. In addition to $\pi_i \geq 0,\ \pi, M$ satisfy further constraints of the form:

$$\sum_{paths\ i} b_i^k \cdot \pi_i \leq d_k \qquad \text{for } k = 1,2, \dots K_1 \quad \text{(inequality constraints)} \tag{3}$$

$$\sum_{paths\ i} c_i^k \cdot \pi_i = f_k \qquad \text{for } k = 1,2, \dots K_2 \quad \text{(equality constraints)} \tag{4}$$

$$\sum_{\substack{paths\ i \\ from\ j\ to\ l}} \pi_i = M_{jl} = M_{jl}(\pi) \qquad source\ \ j \neq sink\ l \tag{5}$$

We assume the constraints imply that the path flow space is bounded and there is at least one path from every source to every sink but no circuits.

In matrix notation the above constraints become $\pi \geq \mathbf{0}, \boldsymbol{B}\pi \leq \boldsymbol{d}, \boldsymbol{C}\pi = \boldsymbol{f}$, and $\boldsymbol{J}\pi = M(\pi)$. We can assume, by suitable equation reductions, that $K_2 \leq n$ and that $\boldsymbol{C}$ is of full rank. Similarly $\boldsymbol{J}$ is of rank $m \leq n$. We rewrite (4) and (5) in block matrix form as $\boldsymbol{A} \cdot \boldsymbol{z} = \boldsymbol{t}$, with:

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{J} & -\boldsymbol{I} \\ \boldsymbol{C} & \boldsymbol{O} \end{bmatrix} \qquad \boldsymbol{z} = \begin{bmatrix} \pi \\ M \end{bmatrix} \qquad \boldsymbol{t} = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{f} \end{bmatrix}$$

where $\boldsymbol{I}$, $\boldsymbol{O}$ are identity, zero matrices respectively and $\mathbf{0}$ is the zero vector. The steps of the algorithm are then as follows:

Step #1: Start with an OD matrix M_0 together with a feasible path flow π_0. We assume that M_0 is in the relative interior of $\mathfrak{M}_0$; that is, in the (possibly lower dimensional) space described by (4) and (5), a neighborhood of M_0 is contained in $\mathfrak{M}_0$.

Step #2: Next pick a random direction of change, u, in the OD space.

1. If there are no equality constraints (4) (i.e. $\mathfrak{M}_0$ has an interior), this is achieved by first generating m independent standard normal random real values Z_{jl} and forming the matrix u, which we write as a vector, with 0 at pre-specified locations and whose other components are Z_{jl} divided by the square root of the sum of all Z_{jl}^2.
2. If there are equality constraints of the form (4), we first generate a direction vector v in $(n + m)$ dimensional space $\mathbb{R}^{n+m}$ from a sequence of normals as in step (2.1). The first n components are potential changes in π_0 and the last m are potential changes in M_0 . To get u project v onto the linear subspace of $\mathbb{R}^{n+m}$ described by (4) with each f_k set equal to 0 and (5). From the block matrix description we want the solution z to $\boldsymbol{A}z = 0$ which is closest to v. This is given by $z^* = [\boldsymbol{I} - \boldsymbol{A}^t(\boldsymbol{A}\boldsymbol{A}^t)\boldsymbol{A}]v$. To verify this we only need show that $\boldsymbol{A}z^* = 0$ and $(v - z^*)$ is orthogonal to $(z^* - z)$ for any solution z of $\boldsymbol{A}z = 0$. This is demonstrated by a straight forward evaluations of $\boldsymbol{A}z^*$ and $(v - z^*)^t(z^* - z)$. Finally we take u to be the vector consisting of the last m components of z^* which is the desired random direction in OD space.

Step #3: We move both in the u direction and the –u direction until we hit the boundary of $\mathfrak{M}_0$ at $M_0 + p_1 u$ and $M_0 - p_2 u$, where p_1 and p_2 are positive reals. Then we pick a point uniformly at random on the line segment connecting these two points in $\mathfrak{M}_0$. Determining the two boundary points requires solving two linear programs:

$$p_1 = \arg\left\{ \max_{\substack{M(\pi)=M_0+pu,\\ satisfying\ (3),(4)\\ \pi\geq\mathbf{0}}} p \right\}$$

and

$$p_2 = \arg\left\{ \max_{\substack{M(\pi)=M_0-pu,\\ satisfying\ (3),(4)\\ \pi\geq\mathbf{0}}} p \right\}$$

where $M(\pi)$ is the OD matrix arising when the path flow vector is π.

The random OD matrix chosen is a convex combination of $M_0 + p_1 u$ and $M_0 - p_2 u$. It has a feasible flow vector which is the same convex combination of the feasible flow vectors for the latter two matrices. Now replace M_0 and π_0 in **Step #1**. with the randomly generated OD matrix and its feasible flow vector. Repeat the steps and perform N iterations. Results in Lovasz and Vempala (2004) indicate that the randomly chosen OD matrix after N repetitions is nearly uniformly distributed for $N = O(m^2)$. Thus we can generate a sample of size q of uniformly distributed OD matrices in $\mathfrak{M}_0$ by choosing every N^{th} matrix in $(q \cdot N)$ repetitions. To generate matrices randomly in $\mathfrak{M}$, accept/reject every N^{th} matrix if it does/does not have an equilibrium flow satisfying the constraints. For moderate sized networks sufficient numbers of acceptances can be rapidly generated. In the case study described in the next section, roughly one out of 168 in $\mathfrak{M}_0$ are also in $\mathfrak{M}$. If the acceptance rate becomes too small, as might be the case with very large networks, methods of regression/attraction from Jones and O'Neil (2002) may be implemented to much more efficiently generate tables uniformly in $\mathfrak{M}$.

For a given OD matrix M we write total cost for signal settings x as $\sum_{\text{arcs } a}\{V_a \cdot C(x_a, V_a)\}$, where the V_a are system optimal or equilibrium flows. See Appendix where it is proved that $C(x_a, V_a)$ is increasing in V_a and is separately convex in x_a and V_a. It follows that $V_a \cdot C(x_a, V_a)$ is separately (but not jointly) convex in x_a and V_a. Hence total system cost, $\sum_{\text{arcs}} \{V_a \cdot C(x_a, V_a)\}$, is convex separately in x and π. These properties demonstrate that alternating minimization procedures could be used to minimize, given a fixed OD matrix, system cost for system optimal flows as a function of controls.

The nature of such minimization for a given OD matrix M, in the case of equilibrium optimal flows, remains an interesting question. If the vector function, whose i^{th} component is the cost per vehicle to traverse path i, is strictly monotone, then there exists a unique equilibrium solution for each control vector x. If this unique solution has reasonable smoothness properties, then there are unconstrained optimization algorithms for minimizing total cost under equilibrium assignments as a function of x. For many practical networks we believe the associated cost vector function using our delay functions is not strictly monotone. In these cases, various constrained optimization methods have been studied, see Patriksson and Rockafellar (2002) for more details.

For the case study network described in Section 4, given an x, we average the total cost over a Monte Carlo sample of OD matrices where flows for each matrix are approximately (using MATLAB convex programming) equilibrium for that matrix and the given x. Then we minimize this average over x using general MATLAB optimization to calculate the Bayes controls. We also compute the 90$^{\text{th}}$ percentile of these costs and minimize over x (to calculate 0.9-quantile controls). Finally, we minimize the relative maximum of two functionals over x (Near-*Bayes* Near-*Quantile*, *NBNQ*).

4. Case Study: City of Lowell network

The methodology described in this paper is applied to a medium-sized network representing the City of Lowell, Massachusetts with 5 origins, of which each one is also a possible destination, 7 signal-controlled intersections and 21 bi-directional arcs forming several alternative routes from each origin to each destination (Fig. 3).

In previous work (Jones et al., 2013), optimization was over spaces of OD matrices for which $\mathfrak{M} = \mathfrak{M}_0$. Monte Carlo methods were used to generate very large samples of matrices from $\mathfrak{M}_0$ to estimate expected costs over $\mathfrak{M}_0$ as functions of the controls from the samples. These same expectations are computed here for the more general $\mathfrak{M}$ under consideration. The chosen network, which has a limited number of constraints, is used to illustrate the approach. The acceptance rate – admissible matrices *vs.* feasible matrices – is as many as one in 168. For larger networks, with many constraints, the acceptance rate may be quite small so that the earlier referenced regression/attraction methods of Jones and O'Neil (2002) might be used.

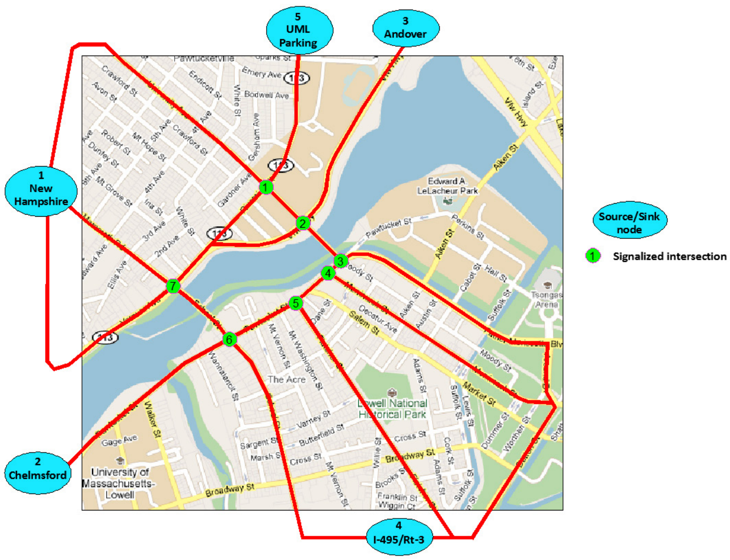


Fig 3. City of Lowell network (all route segments are bi-directional).

Based on sensor data, flows on the following four links are constrained to be within + or - 10% of the observed flows given in Table 1. This yields 8 inequality constraints.

Table 1: Lowell Network, Observed Traffic Flows

Initial Node	Terminal Node	Observed Flow (veh/hr)
New Hampshire	Node 1	1947
New Hampshire	Node 7	1697
495/Rt. 3	Node 3	1280
Node 6	495/Rt. 3	2285

The total flow in the network, based on survey data, is always below 19,500. Hence we have a total of 9 inequality constraints

Since the OD matrix of the network has five origins and five destinations, it contains $m = 20$ non-zero elements. To obtain a uniformly distributed OD table in $\mathfrak{M}_0$ steps 1 to 3 of the algorithm are repeated N = 400 times, thus satisfying the required $O(m^2)$requirement. Examination of the correlation between the tables at the starting points and the one generated 400 steps later indicates that they are approximately independent and uniformly distributed.

In order to obtain a sample of 100 matrices in $\mathfrak{M}$ a total of $q = 16860$ matrices in $\mathfrak{M}_0$ were generated. Signal controls are optimized for these 100 matrices for which equilibrium exists, as well as for 100 matrices randomly selected out of the 16860, using the *Bayes, min 0.90 quantile,* and *NBNQ* strategies. The total costs estimated for each of these cases are listed in Table 2 and Table 3.

The computational results are calculated as follows:

- The *Bayes* Mean and .90 quantile costs.
  The *Bayes* Mean is the average cost over the given Monte Carlo sample of matrices. The calculation is executed as follows:
  1. For fixed controls x, determine an equilibrium assignment and its associated cost for each matrix;
  2. Calculate the average cost C_{avg} for all the matrices in the sample;
  3. Repeat steps (1.) and (2.) for new controls until the minimum average cost $C_B = \min_x C_{avg}$ is reached. This is the *Bayes* Mean (this is the cost in column 1, row 1).
  4. For the optimal control strategy determined in step (3.), calculate the 90th percentile cost; i.e., the 90th lowest cost in Table 2 and Table 3 (this is the cost in column 2, row 1).
- The min .90 quantile cost. The steps are as follows:
  1. For fixed controls x, determine an equilibrium assignment and its cost for each matrix;
  2. Calculate the 90th percentile of the costs for the sample, $C_{0.90}$;
  3. Repeat steps (1.) and (2.) for new controls until the minimum of the 90th percentile cost is reached, $C_Q = \min_x C_{0.9}$ (this is the cost in column 2, row 2).
  4. For the optimal control strategy determined in step (3.), calculate the average for the sample (this is the cost in column 1, row 2).
- The *NBNQ* (near-*Bayes*, near-*Quantile*) cost. In this case we minimize the maximum of two relative quantities. The steps are as follows:
  1. For fixed controls, determine an equilibrium assignment and cost for each matrix;
  2. Determine the maximum of the following two relative costs for the sample at these controls:

     $$C_{avg}/C_B \quad \text{and} \quad C_{0.9}/C_Q$$

     Thus we have for the controls for the *NBNQ* strategy:

     $$x_{NBNQ} = \arg\min_x \max\left\{C_{avg}/C_B\,;\,C_{0.9}/C_Q\right\}$$
  3. Repeat steps (1.) and (2.) for new controls until the minimum of the costs corresponding to the controls calculated in step (2.) is reached. Calculate the average cost for all matrices at these controls (this is the cost in column 1, row 3).

4. For the optimal control strategy determined in step (3.), calculate the 90th percentile cost; ; i.e., the 90th lowest cost in Table 2 and Table 3 (this is the cost in column 2, row 3).

The *NBNQ* strategy is a compromise between the *Bayes* and the *.90 quantile* strategies and, as such, is a conservative approach the controls of which can provide robust or risk-averse performance.

Table 2: Costs for Sample of 100 Feasible Matrices (cost in 10^6 sec)

Strategy	Mean values (Relative Error)	.9 quantile values
Bayes	18.31 (2.0%)	22.96
min .9 quantile	18.32 (2.0%)	22.81
NBNQ	18.33 (2.0%)	22.87

Table 3: Costs for Sample of 100 Admissible Matrices (cost in 10^6 sec)

Strategy	Mean values (Relative Error)	.9 quantile values
Bayes	15.87 (2.1%)	20.09
min .9 quantile	16.14 (1.9%)	19.62
NBNQ	15.91 (2.0%)	19.66

The corresponding green times at the seven signalized intersections of the network, obtained for each of the above strategies, are given in Tables 4 and 5. The following data were used: cycle time (c) of 100 sec; lost time of 10 sec and min/max green splits of 20/70 sec, respectively. The green time values in the tables are listed for the phase in the NE/SW direction only.

Table 4: Green Times at Signalized Intersections (sec) – Sample of 100 Feasible Matrices

	Intersection						
Strategy	1	2	3	4	5	6	7
Bayes	56.9	53.2	20.0	70.0	58.3	47.6	22.3
min .9 quantile	56.5	53.6	20.0	70.0	59.1	46.0	20.5
NBNQ	57.4	53.0	20.0	70.0	63.5	45.6	22.3

Table 5: Green Times at Signalized Intersections (sec) – Sample of 100 Admissible Matrices

	Intersection						
Strategy	1	2	3	4	5	6	7
Bayes	55.8	57.0	20.0	70.0	65.6	51.2	22.0
min .9 quantile	65.4	58.0	20.0	70.0	68.0	49.0	34.7
NBNQ	59.9	57.5	20.0	70.0	68.1	49.0	25.4

Results indicate that the *min 0.90 quantile* and *NBNQ* strategies provide a more robust solution than the *Bayes* strategy. In the admissible set, this improvement is approx. 2.3% of costs (*min .9-quantile* over *Bayes*, table 3, column 3). Furthermore, it is evident that the equilibrium admissible matrices provide a more robust solution than the global set of feasible OD matrices. For example, for the *min .9-quantile* we get a cost of 19.62 for the admissible set (Table 3, column 2, row 2), which is 14% lower than the 22.81 cost for the feasible set (Table 2, column 2, row 2). Similar improvements are obtained for the other strategies. Further details of the computational results are given in Shubov (2016).

5. Conclusions

This paper introduces robust optimization techniques that address the problem of variability and uncertainty in transport network analysis and control. The approach taken in this paper is to model the uncertainty space as a set of *feasible* or *admissible* OD matrices for the network. The set of *feasible* matrices are those matrices for which there exist flows satisfying the demands that are consistent with available observed data (such as those obtained from sensors and/or surveys). The set of *admissible* matrices are those for which there exist such flows which are additionally equilibrium. Since both sets are very complex, we resort to Monte Carlo techniques for generating random matrices from either set in order to estimate various expectations of network cost over the given set. In general, lack of connectedness of the admissible matrices requires that one generate an admissible matrix by first generating a feasible matrix and then accepting it only if there is an associated equilibrium flow which is data consistent. In the test case described in this paper as many as 1/168 of the feasible matrices generated were admissible. For very large networks such acceptance rates might be much smaller, in which cases alternative, more efficient, methods can be implemented.

Which space to employ, feasible or admissible, depends on the assumptions during data collection. Once a reasonable number of matrices are generated from the desired space, one can develop robust control strategies and estimate future network performance. We computed controls for three such strategies: *Bayes, 0.90 quantile and NBNQ* (near *Bayes*, near *Quantile*). Each of these strategies provides a different measure of robustness. The *Bayes* strategy minimizes the average costs over the sample set of OD demands. The *0.90 quantile* strategy minimizes the 90th percentile cost over the sample set. The *NBNQ* strategy is a compromise among the two prior strategies: It is designed to provide performance that is close to the best that can be obtained under *Bayes* conditions yet does not depart too far from the most beneficial controls under the 90th percentile most costly origin-destination demands. As such, this is a conservative approach whose controls can provide robust or risk-averse performance.

Specific contributions in this paper include: (1) characterization of the space of admissible OD matrices in a network for which flow information is available for selected links; (2) demonstration that this space is not necessarily connected, in fact possibly highly disconnected – which, naturally, has implications for optimization over the space; (3) development of efficient algorithms for generating feasible and admissible OD tables for arbitrary networks even if the space is highly separated; and (4) application of the approach in a moderately sized network. Although the table generation was done with a uniform distribution, cost expectations may be estimated from such a sample for any distribution with a given analytic formula. One should note that the approach of dealing with uncertainty described in this paper would apply similarly in a dynamic case given suitable assumptions on the distribution of the OD demands.

Appendix A: Convexity of Delay and Cost Functions

To calculate the costs, we use the HCM2010 delay formula introduced in Section 3 above. This formula is an outgrowth of the classic Webster (1958) model adopting the Kimber-Hollis (1979) transformation. A detailed description of the analytics is given in Rouphail et al. (2001). The following expression

$$C(x_a, V_a) = C^* \cdot length_a + D_a(V_a) =$$

$$= t_a + d_{3,a} + \frac{0.5 \cdot c \cdot (1 - x_a)^2}{1 - \min\{x_a, s^{-1} \cdot V_a\}} + 900 \cdot T \cdot \left[\frac{V_a}{s \cdot x_a} - 1 + \sqrt{\left(\frac{V_a}{s \cdot x_a} - 1\right)^2 + \frac{KD_a(V_a)}{(s \cdot x_a)^2} \cdot V_a} \right] \tag{6}$$

is the link traversal cost (i.e. time) per unit flow plus the costs (times) per unit flow due to delays represented by HCM2010 (see also Appendix A in Jones et al. (2013)).
$t_a = C^* \cdot length_a$ is the time of traversal of link a per unit flow.
$d_{3,a}$ is the initial queue delay experienced by newly arriving vehicles when a queue from the previous period is present. To this is added a fixed queue delay (at the head of a) per unit flow.
c equals the cycle length.(note that the progression factor *PF* is set to 1);
x_a is green time divided by the cycle length (g_a/c);
s is the saturation flow (at a);
$K = 8 \cdot k \cdot I / T$ which is non-negative, where k, I, T are constants ass in HCM2010.
Finally, c_a is the capacity of a which equals $s \cdot x_a$ and $X_a = V_a/c_a$ is the degree of saturation.

A.1. Proof of convexity in x_a

Clearly the third (quadratic) term on the right hand side of (6) has a positive second derivative (is strictly convex) in x_a (since c is positive). We only need to show that the second derivative with respect to x_a of the fourth term in (6) is non-negative when $(sx_a)^{-1}V_a \leq 1$.

Let $z = (sx_a)^{-1}V_a$. We first see that $f(z) = (z-1) + ((z-1)^2 + \rho z^2)^{1/2}$ has nonnegative first and second derivatives in z for $z < 1$, where ρ is non-negative:

$$f'(z) = 1 + ((z-1)^2 + \rho z^2)^{-1/2}(z - 1 + \rho z) \text{ and}$$

$$f''(z) = [(1+\rho)((z-1)^2 + \rho z^2) - (z - 1 + \rho z)^2]((z-1)^2 + \rho z^2)^{-3/2}$$

It is easy to see that that the f'' term in brackets is non-negative. f' is increasing and, since $f'(0) = 0$, we have demonstrated the claim for $f(z)$. The third term in (6) is just $900T\, f((sx_a)^{-1}V_a)$ with $\rho = K/V_a$ and its second derivative with respect to x_a is

$$900T[2s^{-1} \cdot x_a^{-3} \cdot V_a \cdot f'((sx_a)^{-1}V_a) + s^{-2} \cdot x_a^{-4} \cdot V_a^2 \cdot f''((sx_a)^{-1}V_a)]$$

which is non-negative.

A.2. Proof of convexity in V_a

Now it follows easily by differentiating twice that the third term in (6) is convex in V_a. (Note that $(sx_a)^{-1}V_a < 1$ implies $s^{-1}V_a < 1$). Repeating the substitution $z = (sx_a)^{-1}V_a$, we see that $g(z) = (z-1) + ((z-1)^2 + \gamma z)^{1/2}$ has a non-negative second derivative in z for $z < 1$ and $\gamma < 4$ where $\gamma = K/sx_a$:

$$g'(z) = 1 + ((z-1)^2 + \gamma z)^{-1/2}(z - 1 + \gamma/2) \text{ and}$$

$$g''(z) = [((z-1)^2 + \gamma z) - (z - 1 + \gamma/2)^2]((z-1)^2 + \gamma z)^{-3/2}$$

The g'' term in brackets reduces to $\gamma - \gamma^2/4$ which is non-negative if $K \leq 4sx_a$. Hence, $g''(z) > 0$.

Now the fourth term in (2) is just $900Tg((sx_a)^{-1}V_a)$ and its second derivative with respect to. V_a is $900T[(sx_a)^{-2}g''((sx_a)^{-1}V_a)]$ which is positive.

REFERENCES

Bell, M.G.H. (1991). The estimation of origin-destination matrices by constrained generalized least squares. *Transp. Res.* 25B, pp. 13-22.

Bell, M.G.H. and Y. Iida (1997). *Transportation Network Analysis*. J. Wiley & Sons.

Bertsimas, D. , D.B. Brown and C. Caramanis (2011). Theory and Applications of Robust Optimization. *SIAM Review, Vol, 53, No. 3*, pp. 464-501.

Bierlaire, M. (2002). The total demand scale: a new measure of quality for static and dynamic origin–destination trip tables. *Transportation Research B*, Vol. 36, 837–850.

Castillo, E., Zaharias Grande, Aida Calvino, W.Y. Szeto and Hong K. Lo (2015). A state-of-the-art review of the sensor location, flow observability, estimation, and prediction problems in traffic networks. *Journal of Sensors*, Volume 2015, Article ID 903563,

Diaconis, P. and B. Efron (1985). Testing for independence in a two-way table: new interpretations of the chi-square statistic. *Annals of Statistics 13(3),* 845–874.

Florian, M. and D. Hearn (1995). Network Equilibrium Models and Algorithms. *Handbooks in* OR & MS, Vol. 6 (M.O. Ball et al., editors), Elsevier Science.

HCM2010. *Highway Capacity Manual*. Chapter 18: Signalized Intersections. Transportation Research Board, Washington, D.C.

Holmes, R.B. and L.K. Jones (1996). On uniform generation of two- way tables with fixed margins and the conditional volume test of Diaconis and Efron. *Annals of Statistics 24(1),* 64–68.

Jones, L.K. and P.J. O'Neil (2002). On Markov Chain Monte Carlo algorithms for computing conditional expectations based on sufficient statistics. *Journal of Computational and Graphical Statistics 11(3),* 660–677.

Jones, L.K., R. Deshpande, N.H. Gartner, C. Stamatiadis and F. Zou (2013). Robust controls for traffic networks: the near-*Bayes* near-Minimax strategy. *Transp. Res. Part C* , Vol. 27, pp. 205-218.

Kimber, R. and E. Hollis (1979). *Traffic Queues and Delays at Road Junctions*. TRRL Laboratory Report, 909, U.K.

Lovász L. and S. Vempala (2004). Hit-and-Run from a Corner. *Proc. of the 36th ACM Symposium on the Theory of Computing (STOC '04)*, Chicago; *SIAM J. Comput., 35(4)*, 985–1005.

Marcotte, P. and M. Patriksson (2007). Traffic Equilibrium. *Handbooks in OR & MS, Vol. 14* (C. Barnhart and G. Laporte, editors), Elsevier Science.

Mitsakis, E., Salanova, J. M., Giannopoulos, G. (2011). Combined dynamic traffic assignment and urban traffic control models. 14th EWGT, *Procedia - Social and Behavioral Sciences 20*, 427–436.

Patriksson, M. (1994). *The Traffic Assignment Problem – Models and Methods*, VSP BV, Utrecht, The Netherlands.

Patriksson, M. and R. T.Rockafellar (2002). A Mathematical Model and Descent Algorithm for Bilevel Traffic Management. *Transportation Science 36(3):*271-291.

Rouphail, N., A. Tarko and J. Li (2001). Traffic Flow at Signalized Intersections. Ch. 9 in *Traffic Flow Theory – A state-of-the-Art Report* (N. Gartner et al., editors), Transportation Research Board, Washington, D.C.

Shubov, M. (2016). *Lowell Network Traffic Cost Minimization.* Technical Report, University of Massachusetts Lowell, July 2016.

Simonelli, F. Vittorio Marzano, Andrea Papola and Iolanda Vitiello (2012). A network sensor location procedure accounting for o-d matrix estimate variability. *Transportation Research B, Vol. 46(10),:* 1624-1638

Ukkusuri, S.V., G. Ramadurai and G. Patil (2010). A robust transportation signal control problem accounting for traffic dynamics. *Comp. & Opers. Res.*, Vol. 37, pp. 869-879.

van Zuylen, H.J. and L.G. Willumsen (1980). The most likely trip matrix estimated from traffic counts. *Transp. Res.* 14(B), pp. 281-293.

Webster, F. (1958). *Traffic Signal Settings*. Road Research Technical Paper No. 39.

Xie C., K.M. Kockelman and S.T. Waller (2010). Maximum entropy method for subnetwork origin-destination trip matrix estimation. *Transportation Research Record 2196*, J. of the TRB, pp. 111-119.

Xie C., K.M. Kockelman and S.T. Waller (2011). A maximum entropy-least squares estimator for elastic origin-destination trip matrix estimation. 19th International Symposium on Transportation and Traffic Theory, *Procedia Social and Behavioral Sciences 17*, pp. 189–212.

Yang, H., Y. Iida and T. Sasaki (1994). The equilibrium-based origin-destination matrix estimation problem. *Transp. Res.* 28B, pp. 23-33.

Yang, H. (1995). Heuristic algorithms for the bilevel origin-destination matrix estimation problem. *Transp. Res.* 29B, pp. 231-242.

Yin, Y. (2008). Robust optimal traffic signal timing. *Transp. Res.*, Vol. 42B, pp. 911-924.