NISS

Analysis of Probe-Based Information on Signalized Arterials

Piyushimita Thakuriah, Ashish Sen,
and Alan F. Karr

Technical Report Number 48
June, 1996

National Institute of Statistical Sciences
19 T. W. Alexander Drive
PO Box 14006
Research Triangle Park, NC 27709-4006
www.niss.org

Analysis of Probe-Based Information on Signalized Arterials *

Piyushimita (Vonu) Thakuriah
Research Fellow, National Institute of Statistical Sciences
Research Triangle Park, NC 27709

Ashish Sen
Professor, School of Urban Planning and Policy
University of Illinois at Chicago, Chicago, IL 60608
Senior Fellow, National Institute of Statistical Sciences

Alan Karr
National Institute of Statistical Sciences
Research Triangle Park, NC 27709

June 14, 1996

Abstract

The usefulness of the service provided by an Advanced Traveler Information System and Advanced Traffic Management System would depend on the quality of data that the system generates and the ability of system designers to tap into the properties of the data. Probe vehicles are a major source of data in ATIS and ATMS. We investigate the quality of data generated by probe vehicles on signalized arterial streets in terms of measurement errors and analyze the properties of travel time data. A dominant property is the presence of dependence in the data. We then explore the nature of travel time estimates. Given that travel time processes have certain properties and the realizations of travel times are affected by a large number of covariates, the estimates may be of high variance. Also unless one states the travel time estimation model carefully, some estimates may be biased. The dependence property of the data further makes statistical inference difficult. We point out invalid conclusions that one may make under an independence assumption.

1 Introduction

The direct observation of transportation network conditions via specially-equipped 'probe' vehicles has opened up an interesting area of traffic research and measurement. Probe vehicles form the heart of several deployed and proposed mobile detection and monitoring systems within Advanced Traveler Information Systems [ATIS] and Advanced Management Information Systems [ATMS]. While floating vehicles as a technique for measuring speeds and travel times has been in use for a long time, most measurements were made manually by observers in the vehicles. Probe vehicles use modern communications technology to

*The authors would like to thank the ADVANCE project. This work was partially supported by the National Science Foundation's Division of Mathematical Sciences Grant Numbers 9313013 and 9208758. We would also like to thank Siim Sööt, Xioquon Zhu and Todd Graves for their help and suggestions.

transmit real-time information either to a centralized traffic information center or to distributed road-side data-gathering points. Typically, these measurements are on travel times and average speeds.

Thus far, because of the absence of area-wide travel time data from network surveillance sources, the quality of the data issues, in terms of measurement errors and potential effects on information quality, have remained virtually unexplored. Also, the statistical properties of the estimates of travel time obtained from probes and other network performance detection technologies have not been studied in any detail.

In this paper, we have three major objectives:

1. To investigate the quality of travel time measurements made by probe vehicles and to illustrate the kinds of measurement errors that may be evident in the data.
2. To analyze the properties of travel time data.
3. To explore the properties of travel time estimates obtained on the basis of travel time data.

The unifying theme of the paper is that travel time observations have certain unique properties and in order to develop useful estimates of travel times, one needs to take these properties into account. While we focus our attention on measurements made by probe vehicles, our comments would be true about other surveillance technologies that generate direct measurements on travel times (as opposed to technologies and methods that allow the inference of 'synthetic' travel times either from data from other traffic stream variables, as one could do with data from loop detectors or via traffic assignment procedures).

The ability to predict link travel times accurately is a core function of dynamic route guidance. A beginning point of travel time prediction is the accurate estimation of expected travel times. We are ultimately interested in a conditional expectation, since the estimate of an unconditional expectation would have a large variance, which would not be very useful for predictive, and ultimately, the route guidance purpose.

Throughout the paper, we present empirical evidence by using data obtained via probe vehicles as a part of ADVANCE , an ATIS project in suburban Chicago (Boyce, *et al.*, 1994). The authors owe much to the targeted deployment of ADVANCE, which was conducted to test the components of ADVANCE in the real world and which yielded a dataset with measurements of travel time and other variables along certain pre-determined routes. The availability of rich real-world data has allowed us to understand better the nature of the underlying travel time processes. This knowledge could improve our efforts to devise travel time estimation that take into account the properties of the travel times. The empirical evidence presented in this paper uses part of the entire dataset, (described in Section 2) — that part being from signalized arterials. We focus on conditions of recurrent (or incident-free) congestion.

Thus far, the major data quality issue in probe-based ATIS and ATMS has been a data 'quantity' issue — that under low deployment rates, data from the system would be sparse with detrimental effects on the quality of route guidance (Thakuriah and Sen, 1995, Koutsopolous and Xin, 1993). However, the data itself may have measurement errors and because probe-based ATIS/ATMS is a relatively new concept, one needs to understand the nature of measurement errors, if any, in the data. This is because the errors could have unknown, yet devastating effects on the estimates developed from the data. Measurement errors are present in to some extent most data; one needs to know the sources of these errors and their structure. In this paper, we devote some attention to this important issue.

Probe-generated data have various properties; since our analysis is confined to signalized arterials, we will confine our discussion to those properties that are directly relevant to the case of signalized arterials. Periodicity and dependence are two major problems and if not handled carefully, can lead to wrong statistical inference. Dependence can occur between travel time observations on the same link or between travel time observations on different links along a vehicle's route.

One ultimate use of probe data is to obtain estimates of expected link travel times. Again, careful model building should enable us to obtain good estimates of expected travel time; otherwise, the estimates may be of high variance, and in some cases, biased. Moreover, because of dependence in the data, the computation of several key statistics is not straightforward.

The paper is structured as follows: in Section 2, we describe the data used for the empirical evidence presented. We point out some quality of data issues associated with probes in Section 3. We present a detailed analysis of the properties of travel time data and estimates of travel time in Section 4. This section has various subsections, and we will make detailed introductions as we go along. Finally, we draw our conclusions in Section 5.

2 Data Collection Experiment

As already mentioned, the data used for the entire analysis were collected as part of the evaluation of ADVANCE, during the summer of 1995. Travel times on signalized arterials depend on several factors or covariates. These factors include volume and network control factors, physical network attributes, vehicle route (turning movement executed to enter the link) and microscopic vehicle properties or driving style idiosyncrasies. Naturally, the more covariate information we have, the better would be the precision of the estimate. The data collection effort was preceded by careful planning, an objective of which was to design the study such that we could collect as much data on these important covariates as possible. Moreover, since we are interested in estimating a conditional expectation of link (and ultimately route) travel time, the data collection was designed to enable analyses that would allow us to vary the travel time estimation conditioning information.

The field tests were conducted in the northwest suburbs of Chicago. ADVANCE links are one-directional and turning-movement specific, with a segment that may be common to more than one link. The link travel time data were transmitted in real-time from ADVANCE probe vehicles that were driven by paid drivers down pre-specified routes in three 'networks' that were identified for the purpose of evaluating different algorithms in the Traffic Related Functions of ADVANCE. The three networks are presented in Figure 1. The first of these networks consisted of 12 ADVANCE links and is labeled Network 1. The second, Network 2, consisted of 7 ADVANCE links. Network 3 consists of a total of 32 links.

All links in the three networks (except for link 5 in the first network) are part of a Closed Loop Signal Control system. All approaches (two approaches in each direction) to intersections labeled X and Y in Network 1 are detectorized. Intersection X is also covered by Networks 2 and 3.

The purpose of driving probe vehicles over a small number of links was to simulate fairly high levels [1–2 percent] of deployment of such vehicles, using very few equipped vehicles, by concentrating them on these links. The probes were released at the beginning of the route at clock times that allowed the formation of randomized headways. For the experiments on Network 3, the major objective was to gather extensive amounts of data on three study links: Af (right turn), AE (through movement) and Ag (left turn). We designed experiments such that vehicles were driven on randomly designed routes drawn over the 32 links. The objectives of the randomization were the following: (i) to maximize the number of traversals on the study links Af, AE and Ag and minimize traversals on supporting links (ii) to prevent the same driver from executing the same turning movement repeatedly to control for extraneous dependence and (iii) to ensure that all supporting links get comparable coverage.

The major variables on which data were collected are (i) link ID (ii) link exit time (iii) link travel time (in seconds) (iv) congested time (in seconds), that is, the amount of time spent on a link during which the vehicle traveled at or below 2 meters per second (4.4 mph) (v) congested distance (in meters), that is, the distance on the link covered at speeds less than 10 meters per second (22.5 mph).

The entry and exit points of the three study links were also monitored by video surveillance cameras, that

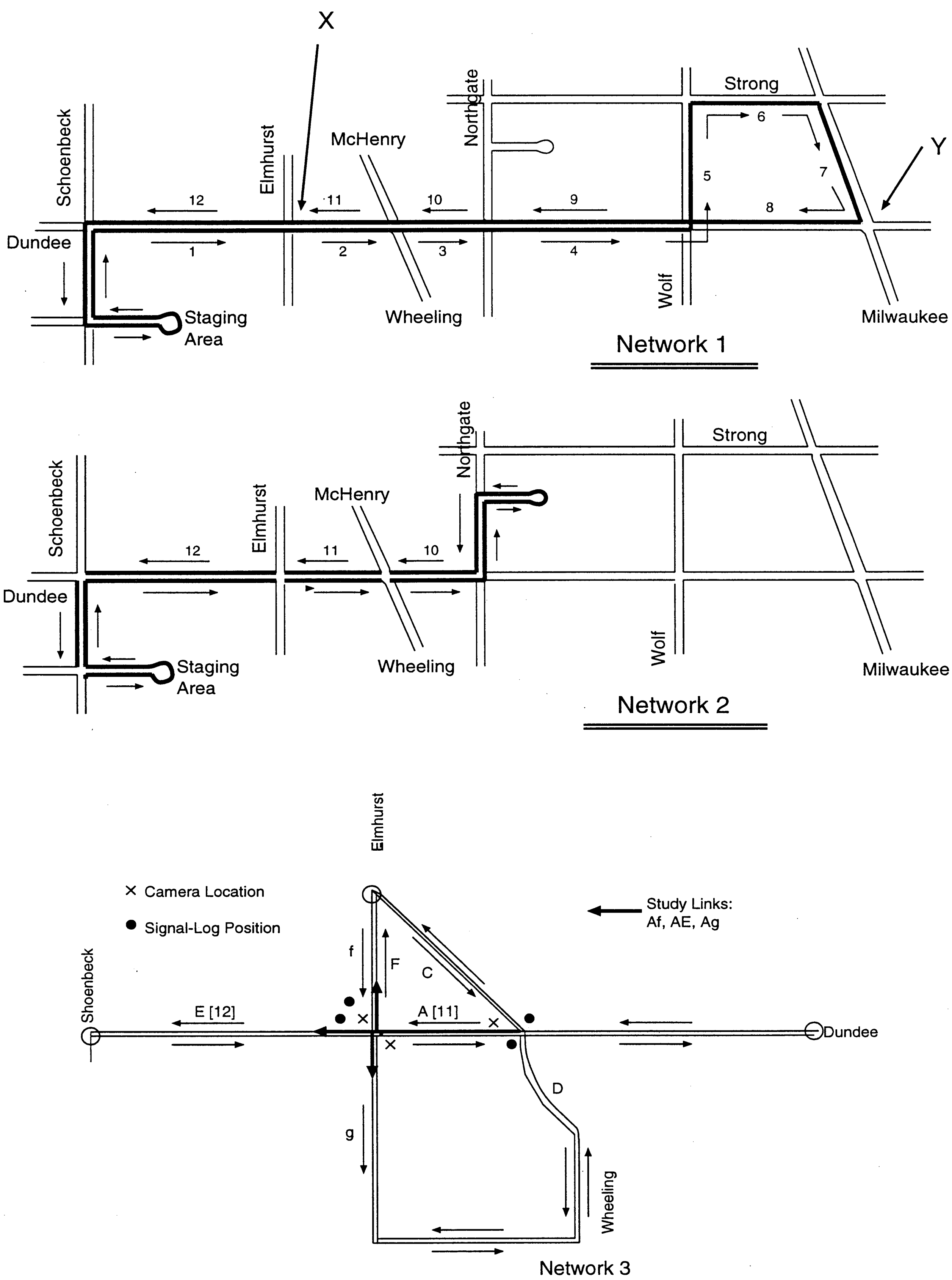


Figure 1: Three networks in the targeted ADVANCE study.

filmed the signal control at the both downstream and upstream intersections of the three study links (data on time-varying signal control were not available from any other source), as shown in Figure 1. It would ultimately be possible to obtain counts of vehicular demand at entry and exit points and also the travel times of all vehicles, by matching vehicles at the two points. Data on signals were also manually recorded at the locations indicated in Network 3 in Figure 1.

During the experiments, care was taken at all times to ensure that all units providing data used in this paper were time-synchronized according to the same clock. This included the video cameras, the clocks according to which the signal data were manually recorded and the Traffic Information Center. The importance of clock synchronization in ATIS and ATMS are elaborated upon in Section 3.

3 Probe Data Quality — Some Issues

The quality of any estimate would depend, to some extent, on the noise present in the data. Although most deployed and proposed ATIS/ATMS utilize state-of-the art hardware and software in communications and electronics technology, it is still essential to explore the sources of measurement errors, if any, and their structure. This is especially the case with new technology. In section, we devote some attention to this very important issue in the context of probe-based systems, and provide some illustrations about the nature of the errors that might arise.

There are three issues which deserve attention in an ATIS that involves forecasting of link travel times, although the first two issues are germane to all aspects of travel time measurements by probe vehicles. The first is one of locating vehicles precisely on the network because of errors induced by Global Positioning System. The second is one of synchronizing the clocks of all measurement devices within the ATIS/ATMS correctly and maintaining this synchronization over extended periods of time. The third is a little obscure but critical for the forecasting problem. It is an issue of the relationship between the amount of processing time needed by the system and the forecasting interval of the system.

At the current time, the United States Department of Defense introduces an error that does not allow precise location establishment for civilian purposes. Differential GPS (which was used in the ADVANCE study) reduces the location error to 10 meters with 90% probability and to 5 meters with 50% probability. However, considering the fact that urban arterial streets are usually a few hundred meters in lengths, even the use of differential GPS may lead to noticeable location errors.

In the ADVANCE evaluation, an in-depth study was done to capture the effects of probe vehicle location error by using observers in the probe vehicles who recorded the exit times from each link that the vehicle traversed. The differences in exit times between each two consecutive links along a route yielded the travel time on the downstream link. The analysis showed that 87.6% of the probe-generated travel times were within $\pm$ 5 seconds of those recorded by observers. Figure 2(A) gives a histogram of the differences between probe-recorded travel times and observer recorded travel times for all links in Network 1 in Figure 1, within this $\pm$ 5 seconds interval.

Measurement errors could lead to bias in estimates of travel time if they are systematic based on different conditions. In the ADVANCE study, the distribution of differences was similar for different links with different lengths and turning movements. Although there were tests done to validate the congested distance and congested time measurements, we did not directly investigate whether measurement errors vary with vehicle speeds. However, the distribution of differences were similar for congested links where virtually all vehicles incurred stopped or queuing delay compared to other links with good progression and low congestion, where vehicles rarely incurred any delay. Hence we have no reason to believe that location error problems with different link lengths, turning movements or vehicle estimates, which could otherwise bias estimates.

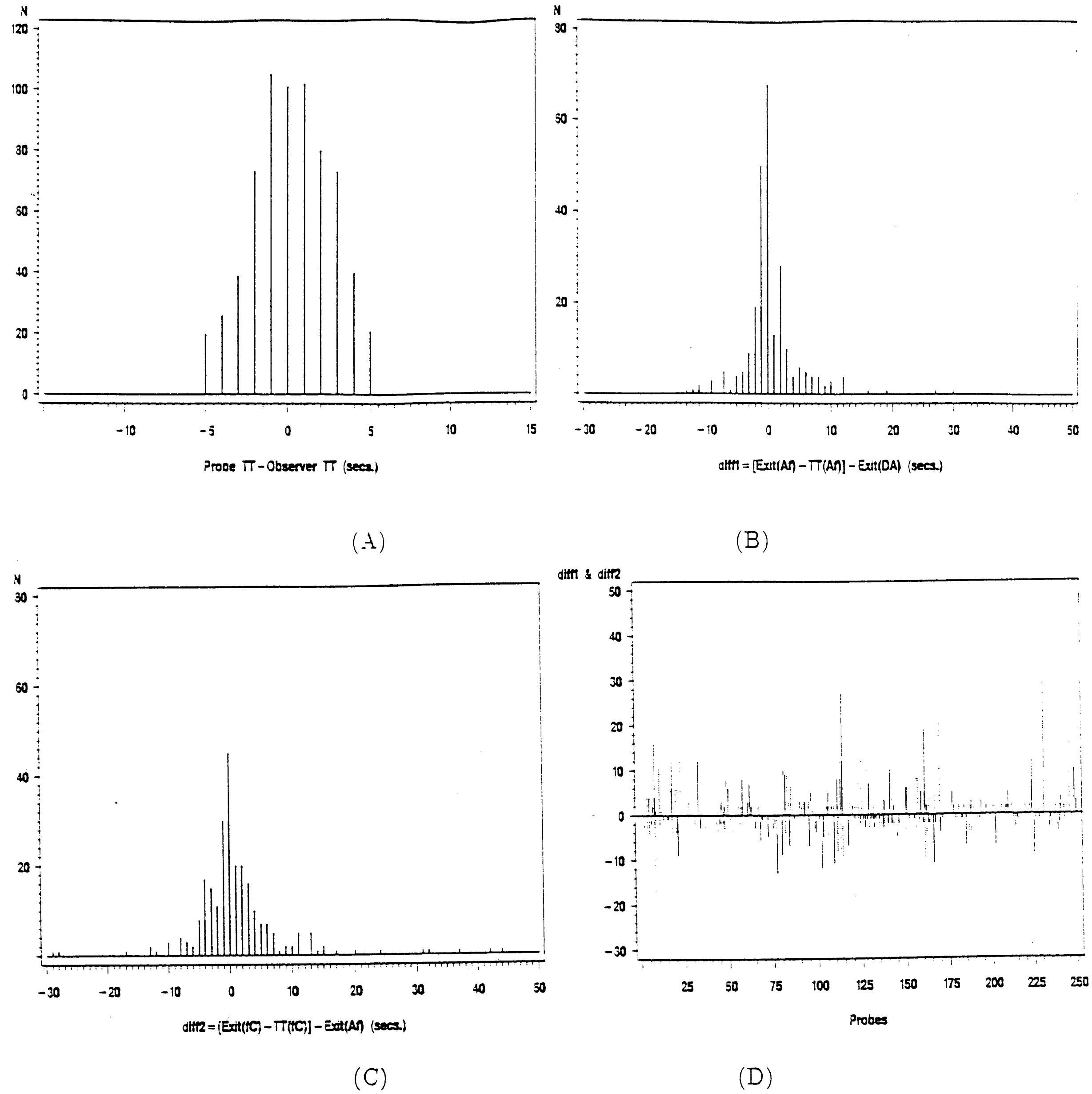


Figure 2: Quality of Probe Reports

A difficulty with this kind of observer-verification is that observers may have difficulty knowing the exact location where a link begins and ends, as defined in the ATIS map database. Hence, observer-recorded exit times may also have a location-error problem.

The second issue is one of the clock time of recorded events. A problem could occur due to two reasons — the first is that the clocks of different data gathering devices are not perfectly synchronized. This is a long-run maintenance issue. The second reason is more 'correctable' and is more of a design issue. A number of currently deployed ATIS/ATMS records the clock times at which events are recorded by the system as opposed to when an event actually occurs, usually to prevent oversaturation of transmission capacity. For example, in a probe-based system, the difference between the two clock time of the time of event and the clock time of recording could occur due to processing time by in-vehicle system hardware or due to lags in transmission or receiving time at the centralized or distributed information centers(s) due to 'radio-congestion'. In a system that measures travel times by video cameras at two different locations, measurements in travel times may be in error due to the recording speed of two video units. In such a case, two camera locations may need to be connected by hard-cable or by some other means to ensure that both cameras are indeed recording at the same speed (this was done in the case of the ADVANCE study). Inconsistency in recorded clock times would especially lead to a problem if data from multiple sources have to be integrated in real-time or even off-line.

The exit time from a link is the clock time stamp available in the ADVANCE data. However, the time that is actually recorded is the time at which this message is received at the Traffic Information Center. This message is transmitted 'instantaneously' after the vehicle exits a link, but due to reasons outlined above, there could be a difference between the actual exit time and the recording time. The difference between the exit time and the travel time on one link should yield exactly the exit time from a link immediately upstream along a vehicle's route. We present here an example from three consecutive links, DA, Af and fC shown in Network 3 in Figure 1. Figure 2(B) and (C) show the distribution of differences, $diff1$=[exit time(link Af) - travel time(link Af)] - exit time(link DA), and $diff2$=[exit time(link fC) - travel time(link fC)] - exit time(link Af) respectively. The shape of the distributions are similar for $diff1$ and $diff2$ and approximately normal, except for the fact that the tails are long on both the positive and the negative side. Figure 2(D) provides some insights into the structure of these long-tailed structure of the distributions. The quantities $diff1$ (black lines) and $diff2$ (grey lines) are shown in Figure 2(D) for each probe vehicle considered in Figures 2(B) and (C). For instance, for each probe, most large positive values of $diff1$ go with large negative values in $diff2$. This indicates that the long tail on the positive side in Figure 2(B) corresponds to the negative tail in Figure 2(C), which means that for the vehicles forming these tails, there was a lag in recording the exit time from Af but little or no lag in recording the exit time from fC and DA. This implies the possibility of cancellation of errors over a vehicle's route. This means that if there was a lag in recording the exit time from link Af, then there was little or no lag in recording the exit times from fC. If Figure 2(D) had instead shown that $diff1$ and $diff2$ are both positive, then the implication would have been that there were lags in recording the exit times from both Af and fC.

The third issue is relevant to probe-based systems with a real-time forecasting function. Typically, for dynamic route guidance purposes, in order to obtain a shortest path, forecasts of link travel times are needed a short horizon into the future. Some studies have shown that the usefulness of prediction deteriorates with increase in the length of the forecasting interval (for example, Thakuriah and Sen, 1995). Therefore, one would like to give predictions for a short time horizon. The system also has a 'cycle', called the processing cycle, during which the system processes incoming information. A problem could arise if there are inconsistencies between the two system 'cycles' — that of the forecasting cycle and the system processing cycle. This is because some time must be allowed for the system to process incoming information. Depending on the scale of the deployment of the system, the time needed for processing information may be quite large. If the forecasting time horizon is smaller than the processing time needed, then a forecasting interval may well retain an estimate of travel time from a previous system processing cycle. This could bar a real-time ATIS from actually using forecasts altogether in its shortest path calculations.

We conclude our comments on the quality of information via probe-based systems by stressing that one needs to be careful to check if ATIS or ATMS surveillance units yield data that have measurement errors that are not understandable or that are not easily correctable. A way to assess the size of measurement errors and their structure is by detailed field tests during the design phase of the system so that one can have some reasonable assurance about the quality of information, once the system is fully deployed.

4 Properties of Travel Time Data and Estimates

In this section, we present some analysis on the properties of data from probe vehicles and the estimates based on these data. The 'all-encompassing' property of data on signalized arterials is the presence of dependence in the data which is important to consider because this property would complicate statistical inference made on the basis of the data. We treat dependence in two ways: (i) situations when the measurements made by a vehicle is affected by another vehicle and (ii) situations where correlations among observations are non-zero and observations are related by either long-memory or short-memory trends or periodicities. The latter case can easily occur even when one observation is not affected directly by another observation.

To motivate the discussion on dependence among travel times and the implications on travel time estimates, we start by laying out the components of dependence in travel time measurements. In Section 4.1, we show that data generated by probe vehicles can be used to decompose the total travel time measurement into components that allow a better understanding of the structure of travel times. We present evidence of the short-lived cyclical pattern in travel time observations that is induced by signal control in Section 4.2. This trend imposes a profound effects on estimates of travel time if not handled carefully, as we will elaborate in Section 4.3. We treat two different cases of dependence — among observations on the same link (in Section 4.4) and among observations along links in a route (in Section 4.6). Then in Section 4.5, we address the issue of potential bias in estimates of travel time.

4.1 Decomposition of Travel Time

The travel time incurred by a vehicle on a signalized arterial may be decomposed into two components: cruise time and delay (intersection delay resulting in queuing and/or stopped delay). This decomposition has been used in some traffic simulations (Thakuriah and Sen, 1995). Volume affects both components by reducing the distance which a vehicle can travel at free-flow speed and increasing delay time. In this section, we investigate whether this decomposition is empirically verifiable.

Probe vehicles allow the empirical verification of this notion by recording the time spent at or below predetermined speeds. In the ADVANCE probe vehicles, as described in Section 2, three measurements were made by the vehicles with respect to this issue: the time spent at free-flow speed (when congested time (*ct*) and congested distance (*cd*) are both 0); the time spent at speed 0 or very low speed (less than 4.4 miles per hour, when *ct* and *cd* are both positive) and the third is a measure of the distance at a speed at or below 22.5 miles per hour, which is less than free-flow speed but greater than stopped or low speed (when *cd* is positive but *ct* is 0).

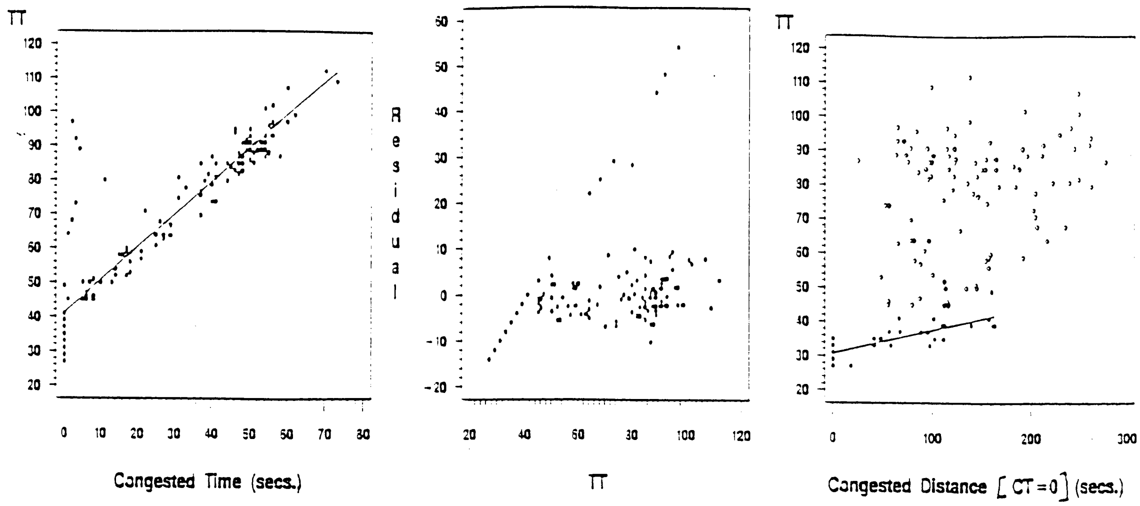
Figure 3(A) presents a scatterplot of *tt* against *ct* on Link 11 on Network 1 of Figure 1, with a linear fit via Model [A] of (1):

$$[A.] \quad tt_i = \alpha + \beta ct_i + \epsilon_i \quad \hat{\alpha} = 42.06 \quad \hat{\beta} = 0.97 \quad s.e.(\hat{\beta}) = 0.09 \quad s = 11.36 \quad R^2 = .88 \qquad (1)$$

One sees a large amount of variability of travel times at and around $ct = 0$, which indicates that in some cases, high travel times are incurred without slowing down to below the threshold speed of 4.4 miles per hour, whereas in other cases, vehicles do not slow down simply because they encountered free-flow conditions (concurred by low total link travel times). Except for this variability at $ct = 0$, the relationship between travel time and congested time is well-described by linear fits, indicating that most of the variability in travel time is accounted for by speeds less than 4.4 miles per hour, as supported by estimating the model

$$[B.] \quad tt_i = \alpha + \beta ct_i^+ + \epsilon_i \quad \hat{\alpha} = 42.69 \quad \hat{\beta} = 0.91 \quad s.e.(\hat{\beta}) = 0.04 \quad s = 5.23 \quad R^2 = .92 \qquad (2)$$

where ct_i^+ are positive values of *ct*. We can expect the structure of the estimated errors to reveal the pattern left in travel times after taking out the effects of queuing or stopped delay. These effects on signalized links are mainly due to the effect of signalization or red phase delay. The residuals are plotted in Figure 3(B), which shows that there is virtually no pattern left, except at $ct = 0$, indicating that the residual time (travel times without queuing or stopped delay) is almost a constant. However, even vehicles that do not incur stopped delay incur lower than free-flow speeds especially if they join the end of a moving queue. This part of the travel time observation would be affected by volume because higher volume would lead to higher queue lengths. This speed regime consists of vehicle acceleration from the critical *ct* speed of 4.4 mph to free-flow speed and de-acceleration from free-flow speed to below the *ct* threshold. This part of the



(A) (B) (C)

(A) Relationship between tt and ct.

(B) Residuals versus predicted values from Model [A] in expression (1).

(C) Relationship between tt and cd^0 where cd^0 are observations with $ct = 0$. $\diamond$: observations with positive ct values.

Figure 3: Decomposition of total travel time of each vehicle via measurements at different speed levels.

speed regime can be explained by means of Figure 3(C), where we have shown the relationship estimated via the model

$$tt_i = \alpha + \beta cd_i^0 + \epsilon_i \quad \hat{\alpha} = 30.71 \quad \hat{\beta} = 0.07 \quad s.e.(\hat{\beta}) = 0.008 \quad s = 2.99 \quad R^2 = .68 \tag{3}$$

where cd_i^0 are congested distances of those observations with $ct = 0$, indicating that slowing down below the free-flow speed is a significant contributor to variability in travel times for vehicles that does not incur stopped delay (recall that cd is the distance spent at speeds between 22.5 mph and 4.4 mph. Hence, the delay component includes the delay time that stopped and also the 'slow-down' time incurred by having to wait for the stopped queue to clear.

The analysis shows that the decomposition of vehicle travel times into cruise time and delay holds quite well on signalized arterials. An implication of this empirical verification is that simulations that use this simple decomposition would approximate vehicle travel times with reasonable accuracy.

4.2 Inferring Signalization Periodicity in Probe Travel Times

In this section, we explore the periodicities imposed by signalization on travel times by using probe data. Let $tt_{d,t,i}$ be the travel time experienced by the ith probe vehicle during day d, that exited a link during time-period t. Now consider the following model

$$\begin{aligned} tt_{d,t,i} &= \gamma + \sum_d \alpha_d I_{d,i} + \sum_t \beta_t I_{t,i} + \epsilon_i \\ \text{s.t.} \quad & \sum_d \alpha_d = 0 \\ & \sum_t \beta_t = 0 \end{aligned} \tag{4}$$

where

$$I_{d,i} = \begin{cases} 1 & \text{if} \quad \text{tt was incurred on day} \quad d \\ 0 & \text{otherwise} \end{cases}$$

and

$$I_{t,i} = \begin{cases} 1 & \text{if} \quad \text{tt was incurred during time interval} \quad t \\ 0 & \text{otherwise.} \end{cases}$$

The parameters α_d and β_t are respectively day effects and time-of-day (used as a surrogate for a link travel time diurnal pattern) effects experienced by the ith probe. In case of day d and time interval t, under the condition that $\mathrm{E}[\epsilon_i] = 0$, the model in equation (4) is equivalent to

$$\mathrm{E}[tt_{d,t,i}] = \gamma + \alpha_d + \beta_t \tag{5}$$

To estimate the model in (4), the independent variables corresponding to the α_d's and β_t's were coded indicator variables, with one indicator variable for each time interval t (each t is five minutes long) and one for each d used in the analysis. We estimated this model by least squares. If $\mathrm{E}[\epsilon_i, \epsilon_j] = 0$, then $\mathrm{Cov}\,(tt_i, tt_j) = \mathrm{Cov}\,(\epsilon_i, \epsilon_j)$, $\forall\, i \neq j$. Then the error structure of this model would allow one to be able to examine the effects of those variables that are not included in the model. Assuming the time-of-day effect lets us subtract out one of the major contributors to variability in travel times, that is volume, the error structure of (4) should reflect the contribution of the periodic structure imposed on estimated link travel times by signalization.

Figure 4(A) shows a plot of the time headway or the difference in exit time (on the same day d) between two probe vehicles i and j on the horizontal axis. The vertical axis givès an empirical covariance, $\mathrm{Cov}\,\widehat{(tt_i, tt_j)}$, (that is, it is the product of the residuals, $e_i \times e_j$, $i \neq j$, obtained from estimating the model given in (4)). The pattern shows that two pairs of vehicles that exit at certain headways are more likely to have the same estimated covariance, than those pairs that exit at other headways. Clearly, there is a cyclical pattern left in the estimated covariance, that cannot be accounted for by subtracting out the time-of-day effect only from the travel time observations.

A periodogram of the estimated covariance (presented in Section 4(B)) lets us see that the largest periodogram ordinate is estimated at 135 seconds. This is, in fact, the average signal cycle length, which we were able to corroborate with the information on signals obtained from the video data. Figure 4(C) shows a plot of the estimated covariance from a model

$$\begin{aligned} ct_{d,t,i} &= \gamma + \sum_d \alpha_d I_{d,i} + \sum_t \beta_t I_{t,i} + \epsilon_i \\ \text{s.t.} \quad & \sum_d \alpha_d = 0 \\ & \sum_t \beta_t = 0 \end{aligned} \tag{6}$$

where $ct_{d,t,i}$ is the congested time of the ith probe vehicle on day d and time interval t and $I_{d,i}$ and $I_{t,i}$ are as defined above. Clearly, the same periodicity in evident whether we use the estimated covariances of travel times or congested times. Figure 3 shows $\mathrm{Cov}\,(tt_i - \widehat{ct_i, tt_j} - ct_j)$, $i \neq j$, that was estimated via a model similar to those given in (4) and (6). The estimated covariance shows the approximate trend in the

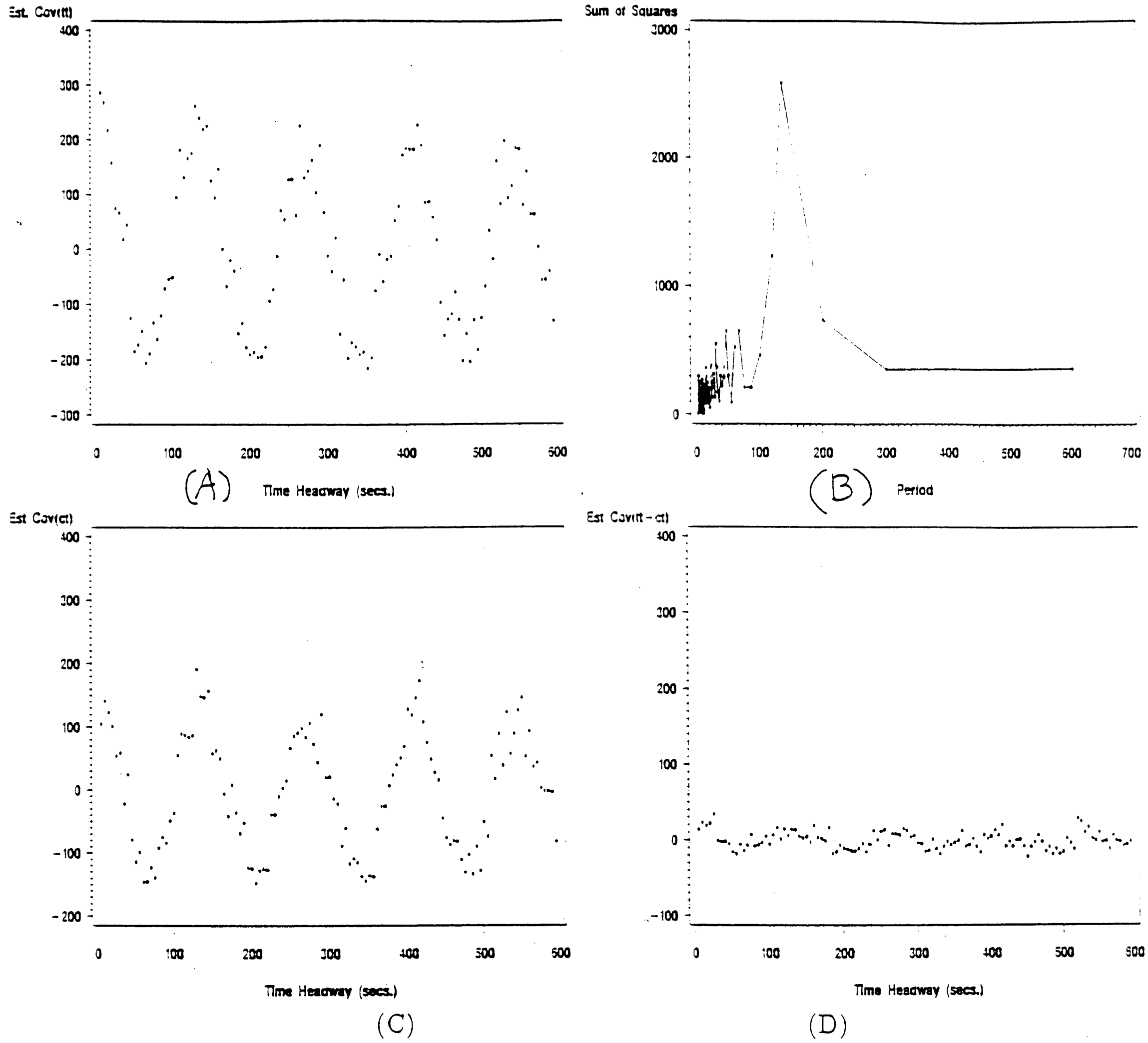


Figure 4: Estimated Covariances and Signal Cycle Periodicity Apparent in Error Structure

difference, indicating that at periods of close to the average signal length, there is some trend left owing to vehicle acceleration and deacceleration, (as discussed in Section 4.1) even after taking out the time-of-day effects and delay. To understand the effect of signalization on travel times, one needs to be able to account for the vehicle arrival process within each cycle.

Figure 5 shows an estimated relationship between the vehicle's exit time within a signal cycle with its travel time. The model is

$$tt_i = f(\rho_i) + \epsilon_i \quad (7)$$

where tt_i is the travel time of the i probe vehicle and $\rho_i = t_i \mathrm{mod}(r_i)$, t_i being the exit time of the ith probe from the link and r_i is the start of the red phase of the cycle in which the vehicle exited from the link (a cycle is assumed to start with the start of the red phase and end with the end of the green phase). The scatter plot indicates that there is a lot of noise in the data. This is because the plot includes data from

1. signal cycles with several different volume levels; for example, the points in the upper right side are travel times of probe vehicles that exited the link in cycles of high congestion, so that even though these vehicles exited late in the cycle, they still incurred high travel times
2. cycles with low congestion and from probes that were able to take advantage of good progression by incurring low travel times even though they exited at the start of the green phase; these are indicated

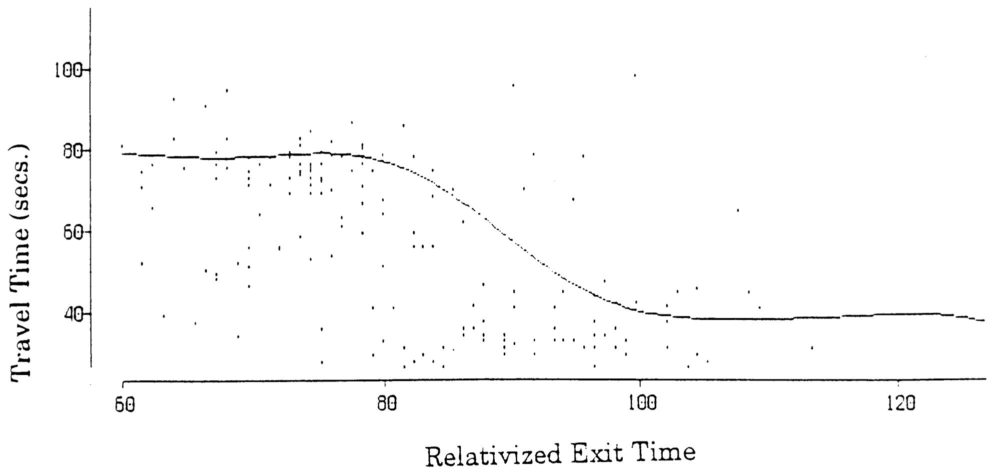


Figure 5: Travel Time and Relativized Time ρ

by the group of points at the lower left side of the plot

3. from cycles of varying lengths so that the pattern is shifted to fit the 'average' cycle'.

The model was estimated by a non-parametric Gaussian kernal estimator (Hastie and Tibshirani, 1990) with a bandwidth, $\lambda = 4.98$ ($s = 17.02$ and $R^2 = .53$). Figure 6 shows the residuals plotted against ρ — there is virtually no pattern left in the residuals. However, the spread in the points is slightly higher for higher predicted travel times because of the role of progression effects in different cycles.

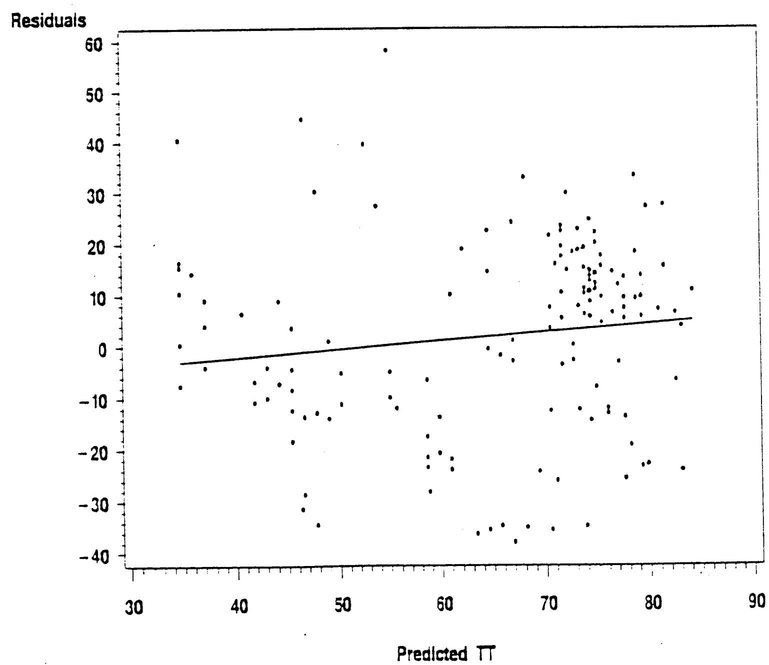


Figure 6: Estimated error structure after subtracting out the effect of signalization.

The function $f(\rho)$ is volume-specific and at least a non-increasing function of volume. In Figure 5, the estimated function is fitted for conditions of signal delay or uniform delay only (which occurs because of differences in vehicle arrival times at the intersection during the red or the green phase) and not for situations of overflow delay (when vehicles incur delay due to sustained conditions of demand exceeding available capacity). However, the estimated variance of this model is the lowest of the models we had considered so far (a piece-wise linear fit of the relationship between travel time and relativized time had yielded an estimated $s = 18.49$, whereas the model in equation (4) relating to day and time-of-day effects yielded an estimated $s = 25.75$). This indicates that the pattern imposed by signal cycles is perhaps the major covariate of travel time on signalized arterial and that in order to subtract out the effect of signalization from travel time observations, one needs to consider the relationship of the vehicle's entry/exit time vis-a-vis the signal cycle. However, all the factors itemized in the previous section contribute to high variance in the estimates of travel time.

4.3 Dependence among Travel Time Measurements

We now investigate further the lack of independence among travel time observations and its effects on travel time estimates. In Section 4.4, we discuss the dependence structure among travel time observations on the same signalized arterial and the nature of the nature of the covariances among observations. In Section 4.6, we explore the dependence structure among observations on multiple links along a route and the implications for route travel time estimates.

4.4 Dependence — Case of a Single Signalized Arterial

In the case of travel time observations on a single signalized arterial street, the lack of independence could occur due to several reasons, the crucial factors contributing to this dependence being signalization and volume levels. Congestion imposes a periodicity via a long-range diurnal trend whereby, even though separated by a long term period, vehicles at the rise of a peak period would incur travel time that are similar to those entering the link during the decay of a peak period. But the short-lived cyclical pattern imposed by signalization is the more pronounced contributor of dependence among observations generated by different vehicles. Signalization causes dependence by (i) inducing similarity in entry times from upstream links during the green phase of upstream links, which leads to similarity in arrival times within the signal cycle of the downstream link; for example under otherwise similar conditions, two vehicles arriving at a traffic signal ten seconds after the onset of the red phase will have similar link travel times (ii) inducing progression effects so that vehicles that enter from the same upstream link incur travel times are similar compared to those that arrive after executing other turning movements. Thus, not only would one conjecture that link travel times are correlated but, as we showed in Section 4.2, that covariances are functions of time headways (difference between exit times between two probes) and are variable.

The variance of the estimated mean $\bar{tt} = n^{-1}\sum_{i=1}^{n} tt_i$ of probe travel times, $tt_1, tt_2, \ldots tt_n$, for the same link over a time interval t is

$$\begin{aligned} \text{var}[\bar{tt}] &= \text{E}[\bar{tt} - \text{E}(\bar{tt})]^2 = \text{E}[n^{-1}\sum_{i=1}^{n}(tt_i - \text{E}[tt_i])]^2 \\ &= n^{-2}\,\text{E}[\sum_{i=1}^{n}(tt_i - \text{E}[tt_i])^2 + \sum_{\substack{i,j \\ i\neq j}}(tt_i - \text{E}[tt_i])(tt_j - \text{E}[tt_j])] \\ &= n^{-2}\sum_{i=1}^{n}\text{var}[tt_i] + \sum_{\substack{i,j \\ i\neq j}}\text{Cov}[tt_i, tt_j] \end{aligned} \quad (8)$$

be exactly the same and even a single observation would suffice as a replacement for all the observations. In practice we would expect something in between, where $\eta > \nu > 0$. Even then as $n \to \infty$, $\text{var}[\bar{tt}] \to 0$. Although our focus has been on the estimated sample mean, we would expect this situation to be true for most estimators.

If the tt_i's were uncorrelated, the average covariance ν would be zero. Therefore, in order tp shpw that the observations are correlated, we need to test if $\nu = 0$. We did for each link on Networks 1 and 2 in Figure 1. Also the estimation was conducted separately for peak and off-peak periods.

To estimate η and ν, we used the estimated residuals $e_{d,t}$ from the model in (4) — if the model in (4) holds, a reasonable estimate of $[\bar{tt} - \text{E}[\bar{tt}]]$ is $\bar{e}_{d,t} = \bar{tt}_{d,t} - \hat{\gamma} - \hat{\alpha}_d - \hat{\beta}_t$, and $[\bar{e}_{d,t}]^2$ estimates $\text{var}[\bar{tt}]$. Therefore, the model estimated is

$$[\bar{e}_{d,t}]^2 = \nu + \gamma_1[1/n_{d,t}] + \epsilon_{d,t} \tag{11}$$

where the $\eta = \nu + \gamma_1$ and $\epsilon_{d,t}$ is the error term relating to the tth time interval on day d. These parameters were estimated by weighted least squares, with weights equal to $\widehat{\text{E}[tt_i]}^{-2}$. To test the hypothesis H, $\nu = 0$, against A, we conducted a t-test (the details of the validity of the t-test is given in Sen *et al.* (1996)).

We can see from Tables 1 and 2 that all estimates of ν are positive and are significant at the 5 percent level or better, with the exception of links 5, 6 and 8 in the peak period and 2, 5 and 6 in the off-peak period. Based on the estimates of ν, the hypothesis of no correlation may be rejected, indicating the presence of dependence among observations on the same link.

Links 5 and 6 are not on Dundee Road, which is a major arterial with high progression effects; they are side-streets with very little congestion. Link 5 is also not signalized. To enter Link 8 while driving along the route, vehicles have to execute a permitted right turn that obviates the effects of signalization to a large extent. These results indicate that the effects of both congestion and signalization induce dependence.

An advantage to treating average covariances and variances as model parameters is that one can estimate them via data generated by probe vehicles, as we have illustrated. Once average variances and covariances are correctly estimated, we can use these estimates for the following purposes:

1. Estimation of Sample Sizes: One purpose for which these estimates could be used is to estimate sample sizes and therefore, deployment levels necessary for precise estimation of travel times. In Sen *et al.* 1996, we presented a detailed analysis of sample sizes required for travel time estimation, given dependence in the data. The relationship between the estimated precision of estimates, $s.e.[\bar{x}]$, where $\bar{x}$ is the sample mean link travel time over a five-minute interval and n, the number of probe observations per time interval of five minutes is shown in Figure 7. The results indicate that due to dependence, the variance of the estimate of expected travel time does not decay to 0, so that high levels of deployment may not offer a great deal of improvement in the quality of dynamic route guidance. Therefore, the marginal improvement in the precision of estimates drop off after the use of data from a few probe vehicles and adding more probe coverage on a link does not improve the quality of the estimate.

2. Decision Rules to Give Dynamic Guidance: Recent results obtained by Thakuriah and Sen (1995) by a simulation-based model of the quality of information showed that in a dynamic route guidance system, dynamic travel time estimates had very large variances, especially at low deployment levels. These estimates resulted in routes which were inferior to routes resulting from static estimates. In an attempt to remedy this difficulty, we developed decision rules for providing dynamic estimates to vehicles only when they differed 'adequately' from the corresponding historical estimate.

   In the simulations, we used two criteria for giving dynamic link travel time updates for the purpose of dynamic route guidance: [A.] *Relaxed Criterion [RC]* when we broadcast dynamic estimates if they differed from the corresponding static estimate by 1 'standard error.' [B.] *Strict Criterion [SC]* when we required a difference of 2 'standard error' for broadcast.

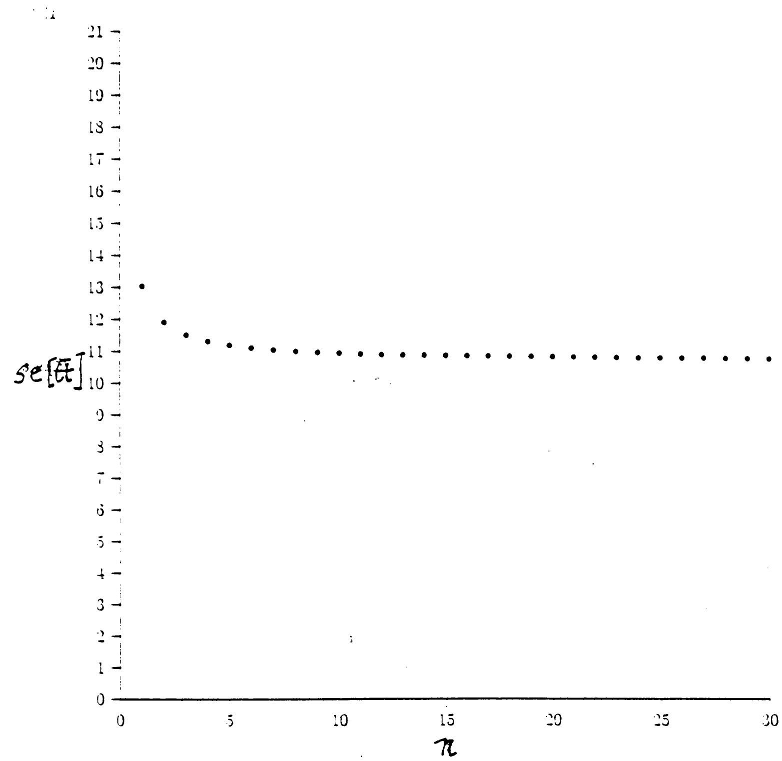


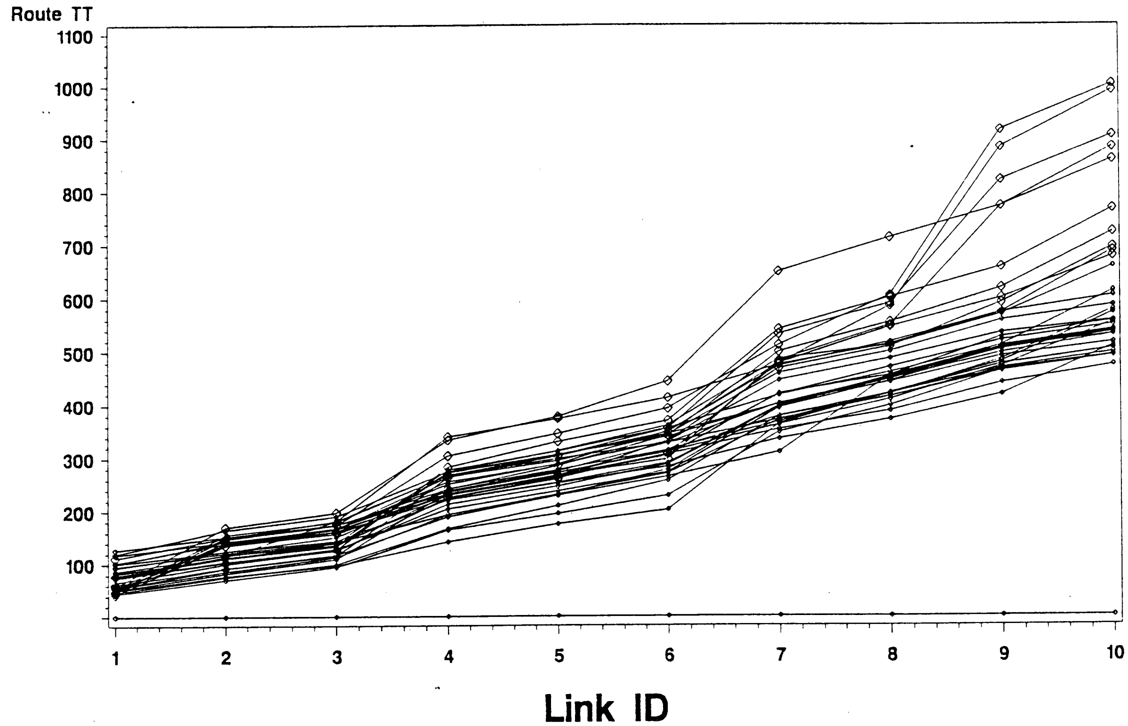
Figure 7: Relationship between standard error $s.e.[\bar{tt}]$ and frequency n of probes on Link 32 during peak period estimated from model $[\bar{e}_{d,t}]^2 = \nu + \gamma_1[1/n_{d,t}] + \epsilon_{d,t}$.

The results showed that the quality of route guidance (in terms of travel time savings to users) improve substantially when these decision rules are used for dynamic guidance, by acting as a "gatekeeper" against "unreliable" dynamic updates. However, if the 'standard error' is computed in the usual way by using the usual sample standard deviation of probe reports divided by the $\sqrt{n}$, where n is the number of probe reports used in constructing a dynamic estimate, then because of the presence of correlations among probe data, the estimate of the variance would be an under-estimate. This would permit gatekeeping decision rules to allow more unreliable dynamic updates, with negative effects on the quality of guidance.

4.5 Bias in Estimates of Travel Times

The analysis presented in Section 4.2 showed that signal control is a major contributor to correlation and that the vehicle's exit time from a link relativized to the signal cycle is an important variable in the estimation of travel times. However, data on the sequences of signal control phases on a continuous clock are unlikely to be readily available in ATIS of the near future. In this section, we explore the effects of non-inclusion of the vehicle's entry/exit time to the cycle in a model that estimates travel times.

If a regression model does not include a term(s) reflecting traffic signal timings, the estimates from that model may be biased. However, not all estimates in such a model will be biased in the same way. A model that estimates travel time as a step-function of the clock time of day is a candidate for such bias. The effects of this bias can be substantial in the case of pre-determined temporal aggregation intervals of travel time on signalized arterials.



◇: Group A.
•: Group B.

Figure 8: Vehicle route trajectories.

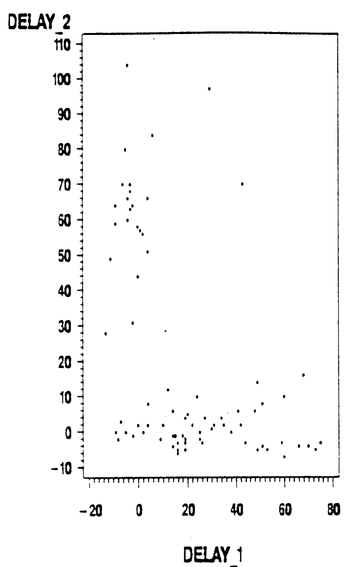
not correct because we have ignored covariance terms, $\sum_{\substack{i,j \\ i \neq j}} \text{Cov}\,[tt_i, tt_j]$ for $i \neq j$ in the same link.

To illustrate a situation of link-to-link dependence, we graphically investigate the effects of link-to-link relationships on 10 links in Network 1 by means of Figure 8. This figure shows the route on the horizontal axis and the cumulative link travel times of each of the 35 vehicles considered, on the vertical axis ($\bar{T}_R = 621$, $s.e.[\bar{T}_R] = 24.89$). We depict two groups of vehicles in the figure; Group A depicted in diamonds with $\hat{T}_R = 918$ seconds and $s.e.[\bar{T}_R] = 39.34$ and Group B depicted in dots, $\hat{T}_R = 585$ seconds and $s.e.[\bar{T}_R] = 4.32$. Both groups incur similar travel times upto Link 3, as indicated by the slope of each line upto Link 3. Vehicles in Group A incur higher travel times on Link 4. Their high travel time condition is exacerbated in Link 7 and in Link 9; however, even on other links, the slope of the grey lines almost never flatten out compared to the black lines, indicating that vehicles in Group A did not get the chance to 'recover' their travel times on subsequent links, so as to incur route travel times comparable to Group B.

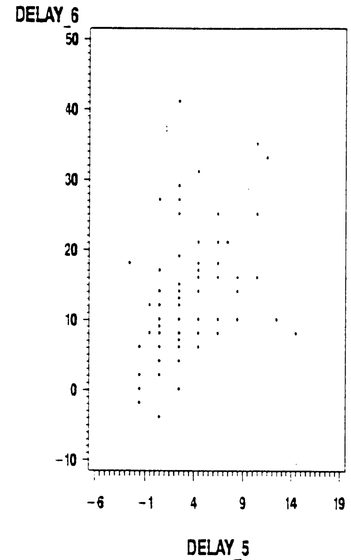
As mentioned earlier, the dependence structure between links depend fundamentally on the signal control parameters and progression factors between links. Figure 9(A) and (B) show scatterplots of

$$delay = traveltime - [link \quad length/speed \quad limit],$$

between consecutive links along the same route as in Figure 8. For example, DELAY-1 means the quantity *delay* on link 1. The variable plotted on the horizontal axis is always the delay of the upstream link. Two patterns become apparent (i) an L shape; for example the pattern between Link 1 and Link 2 and (ii) an



(A)



(B)

Figure 9: Scatterplots of vehicle delays on pairs of consecutive links.

approximate increasing linear shape; for example between Link 5 and Link 6. Pattern (i) indicates poor progression between two links; for instance if a vehicle traverses an upstream links with no delay, then it would almost surely have to stop at the downstream link. Pattern (ii) indicates situations where a high travel time on one link would be followed by high travel times on the downstream link. The same situation holds for the case of low travel times.

Given the dependence structure between links, it would appear to be virtually impossible to predict travel times on downstream links via closed form delay functions. Further research is needed to predict volumes and time-varying sequence of signal control on a continuous clock, in order to estimate travel times on a link-to-link basis for the precise estimation of route travel times, given the dependence structure in the specific network under consideration.

5 Conclusions

In this paper, we commented on the quality of data obtained from probe vehicles and on the properties of the estimates of expected travel time obtained on the basis of these data. Like any new technology, data from probe based systems need to be analyzed to ensure that there are no measurement errors the effects of which are not understood, because these errors could significantly affect the quality of the estimates based on the data. We do so with data from ADVANCE probe vehicles and found the nature of the measurement errors to be understandable and not worrisome. We then analyzed the statistical properties of the estimates of link travel times and discussed issues of high variance and bias. There is dependence among observations on the same link and on links along a route, which has profound implications on statistical inference.

In conclusion, we recommend that direct measurements on travel times generated by probes or other surveillance technologies need careful handling for two reasons:

1. measurement errors may be present, with unknown resultant effects on the estimates of travel time;
2. the data would have certain undesirable properties, that if not carefully attended to, may have

significant negative effects for the purpose of statistical inference.

Thus, any estimation process that uses probe-based data needs to be cognizant of these two issues.

References

1. Boyce, D. E., A. M. Kirson and J. L. Schofer (1994). ADVANCE — The Illinois Navigation and Route Guidance Demonstration Program. In *Advanced Technology for Road Transport*, Ian Catling (ed.), Artech House, Boston.

2. Hastie, T.J. and R.J. Tibshirani (1990). *Generalized Additive Models.* Monographs on Statistics and Applied Probability 43. Chapman and Hall, 1990.

3. Koutsopoulos, H.N. and H. Xin (1993). An Information Discounting Routing Strategy for Advanced Traveler Information Systems. In *Transportation Research* Vol. C, No. 3, pp. 249-264.

4. Sen, A., P. Thakuriah, X. Zhu and A. Karr. (1996). Frequency of Probe Reports and Variance of Link Travel Time Estimates. Accepted for publication in ASCE, The Journal of Transportation Engineering.

5. Sen, A., S. Sööt and S. Berka. (1996). Probes and Detectors: Experiences Gained from ADVANCE. White Paper, Urban Transportation Center, University of Illinois at Chicago, Chicago, Illinois.

6. Sen, A. and M. Srivastava (1990). *Regression Analysis: Theory, Methods and Applications.* New York: Springer-Verlag.

7. Thakuriah, P. and A. Sen (1995). An Investigation into the Quality of Information given by an Advanced Traveler Information System. *In Review.*