FINAL
CONTRACT REPORT
VTRC 07-CR12

PROBE SAMPLING STRATEGIES FOR TRAFFIC MONITORING SYSTEMS BASED ON WIRELESS LOCATION TECHNOLOGY

MICHAEL D. FONTAINE, P.E., Ph.D.
Assistant Professor
Old Dominion University

ANJANI P. YAKKALA
Graduate Research Assistant
University of Virginia

BRIAN L. SMITH, Ph.D.
Faculty Research Scientist
and
Associate Professor of Civil Engineering
University of Virginia

Virginia Transportation Research Council

http://www.virginiadot.org/vtrc/main/online_reports/pdf/07-cr12.pdf

Standard Title Page - Report on Federally Funded Project

1. Report No.
FHWA/VTRC 07-CR12

2. Government Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle
Probe Sampling Strategies for Traffic Monitoring Systems Based on Wireless Location Technology

5. Report Date
January 2007

6. Performing Organization Code

7. Author(s)
Michael D. Fontaine, Anjani P. Yakkala, and Brian L. Smith

8. Performing Organization Report No.
VTRC 07-CR12

9. Performing Organization and Address

Virginia Transportation Research Council
530 Edgemont Road
Charlottesville, VA 22903

10. Work Unit No. (TRAIS)

11. Contract or Grant No.
76464

12. Sponsoring Agencies' Name and Address

Virginia Department of Transportation
1401 E. Broad Street
Richmond, VA 23219

13. Type of Report and Period Covered
Final
March 2005 to December 2006

14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

Transportation agencies have become very interested in traffic monitoring systems based on wireless location technology (WLT) since they offer the potential of collecting travel time data across a wide portion of the road system. Prior tests of WLT-based systems have been unsuccessful, in part because they have treated the road network as a homogeneous entity. This "area-wide" method has inherent limitations, causing congested roadways to be over sampled and uncongested and low volume roads to be under sampled. This project developed a methodology to estimate sampling parameters based on localized traffic conditions in the network, termed a "zonal approach." In zonal WLT systems, the roadway network is disaggregated into smaller areas, termed "zones," based on cellular coverage areas.

In this research, two zonal sampling strategies were examined and tested using three simulated networks. When the road network is complex, the zonal priority sampling strategy was found to distribute probes throughout the network and produced a larger number of speed estimates on uncongested and low volume roads. Moreover, the zonal priority strategy improved speed estimation accuracy by 10 percent over the other two sampling strategies. For networks with simple geometry or uniform congested traffic conditions, there were no significant differences among the sampling strategies. The results of this research indicate that the homogeneous approach used by earlier deployments has limitations, and results could be potentially improved by tailoring sampling parameters to a more localized level.

17 Key Words
Wireless Location Technology, Traffic Monitoring Systems, Probe based Traffic Monitoring, Sampling Strategies

18. Distribution Statement
No restrictions. This document is available to the public through NTIS, Springfield, VA 22161.

19. Security Classif. (of this report)
Unclassified

20. Security Classif. (of this page)
Unclassified

21. No. of Pages
57

22. Price

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

FINAL CONTRACT REPORT

PROBE SAMPLING STRATEGIES FOR TRAFFIC MONITORING SYSTEMS BASED ON WIRELESS LOCATION TECHNOLOGY

Michael D. Fontaine, P.E., Ph.D.
Assistant Professor
Old Dominion University

Anjani P. Yakkala
Graduate Research Assistant
University of Virginia

Brian L. Smith, Ph.D.
Faculty Research Scientist
and
Associate Professor of Civil Engineering
University of Virginia

Project Manager
Catherine C. McGhee, P.E., Virginia Transportation Research Council

Contract Research Sponsored by
the Virginia Transportation Research Council

Virginia Transportation Research Council
(A partnership of the Virginia Department of Transportation
and the University of Virginia since 1948)

In Cooperation with the U.S. Department of Transportation
Federal Highway Administration

Charlottesville, Virginia

January 2007
VTRC 07-CR12

ABSTRACT

Transportation agencies have become very interested in traffic monitoring systems based on wireless location technology (WLT) since they offer the potential of collecting travel time data across a wide portion of the road system. Prior tests of WLT-based systems have been unsuccessful, in part because they have treated the road network as a homogeneous entity. This “area-wide” method has inherent limitations, causing congested roadways to be over sampled and uncongested and low volume roads to be under sampled. This project developed a methodology to estimate sampling parameters based on localized traffic conditions in the network, termed a “zonal approach.” In zonal WLT systems, the roadway network is disaggregated into smaller areas, termed “zones,” based on cellular coverage areas.

In this research, two zonal sampling strategies were examined and tested using three simulated networks. When the road network is complex, the zonal priority sampling strategy was found to distribute probes throughout the network and produced a larger number of speed estimates on uncongested and low volume roads. Moreover, the zonal priority strategy improved speed estimation accuracy by 10 percent over the other two sampling strategies. For networks with simple geometry or uniform congested traffic conditions, there were no significant differences among the sampling strategies. The results of this research indicate that the homogeneous approach used by earlier deployments has limitations, and results could be potentially improved by tailoring sampling parameters to a more localized level.

FINAL CONTRACT REPORT

PROBE SAMPLING STRATEGIES FOR TRAFFIC MONITORING SYSTEMS BASED ON WIRELESS LOCATION TECHNOLOGY

Michael D. Fontaine, P.E., Ph.D.
Assistant Professor
Old Dominion University

Anjani P. Yakkala
Graduate Research Assistant
University of Virginia

Brian L. Smith, Ph.D.
Faculty Research Scientist
and
Associate Professor of Civil Engineering
University of Virginia

INTRODUCTION

Transportation agencies are becoming increasingly focused on acquiring high-quality traffic condition data. These data are used to support a wide variety of functions, including performance measurement, real time control, and traveler information. The current state of the practice in the Virginia Department of Transportation (VDOT) is to collect traffic data using a network of point detectors, usually inductive loops detectors (ILDs). Point detectors are effective at collecting speeds, volumes, and occupancies of vehicles at discrete points, but it is often difficult to generalize that data into link-based measures such as travel time. Furthermore, only major urban freeway facilities use closely spaced sensor networks due to the installation and maintenance costs for point detectors. As system operations continue to develop as a priority for VDOT, the department is becoming increasingly interested in innovative methods to collect travel time data from a broad portion of the roadway system.

Probe-based traffic monitoring systems offer a potential approach to generating travel time information. These systems track probe vehicles as they traverse the roadway system, allowing for direct collection of point-to-point travel time and speed information. In probe vehicle systems, a representative sample of vehicles is monitored and their characteristics are used to estimate overall traffic conditions on a road. Various technologies have been used to support probe-based monitoring systems, including automatic vehicle identification systems (based on electronic toll collection systems),[1] automatic vehicle location (AVL) systems (tracking transit buses or other vehicles equipped with AVL technology),[2] wireless location technology (WLT) (anonymously tracking mobile devices such as cellular phones in vehicles),[3] and future vehicle infrastructure integration (VII) applications that use on-board vehicle equipment as data collectors.[4]

WLT-based traffic monitoring systems are a form of probe-based monitoring that has generated a significant amount of interest in the transportation community. WLT-based traffic monitoring systems approximate the average travel speed and travel time on a road by tracking a series of estimated positions for wireless devices (such as cellular phones) that are located in vehicles. Speeds for these devices can then be derived by looking at a series of position locations over time. With these systems, a private vendor works with a wireless communications provider to generate data that are then sold to a DOT. The traffic data generated become a purchased service, as opposed to a traditional sensor network that the DOT must operate and maintain. Using the wireless infrastructure also allows the monitored network to be more expansive than with point detectors, since, theoretically any road with cellular coverage could be monitored.

Although this technology is conceptually appealing, field deployments were unsuccessful in generating accurate estimates of traffic flow conditions. Field deployments showed that wireless devices can be located with reasonable accuracy, but high quality traffic data could not be generated from the entire network.[5,6,7] Research by the Virginia Transportation Research Council (VTRC) has shown that the fundamental design of WLT-based monitoring systems plays a critical role in determining the quality of the results generated by the system.[3] While the VTRC research demonstrated the feasibility of the WLT-based systems, it also found some major shortcomings in the sampling approaches used in deployments that contributed to the poor quality of traffic data they produced.

Both the VTRC research and early generation field deployments revealed that more vehicles were sampled from congested areas, and fewer vehicles from uncongested and low volume roads. Part of the problem was that early generation WLT systems treated the entire network as a homogenous entity. In other words, a high volume, congested freeway was sampled the same way as a low volume, free-flowing arterial street. The results from these tests showed that specifying sampling parameters on an “area-wide” basis was not an efficient or effective way to operate a WLT-based monitoring system. As a result, the basic design of those early generation systems did not appear to be adequate to support traffic monitoring across a broad cross section of roads.

Ideally, a WLT system would collect a specific number of samples from a particular roadway segment. In practice, this was not accomplished in field deployments. In those systems, vehicles were randomly selected as probes from a network. As a result, vehicles from congested roadways were proportionately more likely to be selected than vehicles traveling on low volume or uncongested roadways. An alternative approach would be to divide the network into smaller areas, termed “zones,” where traffic and geometric conditions would be similar. Sampling parameters could then be tailored to those smaller zones, hopefully ensuring that probes will be available on lower volume roads while at the same time creating a more efficient sampling structure. Zones with congested roadways may need relatively few samples since traffic conditions are relatively uniform, while a zone with high-speed freeways may require more samples to capture the varying traffic conditions.

In cellular telephone systems, the communications network is broken into a number of zones (commonly referred to as "cells") that allow for calls to be handled locally and support system capacity expansion through the addition of new cells. These cells could be utilized as zones for WLT-based monitoring. Figure 1 shows a hypothetical zonal system for WLT-based monitoring, where the zones in Figure 1 have been defined based on wireless telephone cells. In this case, the structure used to support voice communications and traffic monitoring would be identical. By dividing the network into zones, sampling parameters can be defined based on the observed traffic and geometric conditions within a specific zone.

The zonal sampling concept is best illustrated through an example. Figure 1 represents the Springfield area in Northern Virginia, and it is broken into 13 wireless coverage areas (cells) with different roadway characteristics. Cells A, C, and G primarily contain high volume freeways like I-95, while cells I, J, and L mainly contain signalized arterials. Each cell would be assigned different probe sampling parameters based on their characteristics. The number of samples and time between samples would be driven by the total volume of traffic in the cell, the variance of observed speeds, and the geometric complexity of the network. By creating localized sampling parameters, it should be possible to ensure sufficient probe sampling throughout the network while simultaneously operating the system as efficiently as possible.

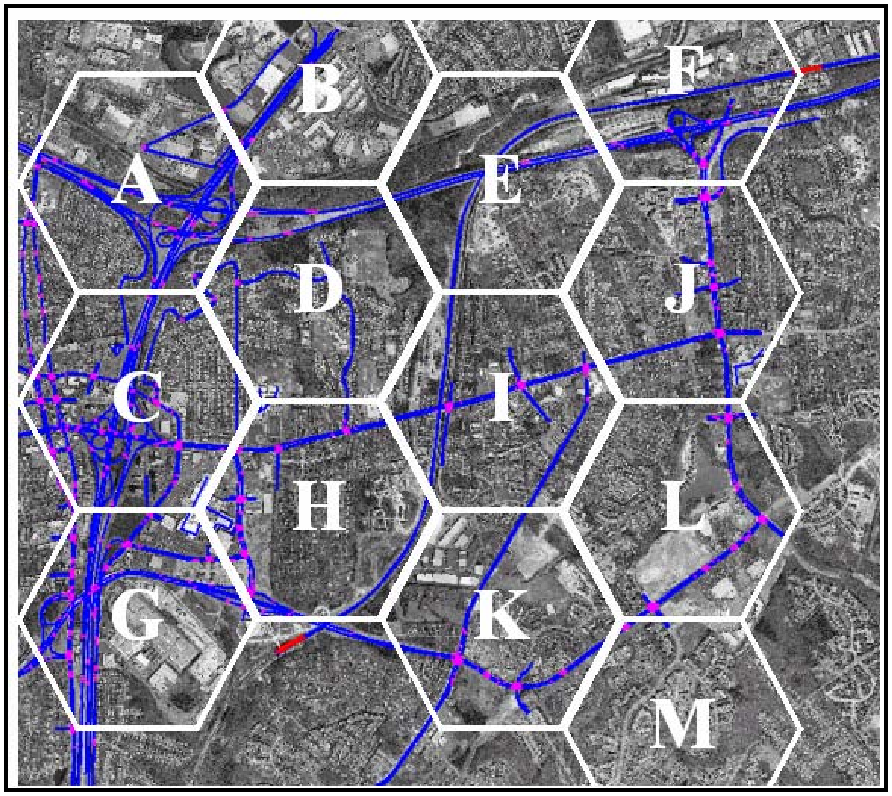

Figure 1. Example of a Zonal WLT System

PURPOSE AND SCOPE

While WLT-based monitoring has generated significant interest within VDOT and other transportation agencies, no field deployments in the United States have been able to generate traffic condition data on a production basis that could support operational or performance measurement activities. Research has shown that the area-wide sampling of the network has inherent flaws. The purpose of this project was to investigate whether the zonal sampling approach could improve the accuracy and availability of traffic data obtained using WLT-based traffic monitoring systems. Specifically, the objectives of this project were:

- Develop a methodology to estimate sampling parameters based on traffic conditions observed within a zone.
- Test and evaluate a zonal WLT-based monitoring system based on cellular coverage areas.
- Compare the effectiveness of zonal and area-wide sampling strategies
- Assess several factors that may influence the effectiveness of zonal systems.

This research considered WLT-based traffic monitoring systems, where the latitude and longitude coordinates of specific cellular phones are produced. The research does not explicitly examine WLT systems that track handoffs between cellular towers, which is the technology being used in a number of recent deployments. The technology considered was able to track the cellular phone even if it is not in use, but is switched on. The project was limited to a simulation evaluation, and no field deployment of WLT-based traffic monitoring was conducted. The simulation was used to emulate the operation of a generic WLT system, and it is possible that the approaches used by individual vendors could vary from what was tested in this research.

While the focus of this research is on WLT-based systems, it is expected that many of the findings could also be applied to AVL-based probe systems that collect location data through global positioning systems (GPS). Systems that use electronic toll tag readers would not be able to directly use these research results, given that data are collected at known, fixed points on the system. That approach differs fundamentally from that used in WLT-based systems.

METHODS

This research consisted of three major tasks. First, the relevant literature was reviewed to lean more about tests of WLT-based systems. Second, a simulation test bed was developed to evaluate zonal systems. Third, the performance of zonal systems was evaluated using the simulation test bed. Each task is discussed in detail in the following sections.

Literature Review

The first task in this research was to review relevant literature on WLT-based traffic monitoring. The VDOT Research Library, the University of Virginia library, and various on-line databases were consulted to identify studies on the following topics:

- Sampling in probe vehicle systems
- Field deployments of WLT-based systems
- Simulation studies of WLT-based systems.

These studies were reviewed and summarized to identify trends in performance of WLT-based systems.

Simulation Test Bed Development

Earlier research by VTRC developed a simulation test bed that could be used to evaluate area-wide approaches to WLT-based monitoring.[3] That simulation had to be modified in order to simulate zonal approaches to WLT sampling. The base simulation involved two major processes, as shown in Figure 2:

1. Traffic simulation of real world networks using VISSIM
2. Simulation of WLT operations using a custom model.

In the first process, the test bed modeled actual traffic flow by simulating roadway networks built using the microscopic simulation model VISSIM. The models were constructed using data obtained from VDOT, and link speeds were calculated based on the travel times produced by the simulation. These "true" link speeds were used as a benchmark to determine error in speeds generated using different sampling strategies.[3]

Next, the operation of a WLT-based system was emulated. VISSIM can record the X-Y coordinate information of every vehicle on the network every second. The location information obtained from actual WLT systems inherently contains error, so the perfectly accurate position coordinates from VISSIM were degraded in the WLT emulation. The position degradation was based on an assumed normal distribution of error with a mean of 0 and standard deviation of 10 m (analogous to errors from GPS). Likewise, an actual system is unlikely to generate data for every vehicle on the network every second, so the number of vehicles tracked and the time that elapsed between position readings was also degraded to more accurately replicate the true operation of a WLT system. This degradation of the frequency of location readings was based on roadway characteristics, and is discussed more fully later in the report. After all degradation was complete, the program generated data that was similar to what would be produced by a WLT-based system, consisting of the X and Y coordinates of the vehicle, the time of the position reading, and a unique identifier for each vehicle.

These degraded positions then had to be matched back to the roadway network. Prior research found that the multiple hypothesis technique (MHT) map matching algorithm produced significantly more estimated positions than other methods for large and complex networks.[3] As a result, the MHT map matching algorithm was used to match the degraded position coordinates back to the roadway network. The MHT method maintains a list of all potentially feasible paths for a vehicle, and re-evaluates the likelihood that a potential path is correct as each new position is added.[8] When a convergence criterion was reached, then the estimated positions of a vehicle were assigned to that path. More detailed information on this algorithm can be found in the VTRC report by Fontaine and Smith.[3]

The new vehicle positions that were matched back to the roadway using MHT were then used to estimate vehicle speeds, which in turn were used to create link speeds. These new link speed estimates were then compared to the actual speeds of the roadway simulated in VISSIM to determine how well the simulated WLT-based system was able to estimate actual travel conditions.

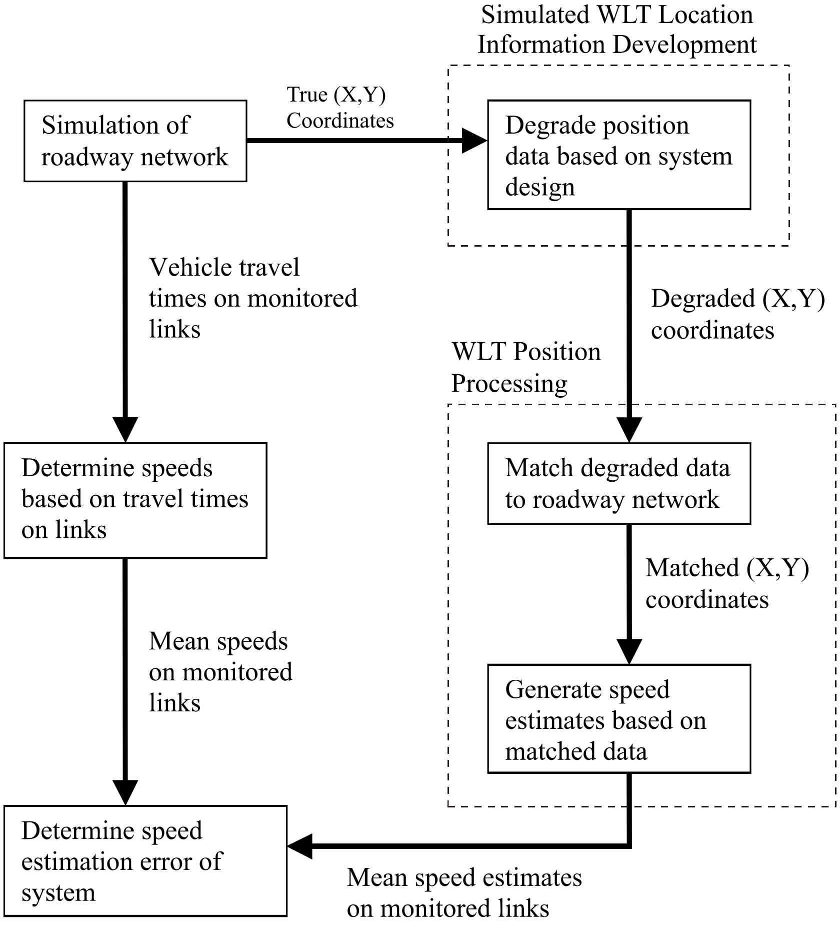

Figure 2. Simulation Test Bed Operation

Modifications to Test Bed to Accommodate New Sampling Strategies

The test bed described in the previous section was set up to perform area-wide sampling, and modifications were needed to allow a zonal system to be simulated. Three sampling strategies were evaluated in this research:

1. Zonal sampling
2. Zonal priority sampling
3. Area-wide sampling.

Before discussing these strategies, it is helpful to define two parameters that were used to determine how vehicles were sampled from the roadway network for all three cases:

1. *Time between position readings (frequency, F, in seconds):* This parameter defines the time that elapses between consecutive location estimates of an individual vehicle. The average travel time on a specific link was examined to determine the appropriate time between position readings in a zone. The time between readings was set at a level where at least 3 position readings were available for each vehicle on each link in a zone. This value was fixed for the entire analysis period.

2. *Number of vehicles tracked (sample size, N):* This parameter sets the number of vehicles that were tracked simultaneously in the system. It was determined based on the standard deviation of speeds on the link and the probability that the vehicle sampled was located on the link, based on random sampling.

These two parameters were set for all sampling strategies tested. In area-wide systems, these two variables were set globally across the network. In the proposed zonal systems, these two variables were set on a zone-by-zone basis. The methods used to determine the values of these two factors and necessary modifications to the test bed are detailed in the following sections for each strategy. The basic simulation process shown in Figure 2 remains the same for all methods, but the way that the cellular network was treated varies.

Zonal Sampling

As noted earlier, treating a large roadway network in a uniform manner has the potential to create problems. The different parts of the region being monitored would likely exhibit a great deal of variation in traffic flow and geometry. In zonal sampling, sampling parameters were set at the level of an individual wireless cell. This allows the sampling characteristics to be varied based on the specific characteristic of the roads within the cell. This system could potentially help ensure that there was adequate probe availability in low volume or uncongested areas, thereby improving the quality of speed estimates from lower volume facilities.

The size of a wireless cell, and thus the WLT zone, varies based on a number of factors related to the communications design of the network. Size generally relates to the expected level of voice traffic that is expected in a zone, so zones will be smaller and more densely spaced in urban areas with high voice traffic and much larger and less densely spaced in rural areas. Proprietary data obtained from a wireless provider shows that the mean cell size varies from approximately 1 square mile in urban areas to 2.7 square miles in suburban areas.[3] Since WLT-based systems are likely to be implemented primarily in urban and suburban areas, the zones used in this research were approximately 2 square miles in size.

In the zonal sampling strategy, each zone is treated independently of the others and sampling parameters were set based on conditions observed within that particular zone. Figure 3 shows an example with the cell outlines overlaid on the roadway network. The dark lines represent the roads being monitored. Each cell has a different sample size "N" and frequency "F" value. Each cell is treated independently so vehicles traveling from one cell to another may not be tracked after they cross cell boundaries in the network. For example, if a vehicle is being tracked in Cell A and then continues to travel to Cell B, this vehicle may not necessarily be tracked in Cell B. The "N" values in each cell are kept constant, so that when a vehicle leaves the zone, a new vehicle is added. When a vehicle is added for monitoring, it could be picked up anywhere in the cell, not just at the cell boundaries.

The simulation test bed was changed so that hexagonal zones could be input into the model. For each zone, separate values could be entered for the number of vehicles being tracked and the time between position readings. In effect, this created a number of independent zones that were simulated.

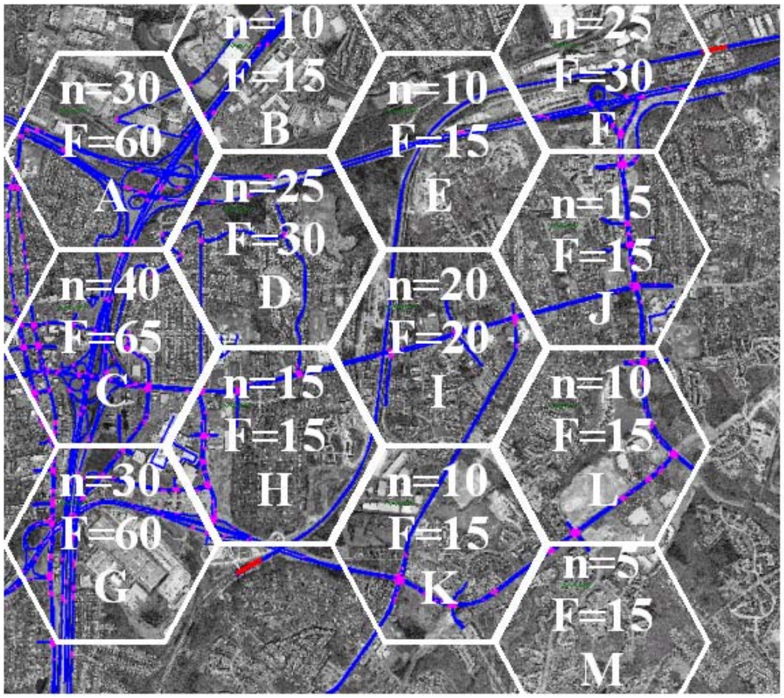

Figure 3. Example of Zonal WLT System with Sampling Parameters

In zonal systems, the time between position readings was determined based on the minimum travel time observed on a link in a zone. In setting the time between position readings, it is desirable to have the vehicle travel as far as possible between readings. This minimizes the contribution of location error to potential errors in speed estimation, helping to improve the accuracy of the speeds generated. To ensure that at least three position readings were obtained on each link, the frequency in a zone was determined by the link with shortest travel time in the zone. This is given by Equation 1, using units of seconds.

$$F_i = \min\left[\frac{TT_{ij}}{3}\right] \qquad \text{(Equation 1)}$$

Where: F_i = Time between position readings for zone i
T_{ij} = Travel time for link j for zone i

A short numerical example can help illustrate the application of Equation 1. A hypothetical zone contains 2 links. Link A is 2 miles long with an average running speed of 40 mph, while link B is 0.5 miles long with an average running speed of 60 mph. This produces and average travel time for link A of (2 miles)(40 mph) = 0.05 hr = 180 sec. Link B's average travel time would be (0.5 miles)(60 mph) = 0.0083 hr = 30 sec. Link B's travel time would control here, so the F_i for the zone would be set to (30 sec)/3 = 10 sec.

The minimum travel time on a link in a zone was calculated from the travel time data obtained from VISSIM. In a real world application, data collection would be required prior to implementing the zonal system to set this parameter.

The minimum number of vehicles to be sampled is dependent, in part, on the underlying speed distribution of the links in the zone. In general, the central limit theorem (CLT) has been shown to hold for estimating mean travel speeds on a link in probe-based monitoring systems.[9] The CLT states that the minimum sample size can be determined if the desired confidence level, speed variance, and allowable error are known. The CLT-based sample size for a specific link in a zone can be determined using Equation 2.

$$n_{ij} = \left(\frac{Z_{\alpha/2}\sigma_{ij}}{d}\right)^2 \qquad \text{(Equation 2)}$$

Where: n_{ij} = CLT based minimum sample size for link j in zone i
$Z_{\alpha/2}$ = Z-value based on normal distribution at significance level (95% confidence level was used in this research, equal to 1.96)
σ_{ij} = Standard deviation in speeds on link j in zone i
d = Allowable error in speed estimates (5 mph was used in this research)

The data collected by VISSIM were used to determine the standard deviation of speeds on a link. For a real world application, this speed distribution would have to be established through data collection prior to proceeding with the design of the sampling parameters.

The n value calculated in equation 2 provides an estimate of the minimum number of samples required to estimate the mean speed for a specific link in a specific zone. To define the total number of samples required in a specific zone, it would have to be inflated further to ensure that the minimum number of samples was collected to cover all roads being monitored in the zone during the analysis interval. Vehicle probes were assumed to be randomly acquired anywhere in a zone. The likelihood of selecting a vehicle on a specific link was directly proportional to the vehicle-minutes of travel on the link versus the total vehicle-minutes of travel in the zone. Equation 3 gives the probability that a newly acquired probe was located on a specific link

$$P_{ij} = \frac{VS_{ij}}{\sum_{j=1}^{L} VS_{ij}} \qquad \text{(Equation 3)}$$

Where: P_{ij} = Probability that a vehicle picked from zone i was from link j
VS_{ij} = Vehicle minutes on link j in zone i
L = Number of links in zone i

Another quantitative example helps to illustrate the application of Equation 3. Consider a hypothetical zone with 2 links. Link A has a total of 3,000 veh-min of travel, while Link B has a total of 10,000 veh-min of travel. Application of Equation 3, provides the following probability, P, for link A:

$$P_{1,A} = \frac{VS_{1,A}}{\sum_{j=1}^{2} VS_{ij}} = \frac{3000}{(3000 + 10000)} = 0.23$$

Applying Equation 3 to link B generates a probability $P_{1,B} = 0.77$.

To ensure that an adequate number of vehicles were tracked on each monitored link in a zone, the sample size for a zone was set as follows:

$$n_i = \max_j \left[\frac{n_{ij}}{P_{ij}}, \sum_{j=1}^{n} n_{ij} \right] \qquad \text{(Equation 4)}$$

Where: n_i = Sample size for zone i
n_{ij} = CLT based sample size for link j in zone i
P_{ij} = Probability that a vehicle picked from zone i was from link j
N = Number of links in zone i

The sample size n_i calculated in equation 4 represents the total number of speed samples required for a zone in a particular analysis period. This sample size must then be translated into the number of vehicles that must be tracked simultaneously based on the time between position readings in a particular zone. Historic data on the MHT map matching algorithm show that only a portion of the location samples could be correctly matched to the network. Some location samples could not be matched due to ambiguity in the true location of the probe, location samples located off of the roadway network being monitored, or a lack of two matched samples for a particular probe that could be translated into speeds. As a result, empirically derived factors from prior VTRC research were used to further inflate the sample sizes based on the expected efficiency of matching.[3] Equation 5 was used to produce the number of vehicles that should be tracked simultaneously in each zone to ensure the minimum number of samples on each monitored link.

$$V_i = \frac{n_i}{\frac{60}{F_i} \times T \times P} \qquad \text{(Equation 5)}$$

Where: V_i = Number of vehicles to track simultaneously in zone i
n_i = Sample size for zone i from equation 4
T = Analysis period (minutes)
F_i = Frequency for zone i (seconds)
P = Percentage of samples matched to the network in decimals.

Another numerical example can help illustrate the application of Equation 5. Consider a zone where the sample size, n, is 100 vehicles being tracked and the calculated frequency, F, is 10 seconds. A 5-minute analysis period is desired, and calibration has shown that about 70 percent of locations can be correctly matched to the network. When Equation 5 is applied, it generates the following:

$$V_i = \frac{n_i}{\frac{60}{F_i} \times T \times P} = \frac{100}{\frac{60}{10}(5)(0.7)} = 4.76$$

This would be rounded up to 5 vehicles to be tracked simultaneously in the zone to ensure that 100 valid data points are obtained per 5 minutes.

Once again, a period of calibration would be required to set these parameters in a real world system. The total amount of travel on the network, as well as the performance of the map matching algorithms would need to be known prior to setting the number of vehicles to be tracked on the network.

Zonal Priority Sampling

In the zonal system, each zone was treated as an independent entity, and there was no special consideration given to tracking vehicles as they transitioned between zones. An alternative approach, termed zonal priority sampling, was developed whereby cells were not treated independently. If a vehicle tracked in one cell travels to a neighboring cell, then it is given priority to be tracked in the next cell. This should create longer vehicle tracks, potentially improving the quality of position estimates and the resulting speed estimates.

As with the zonal system, the number of vehicles being tracked is continuously monitored throughout the analysis period. If the number of vehicles tracked, N, falls below the specified number, then a new vehicle is added to increase N back to the desired level. Vehicles that have entered the zone from neighboring zones within the last second were given priority to be tracked over the other vehicles in the zone. Figure 4 shows an example of how this works. At time *t*, 30 vehicles are being tracked in Cell A. At time *t+1*, 10 of those 30 vehicles move to Cell C, 5 of those vehicles move to Cell B, and 7 of those vehicles move to Cell D. If the N value in Cell D drops to 20 vehicles, then 5 of the 7 vehicles that just left cell A would be selected to fill the needed samples for Cell D. Rather then randomly selecting replacement vehicles, the program will replace those vehicles with vehicles that recently crossed into the cell from a neighboring cell. The time between readings, F, would be held constant in each cell so if a vehicle is tracked across cell boundaries, the time between position readings would change to match the value in the current cell.

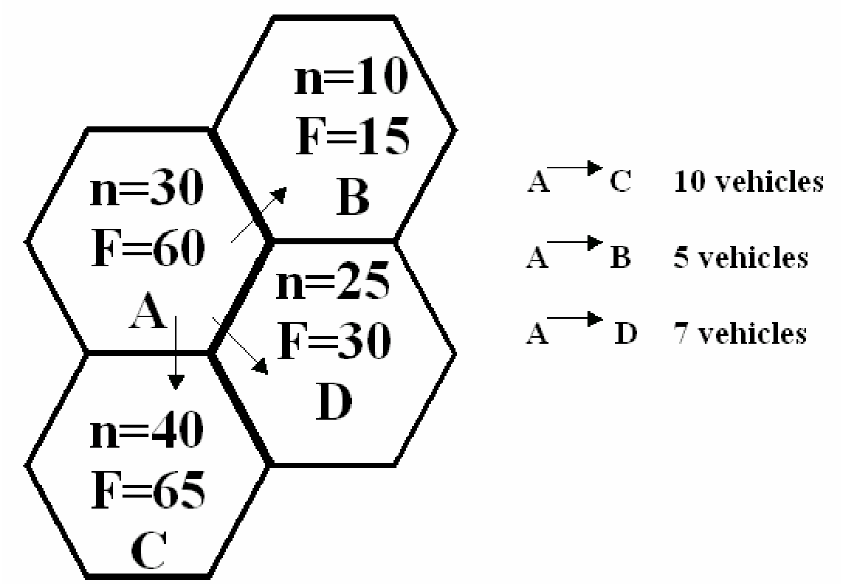

Figure 4. Zonal Priority WLT System Example

The method of probe replacement is different for the zonal priority system than the zonal system, but the method for calculating the time between readings and the number of vehicles to track is the same. This provides consistency in parameter definition, and allows the replacement strategy to be evaluated without confounding the analysis with different sample sizes.

Area-Wide Systems

Area-wide systems represent a base case where the two sampling parameters are set globally. Probe vehicles are randomly selected from anywhere in the network for monitoring. If the number of vehicles being tracked falls below the required value, another vehicle is randomly selected from somewhere on the network.

For the area-wide system, the time between position readings was set equal to the weighted average of the frequency in each zone calculated for the two zonal systems, as shown in equation 6. This provides an aggregate travel time-based estimate for the area-wide system that is comparable to what was calculated for the zonal systems.

$$F = \frac{\sum_{i=1}^{n} F_i \times N_i}{\sum_{i=1}^{n} N_i} \qquad \text{(Equation 6)}$$

Where: F= Frequency for the entire network
F_i = Frequency in zone i
N_i = Sample size for zone i
n = number of zones in the network

The number of vehicles tracked for the area-wide system was set equal to sum of the sample sizes in all the zones in the zonal system, as shown in equation 7. This allowed the relative performance of the area-wide and zonal systems to be compared while tracking the same number of vehicles. In effect, this allows for a direct comparison of how the samples were distributed throughout the network while the number of samples was held constant. Since WLT-based monitoring systems may incur a per-location charge from a wireless provider, the intent was to determine whether the zonal structure could provide a benefit simply through a reallocation of samples over the network. This would keep the potential operations cost consistent between the two approaches.

$$N = \sum_{i=1}^{n} N_i \qquad \text{(Equation 7)}$$

Where: N = Sample size for the entire network
N_i= Sample size for zone i
n = number of zones in the network

Data Processing Techniques

An output file consisting of the estimated speeds of each individual vehicle was produced after the model was run. The final output from the WLT simulation program consisted of the time of the location reading, a unique identifier for each vehicle, position coordinates, and the estimated speed and zone number of the vehicle. This output was processed, and speeds for user defined links were estimated. There were three steps in data processing: elimination of vehicle records with no speed information, data filtering, and link speed estimation. The first two steps minimized errors caused by map matching algorithms, while the last step involves two different approaches in estimating link speeds from vehicle speed data.

Screening of Bad Data Records

The MHT map matching method can successfully match most vehicle positions to the roadway network, but some vehicles cannot be properly matched. If a vehicle cannot be matched to the network or there are not two position readings on the same link, no speed estimate can be calculated. Vehicles with no estimates were deleted from the data file. Likewise, vehicle records with speeds more than 100 mph were also eliminated before link speed estimation. This was done to provide an initial screening of speeds that were likely to be erroneous.

Data Filtering

Roadway networks in the real world are complex with a wide variety of traffic conditions. One specific area where there are known problems in map-matching was when high-speed facilities were located close to parallel, lower-speed facilities. In these cases, a single mismatched point can significantly degrade the quality of speed estimates, as demonstrated in prior research.[3]

A data filtering technique to minimize speed estimation errors due to this problem was applied to the data. If a vehicle had just a single pair of position estimates on a link, then it was not used for estimating link speeds. A thorough analysis of output from the simulation program revealed that many vehicles which were matched to an incorrect road had just a pair of position estimates, so filtering these vehicles would improve speed estimation accuracy. It also ensures that potential vehicles were consistently matched to one route. As a result, a vehicle must have at least three positions recorded on a link before it was used to derive link travel time estimates. The impact of this filtering varied by roadway geometry, with closely spaced, parallel facilities having the most samples that were screened out.

Link Speed Estimation

The last step in the data processing was link speed estimation. There were two potential methods to take the individual speed data and translate them into link speeds. First, all individual speeds contained on a link could be arithmetically averaged together, as in equation 8. This is termed the "average speed method."

$$U_i^A = \frac{\sum_{i=1}^{N} S_i}{N_i} \qquad \text{(Equation 8)}$$

Where: U_i^A = average speed of link i
S_i = individual vehicle speed
N_i = total number of vehicles tracked on link i

The potential disadvantage of this method was that it treats all individual estimates equally. For example, consider a 1-mile long roadway link. If there were two speed estimates, one for a vehicle that travel 0.9 mile at 60 mph and one that traveled 0.1 mile at 30 mph, the average speed method would produce a speed value of 45 mph for the link. An alternative approach was to create a distance-weighted calculation of speed. It would be estimated using equation 9.

$$U_i^S = \frac{S_1 D_1 + S_2 D_2 + \ldots + S_n D_n}{D_1 + D_2 + \ldots + D_n} \qquad \text{(Equation 9)}$$

Where: U_i^S = Distance weighted speed of link i
S_i = Speed of an individual vehicle i
D_i = Total distance of traveled by an individual vehicle on the link i.

Using the data from the prior example, this method would produce an average speed for the link of 57 mph, which obviously differs substantially from the number produced by the average speed method. Speeds were calculated using both of these methods so that their relative merits could be assessed.

Simulation Evaluation of Sampling Strategies

After the test bed modifications were completed, a formal evaluation of the different sampling strategies and key system design alternatives was performed to determine whether the zonal systems could improve the performance of WLT-based monitoring. Three simulation networks that were developed using real-world data were used to examine performance across a wide variety of traffic and geometric conditions. A simulation of the Charlottesville area was used to assess a relatively simple, uncongested network. More complex traffic and geometric conditions were examined using two networks from Northern Virginia: Tysons Corner and Springfield. The specific characteristics of those case studies are discussed in more detail in the Results section.

Factors Evaluated

The case studies were used to examine four different factors that had the potential to impact the performance of a WLT-based traffic monitoring system. The factors investigated included:

- Sampling strategy
- Map matching error
- Method of determining link speeds
- Zone number.

Sampling Strategy

The three sampling strategies discussed earlier were examined. The sampling strategy was the primary factor being investigated, and it played a vital role in determining the sample size and time between consecutive vehicle position readings.

Map Matching Error

The MHT process relies upon knowledge of the estimated location error to make intelligent matches to the roadway network. One potential problem with the MHT method is that it matches to the centerline of the roadway. The X-Y coordinates for a vehicle would be different on a four lane highway depending on what lane a vehicle was traveling in, and matching to the centerline could cause a disproportionate number of vehicles in the outer lanes of multilane highways to be wrongly discarded.

To test the influence of this factor, the assumed position error used by the MHT method was studied at two levels, 10 m and 16 m. The 10-m value conforms to the default estimate of position error used in the location degradation process. The 16-m level would include all potential vehicles on a 4-lane directional multilane highway. The potential disadvantage of the larger assumed error was that it could also increase the ambiguity in the map matching process, potentially resulting in more erroneous matches, particularly at intersections.

Speed Method

The two types of speed estimation methods discussed earlier were evaluated. The average speed method computes link speeds by treating each separate pair of position estimates as a single speed observation. Speeds that were determined over long distances were weighted the same as speeds generated over shorter distances. In the case of the distance-weighted method, link speeds were estimated as the weighted average of individual vehicle speeds based on the distance traveled on a link.

Zone Identification Number

In order to account for the impact of network geometry and traffic conditions on speed estimates, the specific zone identification number was also considered as a factor. The zone identifier was used to isolate the results obtained in a specific zone. By explicitly examining the data on a zone-by-zone basis, it was possible to isolate the impact of localized geometric and traffic characteristics.

Measures of Effectiveness

Three measures were used to analyze the impact of different sampling strategies, system parameters, and data processing techniques. All of these measures were calculated using 5-minute aggregation intervals. Each simulation was performed over a simulated 1-hour period where traffic demands did not change. The measures used were:

- Number of speed estimates
- Speed estimation accuracy
- Distribution of speed errors.

Number of Speed Estimates (Sample Size)

The number of speed estimates was defined as the number of vehicle speed observations used to derive link speeds during a five minute interval. Many early generation WLT-based field deployments were not able to generate traffic speeds due to lack of sufficient samples from low volume roads. As a result, the number of speed estimates generated for each link during an analysis period was considered as a primary criterion for evaluating the effectiveness of sampling strategies and data processing methods. This criterion summarizes both the sufficiency of sample sizes and probe availability.

Speed Estimation Accuracy (Absolute Speed Error)

Speed estimation accuracy was defined as the absolute value of the difference between the WLT estimated link speeds and true link speeds generated in VISSIM. Since the WLT method uses a small sample of total traffic to estimate speeds, it in necessary to check whether the sample was truly representative of actual traffic flow conditions.

Distribution of Absolute Speed Errors

The distribution of the absolute value of speed errors was also examined. Each interval on each link was examined to determine whether it was within 2 mph, 5 mph, and 10 mph of the actual speeds on the link. The percentage of miles that were monitored within each of these accuracy thresholds was then summarized to provide an indication of the distribution of speed errors across the network.

Experimental Design and Analysis

A full factorial design was used to investigate the differences between the impacts of the three major system design factors listed earlier.

Table 1 lists the combinations of factors studied for the three simulated networks.

Table 1. Evaluation Cases for Each Case Study

Sampling Strategy	Map Matching Error	Speed Estimation Method
Zonal Priority	10 m	Average Speed
		Distance Weighted Speed
	16 m	Average Speed
		Distance Weighted Speed
Zonal	10 m	Average Speed
		Distance Weighted Speed
	16 m	Average Speed
		Distance Weighted Speed
Area-wide	10 m	Average Speed
		Distance Weighted Speed
	16 m	Average Speed
		Distance Weighted Speed

The number of speed estimates and the speed estimation accuracy obtained for each case study was analyzed using the generalized linear model (GLM) procedure in the statistical analysis program SPSS. The GLM allows the main effects and interaction effects to be studied in detail. The least significant difference (LSD) post hoc test was used for multiple pairwise comparisons among factor levels. One-way analysis of variance (ANOVA) was used for pairwise comparison if the factor only had two levels. An $\alpha = 0.05$ was used for all statistical tests. It should be noted that while the speed estimation method was included in the analysis of the number of speed estimates, it was only expected to influence the quality of the speed estimates, not their availability.

RESULTS

Literature Review

Researchers have conducted field deployments and simulation studies to study the feasibility and reliability of speed estimates produced by WLT systems. These studies were briefly reviewed to provide additional insight into how WLT-based traffic monitoring has been examined.

Field Deployments

In the past decade there have been a number of deployments of WLT-based monitoring systems in the United States and abroad. The deployments have covered regions with various traffic conditions and technology platforms.[10] While there have been a number of deployments of these systems, none has been successful in creating high-quality traffic condition data. Major deployments that occurred in the United States are reviewed in this section.

Northern Virginia Field Deployment

The first major operational test of wireless location technology was conducted over a 27-month period starting in 1994 on I-66, I-495, and various state routes in Virginia.[6] This project was named CAPITAL (Cellular APplied to ITS Tracking And Location), and was the result of a cooperative agreement between the Federal Highway Administration (FHWA), VDOT, the Maryland State Highway Administration (MDSHA), Raytheon E-Systems, Farradyne Systems Inc., and Bell Atlantic NYNEX Mobile.[6] Raytheon E-Systems provided the equipment to locate and track cellular phones, while Farradyne Systems provided the traffic management information system to convert cellular location data to traffic data. Bell Atlantic provided the communications network infrastructure. The University of Maryland performed an independent evaluation of the system.

The CAPITAL project used Bell Atlantic NYNEX Mobile's cellular network. Call detection and location equipment were physically located on 8 cellular towers in the area. This equipment was used to gather location data on calls handled at each individual tower. Cellular calls were detected when they were initiated in the test area. The location of the phone was then calculated using the signal's line of bearing and time of arrival as seen by multiple eight-element antennae installed on each of the towers. If a cellular phone was estimated to be on a roadway of interest, multiple measurements were performed to calculate the vehicle's speed.

By the end of testing, wireless telephones could be located within an average of 107 m of their actual position. The accuracy of the position estimates improved considerably as the number of towers providing directional information increased. The evaluators noted that accuracies on the order of 5 to 25 m might be needed to perform accurate speed estimation for a network. Although the location estimates were reasonably accurate, speeds could only be determined for 20 percent of all wireless phones that were located. In order to calculate speed, at least four position estimates had to be identified for each phone, and this occurred only 20 percent of the time. As a result, link speed estimates could not be estimated for the network.

The lack of long vehicle tracks appeared to be caused primarily by a lack of sophisticated methods to match vehicles to the roadway network. While the CAPITAL test showed that wireless phones could provide reasonably accurate position data, it was unsuccessful in producing traffic information that would be useful to departments of transportations (DOTs) or motorists.

US Wireless Corporation Tests

The now defunct US Wireless Corporation was a very active vendor of WLT-based systems in the late 1990s, with deployments in Billings, Montana; San Francisco/Oakland; and the Washington, D.C., metropolitan area. Only the San Francisco/Oakland and Washington, D.C., tests included an independent evaluation, however. The US Wireless system relied on their RadioCamera technology, which used location pattern matching technology to recognize signatures of incoming radio frequency (RF) signals and associate them with the specific locations from which they originated.[11] The relative power, direction of arrival, number of dominant reflections, and multipath phase and amplitude were examined and compared to a reference database to determine the likely location of the cellular device on the transportation network. This approach is fundamentally different from the handoff-based approaches being used in more recent deployments. Data were provided on more of a second-by-second basis, similar to GPS measurements.

A deployment of the RadioCamera technology occurred in the San Francisco Bay Area in 2000. This particular test involved the University of California–Berkeley and US Wireless, and focused on I-580 and a major arterial in Oakland.[7] US Wireless provided 44 hours of wireless data to UC-Berkeley researchers to analyze. The researchers found that the position estimates generally had a 60-m accuracy, although 66 percent of all probe vehicle tracks had at least one data point that deviated from the caller's actual position by more than 200 m. The researchers noted that the call lengths were generally very short, with a median call length of only 30 seconds. The short call durations made it very difficult to estimate speeds on links since position estimates were not available for long distances. The researchers were also not able to match 60 percent of vehicles to a roadway link.

Another deployment of the RadioCamera technology occurred in the Washington, D.C., metropolitan area from 2000 through 2001.[12] The deployment was a partnership with the VDOT, MDSHA, and US Wireless. The goal of the project was to prove the feasibility of WLT on congested freeways and arterials.[5] The system was scaled to track 160 phone calls every 2 seconds, generating 4800 data points every minute. The University of Virginia and the University of Maryland were responsible for evaluating the system.

Although more data were generated than in earlier deployments, the RadioCamera system was still unable to generate the quality or quantity of data that would be necessary to support traffic monitoring. No data were generated in approximately 5 percent of the 10-minute intervals analyzed, so no traffic condition estimates could be created. The results from the intervals that did have samples showed wide variations in speed estimates. I-495, a congested urban freeway, had a mean speed error of approximately 8 mph, with some intervals having speed estimates that had more than 20 mph error. The arterials monitored had a mean speed error of 6.8, with a maximum error of 23.2 mph. There was inconsistency in the number of samples generated for different links, resulting in these large variations in speed estimation error.

Hampton Roads, Virginia, Deployment

The vendor AirSage deployed a handoff-based traffic monitoring system in the Hampton Roads region of Virginia starting in 2003. The technology used in this deployment differed from early generation systems in that handoffs between cell towers were used to match vehicles to roads, rather than discrete point estimates of latitudes and longitudes. This deployment was funded by VDOT and FHWA, and the University of Virginia performed an independent evaluation of the system. The WLT system was used to monitor approximately 90 centerline miles of freeways and arterials. The AirSage technology works by mining handoff data that were already collected by cellular service providers. Data on cellular handoffs, as well as transitions between sectors of a cell, was processed to determine a vehicle's location on the roadway network. These transitions were then used to determine speed and travel time information on the network.

The University of Virginia performed an evaluation of the AirSage system in December 2005.[13] The evaluation results show that the under congested conditions, 68 percent of the AirSage speeds have an error greater than 20 mph. Performance also tended to be worse under congested conditions than during free flow. In the original project scope, AirSage proposed that travel time estimates would be produced on the reversible high occupancy vehicle (HOV) facility on I-64 and confidence measures for traffic data records would be produced. As of December 2005, neither HOV travel times nor confidence measures could be produced. Overall, the December evaluation concluded that the traffic monitoring system could not produce accurate travel time estimates in its current form. AirSage cited the lack of access to full cellular network data as the major reason not achieving accurate results.

Simulation Studies

Simulation studies do not represent "real-world" deployments, but they do provide insight information how WLT systems must be deployed and operated to provide good data. The purpose of the simulation studies ranged from a technical evaluation to assessment of sampling characteristics required to support traffic monitoring systems.

French Simulation Study[14]

A simulation study was conducted by the French transportation research organization using a discrete event simulation of traffic flow in order to determine the sample size requirements and accuracy of a hypothetical WLT system. The researchers examined three relatively simple traffic networks. The issues examined in the study were the impact of probe penetration on accuracy of travel time estimates, assuming a location error of 150 m. The results showed that the freeway travel times can be estimated to within 10 percent of their actual value if there was at least 5 percent penetration of wireless devices in the traffic stream. The simulation evaluated relatively simple geometric conditions on small networks with a size of less than a mile. The researchers did not present what type of map matching algorithms they used to match the cell phone data to the network. The study also did not explicitly consider issues like the time between location readings or non-vehicle based probes.

University of California–Berkeley Simulation Study[15]

A simulation study was conducted by Berkeley Institute for Transportation Studies. The researchers examined the impact of location accuracy, time between position readings, and the total number of location estimates determined per square mile per second. The findings from the research indicated that a network-based system could generate sufficient samples on roadways if location accuracy were as low as 100 m. The research did not address whether the sample sizes collected were sufficient to estimate speeds or travel times. It also did not mention the map matching algorithms that were used to match the vehicles to the network.

Virginia Transportation Research Council Simulation Study[3]

The Virginia Transportation Research Council studied the role of system design in dictating the performance of WLT-based monitoring systems. A simulation test bed was developed that combined the microscopic traffic simulation model VISSIM and a custom emulation of WLT-based monitoring. The test bed emulated a WLT-based traffic monitoring system that uses an area-wide definition of sampling parameters. The study investigated the impact of system design, traffic, and geometric characteristics through a combination of tests on simple geometric networks and case studies of simulated complex, real world traffic conditions. The results of the research showed that even though speed estimates could be consistently generated from position data, they were not always accurate. One of the major problems identified was that the sampling strategy that had been implemented in previous field deployments was inadequate. The area-wide sampling approach sampled a disproportionate number of vehicles on congested links, while shorter links and lower volume roads produced fewer samples. The small number of samples on the lower volume links resulted in poor speed estimates in those cases. The distribution of probes was found to be very inconsistent throughout the networks, and the research recommended that methods to improve the distribution of samples throughout the network be explored.

Summary of Studies

Table 2 summarizes the results from field tests and simulation studies. Deployments were unsuccessful in generating accurate speed estimates or traffic conditions in the network. Many of the early generation deployments used area-wide sampling specifications which resulted in few samples from low volume roads. Moreover, the simulation studies showed that sampling parameters such as time between position readings and speed variance affect the speed estimates. It seems that there is a need to develop a sampling strategy that ensures probe availability throughout the entire network, thus improving speed estimation on lower volume roads.

Table 2. Summary of WLT Field Deployments and Simulation Studies

Deployments	Results
CAPITAL project, Washington, D.C.	Speed information was not generated 80% of the time due to lack of sufficient samples on monitored links
US Wireless, San Francisco Bay Area	Sampling methodology used was able to track cell phones for short distances, and speeds could not be estimated over long distances
US Wireless, Washington, D.C.	Very few samples were picked from low volume and uncongested areas
AirSage, Hampton Roads	System produced poor speed estimation accuracy, especially during congestion
Simulation Studies	**Results**
French Transportation Center	Concluded that 5% probe penetration will be sufficient to estimate speeds on links 95% of the time
University of California, Berkeley	Studied the impact of sampling parameters on speed estimation accuracy
Virginia Transportation Research Council	Identified deficiencies in the area-wide approach to sampling. This produced poor quality speed estimates, especially on low volume roads.

Case Studies

The results of the case studies of the Charlottesville, Springfield, and Tysons Corner networks are discussed in this section. The networks are described first, and the impact of various factors on the accuracy and availability of speed estimates is reviewed.

Charlottesville Case Study

The Charlottesville network was the least complex network used in this research. It was intended to provide a best case scenario to test WLT-based systems where geometric conditions were simple and traffic was relatively homogenous. Figure 5 below shows the simulated Charlottesville network with hypothetical zones. The Charlottesville network had an area approximately equal to 4 square miles. The network was divided into two zones, with each zone consisting of a regular hexagon with an area of 2 square miles. The zones were centered along the major roadways, which approximates how cells were laid out in the real world. When dividing the network into zones, borders of each zone were placed in such a way that congested areas were separated from uncongested areas. This represents an ideal situation where the network was divided into relatively homogeneous zones. It is probable that the way cells overlay the roadway network could significantly impact the performance of a zonal system, and different layouts could produce different results. The network has 11.33 centerline miles of roadway with 10 traffic signals and 2 freeway interchanges.

There were two major routes in the network, US 29, and US 250. A total of 10 links were monitored on these two major roads in the network. Interchange ramps were not monitored because the map-matching algorithm had difficulty in rectifying estimated locations to short links where geometry was very complex. These links were considered for calculating the total vehicle minutes of travel in each zone, however. A total of 8 links were monitored in zone 1 and 2 links were monitored in zone 2.

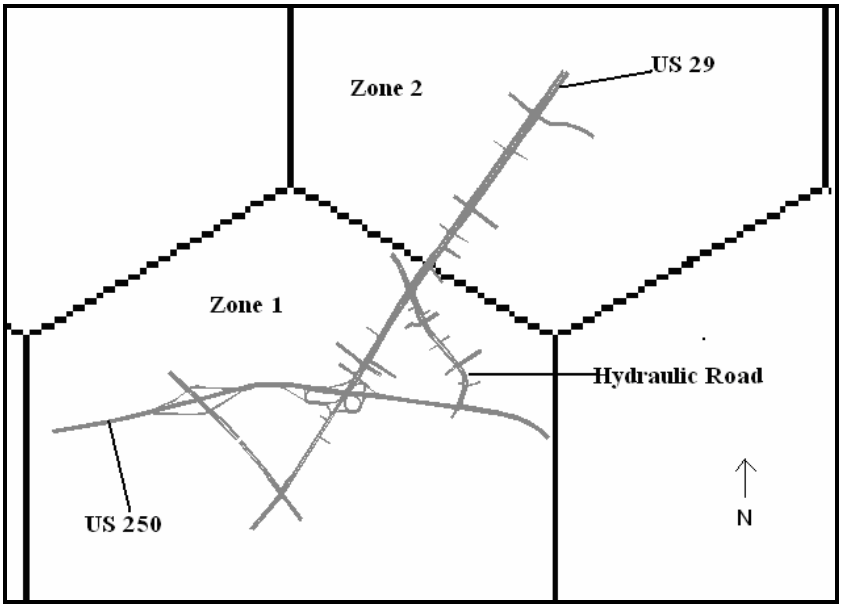

Figure 5. Simulated Layout of Zonal Systems for the Charlottesville Network

Calculation of Sample Size and Frequency

The simulation was run in VISSIM for afternoon peak hour conditions, and travel times were recorded for a 1-hour period. The data collected was disaggregated into eleven 5-minute intervals. The first five minutes were discarded to allow the network to reach equilibrium and hence were not used in the analysis. For each 5-minute interval, the standard deviation of speeds on each link was calculated. Sample sizes and frequencies for each link were determined based on the maximum standard deviation and minimum travel time observed during these eleven 5-minute intervals.

Using the methodology described earlier, sample sizes and frequencies were calculated for zonal and zonal priority sampling strategies. The sample sizes and frequencies for each zone are shown in Table 3.

Table 3. Sample Sizes and Frequencies for Zonal and Zonal Priority System, Charlottesville Case Study

Zone	Frequency Based on Travel Time (F_i, sec)	Number of Vehicles to Track Simultaneously (vehicles)
1	9	42
2	18	20

The differences in the number of vehicles to track were due to disparities in geometric and traffic flow conditions observed within each zone. Zone 1 requires a large number of probes because of the more complex geometry of the network and the high standard deviation of speeds on the links. Zone 2 has relatively uniform geometric and traffic flow conditions so the required sample size was smaller.

For the area-wide system, the number of vehicles to be tracked was set equal to the sum of sample sizes in all the zones in the zonal system. This translated into 62 vehicles tracked simultaneously, network-wide. The frequency was calculated using the weighted average of the frequencies. This corresponded to 12 seconds between position readings across the network for the area-wide system.

Number of Speed Estimates

The number of speed estimates obtained was analyzed using GLM, ANOVA, and LSD. First, significant factors were identified using GLM. Next, if the factor has two levels then ANOVA was used to study the differences between factor levels. If the factor has more than two levels LSD was used to study the differences between factor levels.

Table A-1 in Appendix A shows results obtained using the GLM analysis for the number of speed estimates. That analysis identified only two factors as statistically significant in determining the number of speed estimates: zone number and the interaction of zone number with sampling strategy. The significance of zone number was not surprising, given that the number of vehicles being tracked varies by zone, and the density of cellular probes was likely to vary based on geometry. Table 4 shows that zone 2 generated an average of 12.5 speed estimates more than zone 1. In zone 1, the relatively complex geometry results in a smaller number of speed estimates being matched to the various links.

While the GLM analysis indicated that the interaction of sampling strategy and zone number was significant, further analysis using the LSD test revealed this to not be the case. The positive result using GLM appears to be simply an artifact of the varying number of samples between the zones, not any impact of sampling strategy. Table 5 shows this result. As a result, it does not appear that the sampling strategy produced a significant impact on the average number of speed estimates per link generated in the Charlottesville network. Given that the total number of vehicles being tracked was the same for all three methods, this finding was not surprising.

Table 4. Charlottesville Case Study ANOVA Results for Number of Speed Estimates by Zone

Zone Number	Mean Number of Speed Estimates per Link	Standard Error	F	Significance
1	6.98	0.252	494	0
2	19.5	0.505		

Table 5. Charlottesville Case Study LSD Results for Number of Speed Estimates by Sampling Strategy and Zone

Zone No.	Sampling Strategy	Mean Number of Speed Estimates per Link	Standard Error	Significantly Different From
1	Zonal Priority	7.13	0.43	None
	Zonal	7.16	0.43	None
	Area-wide	6.60	0.43	None
2	Zonal Priority	20.04	0.87	None
	Zonal	17.79	0.87	None
	Area-wide	20.7	0.87	None

Accuracy of Speed Estimates

Next, the speed estimation accuracy was examined. The GLM analysis results are shown in Table A-2 in Appendix A. That analysis indicated that the sampling strategy and zone number main effects have a significant effect on speed estimation accuracy. The interaction between zone number and sampling strategy also was found to have a significant impact on speed estimation accuracy.

Table 6 shows the LSD pairwise comparison of speed estimation accuracy by sampling strategy. All sampling strategies were found to be statistically significant from one another, with the zonal priority performing the best and the area-wide system performing the worst. This indicates that the zonal systems do provide a benefit in speed estimation, even when the number of speed estimates was not statistically significantly different. Likewise, it appears that granting priority to vehicles that were transitioning between zones improved the quality of speed estimates, possibly because the vehicle tracks were longer.

Table 6. Charlottesville Case Study LSD Results for Speed Estimation Accuracy by Sampling Strategy

Sampling Strategy (I)	Sampling Strategy (J)	Mean Difference (I-J) in mph	Standard Error	Significance
Zonal Priority	Zonal	-4.4	0.82	0.0
	Area-wide	-6.5	0.82	0.0
Zonal	Zonal Priority	4.4	0.82	0.0
	Area-wide	-2.05	0.79	0.01
Area-wide	Zonal Priority	6.5	0.82	0.0
	Zonal	2.05	0.79	0.01

Table 7 summarizes the results of the analysis of the interaction of zone number and sampling strategy. The analysis shows that most of the benefits of the zonal priority sampling strategy were accrued in the more complex zone 1. In zone 2, the simple geometry results in very similar speed estimation performance among all three sampling strategies. In zone 1, it appears that providing priority to vehicles that have entered from zone 2 allows better position matches and more accurate speed estimates.

Table 7. Charlottesville Case Study LSD Results
for Speed Estimation Accuracy by Sampling Strategy and Zone

Zone No.	Sampling Strategy	Mean (mph)	Standard Error	Significantly Different From
1	Zonal Priority	8.08	0.69	Zonal, Area-wide
	Zonal	13.63	0.62	Zonal Priority, Area-wide
	Area-wide	16.23	0.62	Zonal Priority, Area-wide
2	Zonal Priority	3.6	1.25	None
	Zonal	3.4	1.25	None
	Area-wide	3.3	1.25	None

The distribution of speed estimation errors for three sampling strategies at two levels of map matching error with two methods of link speed estimation were shown in Tables A-3, A-4, and A-5 in Appendix A. The zonal priority consistently produced speeds within 10 mph accuracy for 74 percent of the monitored miles.

Summary of Charlottesville Case Study Results

The Charlottesville case study provides several interesting findings. First, for small networks with simple geometry and traffic conditions, there was no significant difference in the number of speed estimates generated by the different sampling strategies. Second, the zonal priority sampling strategy produced more accurate speed estimates than either of the other two strategies. This indicates that creating longer vehicle tracks was desirable, particularly in zones that have complex geometry.

Springfield Case Study

The Springfield network was the most complex case study used in this research. This network has 80.76 centerline miles of roadway with 41 traffic signals, 5 freeway interchanges, a heavy rail line, an HOV facility, and 4 locations that simulate the potential impact of pedestrians with handheld phones. The simulation was run for afternoon peak hour traffic conditions. Figure 6 shows the simulated Springfield network with hypothetical zones. The network was divided into four zones, each of which was a regular hexagon with an area of 2.2 square miles. Six major routes in the network were monitored: I-495, I-395, I-95, Franconia Road, Van Dorn Street, and the Franconia-Springfield Parkway. A total of 40 links were monitored on these roads. Once again, interchange ramps were not monitored, although they were included in the estimate vehicle-hours of travel on the network. A total of 9 links were monitored in zone 1, 6 links in zone 2, 15 links in zone 3, and 10 links in zone 4.

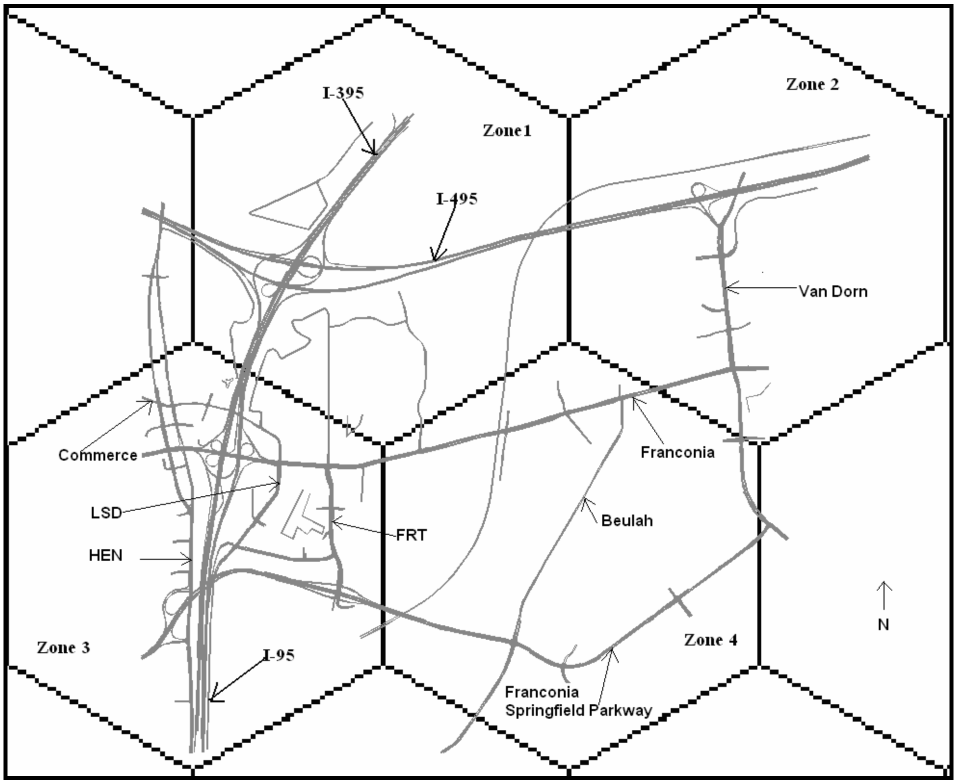

Figure 6. Simulated Layout of Zonal System for Springfield Network

Calculation of Sample Size and Frequency

The simulation was run in VISSIM for afternoon peak hour conditions, and travel times were recorded for a one-hour period. The data collected were disaggregated into twelve 5-minute intervals. For each 5-minute interval, the standard deviation of speeds on each link was calculated. Sample sizes and frequencies for each link were determined based on the maximum standard deviation and minimum travel time observed during these twelve 5-minute intervals. Using the methodology described in methods section, sample sizes and frequencies were calculated for zonal and zonal priority sampling strategies. The calculation of the sample sizes and frequencies for each zone are shown in Table 8 .

Table 8. Sample Sizes and Frequencies for Zonal and Zonal Priority System, Springfield Case Study

Zone	Frequency Based on Travel Time (F_i, sec)	Sample Size Per 5 Minute Analysis Period (vehicles)	Number of Vehicles to Track Simultaneously (vehicles)
1	8	1122	200
2	10	654	145
3	8	1202	214
4	10	586	130

The disparities in the number of vehicles being tracked among the zones are due to differences in geometric and traffic flow conditions observed within each zone. Zone 1 and 3 requires large number of probes because of speed standard deviations as high as 16 mph. The zones 2 and 4 have relatively uniform traffic flow conditions, so the required sample size was small. Likewise, zones 1 and 3 have more links, which requires that more vehicles be tracked in order to ensure that a sufficient sample size is produced on each link. The time between readings was relatively constant across the zones.

For the area-wide approach, the number of vehicles to be tracked was set equal to the sum of sample sizes in all the zones in the zonal system. This translated into 689 vehicles tracked simultaneously, network-wide. The time between position readings was calculated using the weighted average of the frequencies, which equaled 9 seconds for the area-wide approach.

Number of Speed Estimates

Table B-1 in Appendix B summarizes the GLM analysis results for number of speed estimates. That analysis found that the following factors had a significant role in determining the number of data points generated per link, per 5-minute interval:

- Sampling strategy
- Map matching error
- Zone number
- Sampling strategy and zone number interaction
- Map matching error and zone number interaction.

Table 9 shows the LSD analysis for the number of speed estimates generated by sampling strategy. The zonal priority and zonal systems were both significantly different from the area-wide approach, with the area-wide method generating an average of 2 more estimates per link per 5 minutes than the zonal strategies. This was a result of the congested freeways being over sampled by the area-wide method. Matching was more difficult on the arterial system. By better distributing the samples, the zonal approaches actually produced fewer estimates since establishing a good match on those arterial routes was more difficult.

Table 9. Springfield Case Study LSD Results for Number of Speed Estimates by Sampling Strategy

Sampling Strategy (I)	Sampling Strategy (J)	Mean Difference (I-J)	Standard Error	Significance
Zonal Priority	Zonal	-0.01	0.86	0.99
	Area-wide	-2.15	0.86	0.01
Zonal	Zonal Priority	0.01	0.86	0.99
	Area-wide	-2.14	0.86	0.01
Area-wide	Zonal Priority	2.15	0.86	0.01
	Area-wide	2.144	0.86	0.01

Table 10 shows the ANOVA analysis that compares the two levels of map matching error. The 10 m map matching error produced on an average of 4.8 more speed estimates than the 16-m map matching error. In the Springfield network, there were many locations where high-speed links were parallel to low-speed links. In some cases, the gap between these links was as low as 30 m. The 16-m map matching error increased ambiguity while matching vehicles on congested links to the network, resulting in more incorrect map matches between the two parallel routes. The data filtering process eliminates some of these incorrect map matches because they had less than three position estimates on a link. Thus, number of speed estimates for 16-m map matching error was considerably affected by incorrect map matches that were screened out.

Table 10. Springfield Case Study ANOVA Results for Number of Speed Estimates by Map Matching Error

Map Matching Error (I)	Map Matching Error (J)	Mean Difference in Sample Sizes (I-J)	Standard Error	F	Significance
10 m	16 m	4.83	.70	46.0	0.0
16 m	10 m	-4.83	.70		

Table 12 shows the ANOVA results for the interaction of map matching error and zone number. The 10-m map matching error produced a significantly greater number of speed estimates in zones 1 and 2, likely due to the influence of screening the poor matches from the congested roads in those zones. As shown in Table 13, zones 1 and 2 have locations where high-speed links were parallel to low-speed links. The number of speed estimates on these links was considerably affected by 16-m map matching error. While matching vehicles to these links, the 16-m map matching error created ambiguity for MHT, which resulted in more incorrect map matches. Some of the incorrect map matches had less than a pair of position estimates and were filtered out.

Table 11 shows the LSD analysis of the interaction of zone number and sampling strategy. The area-wide system produced significantly more observations in zones 1 and 2 which contained the severely congested, high volume I-495 beltway. The zonal systems produced more estimates in zones 3 and 4, indicating that the zonal approaches were distributing the samples more “evenly” throughout the geographic region. The two zonal strategies were not significantly different from one another.

Table 12 shows the ANOVA results for the interaction of map matching error and zone number. The 10-m map matching error produced a significantly greater number of speed estimates in zones 1 and 2, likely due to the influence of screening the poor matches from the congested roads in those zones. As shown in Table 13, zones 1 and 2 have locations where high-speed links were parallel to low-speed links. The number of speed estimates on these links was considerably affected by 16-m map matching error. While matching vehicles to these links, the 16-m map matching error created ambiguity for MHT, which resulted in more incorrect map matches. Some of the incorrect map matches had less than a pair of position estimates and were filtered out.

Table 11. Springfield Case Study LSD Results for Number of Speed Estimates by Sampling Strategy and Zone

Zone	Sampling Strategy	Mean	Standard Error	Significantly Different From
1	Zonal Priority	27.37	1.22	Area-wide
	Zonal	27.12	1.22	Area-wide
	Area-wide	33.58	1.22	Zonal Priority, Zonal
2	Zonal Priority	25.17	1.50	Area-wide
	Zonal	26.26	1.50	Area-wide
	Area-wide	42.93	1.50	Zonal Priority, Zonal
3	Zonal Priority	18.38	0.95	Area-wide
	Zonal	17.88	0.95	Area-wide
	Area-wide	11.00	0.95	Zonal Priority, Zonal
4	Zonal Priority	17.10	1.16	Area-wide
	Zonal	16.81	1.16	Area-wide
	Area-wide	9.13	1.16	Zonal Priority, Zonal

Table 12. Springfield Case Study ANOVA Results for Number of Speed Estimates by Map Matching Error and Zone

Zone	Map Matching Error	Mean	Standard Error	Significantly Different From
1	10 m	32.00	1.00	16 m
	16 m	26.71	1.00	10 m
2	10 m	36.02	1.22	16 m
	16 m	26.88	1.22	10 m
3	10 m	17.00	0.77	None
	16 m	14.50	0.77	None
4	10 m	15.55	0.95	None
	16 m	13.14	0.95	None

Table 13. Springfield Case Study, High-speed Links Parallel to Low-speed Links

Zone Number	Low-speed Link		Parallel High-speed Link	
	Link Number	Speed (mph)	Link Number	Speed
1	1	2.75	2	64.64
	7	6.30	6	64.82
2	10	2.70	13	63.94
	12	2.68	11	64.09

Speed Estimation Accuracy

Table B-2 in Appendix B summarizes the results of the GLM analysis for speed estimation accuracy for the Springfield network. That analysis shows that the following factors and interactions had a significant impact on speed estimation accuracy:

- Sampling strategy
- Speed method
- Map matching error

- Zone number
- Sampling strategy and zone number interaction
- Speed method and zone number interaction.

Table 14 shows LSD pairwise comparison for the main effect of sampling strategy on speed estimation accuracy. The area-wide and zonal priority strategies did not produce significant different results, but they were significantly different from zonal sampling strategy. The zonal sampling strategy produced an error of 1 mph more than area-wide and zonal priority.

Table 14. Springfield Case Study LSD Results for Speed Estimation Accuracy by Sampling Strategy

Sampling Strategy (I)	Sampling Strategy (J)	Mean Difference (I-J) (mph)	Standard Error	Significance
Zonal Priority	Zonal	-1.25	0.50	0.01
	Area-wide	0.39	0.49	0.43
Zonal	Zonal Priority	1.25	0.50	0.01
	Area-wide	1.64	0.50	0.00
Area-wide	Zonal Priority	-0.39	0.49	0.43
	Zonal	-1.64	0.50	0.00

Table 15 shows the results of ANOVA for the method of speed calculation. That analysis indicates that the distance weighted speed method produced more accurate speeds than average speed method. The difference in speed errors obtained using these two methods differed by a statistically significant margin of 2.75 mph. In examining the data closely, it appears that the distance-weighted method has the tangential benefit of minimizing the impact of vehicles that were matched to the wrong road. When a vehicle was matched to the wrong link, it was often estimated to have only traveled a short distance. When the distance-weighted method was used, the contribution of those matches was smaller versus when the average speed method was used. This had a large impact on the cases where high-speed and low-speed facilities were parallel to one another.

Table 15. Springfield Case Study ANOVA Results for Speed Estimation Accuracy by Speed Method

Speed Method (I)	Speed Method (J)	Mean Difference (I-J) in mph	Standard Error	F	Significance
Average Speed	Distance Weighted Speed	2.75	.40	46.0	0.0
Distance Weighted Speed	Average Speed	-2.75	.40		

Table 16 shows the ANOVA of the two levels of map matching error. The data show that a map matching error of 10 m produced more accurate speeds than the 16 m map matching error. The 16 m map matching error produced more matches to incorrect routes, especially in locations where high-speed facilities were in parallel with low-speed links. This created poor speed estimates.

Table 16. Springfield Case Study ANOVA Results for Speed Estimation Accuracy by Map Matching Error

Map Matching Error (I)	Map Matching Error (J)	Mean Difference (I-J) in mph	Standard Error	F	Significance
10 m	16 m	-.875	.404	4.6	0.03
16 m	10 m	.875	.404		

Table 17 shows the LSD analysis of the interaction of sampling strategy and zone number on speed estimation accuracy. The only zone with significant differences in performance was zone 1, where the zonal system did not perform well. In other zones, no statistically significant differences in performance were observed.

Table 17. Springfield Case Study LSD Results for Speed Estimation Accuracy by Sampling Strategy and Zone Number

Zone Number	Sampling Strategy	Mean Speed Estimation Accuracy (mph)	Standard Error	Significantly Different From
1	Zonal Priority	13.17	0.70	Zonal
	Zonal	17.97	0.73	Zonal Priority, Area-wide
	Area-wide	13.43	0.70	Zonal
2	Zonal Priority	15.28	0.82	None
	Zonal	14.83	0.82	None
	Area-wide	13.92	0.82	None
3	Zonal Priority	8.44	0.55	None
	Zonal	9.03	0.56	None
	Area-wide	8.37	0.56	None
4	Zonal Priority	5.64	0.68	None
	Zonal	5.70	0.68	None
	Area-wide	5.27	0.68	None

Table 18 shows the ANOVA for the interaction of speed method and zone number on speed estimation accuracy, aggregated across sampling strategy. The results indicate that the distance-weighted method produced more accurate speeds in zones 1 and 2, and the average speed method produced more accurate speeds in zone 3. In zones 1 and 2, high-speed freeways were parallel with low-speed roads. Due to map matching errors, vehicles on low-speed links were matched to high-speed links and vice versa. With the distance-weighted method, map matching errors which occurred on small segments of the roadway do not impact speed estimation accuracy as much as the average speed method. In case of zone 3, average speed performed better than distance weighted speed, possibly because of the wider variation in traffic speeds seen in that zone.

Table 18. Springfield Case Study ANOVA Results for Speed Estimation Accuracy by Speed Method and Zone, Aggregated Across Sampling Strategy

Zone	Speed Method	Mean Speed Estimation Accuracy (mph)	Standard Error	Significantly Different From
1	Average Speed	17.12	0.58	Distance Weighted Speed
	Distance Weighted Speed	12.60	0.58	Average Speed
2	Average Speed	19.76	0.67	Distance Weighted Speed
	Distance Weighted Speed	9.59	0.67	Average Speed
3	Average Speed	6.41	0.46	Distance Weighted Speed
	Distance Weighted Speed	10.83	0.46	Average Speed
4	Average Speed	5.91	0.56	None
	Distance Weighted Speed	5.16	0.56	None

The distribution of speed errors for each sampling strategy for the two levels of map matching error and the two speed estimation methods were shown in Tables B-3, B-4, and B-5 in

Appendix B. The percentage of miles that were determined to be within certain accuracy thresholds over the one-hour analysis period was shown in those tables.

Summary of Springfield Case Study Results

The results of the Springfield case study demonstrated that network complexity and traffic volume/congestion have significant impacts on the performance of WLT-based traffic monitoring. Contrary to the Charlottesville case study, the area-wide sampling strategy generated more speed estimates than the zonal priority and zonal sampling strategies. The area-wide approach tracked more vehicles from congested freeways, while zonal strategies distributed samples across freeways and arterials. Matching was more difficult on arterial routes, resulting in lower overall samples sizes in the zonal approaches. Overall, the data availability on links improved with the zonal system, however, since many of the probes in the area-wide system were concentrated on relatively few facilities. The 10-m map matching error generated an average of 4.8 speed estimates more than 16 m map matching error. This was caused by the larger number of bad position matches with the 16-m error.

Several trends in the accuracy of speed estimates were also observed. The distance-weighted speed method was an average of 2.7 mph more accurate than the average speed method. This performance did appear to differ depending on the zone characteristics, with the distance weighted method working better on closely spaced parallel roads, and the average speed method working better on roads with higher travel time variability. On an aggregate basis, however, the accuracy of speed estimates was similar among the various sampling strategies evaluated.

Tysons Corner Case Study

The Tysons Corner region of Northern Virginia was the last case study evaluated. This network has 85.45 centerline miles of roadway, with 32 traffic signals and 7 freeway interchanges. The network also includes two locations that simulate the potential impact that pedestrians with mobile phones may have on system speed estimates. The simulation was run for afternoon peak hour traffic conditions. Figure 7 shows the simulated Tysons Corner network with hypothetical zones. It was an urban network with an area approximately equal to 10 square miles. The network was divided into five hexagonal zones, each of which had an area of 2.5 square miles. The routes that were monitored included: the Dulles Toll Road (SR 276), the Capital Beltway (I-495), SR 7, and SR 123. A total of 26 links were monitored on these four major roads in the network. Zones 1 and 5 contained 5 links, zones 2 and 4 contained 4 links, and zone 3 contained 8 links.

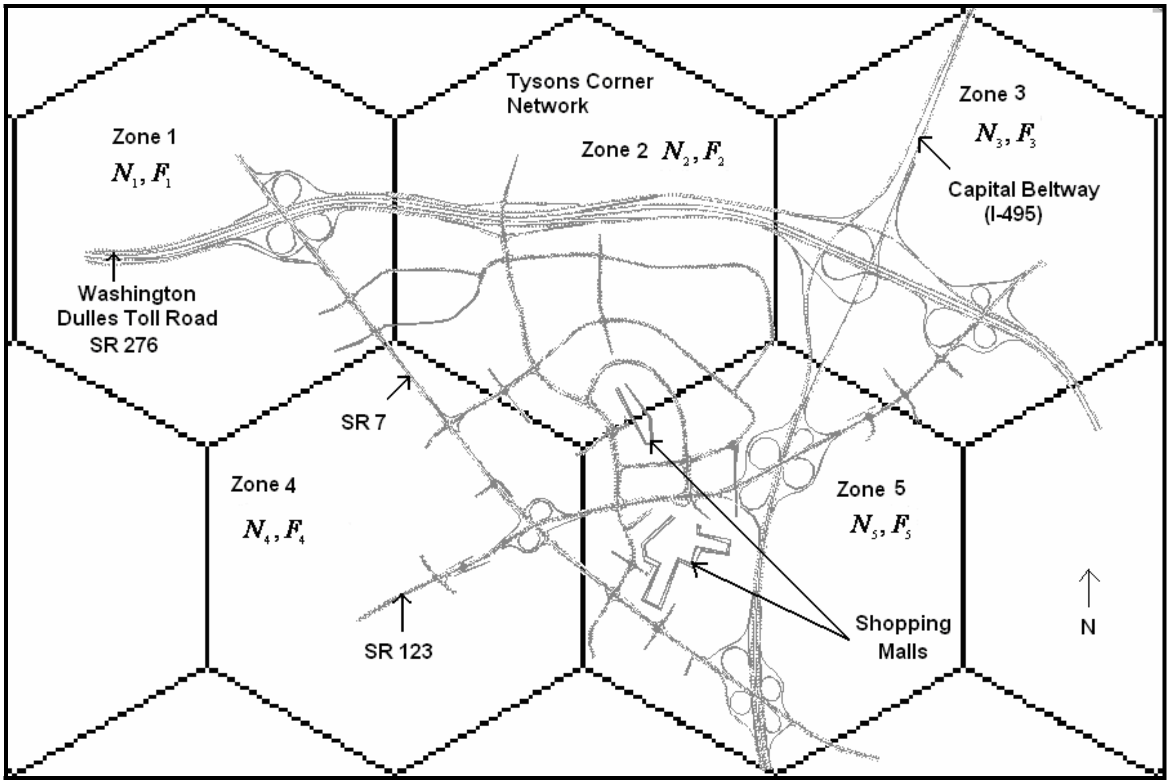

Figure 7. Simulated Layout of Zonal WLT System for Tysons Corner Network

Calculation of Sample Size and Frequency

The simulation was run in VISSIM for afternoon peak hour conditions, and travel times were aggregated into twelve 5-minute periods for each monitored link. For each 5-minute interval, the standard deviation of speeds on each link was calculated. Sample sizes and frequencies for each link were determined based on the maximum standard deviation and minimum travel time observed during these twelve 5-minute intervals. The sample size calculations for zonal and zonal priority sampling strategy were the same as discussed in the methodology chapter. The calculation of the sample sizes and frequencies for each zone are shown in Table 19.

Table 19. Sample Sizes and Frequencies for Zonal and Zonal Priority System, Tysons Corner Case Study

Zone	Frequency Based on Travel Time (F_i, sec)	Sample Size Per 5 Minute Analysis Period (vehicles)	Number of Vehicles to Track Simultaneously (vehicles)
1	26	310	80
2	16	146	22
3	29	842	234
4	74	60	46
5	33	484	154

Zones 3 and 5 required a large number of probes because of the complex geometry of the network and the high standard deviation of speeds on the links. Zone 2 has relatively uniform geometric and traffic flow conditions, so the sample size was small. Zone 4 has highly congested traffic with average speeds around 5 mph, so it requires a small sample size with large times between position readings.

For the area-wide system, the number of vehicles to be tracked was set equal to the sum of sample sizes in all the zones in the zonal system. This translated into 536 vehicles tracked simultaneously, network-wide. The frequency was calculated using the weighted average of the frequencies. This corresponded to 30 seconds between position readings across the network for the area-wide system.

Number of Speed Estimates

Table C-1 in Appendix C shows the results of the GLM analysis of the number of speed estimates generated per link per 5-minute interval. The following main effects and interactions were found to be significant:

- Sampling strategy
- Zone number
- Interaction of sampling strategy and map matching error
- Interaction of sampling strategy and zone number
- Interaction of sampling strategy, map matching error, and zone number.

First, the effect of sampling strategy was investigated further using LSD analysis. Table 20 shows that the zonal priority method generated significantly more speed estimates than either of the other two methods. The zonal and area-wide strategies were not significantly different from one another. It appears that the zonal strategy had problems matching vehicles due to the complex network geometry, resulting in fewer estimates than the zonal priority strategy. The area-wide strategy used a 30 second interval between position readings for the entire network. This resulted in a small number of speed estimates on high-speed links in zone 2.

Table 20. Tysons Corner Case Study LSD Results for Number of Speed Estimates by Sampling Strategy

Sampling Strategy (I)	Sampling Strategy (J)	Mean Difference (I-J)	Standard Error	Significance
Zonal Priority	Zonal	2.56	0.65	0.00
	Area-wide	1.84	0.65	0.00
Zonal	Zonal Priority	-2.56	0.65	0.00
	Area-wide	-0.72	0.65	0.27
Area-wide	Zonal Priority	-1.84	0.65	0.00
	Zonal	0.72	0.65	0.27

Table 21 shows the LSD analysis results for the interaction of zone number and sampling strategy. The results show that zonal priority sampling strategy generated a higher number of speed estimates for zones 1, 2, and 3 than the area-wide approach. The zonal priority system allowed the samples to be distributed more effectively through the entire network and reduced some of the over sampling that was occurring in zones 4 and 5 in the area-wide system. As noted earlier, zone 4 consists of very low-speed, congested traffic. As a result, few samples should be needed in that zone to generate adequate speed estimates. While zonal priority distributed speed estimates, the area-wide tended to concentrate samples in the high volume, congested zones 4 and 5 while simultaneously generating fewer samples in the less congested or lower volume zones 1, 2, and 3.

Table 21. Tysons Corner Case Study LSD Results for Number of Speed Estimates by Sampling Strategy and Zone

Zone Number	Sampling Strategy	Mean Number of Speed Estimates	Standard Error	Significantly Different from
1	Zonal Priority	9.62	1.04	Area-wide
	Zonal	10.32	1.04	Area-wide
	Area-wide	5.63	1.04	Zonal Priority, Zonal
2	Zonal Priority	6.42	1.16	Zonal, Area-wide
	Zonal	1.39	1.16	Zonal Priority
	Area-wide	1.78	1.16	Zonal Priority
3	Zonal Priority	24.46	0.82	Zonal, Area-wide
	Zonal	18.37	0.82	Zonal Priority, Area-wide
	Area-wide	15.06	0.82	Zonal Priority, Area-wide
4	Zonal Priority	2.43	1.16	Area-wide
	Zonal	3.42	1.16	Area-wide
	Area-wide	13.41	1.16	Zonal Priority, Zonal
5	Zonal Priority	8.61	1.04	Area-wide
	Zonal	7.54	1.04	Area-wide
	Area-wide	12.95	1.04	Zonal Priority, Zonal

Table 22 shows that the zonal priority and area-wide approaches produced the same number of speed estimates across the two map matching error levels. The zonal approach produced more speed estimates with a 16 m map matching error. In depth analysis of the data, revealed that the zonal sampling strategy generated more speed estimates in zones 3 and 5 with 16 m map matching error. It appears that this increase in number of data points was due to the larger frequency of incorrect matches that are occurring in these two zones. Since the zonal approach used short vehicle tracks, the influence of map matching error appears to be more pronounced than with the other sampling strategies.

Table 22. Tysons Corner Case Study ANOVA Results for Number of Speed Estimates by Map Matching Error and Sampling Strategy

Sampling Strategy	Map Matching Error	Mean Number of Speed Estimates	Standard Error	Significantly Different From
Zonal Priority	10 m	10.03	0.67	None
	16 m	10.59	0.67	None
Zonal	10 m	6.31	0.67	16 m
	16 m	10.12	0.67	10 m
Area-wide	10 m	10.55	0.67	None
	16 m	9.00	0.67	None

The significance of the zone number main effect appears to be driven by the sampling requirements that are implicit in the zonal systems. The zone number findings are not discussed in detail since the significance was purely a function of the zonal systems distributing samples across the network. Likewise, the three-factor interaction of sampling strategy, map matching error, and zone number does not provide any insight beyond what was seen in the two-factor interactions.

Speed Estimation Accuracy

The GLM analysis results on the accuracy of speed estimates are shown in Table C-2 in Appendix C. The following factors were found to produce a significant effect on speed estimation accuracy:

- Sampling strategy
- Map matching error
- Zone number
- Interaction of sampling strategy and map matching error
- Interaction of sampling strategy and zone number
- Interaction of speed method and zone number
- Interaction of map matching error and zone number
- Interaction of sampling strategy, map matching error, and zone number.

Table 23 shows the LSD analysis of the impact of sampling strategy on speed estimation accuracy. All methods were statistically different from one another, with the zonal priority performing the best and area-wide performing the worst. In the case of zonal priority sampling strategy, samples were distributed geographically throughout the entire network helping to ensure that there were ample data to generate speed estimates. Samples were not distributed as effectively in the area-wide approach, and the zonal approach had problems generating enough data to produce reliable estimates.

Table 23. Tysons Corner Case Study LSD Results for Speed Estimation Accuracy by Sampling Strategy

Sampling Strategy (I)	Sampling Strategy (J)	Mean Difference (I-J) mph	Standard Error	Significance
Zonal Priority	Zonal	-3.40	0.66	0.00
	Area-wide	-4.83	0.68	0.00
Zonal	Zonal Priority	3.40	0.66	0.00
	Area-wide	-1.44	0.73	0.05
Area-wide	Zonal Priority	4.83	0.68	0.00
	Zonal	1.44	0.73	0.05

Table 24 shows the ANOVA of speed estimation accuracy for the two levels of map matching error. The 10 m map matching error proved to be more accurate than the 16 m map matching error. In case of 16 m map matching error, more invalid vehicle positions were wrongly matched to the network, resulting in poor speed estimation.

Table 24. Tysons Corner Case Study ANOVA Results for Speed Estimation Accuracy by Map Matching Error

Map Matching Error	Mean Speed Estimation Accuracy (mph)	Standard Error	F	Significance
10 m	6.21	0.42	54.0	0.0
16 m	10.35	0.36		

Table 25 shows the analysis of the interaction of sampling strategy and map matching error. Both the zonal and area-wide approaches showed significantly different performance between the two map matching levels, with the 16-m error producing worse speed estimates. In case of zonal and area-wide sampling with 16-m map-matching error, vehicles from low-speed links were erroneously matched to nearby high-speed links.

Table 25. Tysons Corner Case Study ANOVA Results for Speed Estimation Accuracy by Map Matching Error and Sampling Strategy

Sampling Strategy	Map Matching Error	Mean	Std. Error	Significantly Different From
Zonal Priority	10 m	4.996	0.606	None
	16 m	6.083	0.606	None
Zonal	10 m	6.933	0.796	16 m
	16 m	10.947	0.624	10 m
Area-wide	10 m	6.715	0.811	16 m
	16 m	14.036	0.658	10 m

Table 26 shows the ANOVA speed estimation accuracy results for the interaction of zone number by sampling strategy. On an aggregate basis, the results show a statistically significant difference among sampling strategies in zone 2 only. Zone 2 was an uncongested area with high-speed freeways. The area-wide approach produced few samples from low volume links in zone 2, resulting in speed errors as high as 30 to 50 mph. The zonal strategy suffered because of short vehicle tracks that generated speed estimates on high-speed links. In case of zone 3, all sampling strategies generated high-speed errors, especially on links on the Dulles Toll Road (DTR). Vehicles from intersections and collector distributor roads where wrongly matched to DTR, and these incorrect map matches resulted in high-speed errors for all sampling strategies.

Table 26. Tysons Corner Case Study ANOVA Results for Speed Estimation Accuracy by Sampling Strategy and Zone

Zone	Sampling Strategy	Mean Speed Estimation Accuracy (mph)	Std. Error	Significantly Different From
1	Zonal Priority	2.26	0.90	None
	Zonal	2.24	0.90	None
	Area	2.20	0.97	None
2	Zonal Priority	6.10	0.98	Zonal, Area-wide
	Zonal	20.60	1.40	Area-wide, Zonal Priority
	Area	27.69	1.88	Zonal Priority, Zonal
3	Zonal Priority	13.37	0.68	None
	Zonal	15.38	0.89	None
	Area	13.63	0.71	None
4	Zonal Priority	1.80	1.27	None
	Zonal	1.99	1.31	None
	Area	3.50	1.06	None
5	Zonal Priority	4.17	0.87	None
	Zonal	4.49	1.06	None
	Area-wide	4.87	0.85	None

Table 27 shows the analysis of speed estimation accuracy by speed calculation method and zone number, aggregated across sampling strategy. Zone 5 was the only zone which exhibited a statistically significant difference between the two methods, with the average speed method producing estimates that were 3.6 mph more accurate than the distance-weighted speed method. The distance-weighted method could not account for the waiting time caused by signalized intersections and congestion at the interchanges.

Table 27. Tysons Corner Case Study ANOVA Results for Speed Estimation Accuracy by Speed Method and Zone

Zone Number	Speed Method	Mean Speed Estimation Accuracy (mph)	Standard Error	Significantly Different From
1	Average Speed	2.04	0.76	None
	Distance Weighted Speed	2.42	0.75	None
2	Average Speed	19.38	1.20	None
	Distance Weighted Speed	16.88	1.20	None
3	Average Speed	13.89	0.63	None
	Distance Weighted Speed	14.37	0.62	None
4	Average Speed	1.85	0.99	None
	Distance Weighted Speed	3.01	0.99	None
5	Average Speed	2.71	0.76	Distance Weighted Speed
	Distance Weighted Speed	6.31	0.76	Average Speed

Much like the findings for the number of speed estimates, the significance of the zone number main effect on the accuracy of the estimates appears to be driven by the sampling requirements that are implicit in the zonal systems. The zone number findings are not discussed in detail since the significance was purely a function of the zonal systems distributing samples across the network. Likewise, the three-factor interaction of sampling strategy, map matching error, and zone number does not provide any insight beyond what was seen in the two-factor interactions. The interaction of zone number by map matching error also appeared to be an extension of the significance of those two main effects, with no other obvious trends apparent.

The distribution of speed errors was calculated by determining the percentage of 5-minute intervals that each mile that was being monitored was within the desired accuracy threshold over the 60 minute analysis period. Tables C-3, C-4, and C-5 in Appendix C show the percentage of monitored miles that were determined to be within certain accuracy thresholds over the analysis period for each sampling strategy.

Summary of Tysons Corner Case Study Results

The results from the Tysons Corner case were consistent with the two prior case studies. The zonal priority strategy produced more speed estimates than other methods. The zonal priority method generated longer vehicle tracks than the zonal method, and more evenly distributed samples geographically than the area-wide method. Both zonal sampling strategies distributed samples throughout the network, especially on low volume and uncongested links in zones 1, 2 and 3. The short tracks produced by the zonal system did create issues with speed estimation, however. The higher map matching error seemed to create problems in the map matching algorithms, causing positions to be matching to the wrong roads. This was particularly problematic at intersections and at locations where high-speed roads paralleled lower speed facilities.

Generally speaking, the zonal priority systems appeared to produce the best results on the Tysons Corner network. Probes were more evenly distributed throughout the network and vehicle tracks were relatively long. Both of these factors helped enable the zonal priority system to produce better estimates of speeds. The data also suggest that the distance-weighted speed methodology does not produce accurate speed estimates on arterial streets.

DISCUSSION

Table 28 shows the p-values from the GLM analysis of speed estimation accuracy for all of the case studies analyzed. Some common trends become apparent when the case studies are examined collectively. First, the sampling strategy was a significant factor in the accuracy of speed estimates for all three cases studies. The zonal priority method was generally found to perform as well as or better than the other methods. It produced long vehicle tracks while simultaneously ensuring that samples were distributed throughout the network. Zone number was also found to be significant factor in overall system accuracy, indicating the importance of localized geometry and traffic conditions in determining the effectiveness of the system. The WLT-based approach functions well in zones that were relatively simple geometrically, but performance degrades if there were a number of parallel facilities or intersections.

Table 28. Speed Estimation Accuracy p-Values for All Case Studies

Number of Factors	Factor Name(s)	Charlottesville	Springfield	Tysons Corner
Main Effects	Sampling Strategy	0.00	0.00	0.00
	Speed Method	0.99	0.00	0.27
	Map Matching Error	0.99	0.03	0.00
	Zone Number	0.00	0.00	0.00
2-Factor Interactions	Sampling Strategy × Map Matching Error	0.87	0.44	0.00
	Sampling Strategy × Zone Number	0.00	0.00	0.00
	Speed Method × Zone Number	0.46	0.00	0.00
	Map Matching Error × Zone Number	0.97	0.48	0.00
3 Factor Interactions	Sampling Strategy × Map Matching Error × Zone Number	0.99	0.95	0.00

The impact of the other factors that were investigated varied somewhat among the case studies. More complex networks like Springfield and Tysons Corner experienced significant effects due to speed method and map matching error that were not observed in the relatively simple Charlottesville case. Both of the more complex cases showed that higher levels of map matching error produced more incorrect position matches, resulting in worse speed estimates. Likewise, the speed method factor showed an impact depending on the zone being analyzed. The distance weighted method generally performed better when there were closely spaced parallel facilities with different traffic characteristics. The distance weighted method often served to minimize the contribution of mismatched vehicles in these situations, helping to improve the overall quality of the speed estimates generated. The average speed method usually produced more accurate results for zones with a large number of signalized arterials or links that had a large speed variance.

These case studies indicate that it is inadvisable to treat sampling the same across a large roadway network. The zonal priority method effectively addresses the issue of geographic probe distribution, but the characteristics of the roadway network still exert a powerful impact on the effectiveness of all simulated WLT-based systems. Closely spaced, parallel facilities with different speeds proved to be particularly problematic. Likewise, the method of calculating link speed can be an important factor in generating accurate speed estimates. It seems clear that local, site-specific conditions will need to be considered when designing and applying WLT-based systems.

CONCLUSIONS

- *Both the availability and accuracy of speed estimates are influenced by the sampling approach used by a WLT-based system.* The results of the case studies show that the zonal priority approach provided speed estimates that were as good as or better than other approaches. When designing a WLT-based system, the sampling method used should ensure that probe vehicles are tracked across all roads that are being monitored. The zonal priority system helped ensure that routes were monitored, and also improved the accuracy of speed estimates by creating long vehicle tracks.

- *The geometric and traffic characteristics of the roadway network play an extremely important role in dictating the accuracy of speed estimates.* All sampling methods performed relatively well on the simple Charlottesville network. The results indicated, however, that the accuracy of WLT-based monitoring is likely to be lower when network geometry is complex or traffic conditions show a lot of variability. While zonal sampling methods can improve availability of probe samples on lower volume roads, available map matching techniques still have limitations when intersections are frequent or parallel routes exist. When high-speed and low-speed facilities exist in close proximity, these matching problems can significantly impact the accuracy of speed estimates.

- *The method used to calculate speeds can be important when networks are complex or congested.* The data from the Springfield and Tysons Corner case studies show that the distance-weighted method provides a level of screening on high-speed facilities, which helps produce more accurate results. The average speed method generally worked better on signalized facilities since it better accounts for control delay.

- *Increasing the map matching error generally degraded the quality of speed estimates.* It was initially hypothesized that a larger map matching error would create better speed estimates by including the vehicles on the inner and outer lanes of multilane freeways. In fact, the larger map matching error results in more erroneous matches, producing worse speed estimates. This effect was most profound on networks with complex geometry and many intersections.

RECOMMENDATIONS

1. *Prior to entering into any agreements with probe-based traffic data service providers, VDOT's Operations & Security Division (OSD) should clearly understand the sampling*

scheme being used by a WLT vendor. Early generation WLT-based systems used an area-wide approach to sampling. Even if a confidentiality agreement is required, the OSD should be sure that they understand clearly how wireless probes will be tracked in any prospective system. Area-wide approaches do not appear to be promising, and the OSD should ensure that the vendor's approach will ensure probe availability on the roads that VDOT wants to monitor.

2. *Prior to entering into any agreements with probe-based traffic data service providers, VDOT's OSD should discuss the roadway network characteristics in detail with a vendor.* The results of this study show that performance can be significantly impacted by the roadway geometry and traffic characteristics. The OSD should be sure to engage the vendor to explain the network in detail, and verify that the vendor will be able to monitor all links of interest.

3. *Prior to entering into any agreements with probe-based traffic data service providers, VDOT's OSD should be sure to understand how the speeds will be derived on a roadway link.* The results of this study show that the average speed and distance-weighted methods both had an area of application. The OSD should be sure that they understand exactly how link speeds will be calculated and be prepared to question the vendor's approach.

COSTS AND BENEFITS ASSESSMENT

If commercial WLT-based monitoring systems can be shown to be reliable, they offer the opportunity to provide significant cost reductions to VDOT over traditional traffic monitoring technology. The Missouri DOT recently embarked on an ambitious program to monitor their entire interstate and primary system using a WLT-based monitoring system.[16] They state that the annual cost per-mile for the WLT-based system will be about $560 per mile, as opposed to $6,000 for an inductive loop detector-based system. Those numbers are speculative as the Missouri WLT-based monitoring system is not yet operational, so it is unclear if these numbers are realistic. If they prove to hold, however, VDOT would stand to save a considerable amount of money on its traffic monitoring expenses. Improving the sampling methods used and probe availability over the network would help make it more likely that a functional, commercial level system could be deployed.

REFERENCES

1. Srinivasan, K.K., and Jovanis, P.P. Determination of Number of Probe Vehicles Required for Reliable Travel Time Measurement in Urban Network. In *Transportation Research Record 1537*. Transportation Research Board, Washington, D.C., 1996, pp. 15-22.

2. Chakroborty, P., and Kikuchi, S. Using Bus Travel Time Data to Estimate Travel Times on Urban Corridors. In *Transportation Research Record* 1870. Transportation Research Board, Washington, D.C., 2004, pp. 18-25.

3. Fontaine, M.D., and Smith, B.L. *Improving the Effectiveness of Traffic Monitoring Based on Wireless Location Technology*. VTRC 05-17. Virginia Transportation Research Council, Charlottesville, 2005.

4. Misener, A. J. Vehicle-Infrastructure Integration and Safety: Rubber and Radio Meets the Road in California. *Intellimotion*, Vol. 11, No. 2, Fall 2005, pp. 1-3.

5. Smith, B.L., Zhang, H., Fontaine, M.D., and Green, M.W. Wireless Location Technology-Based Traffic Monitoring: Critical Assessment and Evaluation of an Early-Generation System. *Journal of Transportation Engineering*. Vol. 130, No. 5, September/October 2004, pp. 576-584.

6. University of Maryland Transportation Studies Center. *Final Evaluation Report for the CAPITAL-ITS Operational Test and Demonstration Program.* University of Maryland, College Park, 1997.

7. Yim, Y.B.Y., and Cayford, R. *Investigation of Vehicles as Probes Using Global Positioning System and Cellular Phone Tracking: Field Operational Test.* UCB-ITS-PWP-2001-9. California PATH Program, Institute of Transportation Studies, University of California, Berkeley, 2001.

8. Pyo, J.S., Shin, D.H., and Sung, T.K. Development of a Map Matching Method Using the Multiple Hypothesis Technique. In *IEEE Intelligent Transportation Systems Conference Proceedings.* Oakland, Calif., 2001, pp. 23-27.

9. Green, M.W., Fontaine, M.D., and Smith, B.L. Investigation of Dynamic Probe Sample Requirements for Traffic Condition Monitoring. In *Transportation Research Record 1870,* Transportation Research Board, Washington, D.C., 2004, pp. 55-61.

10. Hendricks, A.R., Fontaine, M.D., and Smith, B.L. *Probe-based Traffic Monitoring: State-of-the-Practice Report.* National Cooperative Highway Research Program, Washington, D.C., 2005. http://www4.trb.org/trb/crp.nsf/All+Projects/NCHRP+70-01. Accessed January 8, 2007.

11. Kauffman, J. *Wireless 9-1-1 Geolocation: A New Way to Save Lives NOW.* NENA News, Summer 2001. http://www.nena.org/PR_Pubs/enpm_archives/Kauffman_06-01.pdf. Accessed September 26, 2005.

12. Smith, B.L., Pack, M.L., Lovell, D.J., and Sermons, M.W. Transportation Management Applications of Anonymous Mobile Call Sampling. Transportation Research Board 80th Annual Meeting CD-ROM. Transportation Research Board, Washington, D.C., 2001.

13. University of Virginia Center for Transportation Studies. *Wireless Location Technology-Based Traffic Monitoring Demonstration and Evaluation Project.* Charlottesville, 2006.

14. Ygnace, J-L, Remy, J-G, Bosseboeuf, J-L, and Da Fonseca, V. *Travel Time Estimates on Rhone Corridor Network Using Cellular Phones as Probes: Phase 1 Technology Assessment and Preliminary Results.* French Department of Transportation, Arceuil, 2000.

15. Cayford, R., and Johnson, T. Operational Parameters Affecting the Use of Anonymous Cell Phone Tracking for Generating Traffic Information. Transportation Research Board 82nd Annual Meeting CD-ROM. Transportation Research Board, Washington, D.C., 2003.

16. Teel, M. *Missouri DOT's Statewide Traffic Data and Traveler Information Project.* North American Traffic Monitoring Exposition and Conference, Minneapolis, Minn., August 2006.

APPENDIX A
CHARLOTTESVILLE CASE STUDY

Table A-1. Charlottesville Case Study GLM Analysis Results for Number of Speed Estimates

Source	**F**	**Significance**
Corrected Model	21.88	0.00
Intercept	2209.94	0.00
Sampling Strategy	1.87	0.15
Speed Method	0.00	1.00
Map Matching Error	1.82	0.18
Zone Number	494.10	0.00
Sampling Strategy × Speed Method	0.00	1.00
Sampling Strategy × Map Matching Error	0.10	0.91
Speed Method × Map Matching Error	0.00	1.00
Sampling Strategy × Speed Method × Map Matching Error	0.00	1.00
Sampling Strategy × Zone Number	3.17	0.04
Speed Method× Zone Number	0.00	1.00
Strategy × Speed × Zone Number	0.00	1.00
Map Matching Error × Zone Number	0.34	0.56
Sampling Strategy × Map Matching Error × Zone Number	0.07	0.94
Speed Method× Map Matching Error × Zone Number	0.00	1.00
Sampling Strategy × Speed Method ×Map Matching Error × Zone Number	0.00	1.00

Table A-2. Charlottesville Case Study GLM Analysis Results for Speed Estimation Accuracy

Source	**F**	**Significance**
Corrected Model	9.35	0.00
Intercept	394.15	0.00
Sampling Strategy	7.44	0.00
Speed Method	0.00	0.99
Map Matching Error	0.00	0.99
Zone Number	124.51	0.00
Sampling Strategy × Speed Method	0.23	0.80
Sampling Strategy × Map Matching Error	0.14	0.87
Speed Method × Map Matching Error	0.00	0.97
Sampling Strategy × Speed Method × Map Matching Error	0.00	1.00
Sampling Strategy × Zone Number	9.80	0.00
Speed Method× Zone Number	0.54	0.46
Strategy × Speed × Zone Number	1.02	0.36
Map Matching Error × Zone Number	0.00	0.97
Sampling Strategy × Map Matching Error × Zone Number	0.15	0.86
Speed Method× Map Matching Error × Zone Number	0.00	0.99
Sampling Strategy × Speed Method ×Map Matching Error × Zone Number	0.01	0.99

Table A-3. Charlottesville Case Study Percentage of Monitored Miles within Accuracy Thresholds by Speed Method and Matching Error, Zonal Priority Strategy Only

Accuracy Threshold	Zone	Average Speed		Distance Weighted Speed	
		10m	16m	10m	16m
Within 2 mph	Zone 1	18.70	21.03	12.84	17.33
	Zone 2	13.52	9.09	40.34	40.34
	Network	16.91	16.90	22.34	25.28
Within 5 mph	Zone 1	42.19	39.84	38.83	40.52
	Zone 2	63.18	68.07	86.02	86.02
	Network	49.44	49.60	55.14	56.25
Within 10 mph	Zone 1	59.82	61.21	52.85	58.68
	Zone 2	100.00	100.00	95.34	95.34
	Network	73.71	74.62	67.53	71.35

Table A-4. Charlottesville Case Study Percentage of Monitored Miles within Accuracy Thresholds by Speed Method and Matching Error, Zonal Strategy Only

Accuracy Threshold	Zone	Average Speed		Distance Weighted Speed	
		10m	16m	10m	16m
Within 2 mph	Zone 1	23.33	23.55	12.83	12.56
	Zone 2	4.43	8.86	62.95	58.52
	Network	16.80	18.47	30.16	28.45
Within 5 mph	Zone 1	42.03	45.14	35.52	33.25
	Zone 2	63.41	72.50	95.34	90.68
	Network	49.42	54.60	56.20	53.10
Within 10 mph	Zone 1	58.23	58.30	53.03	51.87
	Zone 2	95.57	90.91	95.34	95.34
	Network	71.13	69.57	67.66	66.89

Table A-5. Charlottesville Case Study Percentage of Monitored Miles within Accuracy Thresholds by Speed Method and Matching Error, Area-wide Strategy Only

Accuracy Threshold	Zone	Average Speed		Distance Weighted Speed	
		10m	16m	10m	16m
Within 2 mph	Zone 1	21.13	20.97	20.32	14.47
	Zone 2	59.66	50.34	26.59	31.25
	Network	34.45	31.12	22.49	20.27
Within 5 mph	Zone 1	35.57	34.10	38.14	36.62
	Zone 2	90.91	86.48	58.06	58.06
	Network	54.69	52.20	45.03	44.03
Within 10 mph	Zone 1	52.79	51.40	55.28	53.89
	Zone 2	100.00	100.00	95.34	95.34
	Network	69.10	68.20	69.13	68.22

APPENDIX B
SPRINGFIELD CASE STUDY

Table B-1. Springfield Case Study GLM Analysis Results for Number of Speed Estimates

Source	F	Significance
Corrected Model	14.95	0.00
Intercept	4,142.85	0.00
Sampling Strategy	4.12	0.01
Speed Method	0.00	1.0
Map Matching Error	46.84	0.00
Zone Number	158.84	0.00
Sampling Strategy × Speed Method	0.00	1.00
Sampling Strategy × Map Matching Error	0.13	0.87
Speed Method × Map Matching Error	0.00	1.00
Sampling Strategy × Speed Method × Map Matching Error	0.00	1.00
Sampling Strategy × Zone Number	28.93	0.00
Speed Method× Zone Number	0.00	1.00
Strategy × Speed × Zone Number	0.00	1.00
Map Matching Error × Zone Number	4.31	0.00
Sampling Strategy × Map Matching Error × Zone Number	0.19	0.97
Speed Method× Map Matching Error × Zone Number	0.00	1.00
Sampling Strategy × Speed Method ×Map Matching Error × Zone Number	0.00	1.00

Table B-2. Springfield Case Study GLM Analysis Results for Speed Estimation Accuracy

Source	F	Significance
Corrected Model	13.16	0.00
Intercept	2929.43	0.00
Sampling Strategy	5.91	0.00
Speed Method	46.59	0.00
Map Matching Error	4.70	0.03
Zone Number	127.58	0.00
Sampling Strategy × Speed Method	0.41	0.66
Sampling Strategy × Map Matching Error	0.82	0.44
Speed Method × Map Matching Error	0.30	0.58
Sampling Strategy × Speed Method × Map Matching Error	0.09	0.92
Sampling Strategy × Zone Number	3.25	0.00
Speed Method× Zone Number	60.74	0.00
Strategy × Speed × Zone Number	0.97	0.44
Map Matching Error × Zone Number	0.82	0.48
Sampling Strategy × Map Matching Error × Zone Number	0.18	0.98
Speed Method× Map Matching Error × Zone Number	0.13	0.95
Sampling Strategy × Speed Method ×Map Matching Error × Zone Number	0.08	1.00

Table B-3. Springfield Case Study Percentage of Monitored Miles within Accuracy Thresholds by Speed Method and Matching Error, Zonal Priority Strategy Only

Accuracy Threshold	Zone	Average Speed		Distance Weighted Speed	
		10m	16m	10 m	16 m
Within 2 mph	Zone 1	40.01	38.00	27.30	20.38
	Zone 2	45.86	45.86	15.80	20.66
	Zone 3	27.20	26.21	13.69	9.89
	Zone 4	41.41	41.49	36.58	35.10
	Network	36.71	35.94	23.63	20.98
Within 5 mph	Zone 1	59.04	56.72	65.41	60.31
	Zone 2	47.54	47.54	57.11	52.74
	Zone 3	62.29	59.64	32.89	31.23
	Zone 4	61.66	58.68	55.37	57.42
	Network	59.36	57.06	49.95	48.22
Within 10 mph	Zone 1	67.19	64.71	74.20	68.30
	Zone 2	49.28	51.01	76.41	65.03
	Zone 3	79.04	76.65	50.11	49.12
	Zone 4	74.45	73.38	73.10	72.03
	Network	71.00	69.53	65.72	62.19

Table B-4. Springfield Case Study Percentage of Monitored Miles within Accuracy Thresholds by Speed Method and Matching Error, Zonal Strategy Only

Accuracy Threshold	Zone	Average Speed		Distance Weighted Speed	
		10 m	16 m	10 m	16 m
Within 2 mph	Zone 1	39.45	37.01	21.75	20.75
	Zone 2	46.70	45.86	24.88	21.86
	Zone 3	22.08	23.61	15.95	15.04
	Zone 4	42.00	37.65	39.51	37.05
	Network	35.07	33.70	25.27	23.60
Within 5 mph	Zone 1	50.84	49.76	48.48	45.27
	Zone 2	47.54	47.54	64.06	52.35
	Zone 3	60.64	58.95	31.89	30.64
	Zone 4	54.48	55.11	59.33	56.35
	Network	54.89	54.23	47.91	44.30
Within 10 mph	Zone 1	59.93	57.54	58.53	58.22
	Zone 2	47.54	47.54	71.09	69.80
	Zone 3	78.70	76.99	48.41	46.29
	Zone 4	73.14	76.12	71.93	71.07
	Network	68.65	68.38	60.56	59.32

Table B-5. Springfield Case Study Percentage of Monitored Miles within Accuracy Thresholds by Speed Method and Matching Error, Area-wide Strategy Only

Accuracy Threshold	Zone	Average Speed		Distance Weighted Speed	
		10 m	16 m	10 m	16 m
Within 2 mph	Zone 1	37.27	37.97	31.19	23.69
	Zone 2	45.02	45.02	12.21	14.16
	Zone 3	25.95	28.73	12.61	16.28
	Zone 4	44.36	43.00	33.60	35.13
	Network	36.40	37.14	22.76	23.08
Within 5 mph	Zone 1	50.38	60.00	65.19	62.06
	Zone 2	47.54	47.54	51.44	51.89
	Zone 3	59.19	59.87	29.51	32.13
	Zone 4	67.56	61.71	58.77	54.32
	Network	58.05	58.74	48.91	47.91
Within 10 mph	Zone 1	62.39	80.59	64.80	70.88
	Zone 2	49.28	70.77	49.28	66.32
	Zone 3	77.77	49.79	79.55	51.54
	Zone 4	75.97	71.35	72.03	71.21
	Network	69.92	65.75	69.95	63.55

APPENDIX C
TYSONS CORNER CASE STUDY

Table C-1. Tysons Corner Case Study GLM Analysis Results for Number of Speed Estimates

Source	F	Significance
Corrected Model	12.61	0.00
Intercept	1200.73	0.00
Sampling Strategy	5.35	0.01
Speed Method	0.00	0.98
Map Matching Error	2.98	0.08
Zone Number	126.85	0.00
Sampling Strategy × Speed Method	0.00	1.00
Sampling Strategy × Map Matching Error	8.20	0.00
Speed Method × Map Matching Error	0.00	0.98
Sampling Strategy × Speed Method × Map Matching Error	0.00	1.00
Sampling Strategy × Zone Number	17.90	0.00
Speed Method× Zone Number	0.03	1.00
Strategy × Speed × Zone Number	0.03	1.00
Map Matching Error × Zone Number	2.09	0.08
Sampling Strategy × Map Matching Error × Zone Number	4.07	0.00
Speed Method× Map Matching Error × Zone Number	0.03	1.00
Sampling Strategy × Speed Method ×Map Matching Error × Zone Number	0.03	1.00

Table C-2. Tysons Corner Case Study GLM Analysis Results for Speed Estimation Accuracy

Source	F	Significance
Corrected Model	16.23	0.00
Intercept	867.71	0.00
Sampling Strategy	28.53	0.00
Speed Method	1.22	0.27
Map Matching Error	54.18	0.00
Zone Number	139.23	0.00
Sampling Strategy × Speed Method	0.68	0.51
Sampling Strategy × Map Matching Error	10.72	0.00
Speed Method × Map Matching Error	0.58	0.45
Sampling Strategy × Speed Method × Map Matching Error	0.48	0.62
Sampling Strategy × Zone Number	14.98	0.00
Speed Method× Zone Number	2.71	0.03
Strategy × Speed × Zone Number	0.90	0.51
Map Matching Error × Zone Number	24.17	0.00
Sampling Strategy × Map Matching Error × Zone Number	5.75	0.00
Speed Method× Map Matching Error × Zone Number	0.31	0.87
Sampling Strategy × Speed Method ×Map Matching Error × Zone Number	0.14	1.00

Table C-3. Tysons Corner Case Study Percentage of Monitored Miles within Accuracy Thresholds by Speed Method and Matching Error, Zonal Priority Strategy Only

Accuracy Threshold	Zone	Average Speed		Distance Weighted Speed	
		10m	16m	10m	16m
With 2 mph	Zone 1	56.20	57.67	45.98	45.98
	Zone 2	16.81	16.74	22.98	20.87
	Zone 3	35.78	27.63	32.75	31.01
	Zone 4	32.32	30.32	35.96	33.97
	Zone 5	50.72	54.99	39.04	37.30
	Network	38.68	36.56	35.28	33.78
Within 5 mph	Zone 1	82.67	85.89	79.45	74.19
	Zone 2	44.00	52.45	64.79	58.60
	Zone 3	49.69	49.69	47.89	47.89
	Zone 4	49.80	49.68	53.55	49.68
	Zone 5	93.65	93.82	54.09	54.63
	Network	62.31	64.24	58.38	55.93
Within 10 mph	Zone 1	85.89	85.89	82.67	87.06
	Zone 2	75.18	87.66	81.46	71.17
	Zone 3	60.58	59.47	64.33	64.91
	Zone 4	63.30	57.31	63.30	61.42
	Zone 5	98.42	98.42	73.09	71.64
	Network	74.26	74.90	71.75	70.68

Table C-4. Tysons Case Study Percentage of Monitored Miles within Accuracy Thresholds by Speed Method and Matching Error, Zonal Strategy Only

Accuracy Threshold	Zone	Average Speed		Distance Weighted Speed	
		10m	16m	10m	16m
Within 2 mph	Zone 1	58.94	55.72	46.61	41.37
	Zone 2	4.29	6.40	8.52	8.52
	Zone 3	12.41	25.87	8.44	20.49
	Zone 4	21.32	36.65	19.32	52.15
	Zone 5	26.07	53.07	11.62	33.81
	Network	23.62	34.64	17.90	29.58
Within 5 mph	Zone 1	79.45	79.45	73.01	76.23
	Zone 2	6.40	14.87	23.34	21.22
	Zone 3	18.99	35.75	16.49	29.24
	Zone 4	33.17	63.77	35.04	63.77
	Zone 5	39.99	87.35	26.82	65.15
	Network	34.16	53.54	32.82	48.05
Within 10 mph	Zone 1	85.89	85.89	85.89	85.89
	Zone 2	25.39	23.22	29.62	27.45
	Zone 3	24.34	42.56	24.61	46.08
	Zone 4	38.80	71.28	38.80	71.28
	Zone 5	48.42	96.85	36.34	78.09
	Network	42.37	61.06	41.14	59.84

Table C-5. Tysons Corner Case Study Percentage of Monitored Miles within Accuracy Thresholds by Speed Method and Matching Error, Area-wide Strategy Only

Accuracy Threshold	Zone	Average Speed		Distance Weighted Speed	
		10m	16m	10m	16m
Within 2 mph	Zone 1	63.12	52.52	51.62	41.73
	Zone 2	0.00	0.00	2.12	2.12
	Zone 3	34.84	31.76	23.76	24.59
	Zone 4	53.79	51.91	3.87	5.75
	Zone 5	50.37	48.53	26.33	31.34
	Network	40.27	36.61	23.17	22.67
Within 5 mph	Zone 1	80.21	76.99	77.48	70.55
	Zone 2	0.00	0.00	6.23	4.12
	Zone 3	49.05	44.46	41.01	36.42
	Zone 4	71.16	71.16	36.31	40.29
	Zone 5	88.53	85.24	51.15	49.44
	Network	57.23	54.51	43.57	40.67
Within 10 mph	Zone 1	83.44	79.72	83.44	79.72
	Zone 2	2.12	2.12	10.47	8.35
	Zone 3	59.47	55.51	59.74	54.35
	Zone 4	73.16	73.16	61.43	57.90
	Zone 5	96.85	93.56	74.53	71.65
	Network	63.39	60.79	59.35	55.47