A GIS-BASED TOOL FOR THE ESTIMATION OF NETWORK-WIDE FULL MARGINAL COST OF HIGHWAY TRANSPORTATION

Kaan Ozbay[1], Ozlem Yanmaz-Tuzel[2], Bekir Bartin[3] and Sandeep Mudigonda[4]

ABSTRACT

In this paper, a GIS-based interactive computer tool, developed for the evaluation and analysis of network-wide full marginal costs (FMC) of highway transportation in Northern New Jersey (NJ), is presented. The first part of the paper is concerned with the development of an O-D based multiple-path finding algorithm. The second part of the paper deals with the estimation of various transportation cost categories including, *vehicle-operating*, *congestion*, *accident*, *air-pollution*, *noise*, and *infrastructure costs*. The total, marginal and average cost functions of each cost category are developed using NJ specific data. Since link-based demand loading methodology is unable to realistically capture the idea of unit increase in demand required to estimate marginal costs, a trip-based FMC estimation methodology is proposed. This new approach considers a set of feasible paths attractive to each O-D pair. A GIS-based multiple-path FMC estimation tool based on a Graphical User Interface (GUI) developed in ArcGIS using Visual Basic and C-programming language. The proposed tool will help planners to calculate the true trip costs between different O-D pairs for various user-defined scenarios causing demand/supply changes.

KEY WORDS

network-wide full marginal cost, trip-based estimation, multiple shortest path algorithm, GIS in transportation planning.

INTRODUCTION

Estimation of full highway transportation costs has long been one of the major concerns of the policy makers due to its importance in decision-making and policy considerations. The main objective is to ensure that prices paid by the transportation users reflect the true cost of providing transportation services.

Full Marginal Cost (FMC) is an effective measure to determine the true cost of transportation. It is defined as the cost of an additional unit of output. The term "output"

[1] Dept. of Civil and Environmental Engineering, Rutgers University, 623 Bowser Road, Piscataway, NJ, 08854, USA, Phone 732/445-2792, FAX 732/445-0577, kaan@rci.rutgers.edu
[2] Dept. of Civil and Environmental Engineering, Rutgers University, 623 Bowser Road, Piscataway, NJ, 08854, USA, Phone 732/445-0576, FAX 732/445-0577, yanmaz@rci.rutgers.edu
[3] Dept. of Civil and Environmental Engineering, Rutgers University, 623 Bowser Road, Piscataway, NJ, 08854, USA, Phone 732/445-3162, FAX 732/445-0577, bbartin@rci.rutgers.edu
[4] Dept. of Civil and Environmental Engineering, Rutgers University, 623 Bowser Road, Piscataway, NJ, 08854, USA, Phone 732/445-3162, FAX 732/445-0577, sandeepm@rci.rutgers.edu

is defined as the representation and simplification of the overall utilization of product systems by means of selected units. Berechman et al. (1984) define terms **intermediate** and **final outputs** for transportation systems. The usage of these outputs in the cost calculation depends on the purpose of the analysis. **Intermediate outputs** such as vehicle- miles or vehicle-hours are mainly used to evaluate the technical efficiency of a system. On the other hand, **final outputs** (also called demand oriented measures by Berechman et al. 1984), such as number of trips or number of passengers, are used to analyze the overall efficiency and effectiveness of the system.

The motivation of this study lies in the estimation of the FMC of highway transportation based on **final outputs**. In this study, **trip** is taken as the major output measure that will enable policy makers to better understand the policy implications of additional travelers. The originality of this paper lies in four areas.

- Development of an origin-destination (O-D) based multiple-path finding algorithm considering a set of feasible paths between each O-D pair attractive to the travelers.
- Development of a GIS-based multiple-path FMC estimation tool in ArcGIS, based on Graphical User Interface (GUI), using Visual Basic for Applications (VBA) and C-programming language.
- Estimation of various transportation cost categories including, *vehicle-operating*, *congestion*, *accident*, *air-pollution*, *noise*, and *infrastructure*.
- Application of the proposed methodology to the Northern NJ road network.

This paper has the following sections. The next section presents the previous work regarding the estimation of marginal cost methodologies. The following section explains GIS-based interactive computer tool. Then in the next section, the proposed methodology is applied to Northern New Jersey network after estimating road network specific cost functions for each cost category. Finally, the last section presents the key findings of our analyses and conclusions.

BACKGROUND AND LITERATURE REVIEW

As mentioned earlier, FMC is defined as the cost of an additional unit of output. In transportation facilities there is no unique or exact way of representing the output. Depending on the output definition, two distinct approaches can be followed: (1) distance-based marginal cost estimation, (2) trip-based marginal cost estimation.

The distance-based approach considers intermediate outputs, and estimates marginal cost, based on distance, using a cost function specific to a segment of roadway (Levinson et al. 1996, Mayeres et al. 1996, Anderson and McCullough 2000, Banfi et al., and Zhang and Levinson 2005). This approach assumes that each link is loaded with the same amount of unit demand irrespective of its location, or origin and destination. Thus, in this case, network-wide marginal cost represents the overall effect of all the demand changes in all the links on the network. However, this methodology is unable to realistically capture the effect of unit increase in demand required to estimate marginal costs. It is not possible “to uncouple the individual effects of changes in one link on the other links by this method” (Safirova and Gillingham 2004). This problem encountered by the distance-based approach has been addressed by Jara-Diaz et al. (1992), who proposed the concept of O-D specific marginal transport cost estimation at a zonal level in a trucking system. In fact, Ozbay et al. (2001) further improved this idea, and proposed a trip-based FMC estimation methodology, where FMC of a trip is calculated along the shortest "travel time" path between each O-D pair.

Parameters other than travel-time, such as volume, highway-type, trip-distance etc. also affect users' path choice between a particular O-D pair, and consequently calculation of the FMC. Therefore, not only the shortest "travel time" path but also all feasible paths between each O-D pair attractive to the travelers should be considered while calculating the FMC.

There are a number of several approaches that can be employed to determine multiple paths between O-D pairs, mainly based on the determination of k-shortest paths that satisfy user defined constraints. Dial (2000) and Sherali et al. (2003) adapted a labeling approach, which includes all paths that are optimal with respect to a label, such as time, cost or distance. Alternatively, heuristic methods are deployed by many researchers (see for example Scott et al. 1997, and Akgun et al. 2000). These methods are mainly based on link elimination and penalty rules, where the network is modified after finding the shortest path.

Existing k-shortest path algorithms may be divided in two; algorithms that allow paths to have repeated links (see for example Shier 1979, and Eppstein 1998), and the ones that only consider acyclic paths (see for example Lawler 1976, Hadjiconstantinou and Christofides 1999). A comparative numerical study by Brander and Sinclair (1995), show that within the class of algorithms considering only acyclic paths that are applicable to directed graphs (such as transportation networks), the method proposed by Lawler (1976) offers the best performance. Lawler (1976) provides an exact algorithm for finding the K-shortest paths between a single origin and a single destination. This algorithm first defines the set of all paths in a network and determines the shortest path of this set. Then the remaining paths are divided into mutually exclusive subsets, and the shortest path for each of these subsets is determined. Later van der Zijpp and Catalano (2005) extended this approach by adding constraints related to detour and overlap.

GIS BASED CONSTRAINED K-SHORTEST PATH ALGORITHM

The method proposed in this paper provides an alternative to algorithms mentioned above. The main advantage of the proposed method is that it finds the constrained shortest paths directly, instead of selecting the paths from a large set of overall paths. In addition, it allows defining a path choice set on the basis of objective constraints such as limitation of travel time, minimum required disjoint links and limitation on total number of links.

The algorithm adapted in this paper to find k-shortest paths from an origin to a particular destination in a directed acyclic transportation network is mainly based on iterative application of modified Dijkstra's algorithm. The basic idea in Dijkstra's algorithm is to find the shortest path from one origin to all destinations. However, in our case, the main focus is to find O-D specific shortest paths, thus to reduce the complexity of the algorithm, Dijkstra's approach is modified such that it terminates as soon as a path from the selected origin to the specified destination is found. Once the shortest path between the particular O-D pair is found, the network is modified by randomly deleting two links from the shortest path while keeping the network connected. The modified Dijkstra's algorithm is then reapplied to the modified network to find the next candidate path. The iteration continues until a user defined number of paths has been found, or no more paths which satisfy the required constraints can be found.

Since the main idea of the multiple-path approach is to find set of pre-defined number of the *feasible* paths that are attractive to the travelers between the selected O-D pair, several constraints are introduced into the proposed algorithm. These constraints can be summarized as follows:

Constraint (1) *Travel Time Constraint*: Let t_i be the travel time of the candidate path i, t_1 be the travel time of the first shortest path. Path i is feasible if the following condition holds:

$$t_i \leq \phi_{max} t_1 \tag{1}$$

Path i is infeasible if constraint (1) is not satisfied. ϕ_{max} is a user defined limitation factor on travel time. For illustration purposes a value of $\phi_{max} = 1.3$ is selected.

Constraint (2) *Rate of Disjointness Constraint*: Let $A_i = \{a_{1i}, a_{2i}, \ldots a_{Mi}\}$ denote the links of the i^{th} candidate path where M is the number of links of the candidate path, and $A_1 = \{a_{11}, a_{21}, \ldots a_{N1}\}$ denote the links of the first shortest path, where N is the total number of links of the first shortest path. Then path i is feasible if a sequence of links within path i cannot be found for which the following conditions hold:

(i) $$s_n = \begin{cases} 0 & \text{if } a_{n1} = a_{mi} \quad \forall n \in N, \forall m \in M \\ 1 & \text{otherwise} \end{cases} \tag{2}$$

(ii) $$S = \frac{\sum_n s_n}{N} \tag{3}$$

(iii) $$S \geq \delta_{min} \tag{4}$$

δ_{min} is a user defined limitation factor on disjoint rate. For illustration purposes a value of $\delta_{min} = 0.35$ is selected. Path i is infeasible if constraint (2) does not hold.

Constraint (3) *Link constraint*: Using the variables defined earlier, path i is feasible if the following condition holds.

$$M \leq \theta_{max} N \tag{5}$$

θ_{max} is a user defined limitation factor on the total number of links. For illustration purposes a value of $\theta_{max} = 1.35$ is selected if M is larger than 6, and $\theta_{max} = 1.5$ is selected otherwise. Path i is infeasible if no such constraint (3) is not satisfied.

The pseudo code for the proposed k-shortest path algorithm is shown Figure 1. To estimate network-wide marginal cost of each O-D pair, constrained k-shortest path algorithm is coded using C programming language and implemented in the proposed GIS-based FMC estimation tool. The tool is developed in ArcGIS, implemented using VBA and C-programming language. In the first step, the planner is prompted to select the year (between 1996 and 2006)[5], the time period (a.m. peak, p.m. peak, or off-peak), and the origin and destination of the trip s/he is interested to calculate the FMC for. Next, using the input information, C-program determines user-defined number of feasible paths using constrained k-shortest path algorithm. Finally, for each of the paths, the marginal costs for vehicle operating, congestion, accident, air pollution and noise costs are calculated. Each of these costs and their weighted average for the selected O-D pair are displayed on a Visual Basic form. By selecting the row of a path in the cost output form, the particular path is highlighted on the map. The GIS-based tool enables planners to efficiently identify areas of interest and to visualize results on the study network by taking advantage of powerful graphical capabilities of ArcGIS combined with the algorithm developed in this study.

k = *Total number of paths to find*, *n* = *number of paths found so far;*
s = *origin*, *d* = *destination;* *T* = *travel time matrix ;* *TempT* = *temporary travel time matrix;*
N = *total number of links of the first path;* *M* = *total number of links of the candidate path*

begin
  $tempT = T,\ n = 0,\ s = S,\ d = D;$
  call *modified Dijsktra* $(s, d, tempT);$
    $n = 1;$
  store $A_1 = \{a_{11}, a_{21}, \ldots a_{N1}\};\ t_1 = \sum_{i,j \in A_1} t_{ij};\ N;$ // *link, travel time, number of link information of the* 1^{st} *path*
  while $(n < k)$ {
    $w = random(1, N),\ \ y = random(1, N);$
    $TempT = TempT \setminus (a_{y1}\ \&\ a_{y2});$ // *remove the links* a_{y1} & a_{y2} *from TempT matrix*
    call *modified Dijsktra* $(s, d, tempT);$
    temp store $A_m = \{a_{1m}, a_{2m}, \ldots a_{Mm}\};\ t_m = \sum_{i,j \in A_m} t_{ij};\ M;$
    if (Constraint - 1 = 1) & (Constraint - 2 = 1) & (Constraint - 3 = 1) then
      $n = n + 1;$
      store $A_m; t_m; M;$
    }
    $TempT = T;$ // *reinsert the links* a_{y1} & a_{y2} *to the TempT matrix*
    for $(h = 1,\ h < n,\ h{+}{+})$ {
      $w_h = random(1, M_h);\ y_h = random(1, M_h);$
      $TempT = TempT \setminus (a_{w_h 1}\ \&\ a_{w_h 2});$
    }
  }
end

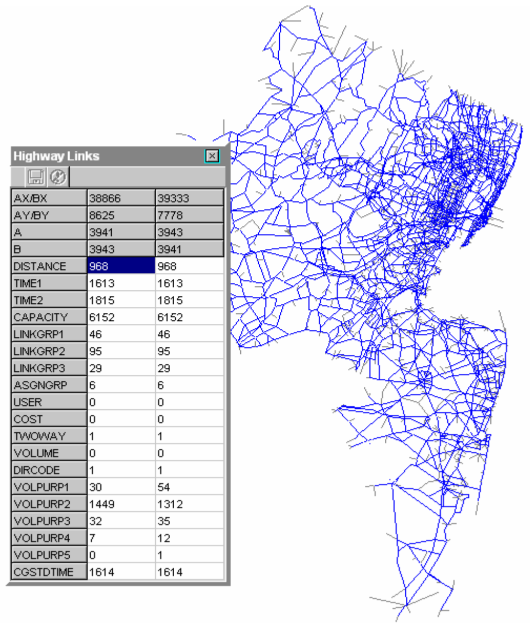
Figure 1 Pseudo Code for Constrained K-shortest Path Algorithm

MODEL APPLICATION

This section presents preliminary results of the proposed GIS-based FMC estimation tool

[5] The planning period is 10 years.

applied to northern NJ highway network. The network, shown in Figure 2, consists of *5,418* nodes, *1,451* of which are zonal nodes[6]- and a total of *15,387* links. The input data required for the estimation process are obtained from transportation planning software, TP+, by the New Jersey Department of Transportation (NJDOT). The input to the developed tool includes; volume, travel time (resulting from the assignment of a specific O-D demand matrix onto the network using TP+), capacity, node and link ID's, highway type, residential area type, distance, number of lanes, free flow speed, and free flow travel time for each link.



AX/BX	38866	39333
AY/BY	8625	7778
A	3941	3943
B	3943	3941
DISTANCE	968	968
TIME1	1613	1613
TIME2	1815	1815
CAPACITY	6152	6152
LINKGRP1	46	46
LINKGRP2	95	95
LINKGRP3	29	29
ASGNGRP	6	6
USER	0	0
COST	0	0
TWOWAY	1	1
VOLUME	0	0
DIRCODE	1	1
VOLPURP1	30	54
VOLPURP2	1449	1312
VOLPURP3	32	35
VOLPURP4	7	12
VOLPURP5	0	1
CGSTDTIME	1614	1614

Figure 2 Northern NJ Highway Network

COST FUNCTIONS AND DATA SPECIFICATION

The cost categories used in this study are (1) vehicle operating costs, (2) congestion costs, (3) accident costs, (4) air pollution costs and (5) noise costs.

Each cost category was estimated using the data obtained from NJDOT and from other sources. It should be noted that data on vehicle operating costs, accident costs, and infrastructure costs are NJ-specific (AAA 2005, USDOT 1991, KBB 2005, NJDOT 2005, FHWA 2005). Whereas, congestion and environmental costs were adopted from relevant studies in the literature but their parameters were modified to fit NJ-specific conditions (Mun 1994, Small and Chu 2003, EPA 1995, Delucchi and Hsu, 1998). For illustration purposes, only the final estimated cost functions are presented here. The detailed derivation of each cost category can be found in a study performed by the authors (Ozbay et al. 2005).

[6] A zonal node here connote to origin-destination zones where trips originate and end.

MODEL RESULTS

This section presents the marginal cost estimation results between New Brunswick, NJ to Princeton Junction, NJ estimated using GIS-based FMC calculation tool. The user first chooses the scenario that is to be analyzed in the scenario selection form (as shown in Figure 3). Upon selection of the year, the time period of analysis, and the origin and destination between which the scenario is to be analyzed. Values of various costs categories are displayed and the selected paths are also highlighted as shown in Figure 3. These costs are the stored in a file for possible use a later stage of the analysis.

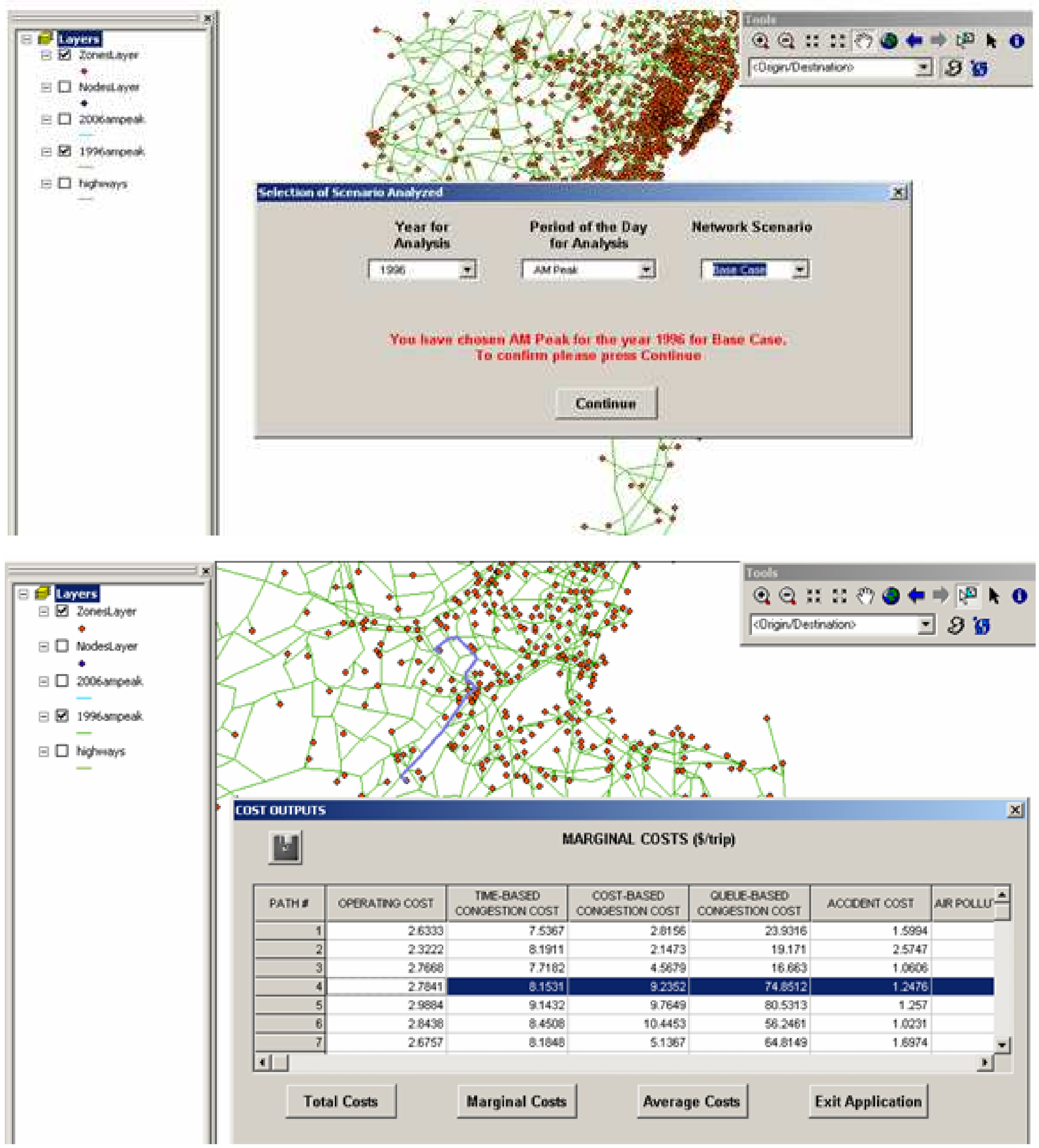


Figure 3 Illustration of the GIS-based FMC tool

Numerical experiments are conducted to investigate the performance of the GIS-based FMC tool and the effect of each cost category on the FMC for various O-D pairs. The main results of this numerical study are as follows:

(1) For the considerably large test network used in this study (nodes and links), GIS-based tool computes user-defined number of paths (5 to 10 paths) in 15 seconds on the average (on a 1.4 GHz Pentium-4 PC with 256 MB RAM), this algorithmic efficiency will allow the users to conduct interactive analysis of various scenarios in real-time.
(2) FMC values are found to differ among different paths connecting a given O-D pair; although some of these paths have similar travel times; this result-emphasizes the importance of considering multiple paths when studying full marginal costs of trips.
(3) Travel time (delay) costs usually constitute the largest portion of the FMC of an O-D pair during peak periods.

CONCLUSIONS

This paper presents the GIS-based efficient and interactive implementation of a novel methodology for the calculation of the network-wide FMC of highway transportation. Unlike previous studies, the methodology presented here estimates trip based FMC, considering all feasible paths between each O-D pair. This approach enables the planners to realistically capture the effect of unit increase in demand. This is essential accurately calculating network-wide marginal costs.

In the first part of the paper, a constrained k-shortest path algorithm is presented. The user-defined constraints include travel time limitation, minimum rate partial disjointness and the limitations imposed on the total number of links. The algorithm is coded in C programming language and implemented in the proposed GIS-based FMC estimation tool. GIS-based FMC estimation tool is developed based on GUI developed in ArcGIS using VBA. While dealing with complex transportation networks, it is important to have a user-friendly tool that will enable planners to efficiently identify the areas of interest and to visualize results on the study network. The GIS-based tool developed in this study achieves both these objectives by taking advantage of visualization capabilities of ArcGIS in tandem with the unique, efficient and versatile algorithm developed by the research team.

In the second part, the proposed GIS-based FMC estimation tool is applied to northern NJ highway network. The input data required for the calculation of FMC are obtained from TP+ output. Vehicle operating, congestion, accident, air pollution, noise and infrastructure costs are estimated based on the data obtained from the NJDOT and other sources.

Experimental results obtained from the Northern NJ network demonstrate that FMC between an O-D pair show differences among various paths that connect this O-D pair. This simple result depicts the importance of analyzing trips based on a number of factors in addition to travel times. This GIS-based tool will help planners to estimate the changes in transportation costs due to a particular transportation demand management measure or supply change such as adding new lanes or improving existing lanes. This is a critical component of transportation planning, because demand patterns experience both spatial and temporal variations due to the changes in demand and supply.

ACKNOWLEDGEMENTS

We would like to thank Professor Joseph Berechman for providing invaluable comments on an earlier version of a research report that contained some of the information presented in this paper. We would also like to thank Mr. Robert Diogo of New Jersey Transportation Planning Authority for proving the test network and some of the data used in the paper. The opinions and conclusions presented are the sole responsibility of the authors and do not reflect the views of sponsors and other participating agencies.

REFERENCES

Akgun, A., Erkut, E., and Batta, R., (2000). "On Finding Dissimilar Paths" *European Journal of Operation Research* 121, 232– 246.

American Automobile Association, (2005). "Your Driving Costs 2005".

Anderson D., and McCullough G., (2000). "The Full Cost of Transportation in the Twin Cities Region" *Center for Transportation Studies*, University of Minnesota, CTS 00-04.

Banfi S., Boll C., Maibach M., Rothengatter W., Schenkel P., Sieber N., and Zuber J., (2000). "External Costs of Transport: Accident, Environmental and Congestion Costs in Western Europe" *INFRAS*.

Brander, A.W., and Sinclair, M.C., (1995). "A Comparative Study of K-shortest Path Algorithms" *11th UK Performance Engineering Workshop for Computer and Telecommunications Systems*.

Berechman, J., and Genevieve G., (1984). "Analysis of the Cost Structure of an Urban Bus Transit Property" *Transportation Research* 18B, 273-287.

Delucchi, M., and Hsu S. (1998). "The External Damage Cost of Noise Emitted from Motor Vehicles" *Journal of Transportation and Statistics* 1 (3), 1-24.

Dial, R.B., (2000). "Bicriterion Traffic Equilibrium: Mode, Algorithm, Software Overview" *Transportation Research Record* 1725, 54–62.

Environmental Protection Agency Publications, (1995). "National Emission Trends".

Eppstein, D., (1998). "Finding the K Shortest Paths" *SIAM Journal Comput*. 28, 652–673.

Federal Highway Administration (FHWA), (2005). "Motor Vehicle Accident Costs", (available at http://www.fhwa.dot.gov/legsregs/directives/techadvs/t75702.htm)

Hadjiconstantinou, and E., Christofides, N., (1999). "An Efficient Implementation of an Algorithm for Finding K-shortest Simple Paths" *Networks* 34, 88–101.

Jara-Diaz, S.R., Donoso, P.P., and Araneda, J.A., (1992). "Estimation of Marginal Transport Costs: The Flow Aggregation Function Approach" *Journal of Transport Economics and Policy*, 35-48.

Levinson, D., Gillen, D., Kanafani A., and Mathieu, J. M. (1996). "The Full Cost of Intercity Transportation - A Comparison of High Speed Rail, Air, and Highway Transportation in California" *UCB-IS-RR-96-3*.

Mayeres I., Ochelen S., and Proost S., (1996). "The Marginal External Cost of Urban Transport" *Transportation Research Part* 2D, 111-130.

Mun, Se-IL, (1994) "Traffic Jams and Congestion Tolls" *Transportation Research* 28B, 365-375.

Ozbay, K., Bartin B., and Brechman J., (2001). "Cost of Transporting People in New Jersey-Phase1" *New Jersey Department of Transportation Report*.

Ozbay, K., Yanmaz-Tuzel O., Mudigonda S., Bartin B., and Berechman J., (2005) "Cost of Transporting People in New Jersey-Phase II" *New Jersey Department of Transportation Draft Report*.

Safirova, E., and Gillingham, K., (2004). "Measuring Marginal Congestion Costs of Urban Transportation: Do Networks Matter?" *83rd Transportation Research Board Annual Meeting*, Washington, DC, USA.

Scott, K., Pabo´n-Jime´nez, G., and Bernstein, D., (1997). "Finding Alternatives to the Best Path" *76th Transportation Research Board Annual Meeting*, Washington, DC, USA.

Sherali H.D., Hobeika A.G., and Kangwalklai S., (2003). "Time-Dependent, Label-Constrained Shortest Path Problems with Applications" *Transportation Science* 37, 278-293.

Shier, D.R., (1979). "On Algorithms for Finding the K-shortest Paths in a Network". *Networks* 9, 195–214.

Small K., and Chu X., (2003). "Hypercongestion" *Journal of Transport Economics and Policy* (37), 319-352.

US Department of Transportation, (1991). "Cost of Owning and Operating Automobiles, Vans and Light Trucks".

US Department of Transportation, (2005). "Crash Records", (available at http://www.state.nj.us/transportation/refdata/accident/rawdata01-03.shtm)

Van der Zijpp NJ., and Catalano S.F., (2005). "Path Enumeration by Finding the Constrained K-shortest Paths" *Transportation Research* 39B, 545-563.

Zhang, L., and Levinson, D. (2005). "Road Pricing with Autonomous Links". *84th Transportation Research Board Annual Meeting*, Washington, DC, USA.