Estimating Nationwide Link Speed Distribution Using Probe Position Data

Iris Lin, Rong He, Alain L. Kornhauser

University of Texas at Austin, Austin, Texas, USA
Department of Civil Engineering, Santa Clara University, Santa Clara, California, USA
Department of Operations Research and Financial Engineering, Princeton University, Princeton, New Jersey, USA

Published online: 08 Feb 2008.

To cite this article: Iris Lin, Rong He & Alain L. Kornhauser (2008) Estimating Nationwide Link Speed Distribution Using Probe Position Data, Journal of Intelligent Transportation Systems: Technology, Planning, and Operations, 12:1, 29-37

To link to this article: http://dx.doi.org/10.1080/15472450701849667

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the “Content”) contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions
Estimating Nationwide Link Speed Distribution Using Probe Position Data

IRIS LIN,1 RONG HE,2 and ALAIN L. KORNHAUSER3

1University of Texas at Austin, Austin, Texas, USA
2Department of Civil Engineering, Santa Clara University, Santa Clara, California, USA
3Department of Operations Research and Financial Engineering, Princeton University, Princeton, New Jersey, USA

The purpose of this study was to investigate different methods of estimating speeds on links in United States’ nationwide transportation network from different sources, such as probe vehicles and loop detectors. The Trip Speed Assignment Technique was developed and applied to the probe vehicle data to create speed distributions by time of day for many of the major roadway segments of the U.S. highway system. The technique can generate average speeds for a large network using a relatively small sample size. The average speeds follow general traffic patterns. Thus it is possible to use the probe vehicle data with the Trip Speed Assignment Technique to provide appropriate travel speed information for a large area. Exhaustive spot loop detector data have been used to compare the accuracy of the speeds derived from the Trip Speed Assignment Technique. Discrepancies are discussed and recommendations are made to improve the Trip Speed Assignment Technique.

Keywords GPS; Loop Detector Data; Probe Data; Speed Distribution; Trip Speed Assignment

INTRODUCTION

Interest in developing Intelligent Transportation Systems (ITS) has been growing, specifically in the area of Advanced Traveler Information Systems (ATIS) due to recent technological innovations and increasing congestion in the U.S.A.’s highway system. Unexpected delays and increased uncertainty in travel times are not only frustrating to motorists but are the source of substantial inefficiencies in the freight transportation system. Highly sought are abilities to plan routes that minimize travel times and the tools to more accurately estimate arrival times. Each of these applications necessarily requires the collection of real-time information about conditions on the road ahead. Such information is also valuable to traffic operation and management, such as ramp metering systems, coordination of traffic control systems, incident detection (Balke et al., 1996), and congestion pricing. When applied to dynamic in-vehicle route guidance, real-time information can help travelers assess alternate routes and make informed real-time routing decisions since they would possess more accurate estimates of route travel times ahead.

The goal of calculating travel times on many feasible routes necessitates active mechanisms that would allow the amalgamation of data that has both breadth of coverage throughout the network as well as depth of observations at each location. Two basic approaches are available that can estimate off-nominal, temporal, and spatial variations in speeds for “all” feasible routes ahead. One is based on the “continuous experience data” that can be made available from a relatively small sample of all the vehicles traveling throughout the network. These sampled vehicles are called “probe” vehicles. They simply report their position and time intermittently. Variables such as actual driver route choice behavior, the frequency of probe reports (Sen et al., 1997), the coverage of probe vehicles (Chen and Chien, 2000), and the characteristics of the links (Ferman et al., 2005) within the network affect the accuracy of the travel time estimates based on probes. Within the probe vehicle techniques, vehicles that utilize the Global Positioning Systems (GPS) provide accurate data with relatively high initial cost but low operating costs. One of the characteristics of probe vehicles is that the drivers can be in the traffic stream without disrupting it. Drivers are not actively involved in the data collection process, and as a result,
the data are less likely to be biased by the driver. So gathering travel times/speeds with moving vehicles produces measures over space that are more representative of highway performance than the point estimates of speed from fixed detectors, particularly for capturing the essentially stochastic travel speeds/times (He et al., 2006). Recently, GPS and cellular probe techniques have been researched and practiced in academic and industrial areas all around the world (Qiu et al., 2007; NCHRP Project 70-01, 2005). While it seems to be technically feasible to implement GPS-type probe techniques for traffic data collection, a few field trials have been conducted with ambiguous results. In 2002, Yim and Cayford at University of California conducted an evaluation of GPS and cellular probe tracking. They found that global positioning systems with 20-meter accuracy can produce data for 99.2% of surface street segments and 98.9% of the freeway segments. The Virginia Department of Transportation (VDOT) and Maryland State Highway Administration (MSHA) study (Smith et al., 2004) determined that only 4.2% of the intervals demonstrated significantly different mean speeds than the actual means for daytime periods. For nighttime periods, 27.3% of the speed estimates were different on the highway and 20.7% on major arterial streets.

The other approach focuses on gathering data on all vehicles at some fixed locations in the network where loop detectors, cameras, or other remote sensing technologies are utilized. The location, accuracy, and number of sensors affect the ability to reliably estimate travel times.

The purpose of this study was to investigate these two different sampling mechanisms for estimating speed distributions along links in U.S.A. nationwide transportation network. The trip speed assignment technique uses a data set collected using a technology similar to GPS in probe vehicles. These vehicles are 27,417 trucks traversing the U. S. highway network from March 25 to March 31, 1998. Figure 1 is a snapshot of those moving trucks; 4.5 million observations were recorded in a tabular log with the following information:

- Unique truck identifier number
- Latitude and longitude of the location
- Time at that location

Traditional techniques require very large amounts of data in order to build a speed distribution for a network as large as the U. S. The trip speed assignment technique circumvents this problem by assigning every link within a particular route the same average speed as the one derived for the entire route. It generates speed distributions at specific locations by deriving average speeds between specific reporting points. This potentially less precise measure of speed estimation has the benefit of breadth of coverage across the network. The focus of the study

Figure 1 Instantaneous location of the 27,417 moving trucks.
was to demonstrate the capability of generating speed distributions throughout the North American highway network and to estimate the degree of precision lost in the estimation of travel speeds at specific points as compared to traditional loop detectors and similar sensing techniques. The ultimate objective was to determine if useful speed distributions could be built for large networks using a relatively small number of probes. The trip speed assignment technique takes advantage of the following two basic assumptions. One, the actual path taken between consecutive observations can be computed. The other is that the average speed across the path between consecutive observations is an accurate representation of the actual spatial distribution of speeds experienced by the probe as it traveled that path. If these assumptions are valid, then the question remains as to how large of a sample size is sufficient to accurately represent the actual speed distributions across the network.

**TRIP SPEED ASSIGNMENT TECHNIQUE**

The trip speed assignment technique takes the average speed along a route and assigns that average speed to every link composing the route. Links can have multiple observations if they are part of more than one route. This methodology aims to provide accurate travel information for a large network given a relatively small sample size. Traditional methods require that the vehicles must traverse the link in order to build the speed distribution. For example, the ADVANCE project estimated 4,000 probe vehicles just to monitor an area of 300 square miles (Boyce et al., 1991). The study in Houston required an average of two vehicles as a small sample size. Traditional methods require that the vehicles would take an extremely large number of probe vehicles to cover the 132,954 links of U.S.A. Thus, the average speed and “nominal time” associated with each O-D pair could be assigned to one of the 50 resulting distinct classes. The trip speed assignment technique created, for each link of the network, the distributions of the number of observations in each of the 10 speed classes for each of the 5 times-of-day. On a particular trip, each link of the route will have a commodity vector in which a “1” is placed in the appropriate

The study in Houston required an average of two vehicles as a small sample size. Traditional methods require that the vehicles must traverse the link in order to build the speed distribution. For example, the ADVANCE project estimated 4,000 probe vehicles just to monitor an area of 300 square miles (Boyce et al., 1991). The study in Houston required an average of two vehicles as a sample size per link for a 95% confidence level and 10% relative error (Turner and Holdener, 1995). Granted that many of these numbers are dependent on the characteristics of the network, it would take an extremely large number of probe vehicles to cover the 132,954 links of U.S.A.

The trip speed assignment technique couples consecutive observations to form a sequence of timed origin and destination pairs. In Figure 2, the two red stars, A and B, are the recorded origin and destination using GPS or similar location reporting devices. The first assumption is that the vehicle was actually in part of more than one route. This methodology aims to provide an accurate representation of the actual spatial distribution of speeds experienced by the probe as it traveled that path. If these assumptions are valid, then the question remains as to how large of a sample size is sufficient to accurately represent the actual speed distributions across the network.

![Figure 2](image-url) Geo-coding to the network.

The distance from A to B would be 0.32\(x_{b,a} + x_{a,q} + 0.5x_{q,c}\).

For each trip composed of an origin and destination, a “nominal time” is also assigned. This “nominal time” is the average of the times recorded at the origin and the destination. The “nominal time” is used to associate the average speed data with a particular time-of-day such as AM peak.

After this technology is implemented for consecutive observations, the average speed and the time assigned to each link are placed in specific speed class and time group, respectively. For the purpose of efficiently generating travel speed distributions, we chose to accumulate the speed data in terms of 10 speed classes and 5 time-of-day groupings. This allowed us to generate coarse travel speed distributions for each link of the network for each time-of-day. Speed classes chosen were:

1. \(< 24 \text{ km/h (15 mph)}\),
2. \(24 \text{ to } 32 \text{ km/h (15 to 20 mph)}\),
3. \(32 \text{ to } 40 \text{ km/h (20 to 25 mph)}\),
4. \(40 \text{ to } 48 \text{ km/h (25 to 30 mph)}\),
5. \(48 \text{ to } 56 \text{ km/h (30 to 35 mph)}\),
6. \(56 \text{ to } 64 \text{ km/h (35 to 40 mph)}\),
7. \(64 \text{ to } 72 \text{ km/h (40 to 45 mph)}\),
8. \(72 \text{ to } 80 \text{ km/h (45 to 50 mph)}\),
9. \(80 \text{ to } 88 \text{ km/h (50 to 55 mph)}\),
10. \(88 \text{ km/h (55 mph)}\).

Time-of-day groupings were:

1. 6am to 10am (AM peak),
2. 10am to 3pm (Mid day),
3. 3pm to 7pm (PM peak),
4. 7pm to 6am (Evening),
5. Weekend.

Thus, the average speed and “nominal time” associated with each O-D pair could be assigned to one of the 50 resulting distinct classes. The trip speed assignment technique created, for each link of the network, the distributions of the number of observations in each of the 10 speed classes for each of the 5 times-of-day. On a particular trip, each link of the route will have a commodity vector in which a “1” is placed in the appropriate
class and forty-nine “0” in all the other classes.

\[ V_j = 1 \text{ if } S_k + 10 \times (T_k - 1), \]  

(1)

where

\[ V_j = \text{commodity vector with 1 in the } j\text{th bucket}, \]
\[ S_k = \text{speed block at observation } k, \]
\[ T_k = \text{time block at observation } k. \]

One of the assumptions that consecutive observations make up a continuous trip can be suspect. In reality, the trucker could have stopped for a length of time between the two points, which would result in a slow average speed. Thus vehicle counts in the first speed class, less than 24 km/h or 15 mph, is representative of “bad data”. Care was taken to eliminate observations at the beginning and end of sequences having multiple observations at the same location. Also, the driver could have diverted to an intermediate location not on the computed path. In such cases, the derived average speed would be slower than reality and there could be a lack of correspondence between the links traveled and the links assigned. However, because time intervals between observations were short, rarely greater than one hour, and the trips were generally long, it was concluded that this did not significantly bias the results.

**SPATIAL AND TEMPORAL CHANGES IN SPEED**

The purpose behind the truck data analysis is to observe how accurate the truck data set collected from GPS-like technology is in estimating travel time over a temporal and spatial distribution. However, travel time is not studied directly. Instead, it is the average speed that is analyzed. The average speed is derived from the location and time of the recordings. Only one more step is required to calculate the travel time, which is to divide the distance of the link by the average speed measured across the link. All the links in the network have different distances, and the travel time is a function of the length. Thus, it would be more meaningful to study the average speed rather than the travel time.

Efforts were made to display the spatial variations of the average speed distributions over the U. S. network graphically. Figure 3 shows only one time class—PM peak from 3pm to 7pm, and one speed class—over 80 km/h (50 mph). Only those links

---

**Figure 3** National highway network during PM peak. (Links are highlighted if average speed >80 km/h.)
on which average speed is over 80 km/h are highlighted. It is found that those links are away from the major cities. Many of them are located in the remote parts of the interstates.

Due to the difficulty of viewing this information on a national level, in particular, in Figure 4, maps are zoomed into the New York City and northern New Jersey area to examine the speed differential by direction. Probe vehicles moving in one direction will be represented by a band on one side of the link while those traveling in the opposite direction will have a band on the other side of the link. These maps do not indicate the volume. All the bands represent one unit volume. Links are highlighted red or green corresponding to the average speed on those links. Red segments, on which average speed is lower than 64 km/h (40 mph), are mainly in the urban areas. Because of the large amount of commuters going into New York City during the morning and leaving during the early evening, there exists an imbalance in the color bands on the links. During the AM peak, the faster moving traffic is located on the side of link directing outward New York City, while directing toward New York City during the PM peak.

In order to find out how many vehicles are moving in which speed class, the volume distribution over speed is necessary. As shown in Figure 5, most vehicles belong to the higher speed classes, such 8, 9, and 10 (45–55 mph). The vehicle counts in

![AM Peak](image1)

![PM Peak](image2)

**Figure 4** Average speed (by direction) around NY/NJ during AM and PM peak.

![Graph](image3)

**Figure 5** Percentage distribution of average speed on national highway network by time of day.
each speed class were summed over all the speed class as follows:

\[ v_{ci, TOTAL, k, t} = \sum_j v_{ci, j, k, t} \]  

(2)

where \( v_{ci, TOTAL, k, t} \) = vehicle count at link \( i \), direction \( k \), time class \( t \) summed over \( j \) speed classes,

\( i = \) link number from 1 to 131,712,
\( j = \) speed class from 1 to 10,
\( k = \) direction 1 or 2,
\( t = \) time class from 1 to 5.

If \( v_{ci, TOTAL, k, t} \geq 20 \), then the link \( i \) was used for analysis. The vehicle count, 20, was chosen because those links, where \( v_{ci, TOTAL, k, t} < 20 \), had a distribution in which there was a large count in the first speed class, less than 24 km/h (15 mph), and in the last one, greater than 88 km/h (55 mph). This is not an accurate reflection of what is occurring on the roads as the majority of the trucks should not be traveling at these two extremes. When \( v_{ci, TOTAL, k, t} \geq 20 \), the distribution on each link looks similar with the majority of the vehicle counts falling in the speed class 8, 9, or 10 as shown in Figure 5.

The travel time, \( tt \), along a link occurs 75% of the time if the 25-percentile speed is chosen to derive \( tt \). The 25-percentile gave the best tradeoff between finding a percentile small enough that the speed occurs with the greatest probability and ensuring that the probe vehicles are traveling at a sufficiently fast speed (64 km/h to 88 km/h or 40 mph to 55 mph). Equation 3 shows the calculation of the 25-percentile speed:

\[ c_{i, p} = p \times v_{ci, TOTAL, k, t} \]

\[ sc_{i, p} = \Phi(c_{i, p}) \]  

(3)

where

\( c_{i, p} = \) vehicle count at link \( i \) at \( p \) percentile,
\( sc_{i, p} = \) speed class for link \( i \) where the \( p \) percentile lies in the cumulative distribution function \( \Phi(c) \).

In Figure 6, the width on the bands indicates the vehicle counts on the links. The band on one side of the link is traffic going in one direction and the band on the other side is traffic going in the other direction. In looking at the map, imbalances reflect the traffic congestion going into the New York City during the AM peak and leaving from the New York City during the PM peak. For instance, the slow speed of the AM peak is reflected in the region of commuters coming in from New Jersey to New York City. Many commuters from New Jersey are traveling into the city via Route 1 and New Jersey Turnpike. The speeds along those roads are between 32 to 40 km/h (20–25 mph). It is also found that during the midday and evening hours, the overall speed increases. In the weekend time class, the volume drops. This represents the lack of truck traveling through the cities during the weekend.

### LOOP DETECTOR DATA AND PROBE VEHICLE DATA

Loop detectors are invasive as they are buried under the pavement. Loop detectors possess the capability to gather comprehensive traffic data. This includes vehicle speed, occupancy, length, axle weight, and vehicle category (Kim et al., 1998). Some types of loop detectors only measure the vehicle counts and lane occupancy. Consequently, spot speeds can only be extrapolated from these parameters. Disadvantage of loop detectors include their high cost and the inconvenience involved in installing and maintaining loop detectors. Loop detector data
was obtained from Phoenix, Arizona, and Seattle, Washington for this study. The spot speeds were compared to the speed derived from the trip speed assignment technique for the same time and same location. In order to compare the two data sets, it was necessary to sum the loop detector data over the different days of the week by the 5 time classes. The reasoning for this is that the truck data was divided into different times of the day, but not into the different days of the week.

The comparison involved looking at the truck volume, total volume, and the average speed from the loop detector data set. By looking at the volumes, the sampling rate for the probe vehicles can be calculated. Depending on the location and time of day, the sampling rate of the probe vehicles ranged from 0.35% to 5.20% among the truck volume.

**Arizona**

The loop detector from Arizona covers the area of the greater Phoenix area, including interstates I-10 and I-17, state roads SR143 and SR51, and loop L202. The loop detector in each lane measures travel speed and volume on both directions of the road. On average, the loop detectors were placed 0.3 miles apart (http://www.azfms.com/Travel/freeway.html). Three loop detectors were chosen from all the installed loop detectors in the greater Phoenix area along I-10 and I-17. The average speeds measured by loop detectors remain fairly constant, between 80 km/h (50 mph) to 96 km/h (60 mph). In contrast, the speed of the probe vehicles covers the entire range of speed classes. The average speeds in the two data sets do not match up very well. In the loop detector data, slight dips are found during the AM peak and PM peak, but they never last very long, while these dips last an hour at the longest in the probe truck data. The reason might be that probe data are only differentiated by time classes, which last more than an hour long. Thus, in order to capture these short dips, the truck data may have to be aggregated differently. If time class 1, from 6 AM to 10 AM, were further differentiated, it is possible that the dips caused by morning congestion would be better captured. Presently, the aggregation of the data is still capturing travel behavior characteristic of other speed classes. For instance, time classes 1 and 3 are the times in which commuters are on the road. The band representing 50 to 55 mph is much smaller relative to the other bands on the link, which means trucks, usually slower-moving vehicles than passenger cars, did not move that quickly near the CBD during rush hours.

The aggregation issue also applies to the loop detector data. The loop detector data was originally aggregated every 15 minutes. The dips in the speed are significant but they are usually not very big. If the aggregation was over a smaller time interval, the dips may be more significant as the time interval would capture only the time when traffic slowed down and not the times when traffic speeds back up again.

### Table 1

**Loop detector and probe vehicle comparison for one site in Seattle**

<table>
<thead>
<tr>
<th>Location</th>
<th>Time Class</th>
<th>Speed (km/h)</th>
<th>Sampling rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Loop detector</td>
<td>Probe vehicle</td>
</tr>
<tr>
<td>Northgate (SB)</td>
<td>6am–10am</td>
<td>95.36</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>10am–3pm</td>
<td>92.48</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>3pm–7pm</td>
<td>73.92</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>7pm–6am</td>
<td>95.04</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Weekend</td>
<td>95.52</td>
<td>24</td>
</tr>
<tr>
<td>Northgate (NB)</td>
<td>6am–10am</td>
<td>90.88</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>10am–3pm</td>
<td>91.52</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>3pm–7pm</td>
<td>84.32</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>7pm–6am</td>
<td>96.64</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>Weekend</td>
<td>95.52</td>
<td>32</td>
</tr>
</tbody>
</table>

**Washington**

The Washington loop detector data is from the Puget Sound area and it covers the time from January 1 to March 31, 1998 (http://www.wsdot.wa.gov/PugetSoundTraffic/). The loop detector does not measure every vehicle's speed during times of heavy congestion. One link, 66224, in the Puget Sound area was chosen for the comparative study between the truck data and the loop detector data. This particular link is know for heavy congestion and was initially chosen for the purpose of studying the travel patterns. In the truck data, the vehicle count was lacking. The mean vehicle count in both directions and all time classes was 14. This is less than the 20 vehicle count minimum criteria for analysis. A possible reason for this is that trucks may avoid routes containing this link so that they are not caught in traffic. Thus, while information can still be derived from the truck data set for this link, there might be a potentially significant bias.

In looking at Figure 7, the direction for the bands on the left side of the link is towards the CBD. In the first time class, the large purple band along with the other blue, green, yellow, and orange bands on the left side of the link represent the slow traffic moving into the CBD during the AM peak at a rate of 16 to 56 km/h (10 to 35 mph). The speed increases with the larger yellow and green bands in time class 2, which means traffic moves faster after the morning commuters arrive at work. The PM peak is shown by the slow speeds as represented by the large purple bands on the right side of the link. People are moving away from the CBD as they are leaving work. In time class 4, there is a significant amount of vehicle probes moving at higher speeds of 72 to 80 km/h (45 to 50 mph). The purple band on the right side of the link is “bad data.” There is no reason for that many vehicles to be traveling less than 56 km/h (35 mph) overnight. Truckers are not likely to be in continuous movement overnight, hence, the derived average speed is slow. Finally, during the weekend, there are vehicles traveling between 64 to 88 km/h (40 to 55 mph). There is also much data below 48 km/h (30 mph). This could be due to people traveling into the city during the weekend for recreational purposes or this could...
be more “bad data” as there are some truckers that stay at home for part of the weekend.

Table 1 shows that the 25-percentile average speed is much lower than the spot speed, but as mentioned above, the lack of vehicle counts might result in the inconsistency. Also, this is just the speed at the 25 percentile. At a large percentile, the average speed will increase.

CONCLUSIONS AND FUTURE RESEARCH

The trip speed assignment technique has demonstrated that it can generate average speeds for a large network using a relatively small sample size. These average speeds follow general traffic patterns as mentioned, such as, most interstate highway routes in the U.S. are traveled at high speeds; drivers travel more slowly in major cities; average speed decreases during morning peak hours and evening peak hours; and the amount of traffic traveling in each direction of the road is dependent on the time of day. Thus it is possible to use the probe vehicle data with the Trip Speed Assignment Technique to provide appropriate travel speed information. However, when compared to loop detector data (that provides average speeds), the probe vehicle data are not readily comparable. This is due to a combination of factors, including the issue of aggregation for both data sets. For example, time class 1 will not only capture morning congestion but possibly part of the post–rush hour traffic. Furthermore, the loop detectors aggregate information by 5-minute (Seattle) and 15-minute intervals (Phoenix), requiring this loop detector data set be further aggregated in order to be compared to the truck data set. As a result, the drop in average speed due to congestion as measured by the loop detectors is often “lost” due to the otherwise fairly constant average speed. Further studies can be done to see what the appropriate time interval for aggregation should be. While a smaller time interval would be needed for more accurate results, a tradeoff does exist between accuracy and cost. While this is computationally feasible, it may not be fast enough to be beneficial in real time.

As the analysis from three months of Seattle data has shown, weekly seasonal effects do indeed exist in traffic patterns. In addition to differentiating the data by speed and time of day, the data should also be split into the days of the week. Present analysis does take the day of the week into account by splitting the information between weekend and weekdays. However, some of the preliminary analysis from Phoenix suggests that there may be a slight peak on Fridays as drivers are hurrying about to finish up their deliveries before the weekend drive home. Only by further differentiating the data can this effect be captured.

Finally, different criteria should be tested for cleaning up the data set. Lower bounds should be placed for the time between consecutive readings so as to exclude drivers who have stopped for long periods of time. However, a tradeoff exists in that slow speeds caused by congestion still need to be captured. Thus, while the trip speed assignment technique seems to produce results that reflect reality, the method can be made more accurate through further knowledge of travel behavior. Also with the

Figure 7 Speed distribution of Link 66224. (The bandwidth is proportional to the flow volume.)
availability of probe data from passenger cars, more accurate link speed distributions can be achieved using the proposed trip speed assignment technique.

REFERENCES


