

Users' Perceptive Evaluation of Bus Arrival Time Deviations in Stochastic Networks

*Nikolaos G. Daskalakis, Anthony Stathopoulos
National Technical University of Athens*

Abstract

This work is a report on research concerning transit service characteristics as seen from the users' point of view. Users of two separate bus lines, operating in a shared/common urban infrastructure, were interviewed at bus stops about their perception concerning headways of bus lines operation. An analysis was made regarding deviations between actual and scheduled bus arrival headway. Further statistical analysis was carried out to check factors giving rise to different perceptions. The operation of each bus line was registered, and corresponding service characteristics were compared with those perceived by the users. Based on these results, a model for bus line headways was proposed, incorporating the perception of deviations by the users. In conclusion, a reliable service, meaning smaller deviations, is more appreciated by the public than any service of shorter headways and less reliability.

Introduction

In recent studies regarding travel time and reliability, it has been found that travelers are not only interested in saving travel time but also in reducing travel time variability. Their attitudes and "choice of way and route" strongly depend on this perception. Variability causes uncertainty, as they do not know arrival time at the destination. Thus, variability is considered by travelers as an additional

cost. Research is rather limited on demand-and-supply variation effects on travel time reliability in an uncertain environment. For example, a recent report was examined how individual travelers with different risk-taking attitudes responded to such changes (Chen et al. 2001). Route choice models that were originally proposed were based on the assumption that all travelers are aware of the travel times of a certain network (deterministic models [Beckmann et al. 1956] as well as stochastic models [Daganzo and Sheffi 1977; Fisk 1980; Sheffi and Powell 1982]). However, both categories of models tend to disregard network uncertainty (stochasticity) and assume that the network is deterministic, an assumption that is not true, especially during rush hours.

This paper, based on a recent university study (Daskalakis 2002), focuses on how passengers perceive reliability of bus line operation, and models the relationship between bus line operation characteristics and this perceived measure of reliability.

Travel Time Perception and Evaluation by Passengers

Travel time is a natural measure of the effectiveness of a bus service. The purpose of bus service is to transport people to their destination with safety and convenience, offering easy access and providing service information. However, most people rate travel speed and, consequently, travel time above all quality characteristics (Chen et al. 2002). This time often varies considerably, primarily during rush hours in everyday commuting.

Waiting time deviation is an indicator of how passengers experience the operation of a bus line, while waiting at the same stop, around the same period of the day, when headway schedules are the same. Studies concerning time cost have shown that passengers would rather wait than pay for a more frequent service, though not for long (Hess et al. 2003). Bus operators aim at offering services that best suit passenger needs. Does this mean they have to provide more frequent bus schedules, which is something evidently expensive, or is there any other way of keeping passengers sufficiently satisfied while waiting for the bus? The answer should be regular bus transportation, leading to an increased quality of service.

Collection and Analysis of Data

To acquire data that would assist in this evaluation of user perception of bus schedule variability, a survey was conducted downtown Athens, Greece. Frequent-

cies of bus lines were registered on site and were compared to the original scheduled frequencies by OASA (Athens Urban Transportation Organization).

Two lines were selected: line A and line B (originally code named "A14" and "730"), linking central Athens to western suburban districts. The survey was carried out from May to June (on usual weekdays, primarily Tuesdays and Thursdays) from 07:00 to 17:00, covering both morning and afternoon commuting. More specifically, the time period 07:00- 09:30 was chosen to cover traveling to work from the suburbs to central Athens. This period was called the inbound direction "I." The period 13:45-16:45 was chosen primarily for passengers returning to their homes, following the outbound ("O") direction. It is noted that, during the days and the times of the interviews, weather conditions were normal, no major events affected the usual operation of bus lines, and the interviewed passengers were chosen randomly. These two bus lines are operated by ETHEL (an OASA partner, responsible for operating thermal buses). Both are radial-shaped, linking the commercial center of Athens with suburban districts and run along signalized arterials (see Tables 1 and 2). By the time of the research, schedule information was not posted at bus stops. Passengers had to find out the scheduled bus line frequency usually by asking bus drivers. For survey needs, such data were derived from original scheduling timetables of OASA.

Table 1. Characteristics of the Surveyed Bus Lines

	Line A		Line B	
	<i>Inbound (I)</i>	<i>Outbound (O)</i>	<i>Inbound (I)</i>	<i>Outbound (O)</i>
Route length (m)	7631	7270	7555	7097
No. of stops	27	26	25	26
Average length between stops (m)	293	280	302	273
Scheduled round trip time (min)	85		80	

Source: Athens Urban Transportation Organization

Table 2. Scheduled and Observed (Mean) Bus Headways

Line	A		B	
	<i>I</i>	<i>O</i>	<i>I</i>	<i>O</i>
<i>Direction</i>				
Scheduled time (min)	7.9	6.9	17.0	16.0
Mean observed time (min)	8.2	7.7	21.0	16.6

Basic questions asked were the following:

- 1) How often do you use this particular bus line (weekdays)? (daily—4+ days a week, 2-3 times a week, 1-2 times a month, less than once a month)
- 2) In your opinion, what is the usual delay? (no delay, considerable, much, too much) (i.e., magnitude of the delay)
- 3) In your usual schedule, how long would you be willing to wait for the bus? (0, 5, 15, 20, 20+ min)
- 4) From your experience, how long (in minutes) is the usual bus latency? (0, 5, 15, 20, 20+ min)
- 5) How long do you usually wait at the bus stop before concluding that the bus is late? (0, 5, 15, 20, 20+ min)
- 6) How long (in minutes) would you be prepared to wait for the bus? (0, 5, 15, 20, 20+ min)
- 7) You arrive at the bus stop. When would you decide that your schedule has been seriously affected? (0, 5, 15, 20, 20+ min)
- 8) What is the purpose of the particular trip? (work, returning home, education, shopping, recreation, other)
- 9) After how long (in minutes) would you consider the delay unjustified?
- 10) Suppose that the exact arrival time is indicated at the bus stop. How would you describe a bus delay of
 - a) 5 minutes? (short, average, long, unacceptable)
 - b) 10 minutes? (short, average, long, unacceptable)

A total of 300 valid questionnaires were collected. The resulting data were subjected to a series of statistical tests and analysis.

T-tests were conducted for each travelling direction with regard to different expressions of waiting time perception, as addressed in questions 3, 4, 5, 6, 7, 9 and 10. Table 3 shows the statistical results for the means, standard deviations and the significance levels for the null hypothesis H_0 of equal means.

No assumption of equal means, except that for question 5, was found to be statistically significant with a confidence coefficient in excess of $(1-\alpha)_5 = 95\%$, since all t-statistics were over 0.05. Homogeneity of variance test was performed by calculating Levene's statistic to verify the assumption of homogeneity of variance

Table 3. Statistical Results (t-test)

<i>Question</i>	<i>Direction</i>	<i>Mean Waiting Time (min)</i>	<i>Standard Deviation</i>	<i>Significance Level, α</i>
#3	Inbound	13.8	4.5	0.12
	Outbound	12.8	4.8	
#4	Inbound	12.7	5.2	0.19
	Outbound	11.8	5.5	
#5	Inbound	13.0	4.9	0.01
	Outbound	11.4	5.3	
#6	Inbound	6.6	3.7	0.52
	Outbound	6.9	4.5	
#7	Inbound	14.3	5.9	0.97
	Outbound	14.4	4.8	
#9	Inbound	16.9	5.6	0.12
	Outbound	15.9	5.3	
#10(a) Delay of 5 min	Inbound	1.2	0.63	0.05
	Outbound	1.4	0.78	
#10(b) Delay of 15 min	Inbound	2.7	0.87	0.13
	Outbound	2.8	0.85	

that would certify the performance of further tests, such as ANOVA and Discriminate Analysis. Levene's test is an alternative to Bartlett test (Bartlett 1937), testing also observations originating from populations showing the same variance that depend heavily on the assumption that these observations refer to normal distributions. Since in our case is that no such evidence exists, Levene's test was considered as preferable.

Regarding questions 6 and 10, the test confirmed heteroscedasticity with confidence coefficient $(1-\alpha)_{6,10} = 99.8\%$. That is, the hypothesis that waiting time (as specified in questions 6 and 10) is of equal levels of variance for both directions "I" and "O" is not accepted. Consequently, the answers to questions 6 and 10 do not explain the same proportion of the variance by direction.

Results of the t-tests for the rest of the questions 3, 4, 5, 6, 7 and 9 did not indicate significant differences concerning the means of the given answers for a confidence coefficient $(1-\alpha)_{3,4,6,7,9} > 90\%$, and no definite conclusion may be drawn about the effect on the answers of any of the two directions. In question 10, for a waiting time of 5 minutes, the mean value of the answers given by the passengers of direction "O" is greater than that of direction "I" with a confidence coefficient of $(1-\alpha)_{10} = 95\%$. It should be noted that the ordinal scale of question 10 was transformed to a numerical, using the following convention: small = 1, medium = 2, large = 3, unacceptable = 4.

Similar tests were performed for each line separately, combining the two directions. Relevant results are shown in Table 4. All the differences of the means for the two bus lines, except that of question 7, are statistically significant with a confidence coefficient $(1-\alpha)_{3,4,5,6,9,10(5)} > 95\%$. In questions 5, 6 and 10, Levene's test gives $\alpha < 0.05$, which reveals heteroscedasticity. That is, the null hypothesis that the variable (waiting time as specified in questions 5, 6 and 10) has equal levels for both bus lines A and B does not hold and the variance cannot be explained at the same degree.

There is also a direct correspondence between the answers given by passengers to questions 3, 4 and 9 and the type of bus line. In particular, passengers of bus line A stated at the above questions significantly (with a confidence coefficient $(1-\alpha)_{3,4,9} > 99\%$) shorter mean time.

The mean (waiting time) based on the samples answering questions 5 and 6 for bus line A is shorter than waiting time of bus line B. However, existence of heteroscedasticity in the sample does not allow concluding that the type of bus line is a significant factor. In question 10 (a), mean waiting time in the case of bus line B is longer than that of line A, indicating less tolerance by the users as long as waiting is concerned. In question 7, bus line type does not affect the answers.

In contrast to passengers of line B, passengers of line A spend everyday shorter waiting time at the bus stop. It appears from the answers, that passengers perceive scheduled headway of each bus line rather accurately, even if they do not have direct information about it. They evaluate the degree of schedule adherence and adapt their own trip schedule to the mean headway for each line. The time they are prepared to wait at the bus stop is not related to the mean headway of the bus line. Deviations between actual bus lines operation and scheduled headways create problems and affect their activities. In cases where these activities require

Table 4. Combined (Inbound-Outbound) Statistical Results (t-test)

<i>Question</i>	<i>Bus Line</i>	<i>Mean Waiting Time (min)</i>	<i>Standard Deviation</i>	<i>Significance Level, α</i>
#3	A	11.3	4.1	<0.001
	B	15.7	4.2	
#4	A	10.7	5.0	<0.001
	B	14.0	5.2	
#5	A	10.5	4.5	<0.001
	B	14.0	5.2	
#6	A	5.2	3.4	<0.001
	B	8.1	4.2	
#7	A	14.7	5.0	0.32
	B	13.9	5.9	
#9	A	15.4	5.3	0.003
	B	17.3	5.4	
<i>Question 10</i>				
(a) Delay of 5 min				
	A	1.2	0.54	0.03
	B	1.4	0.82	
(b) Delay of 15 min				
	A	2.7	0.76	0.84
	B	2.7	0.94	

a precise schedule, passengers begin to consider alternative bus lines, taxi service, a combination both, or even walking.

Usually, after carrying out a test of statistical importance, it is desirable to know which factor contributed to the results. In our case, we pay special attention to the differences between the answers. Analysis can, of course, be limited to the simple t-tests, in order to compare all possible pairs of the sample means. However, such a procedure would depend on chance.

Post hoc comparison techniques, on the other hand, take into account specifically the fact that more than two sample means may be examined. Post hoc stands for the logical error of believing that temporal succession implies a relation. These post hoc comparisons were made using Scheffé's and Duncan's tests. Scheffé's test performs simultaneous joint comparisons in pairs for all possible combinations of means in pairs using the F sampling distribution. This test is considered to be "conservative" (Clarke and Cooke 1998); therefore, its usage helps to find out significant (at a level $\alpha = 0.05$) errors occurring in multiple comparisons.

At the same time, there is a chance that important differences, possibly existing, may not occur. To limit this possibility, a more tolerant test (Duncan) is performed. Duncan's test makes comparisons in pairs using a stepwise order of comparison, setting a protection level for the rate of error regarding the collection of data sets, rather than rate of error for individual tests.

Tests mentioned above made it clear that a significant factor differentiating the answers is the headway of each bus line. To find a quantitative expression (function) of that differentiation, Discriminant Analysis was used. This type of analysis describes the differentiating features from observing known populations and tries to find "discriminants" of which numerical values are such that the observations (responses) are as distinct as possible (Fisher 1936).

Responses to questions 3, 4, 5, 6 and 9 were treated as quantitative variables. Analysis indicates that 57-74 percent of the answers at the basis of bus lines A and B with confidence coefficient of $(1-\alpha)_{3,4,5,6,9} > 95\%$. Wilk's Lambda found to be ranging between 0.78 and 0.88, depending on the type of question. Wilk's Lambda ($\in [0,1]$) is a multivariate test of significance, sometimes called the U-statistic, with values close to 0 indicating that the group means are different and values close to 1 indicating the group means are not different.

The most felicitous analysis was found to be the one referring to question 3 ("In your usual schedule, how long do you estimate you will be willing to wait for the bus?"), in which the discriminant percentage of the questions was 74 percent, with Wilk's Lambda 0.78 with a confidence coefficient of $(1-\alpha)_3 > 99\%$. The linear discriminant function for each of the two bus lines was:

$$\text{Bus line A: } y_{3,A} = -4.415 + 0.661x \tag{1}$$

$$\text{Bus line B: } y_{3,B} = -7.878 + 0.918x \tag{2}$$

where: y is the classification variable of bus line

x is answer to the question no.3, in minutes

For a specific x , the larger of the two classification variables $y_{3,A}$, $y_{3,B}$ classifies the user to the one or the other bus line. The discriminant line is made up by the parts of the two functions that give higher scores before and after the point of intersection (dashed line, Figure 1). The two lines intersect at a point with an abscissa equal to 13.45 min. Thus, a user whose response is less than 13.45, is more likely to use

line A. Users of line A, compared to those of line B, spend less of their time at the bus stop. This indicates that the users of the more frequent line A perceive their waiting time through the assumption that line A has a higher headway. Passengers of line B, on the contrary have a better perception of the actual headway of the particular line.

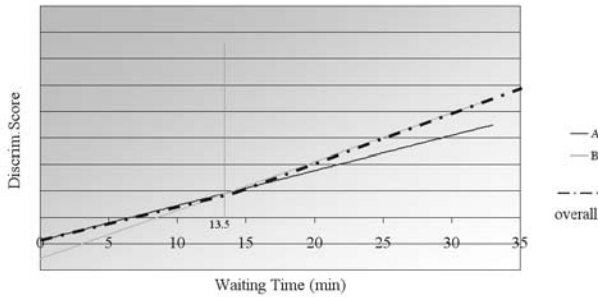


Figure 1. Bus Line A & B Discriminant Functions

Results of previous test-controls are summarized as follows:

- Passengers of both lines in the outbound direction, that is, those mainly returning home in the afternoon, answered that they are less tolerant compared to answers given to the same question while making the morning inbound trip. For passengers returning home, the reliability of service is evaluated (perceived) as more important than is in the inbound trip.
- Headway analysis showed a large degree of schedule deviation (up to 90%–95%). This implies about ± 7.5 minutes for line A and ± 16.5 minutes for B (extreme negative signs indicate a bunching). Passengers perceived average times of: 10.7 and 14 minutes, respectively. Passengers of the most frequent line (A) perceive greater delays than actual ones, while passengers of the less frequent line (B) perceive smaller schedule deviations than actual ones. This phenomenon is known as “time drag,” in which waiting time seems longer than it actually is (Moreau 1992). A possible explanation for this would be that the perception of time from an unreliable bus service follows a logarithmic trend (i.e. during the first waiting minutes, time “runs” faster).

Model Proposal and Development

To investigate further the claim of the logarithmic-like relationship as suggested in the previous chapter, the following simple calculations were undertaken. The basic relationship is expressed as follows:

$$T_i = k_i H_i^L \tag{3}$$

where: T_i : the users' perception of deviation as stated in the interview for bus line i.

H_i : headway of bus line i

k_i : coefficient of proportion, independent of the bus line's headway, related to bus line i user's characteristic, the purpose of traveling, the frequency of bus usage, travel time, etc.

L : unknown numerical variable

For bus line A, eq(3) becomes: $T_A = k_A H_A^L$ and for bus line B, eq(3) becomes: $T_B = k_B H_B^L$

In our case, users of bus lines A and B have similar characteristics (purpose of travelling, frequency of bus usage, etc) and this means $k_A = k_B$

$$\begin{aligned} k_A = k_B &\Rightarrow \frac{T_A}{H_A^L} = \frac{T_B}{H_B^L} \Rightarrow \frac{T_A}{T_B} = \frac{H_A^L}{H_B^L} \Rightarrow \left(\frac{T_A}{T_B}\right) = \left(\frac{H_A}{H_B}\right)^L \Rightarrow \\ &\Rightarrow \ln\left(\frac{T_A}{T_B}\right) = \ln\left(\frac{H_A}{H_B}\right)^L \Rightarrow \ln\left(\frac{T_A}{T_B}\right) = L * \ln\left(\frac{H_A}{H_B}\right) \Rightarrow \\ &\Rightarrow L = \frac{\ln\left(\frac{T_A}{T_B}\right)}{\ln\left(\frac{H_A}{H_B}\right)} \end{aligned} \tag{4}$$

Observed $H_A=7.68\text{min}$ and $H_B=17.65\text{min}$. T_A and T_B can be derived from the answers collected in the survey and refer to the perception of schedule deviation of bus lines A and B, respectively.

By substituting in (4) H_A, H_B and the values given in question no.3 (Table 4), i.e. $T_A=11.27$ and $T_B=15.66$:

$$L = \frac{\ln\left(\frac{11.27}{15.66}\right)}{\ln\left(\frac{7.68}{17.65}\right)} = \frac{\ln(0.719)}{\ln(0.435)} = \frac{-0.33}{-0.83} = 0.4 \Rightarrow L = \frac{2}{5} \quad (5)$$

And finally:

$$T = kH^{2/5} \quad T, H \text{ in min} \quad (6)$$

The above model was iteratively fitted to the survey responses, i.e. the values of the responses to those questions that combine on the same basis, passenger perception with actual bus line performance. These are questions no. 3,4 and 5 (Table 4). Coefficient of proportion k for passengers of bus lines A and B is then derived as the minimum square root error solution to (6), using the survey data for each pair of $T^{\text{quest no.}(3),(4),(5)}_{(A),(B)}$ and $H_{(A),(B)}$.

The resultant form of the model is then:

$$T(H) = 4.7H^{2/5} \quad T, H \text{ in min} \quad (7)$$

The rate of T vs. H, $r(h)$ derives from:

$$\begin{aligned} T(H) = 4.7H^{2/5} &\Rightarrow \frac{\partial T(H)}{\partial H} = 4.7 \frac{2}{5} H^{\frac{2}{5}-1} \Rightarrow \\ &\Rightarrow r(H) = 1.88H^{(-\frac{3}{5})} \end{aligned} \quad (8)$$

H in min, r in $\text{min}^{(-3/5)}$

where: $r(H)$ is a decaying function of H, signifying the diminishing impact of a headway increase on the perception of schedule deviation.

The plotted results of the (7) and (8) are shown in Figure 2.

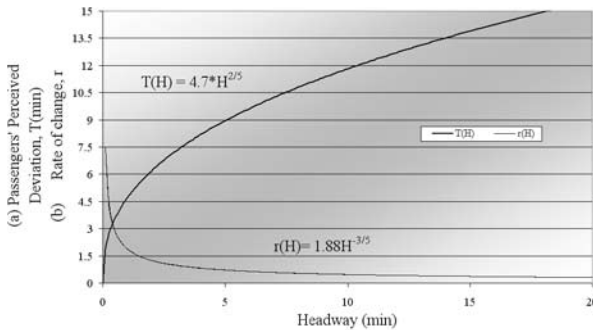


Figure 2. Perceived Deviation and its Rate of Change vs. Headway

(Factors 4.7, 1.88 concern users of specific bus lines A & B as described above)

Conclusions

Bus line users traveling in stochastic transportation networks, having little knowledge of the exact schedule timetables, usually base their travel decisions on the empirical perception of time in order to organize their own time schedules. The perception of the mean waiting time and its variances determines how service reliability is evaluated by the user and, subsequently, the user’s attitude towards the way of traveling. Knowing the way passengers perceive schedule deviations and the resultant variations of their waiting time would help the management of transportation in achieving operational effectiveness. Having interviewed passengers waiting at bus stops, it was verified that a (the) significant factor related to the perception of waiting time deviations was the headway, yet not linearly.

On the basis of the proposed model, deviation is perceived as a function of the headway H raised to the number of $2/5$. The curve expressing the relation between perceived deviation (leading to the so-called “time drag”) and headway has a logarithmic shape, while the curve expressing the rate of the perception of deviation has that of a negative exponential. The greater the headway, the greater deviation the users perceive, but at a diminishing rate. So, if an operator wishes to upgrade the quality of those services related to passenger waiting time, it is important to keep bus lines even with greater headways reliable and then try to achieve shorter headways. Once a bus line with shorter headway is in operation, it should be strictly reliable, as passengers become indignant about unreliability of bus lines with shorter headways.

Limitations and Future Work

Natural limitations in this research concern the basic sample of only two specific bus lines, the fact that bus operators may have incorporated some manual headway control into schedules—a thing unknown to us, and that passengers had no credible source available of information about bus line scheduled operations. By implementing intelligent transport systems such as real-time information at bus stops and automatic headway control methods, new and challenging conditions appear in transportation environment.

Suggestions for further investigation on the subject could involve research on different types of bus lines and networks, such as peripheral instead of radial bus lines or bus lines using exclusive lanes. The question of how reliable (in quantitative terms) a service should be before it is made more frequent, regarding cost and benefits of alternative operational strategies, is also another interesting direction of research.

Acknowledgment

The authors wish to express their appreciation and thanks to Anthony Stamatogiannakis, MSc in Marketing, for his help with the analysis of the surveyed data and Dr. Alexandros Fragkiadakis for his useful remarks.

References

- Bartlett, M.S. 1937. Properties of sufficiency and statistical tests. *Proceedings of the Royal Statistical Society Series A* 160, 268–282.
- Beckmann, M. J., McGuire, C.B. and Winstein, C.B. 1956. *Studies in the Economics of Transportation*. New Haven, Connecticut:Yale University Press.
- Chen, C., Skabardonis, A., Varaiya, P., 2003. Travel Time Reliability as a Measure of Service. *TRB 82nd Annual Meeting CD-ROM*, Transportation Research Board, Washington, D.C.
- Chen, A., Zhaowang J., and Recker, W. 2001. Travel time reliability with risk-sensitive travelers. Institute of Transportation Studies, University of California, Irvine.
- Clarke, G.M., Cooke, D. 1998. *A Basic Course in Statistics*. 4th edition, (Chapter 22) pp520–546, Arnold, London, UK.

- Daganzo, C., and Sheffi, Y. 1977. On stochastic models of traffic assignment. *Transportation Research* 14B, 243-255.
- Daskalakis N. 2002. Investigation of system operation reliability – Application to bus lines in Athens. *Diploma Thesis*, Department of Transportation Planning and Engineering, National Technical University of Athens, Greece.
- Fisher, R.A. 1936. The use of multiple measurements in taxonomic problems. *Annals of Eugenics* 7, 179-188.
- Fisk, C.S. 1980. Some developments in equilibrium traffic assignment. *Transportation Research* 14B, 243-255.
- Hess D.B., Brown, J., and Shoup, D. 2003. Waiting for the bus. *TRB Annual Meeting CD-ROM*, Transportation Research Board, Washington, D.C.
- Moreau, A. 1992. Public transport waiting times as experienced by customers: Marketing research involving the Grenoble System. *Public Transport International* vol. 41, no. 3, 52-68.
- Sheffi, Y., Powell, W.B. 1982. An algorithm for the equilibrium assignment problem with random link times. *Networks Journal*, vol.12, issue 2, pp.191 – 207. Wiley Periodicals, Inc, USA.

About the Authors

NIKOLAOS DASKALAKIS (nikolasmail@gmail.com) obtained a degree in Civil Engineering (Dipl-Ing) and received a Master's degree in GIS, both from National Technical University of Athens. He has participated as scientific collaborator in various projects, mainly concerning urban transportation research, analysis and organization. He is a member of Technical Chamber of Greece & Hellenic Association of Civil Engineers.

ANTONY STATHOPOULOS (a.stath@transport.ntua.gr) is Professor at the Department of Transportation Planning and Engineering, School of Civil Engineering, NTUA. His research and teaching activities include transportation systems planning and engineering, traffic network analysis and Intelligent Transportation Systems. He has an extensive involvement in ITS research and an active role in demonstration programs.