SYNTHESIZING ROUTE TRAVEL TIME DISTRIBUTIONS FROM SEGMENT TRAVEL TIME DISTRIBUTIONS

Isaac Kumar Isukapati
ikisukap@ncsu.edu

George F. List
gflist@ncsu.edu

Billy M. Williams, Ph.D., P.E.
billy_williams@ncsu.edu

North Carolina State University
Department of Civil, Construction, and Environmental Engineering
208 Mann Hall
2501 Stinson Drive, Campus Box 7908
Raleigh, NC 27695-7908 USA

Alan F. Karr
karr@niss.org
National Institute of Statistical Sciences
19 T.W. Alexander Drive
Research Triangle Park, NC 27709-4006

Submission Date: August 1, 2012
Word count: 5,957 + (7 x 250) = 7,707
ABSTRACT

This paper examines a way to synthesize route travel time probability density functions (PDFs) based on segment-level PDFs. Real-world data from I-5 in Sacramento, CA are employed. The first finding is that careful filtering is required to extract useful travel times from the raw data because it is trip times that are observed not travel times – i.e., the movement of vehicles between locations. The second is that significant correlations do exist between individual vehicle travel times for adjacent segments. Two analyses are done in this regard: one predicts downstream travel times based on upstream travel times; and the second checks for correlations in travel times between upstream and downstream segments. Results of these analyses suggest that strong positive correlations exist. The third finding is that comonotonicity, or perfect positive dependence, can be assumed when generating route travel time PDFs from segment PDFs. Kolmogorov-Smirnov tests show that travel times synthesized from the segment-specific data are statistically different only under highly congested conditions. And even then, the percentage differences in the distributions of the “synthesized” and “actual” travel times are small. The fourth finding, somewhat tangential, is that there is little variation in individual driver travel times under given operating conditions. This is an important finding since such an assumption serves as basis for all traffic simulation models.
1. INTRODUCTION

Travel times are an important aspect of transportation systems (1-3). Travelers and shippers want to know how long it will take to get from point A to B and whether the trip will be completed on time. Users want to select routes that have reliable travel times; and if possible, depart when the likelihood of arriving on time is highest. In addition, system operators want to minimize variations in travel times at the segment and route level. The sources of variability can be traveler behavior or influencing factors such as capacity reductions, traffic control devices, weather, incidents, work zones, special events, and demand fluctuations.

The recent studies of travel time reliability have produced significant new findings. Seminal work in this regard was conducted by Lomax et al. (4), who studied possible travel time reliability measures. Follow-on efforts focused on finding ways to improve reliability (5) and mitigate congestion (6). Margiotta et al. (7, 8) developed a guide to effective freeway performance measurement, and created analytical procedures for determining the impacts of reliability mitigation strategies.

Vehicle-based sensors in the form of automatic vehicle identification (AVI) and automatic vehicle location (AVL) systems have made it possible to collect data on individual trips. Though AVI technology is as old, if not older than AVL, it has only recently been used to monitor travel times. The use of AVI technology gained popularity when toll tags were put into service by agencies like the New York State Thruway and companies like Mark IV (9). As with AVL, AVI probes make it possible to monitor point to point travel times on both freeways and arterials. Li et al. (10) used AVI data to gain insight into travel time variability and its causes. Wasson et al. (11) were first to suggest using Media Access Control (MAC) ID matches to estimate travel times. The advent of Bluetooth sensing technology and the increasing density of MAC IDs in the traffic system made this approach more attractive.

The question is how to generate travel time probability density functions (TT-PDFs) at the route level based on this AVI and AVL data. If the number of observations for a given origin-destination (OD) pair is large, the PDF can be developed directly. If the observations are too sparse, it ought to be possible to construct route related travel times and travel time distributions based on segment level observations. But, to do so, correlations between the travel times have to be taken into account. In general, if X and Y are independent random variables, then the distribution h(Z) for Z = X+Y can be constructed by convolving f(X) with g(Y). However, if X and Y are correlated, then convolution cannot be used; a different technique must be employed. And this paper shows that those correlations are significant.

Hence, this study explores an idea for synthesizing route TT-PDFs that takes into account the possible correlations between adjacent segments. This paper presents findings from three months of Bluetooth data collected on I-5 in Sacramento, California.

2. PRIOR WORK

The ADVANCE project in Chicago was one of the first to use probe data to study travel times (12, 13). Vehicles equipped with dead reckoning equipment recorded time stamps at specific locations (Evanston and O’Hare airport) as they traveled across northeast Chicago.
These travel times were then used to help motorists determine the best paths to use when traversing this largely arterial-based urban network.

Morris, Kay, and Kornhauser (14) reported that trucking companies use AVL technologies to manage their fleets. Information about truck movements is retrieved in real time; the dispatcher switches load assignments, re-routes the trucks to maximize quality of service while minimizing the impacts of congestion. Transit operators use AVL technology to track buses, which improves their ability to keep the buses on schedule, to alert riders to bus arrival times and locations, and to make it easier for the system to meet rider demands. Quiroga and Bullock (15, 16) developed a methodology for conducting travel time studies using GPS-based equipment, and created guidelines for determining appropriate sample sizes.

List et al. (17) used a peer-to-peer system of AVL probes to collect real-time travel times. Two hundred probes were deployed in upstate New York. The main focus was on journey-to-work trips for people traveling to a cluster of businesses in a single area. The experiment demonstrated that time data from such probes would help drivers determine the best routes to use through congested and incident-impacted networks. The findings from the experiment suggested that probe data could be used to create real-time route guidance systems (17-22).

There has also been experimentation with using transit vehicles and taxis as AVL probes. Chakroborty and Kikuchi (23) developed a model for predicting travel times for automobile using transit vehicles as probes. Hall and Nilesh (24) compared automobile and transit vehicle trajectories to explore alternative methods for detecting congestion on arterials. Bertini et al (25) demonstrated that buses can be used as probes to evaluate traffic conditions on arterials. Pan, Wang, and Ran (26) indicate that taxis have been used to collect traffic data in Japan, Germany, and Malaysia. Ma and Koutsopoulos (27) indicated that in Berlin the taxis served as floating cars and their travel time observations were processed into traffic information that other services then offered to clients.

Methodologically, Hellinga et al. (28) developed insights into reducing the bias of link travel time estimates from probe data. Cetin et al. (29) suggested the factors affecting minimum number of probes required for reliable average travel times. Yamamoto et al. (30) developed ways to analyze the variability of travel time estimates using probe vehicle data. Byon et al. (31) developed a GISTT (GPS-GIS Integrated System) for travel time surveys. This system enables one to match GPS data with spatial map features using GIS for monitoring traffic conditions on specific links. Dion and Rakha (32) developed ways to construct average travel times using AVL-based sampling. Pan et al. (27) proposed a GPS based methodology for collecting historical travel time data that includes link travel time and information on intersection signal delay, for an arterial. Furthermore, they developed a post-trip map-matching algorithm to project GPS data onto an arterial network. Berkow et al. (33) developed techniques for constructing the shape of the congested regime in time and space along urban arterial, combining signal system detectors and buses as probe vehicles. Feng et al. (34) explored the ability of GPS-equipped probe vehicles to provide characterizations of arterial travel times and real-time traffic conditions.
Researchers’ also explored statistical ways to represent the distributions of observed travel times. Van Lint et al. (35) suggested using variance as a measure to portray travel time unreliability on freeways. Jintanakul et al. (36) experimented with using Bayesian mixture models to estimate freeway travel time distributions based on small samples of probe vehicles on multiple days. Guo et al. (37) experimented with using a multi-state (multi-mode) model to portray the distributions of travel times being observed. Susilawati et al. (38) have experimented with the use of multi-modal Burr density functions. This work shows promise and is intriguing because the analytical functions are simple and the density function allows for skew.

The most recent efforts continue to advance these lines of investigation. Mahmassani et al. (39) explored ideas about robust relationships for reliability analysis. Barkley et al. (40), as a result of involvement in SHRP-2 L02 explored the use of multi-state models to develop relationships between travel time reliability and non-recurrent events. Ernst et al. (41) developed sampling guidelines for using probe vehicles to characterize arterial travel times. Ramezaniet al. (42) explored the use of Markov chains to estimate arterial route travel time distributions; Guo et al. (43) extended their investigations of using multi-state models to represent travel time distributions by exploring the use of mixed skewed models, and Feng et al. (44) explored the use of Bayesian models to construct arterial travel time distributions based on data from GPS-equipped probe vehicles.

3. FIELD DATA COLLECTION

For this study, time stamps were collected for Bluetooth MAC addresses at four monitoring stations on I-5 just south of US-50 in Sacramento, CA. Data were collected both in the northbound and southbound directions for approximately three months (January to April) in 2011. These data are very useful in exploring ways of generating route TT-PDF’s from adjacent segment TT-PDF’s.

3.1. Data Source

The data were obtained on a 5-mile section of I-5 in Sacramento, CA. Figure 1 shows where the four Bluetooth readers were stationed. The distances between the stations (going southbound) were 3.6 miles, 1.1 miles, and 0.9 miles, respectively. The Bluetooth data included the MAC address, signal strength, and UNIX time stamps. These data were processed using BluSTATS software developed by TRAFFAX, Inc. to create reader-to-reader observations. The output records include: the MAC address, timestamps at the upstream and downstream data collection devices, and the associated estimated “trip” time (computed by differencing upstream and downstream timestamps) in minutes. Around 150,000 unique vehicle trip times were created in each direction.
An important detail is that nothing in the dataset indicates whether the vehicle made intermediate stops. Hence, even though the times-between locations can be observed, there is no guarantee that the times represent travel times. Instead, only some of the observations are travel times, the other observations are trip times where the vehicle made an intermediate stop, etc. Hence, the following section explains how travel times were distinguished from trip times.

3.2. Filtering Travel Times From Trip Times

Figure 2(a) shows all of the southbound route “trip” times for the 5.6 miles. It is obvious that more than just travel times are being captured. The times range up to 240 minutes – the limit on timestamp connections employed by the processing software – not necessarily the greatest differences in time between the timestamps.

Clearly, many observations are in the 4-10 minute range. This is highlighted in Figure 2(b) where the upper bound on the times displayed is 40 minutes. It is also clear that travel times are embedded in the trip times – ranging from 4-7 minutes – and that for certain situations (when non-recurring events occur) the times are 10 minutes or more, up to about 35-40 minutes.

Therefore, an important task is to extract the travel times from the trip times. This is non-trivial, because legitimate travel times do exist under certain (non-recurring) conditions that are well within the range of values that would otherwise represent trip times.

The methodology employed here stitches together small batches of legitimate observations, working from one batch to the next. It picks acceptable travel times from batch $n$
and then examines batch \( n+1 \) to find additional acceptable values. The procedure begins by examining all the times to find a reasonable lower bound – technically, the 0.1% trip time. This eliminates outliers well below times that are physically feasible. It then works with batches of \( n \) data points to see which ones should be kept (in this instance, the procedure uses \( n = 10 \)). From among the \( 4^{th} \) through the \( 7^{th} \) of 10 values it finds the smallest one that is above the 0.1% threshold. It checks the remaining nine values to see if any of them should be rejected. In doing so, two tests are performed: 1) is the value less than a local minimum or 2) is it greater than a local maximum? The algorithm dynamically adjusts its asymmetric acceptance window to accept or reject new observations. If the value is within 10% of the upper bound, the allowable increment to the upper bound is increased by 20%; otherwise it is decreased by 10%. Similarly, if the value is within 10% of the lower bound, the allowable decrement to the lower bound is increased by 10%; else it is decreased by 10%.

The results from applying this procedure can be seen in Figure 2(c). Notice that the accepted travel times rise dramatically for a short period of time during a PM peak when an incident occurred. It is clear that the values representing noise have been removed and that acceptable observations of travel times have been retained.
Figure 2: Vehicle Trip Times on I-5 Southbound in Sacramento

(a) Observations of Vehicle Trip Times on I-5 Southbound in Sacramento

(b) Observations of Vehicle Trip Times 40 minutes or Shorter on I-5 Southbound in Sacramento

(c) An Example of Cleaned up AVI Data for Travel Times Southbound on I-5 in Sacramento
3.3. Categorizing Observations by Regimes

The operating conditions that are extant on the segments must be taken into account to develop meaningful travel time PDFs. Therefore, the next step is to label each observation – all 150,000 in this case – with the regime that was operative when the travel time was observed. A regime is a combination of a nominal system loading and a non-recurring event condition. Ultimately, it is for these regimes that the TT-PDFs are developed and through which the effect of these conditions is understood.

The labeling is accomplished in two steps. The first labels each observation with the non-recurring event that was underway when the observation was obtained, including “none”. The second adds a second label that indicates what congestion level was (or should have been) extant when the observation was obtained.

3.3.1. Adding the Non-Recurring Event Labels

The label for the non-recurring event (including “none” or “normal operations”) should be consistent with FHWA’s ideas about the sources of congestion. For the data used here, PeMS was used to identify incidents, weather, special events, and unusually high demand events. “Insufficient base capacity” was captured by the nominal congestion condition categories described above (i.e., situations where the D/C ratio was high enough that sustained queuing occurred). The high demand category was equivalent to “fluctuations in demand”. Work zones were not present in any of the networks for which analyses were conducted, but it is clear from the congestion-related analyses that they would have an impact on reliability.

3.3.2. Adding the Congestion Condition Labels

The congestion level labels were created by analyzing the observations that remain once the “non-normal” observations have been removed. While average speeds and/or travel times are often used as the basis for congestion assessment, here a semi-variance measure was used instead because the study was focused on travel time reliability not average delays. A semi-variance is a one-sided metric that uses a reference value \( r \) (instead of the mean) as the basis for calculating the sum of the squared deviations, and only observations \( x_i \) that are greater than (or less than) that reference value \( r \) are included in the calculation.

\[
\sigma_r^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - r)^2 \quad \text{and} \quad \sigma_r = \sqrt{\sigma_r^2} \quad \exists \ x_i \geq r
\]

For each five-minute time period, here, \( x_i \) was the average travel time observed on a specific day and \( r \) was the minimum of those values. The values of \( \sigma_r^2 \) were computed for each of the 288 five-minute time periods of the day.

The \( \sigma_r^2 \) values were used to classify the 5-minute time periods into one of four categories of operation: uncongested, low, moderate, and high. Uncongested meant that the \( \sigma_r^2 \) value was low: vehicle interactions were not creating variations in the average value from one day to another. Low meant there was some variation in the average travel times. Moderate meant there was more variation; and high meant there was a lot.
4. SYNTHESIZING ROUTE TRAVEL TIME PDFs

Since there are likely to be instances when too few route-specific observations are available to support direct development of route travel time PDFs, route-based travel time distributions have to be synthesized from segment level distributions.

4.1. Evidence of Positive Correlations Between Adjacent Segments

Creating the route-level distributions might not be too difficult if the travel times were highly correlated. Suppose the upstream travel times can be used to predict the downstream travel times. If this were the case, an equation could be developed that predicts the downstream travel time based on the upstream travel time.

\[ \tau_{j+1}^i = \alpha \times \tau_j^i + \beta \]

Where \( \tau_{j+1}^i \) = Downstream travel time of vehicle ‘i’
\( \tau_j^i \) = Upstream travel time (rate) of vehicle ‘i’
\( \alpha, \beta \) = regression parameters

To see if this might be possible, regression equations were developed for temporally sequential sets of 50 pairwise observations of upstream and downstream travel times. Figure 3 presents the analysis results. The \( R^2 \) values tend to be quite high during off-peak periods and drop during the peaks. This trend makes sense in that drivers are able to achieve their desired speeds (travel rates) when the congestion is “not high”; however when congestion is “high” the interactions with other vehicles make it difficult for the desired speeds (travel rates) to be achieved. These analyses demonstrated that the times are correlated.

Figure 3: Scatterplot representing \( R^2 \) Vs. Time of the day

A broader brush analysis, in which the travel times were grouped on the basis of similar flow rates led to the same conclusion. The data were classified into one of four groups (based on hourly volumes) and a regression analysis was conducted for each group. (Note that in this case the chronology of observations is not preserved.) The \( R^2 \) values within each group varied
between 0.4 – 0.7. Moreover, the $R^2$ value diminished as the flow rate increased.

Evidence of these upstream-to-downstream relationships can also be seen in Figure 4. Correlation coefficients are plotted for non-overlapping sets of 50 sequential combinations of upstream and downstream travel times (paired by vehicle ID). That is, if $\tau_{ij}$ is the upstream travel time for vehicle $i$ then $\tau_{ij+1}$ is the downstream observation for a vehicle ‘i’.

The correlation coefficient was then computed via:

$$r_{j,j+1} = \frac{1}{(n - 1)} \sum \left( \frac{\tau_{ij} - \bar{\tau}_j}{\sigma_j} \right) \left( \frac{\tau_{ij+1} - \bar{\tau}_{j+1}}{\sigma_{j+1}} \right)$$

Where $r_{j,j+1}$ = correlation coefficient between segments ‘j’ and ‘j+1’

- $n$ = sample size
- $\bar{\tau}_j$ = mean travel time (rate) of the sample upstream
- $\sigma_j$ = standard deviation of the sample upstream
- $\bar{\tau}_{j+1}$ = mean travel time (rate) of the sample downstream
- $\sigma_{j+1}$ = standard deviation of the sample downstream

Several inferences can be drawn from Figure 4. The first is that there are strong correlations between adjacent segments from mid-night to about 3:00 pm (please note that the ‘r’ values are consistently in the range of 0.8 - 0.9 for segments 39-9 & 9-10, whereas the range is between 0.6 – 0.8 in the case of 9-10 & 10-11). The second is that the correlation coefficient declined between 3:00 pm – 6:00 pm. This is due to reduced maneuverability of the drivers under highly congested network conditions. Also, notice that the correlation values are again high from 6:30 pm onwards. Therefore, adjacent segments exhibit strong positive correlations except under highly congested conditions.

**Figure 4: Scatterplot representing ‘r’ Vs. Time of the day**
4.2. Synthesis Options

Since correlations do exist, convolution cannot be used to construct route-level travel time density functions. Alternative techniques are necessary.

Two approaches are discussed here. The first uses simulation. The second uses a concept called comonotonicity (45, 46), which is an extreme form of positive, but not necessarily linear, dependence.

Monte Carlo simulation uses conditional sampling of the CDFs from one segment to the next to create the route-level density function. Technically, if a random vector \( T = (\tau_1, \tau_2, \ldots, \tau_n) \) is the sum of a set of correlated random variables, then it has a joint CDF:

\[
F_{(\tau_1,\tau_2,\ldots,\tau_n)}(\tau_1, \tau_2, \ldots, \tau_n) = P(T_1 \leq \tau_1, \ldots, T_n \leq \tau_n)
\]

Here the joint CDF represents the CDF for the route-level travel time. The critical element is that the correlations must be taken into account. The method of conditional distributions needs to be used. The generation of \( T \) is explained below for \( n = 2 \).

\[
F_{(\tau_1,\tau_2)}(\tau_1, \tau_2) = P(T_1 \leq \tau_1, T_2 \leq \tau_2) = P(T_1 \leq \tau_1)P(T_2 \leq \tau_2/T_1 \leq \tau_1) = F_{(\tau_1)}(\tau_1)F_{(\tau_2/\tau_1)}(\tau_2/\tau_1)
\]

To generate \((T_1, T_2)\), one first generates \( T_1 \) from \( F_{(\tau_1)}(\cdot) \), and then \( T_1 \) from \( F_{(\tau_2/\tau_1)}(\cdot) \).

The comonotonicity-based idea adds the travel times at matching percentile values for successive segments. Comonotonicity has been employed in calculating portfolio risk. Random variables \( X \) and \( Y \) are comonotonic, if there is a third random variable \( Z \) as well as increasing functions \( f \) and \( g \) such that \( X = f(Z) \) and \( Y = g(Z) \). As one would expect, the concept generalizes to three or more random variables. It implies that individual percentile values from each of the distributions can be added to obtain the percentile value for the distribution of the sum.

I-5 data for successive segments strongly suggest that comonotonicity holds. Table 1 provides a comparison of the southbound travel times (for four regime conditions) on I-5 predicted by summing the segment travel times against the overall route travel time.
Table 1: A Comparison of actual percentile travel times for a given route against values obtained by summing the travel times for the same percentile on the individual segments

For example, the second, third and fourth columns show the percentile travel times for segments 39-9, 9-10, and 10-11 based on the travel times for each individual segments. The fifth column shows travel times obtained if these percentile values are simply summed. That is, the values in this column do not represent the percentiles of any underlying distribution. They are simply the algebraic sums of the percentile-based travel times shown to their left. The sixth column then shows the percentile travel times for the overall route based on actual observations of the route travel time. Finally, the last column shows that the differences between the percentile value sums and actual observed values. It is clear that the values are nearly identical for the Uncongested, Low, and Moderate congestion conditions. Moreover, when congestion is high, the differences for every percentile are more than 1%, and all percentiles greater than 70% have differences between 2%-6%.
Therefore, not only does comonotonicity seem to pertain, but it does so in spite of the fact that the density functions are multi-modal. Figure 5 shows both the density function for four operating regimes. In all four instances, the density functions are bi-modal. Except for highly congested conditions, the match is strong between the density function obtained by adding the percentiles and that which was actually observed. The explanation for this lies in the fact that dependencies between adjacent segments are weak when network operating conditions are highly congested.

To examine the comonotonicity idea more closely, Kolmogorov-Smirnov tests were performed (for \( \alpha = 0.01, 0.05, 0.1, 0.15, \) and 0.2) to see if statistically significant differences existed between synthesized and actual travel time distributions. The results suggest that the density functions are only significantly different for the highly congested conditions. The results are summarized in Figure 5.
Figure 5: Comparisons between the travel time density functions synthesized from individual segment percentiles and the observed values.
4.3. Driver Consistency

There is a side issue about driver consistency. A driver would be consistent if he or she chose to target the same travel speeds (perhaps by facility type) when making trips. This is a fundamental principle upon which all traffic simulation models are based. Driver consistency would not have to exist across days for the results presented earlier to pertain. Drivers could be randomly selecting desired speeds on a per-trip basis and then following those desired speeds for the duration of the trip. This would produce the results presented above. However, intuition suggests that drivers might be consistent more generally about their desired speeds.

Since the data were available, the authors elected to see if there was any evidence of consistency in desired speeds for specific drivers across all available observations. To be clear, the dataset included the MAC IDs for all the vehicles that were observed. (So in truth what follows is an analysis in the consistency of travel rates for specific MAC IDs (whatever they happen to pertain to, not drivers per se.)

The analysis involved examining consistency in the travel rates that were observed. The travel rate observations were broken down into regimes and then grouped by MAC ID. Where a specific MAC ID appeared more than once for a given regime condition, its average travel time ($\bar{\tau}_n^i$), standard deviation of travel times ($\sigma_n^i$), and coefficient of variation ($C_v^i = \frac{\bar{\tau}_n^i}{\sigma_n^i}$) were computed.

Where $\bar{\tau}_n^i$ = average of ‘n’ observed travel times for a specific MAC ID ‘i’
$\sigma_n^i$ = standard deviation of ‘n’ observed travel times for the corresponding MAC ID ‘i’
$C_v^i$ = coefficient of variation for the corresponding MAC ID ‘i’

Each ‘dot’ in the Figure 6 represents a specific MAC ID, it’s x-value represents average travel rate in seconds per mile, and the corresponding y-value represents coefficient of variation. Please note that variation in individual driver travel times under Normal-Uncongested, and Normal-Low congested is almost negligible. The variation in travel times grows as network operating conditions grow worse.
Figure 6: Average Vs. Coefficient of Variation Plot for Different MAC ID's Under Various Regimes

5. CONCLUSIONS AND RECOMMENDATIONS

This paper has examined the distributions of travel times for individual vehicles and segments with a clear focus on how those distributions could subsequently be used for synthesizing route travel time PDFs. Four observations have been highlighted. The first is that careful filtering is required to extract the travel times from the raw data because it is trip times that are observed not travel times – i.e., the movement of vehicles between locations. The second is that significant correlations might exist between the travel times for adjacent segments – two analysis are done in this regard 1) upstream travel times (rates) are used to predict downstream travel times (rates); 2) correlation analysis is done to check for relationships in travel times (rates) on upstream-to-downstream segments, results suggest that there are strong positive correlations between adjacent segments. The third is that the concept of comonotonicity, or perfect positive dependence, can be assumed when generating route travel time PDFs from segment level observations. Kolmogorov-Smirnov tests are conducted to check for statistical differences between “synthesized” and “actual” travel time distributions; results suggest that these CDF’s are statistically different only under highly congested conditions. However, concept
of comonotonocity can still be used even under these conditions as the percentage errors between
the “synthesized” and “actual” percentile travel times are small. The fourth is that there is little
variation in individual driver travel times under given operating conditions. This is an important
and pleasing finding given that an assumption of driver consistency serves as basis for traffic
simulation models.

Recommendations from these findings are twofold. First, careful filtering is required to
extract travel times from the trip times reported by AVI and AVL-equipped vehicles. Simple cut-
off criteria are inadequate, especially when one has an interest in exploring the impacts of non-
recurring events and other abnormal conditions. Second, the strong correlations among travel
times on adjacent segments strongly suggests that synthesis of route-level travel times needs to
be predicated on methodologies that explicitly take into account those correlations. Adding the
percentile travel times for adjacent segments appears to have promise in reproducing the
empirically observed route travel time distribution.

6. ACKNOWLEDGEMENTS

The authors would like to acknowledge the support and encouragement provided by the
SHRP-2 staff during project L02, which forms the basis for much of the material presented here.
Of special note is William Hyman who served as the contract manager through much of the
development work. Also of note are the many other individuals that were part of the L02 project
team whose insights contributed to the materials presented here.

7. REFERENCES

Board, 14, p. 468, 1934.
5. Cambridge Systematics, Texas Transportation Institute, University of Washington, and
Prepared for the Future. Strategic Highway Research Program NCHRP Project 20-58,
Transportation Research Board, Washington, DC.
6. Cambridge Systematics and Texas Transportation Institute (2005) Traffic Congestion and
Reliability: Trends and Advanced Strategies for Congestion Mitigation. Prepared for the
FHWA Office of Operations, Washington, DC.
7. Margiotta, R., Lomax, T., Hallenbeck, M., Turner, S., Skabardonis, A., Ferrell, C., and Eisele,
ect 3-68, Transportation Research Board, Washington, DC.
Mitigation Strategies. SHRP-2 L03 Final Report Transportation Research Board,
Washington, DC.
to a new tollway technology. *Marketing Research* 11(2), 5-16.


of the Transportation Research Board, Washington, DC.
