Exploring the Effects of Sensor Data Aggregation on Measuring Arterial Performance

Paper 13-2707

Michael Wolfe
4403 SE 37th Ave
Portland, OR 97202
Phone: 503-709-9614
Email: cyclotouriste@gmail.com

Christopher Monsere (*corresponding author)
Department of Civil and Environmental Engineering
Portland State University
P.O. Box 751
Portland, OR 97207-0751
Email: monsere@pdx.edu
Phone: 503-725-9746
Fax: 503-725-5950

Robert L. Bertini
Department of Civil & Environmental Engineering
Portland State University
P.O. Box 751
Portland, OR 97207-0751
Phone: 503-725-4282
Fax: 503-725-5950
Email: bertini@pdx.edu

Submitted for presentation and publication to the
92nd Annual Meeting of the Transportation Research Board
January 13-17, 2013

Revised November 15, 2012

5,249 words + 5 figures x 250 + 4 tables x 250 = 7,499 words
ABSTRACT

Modern loop detector technology offers operations staff a wealth of data about their facilities. It should be possible to use these data to monitor, analyze, and communicate an arterial's performance, but dealing with gaps in the detector network, aggregated vehicle counts, and lags in receiving data can pose a challenge. This paper proposes six methods of using data from loop detectors to derive useful performance measures. A simulated arterial is employed to investigate each method's accuracy under various conditions. Estimates are compared to simulated data visually, with input/output diagrams; and statistically, with simulated ground truth travel times. Methods for estimating travel time are applied to aggregated data and to varying detector densities. It is found that data from detectors combined with information about signal timing, saturation headways, and free flow travel times can be used by two of the methods described in the paper to provide accurate and useful estimates of average travel time and average vehicle delay, even in conditions where detectors are missing from intersections or detector data are aggregated.
INTRODUCTION

Methods for measuring the performance of limited access facilities (i.e., freeways) have been broadly implemented. However, in the U.S., over 40% of vehicle miles are traveled on arterials (1). This leaves a large gap in the information that is available about transportation system performance. Arterials present far more options for travelers in choosing a route, so information about the relative levels of congestion of alternatives would be quite valuable to users. Furthermore, making archived arterial performance data available to researchers and planners would have substantial benefits.

Newer probe-based and crowd-sourced data are available in many large urban areas but fixed sensors (e.g. system loop detectors) are common and are more readily associated with traffic signal operations. Using inductive loop detectors to monitor vehicle performance on arterials is complicated by many factors. Most prominently, there is no vehicle conservation between detectors, and traffic control devices (i.e., signals) are designed to interrupt the flow of vehicles. Couple these considerations with the realities of sparsely installed loop detectors and detector management software that typically only makes data available in aggregate, and the challenges of generating arterial performance information from these data are apparent.

The objective of this paper is to present methods to measure arterial performance using existing detector infrastructure. While there are many metrics to measure arterial performance, this paper considers vehicle travel time and delay, as many other measures can be derived from them. Methods to measure travel time and delay are proposed and compared. The effects of varying densities of detectors and levels of data aggregation are investigated.

Previous research has addressed the problem of assessing freeway performance and schemes for doing so have been implemented widely around the world. The consensus appears to be that using loop detectors to measure speed and extrapolating that speed to areas not covered by detectors is the best method for measuring freeway performance – as shown, for example, in a study of I-4 in Florida (2).

The complexity of arterial operations calls into question the validity of extrapolating speed measurements over arterials, however. One study at the University of Washington examined the feasibility of using stop bar detectors to monitor arterial performance (3). The authors tested several variables, including detector placement, traffic volume, and site specific factors (such as turning volumes and signal coordination). It was found that stop bar detectors taking readings during the green and amber phases yielded good information on the amount of congestion on the arterial. This supports the idea that data from stop bar detectors can be used to generate relevant information, but does not suggest how one might compute delay, speed, or other traditional measures from such data. The City of Bellevue developed a method to communicate arterial performance simply by displaying colors that correspond to occupancy times on its detectors (4). This method was satisfactory as a traveler information tool and for use in the operations center, but required site-specific calibration and, again, lacked the ability to be quantified.

A study in Singapore attempted to use speeds measured by loop detectors at signalized intersections to estimate travel time (5). The model required no site-specific calibration, using measured speeds and signal timing as its inputs. The algorithm was tested in a simulation, and found to be highly accurate when detectors were closely spaced.

One paper attempted to leverage transit vehicles as probes to supplement loop detectors as sources of arterial performance information (6). In that study, two approaches were tried. In the first, loop detectors were used much as they are on freeways: speeds were measured and then extrapolated halfway to the next detector location. This yielded implausible results, in part because detector density (4 detectors on a 4.5 mile segment of arterial) was low. The second method limited the area that the detectors described to the segment of roadway immediately around them and used transit vehicles as probes to fill in the gaps. This method represented an improvement, but there was reason to doubt the accuracy of
speeds measured at the detector, and to question the usefulness of data that was aggregated into 5 minute bins.

A common approach to measure arterial performance has been to break the problem of travel time into discrete delay components. Typically, these have included signal delay, accelerations/decelerations at signals, shockwave propagation, cycle failure, and platoon interactions. The interactions of vehicles between signals have been computed using platoon dispersion algorithms or by iterative use of “virtual probes” (7,8). Arterial performance measurement methods of this type require calibration of a number of site-specific parameters, employ a great deal of computation, and are not effective when supplied with aggregated data, which makes them difficult to employ on a widespread basis.

Another paper investigated two methods for estimating signal delay using inductive loop detector data: the Input-Output (IO) technique and the Hybrid technique (9). Both counted the number of vehicles entering the queue at the signal with an advance detector, and both computed their estimates based on available queue storage and signal phase data. In the IO method, queue egress was estimated by using the intersection’s saturation headway. The Hybrid method used a stop bar detector and vehicle signature identification to more directly measure queue egress. Delay was calculated by superimposing the arrival and departure profiles and finding the area under the curve. These techniques again required event based data. The IO technique performed even better than the Hybrid technique in this study, suggesting its default use in this application, though the Hybrid technique would be useful in situations where saturation headway tended to change due to weather, or where there were a large number of access points between the advance detector and the stop bar.

The remainder of this paper is organized as follows. First, the site and data used to create the simulation are described, along with the simulation’s calibration and the data extracted from it. Then the methods under investigation for generating travel time information from the data are presented. An analysis comparing these methods when applied to the simulation data follows. First with disaggregated data and one detector per stop bar then with varying levels of data aggregation and detector density. Finally, the findings are summarized in the conclusion.

SITE DESCRIPTION AND DATA

The study location is a 3-mile segment of East Burnside Road in Gresham, Oregon between Eastman Parkway and Kelly Avenue (see Figure 1). Burnside Road is a major 5-lane arterial carrying 38,000 vehicles per day. The corridor features 11 instrumented signalized intersections and is currently running TransCore’s SCATS adaptive signal system. All intersections in the study area are signalized, though there are numerous other access points. Data collected by GPS units in probe vehicles from a previous study and data archived by the SCATS system from the same time period were used to create a VISSIM simulation of the corridor (10). In order to make the simulation as authentic as possible, it was created and calibrated with extensive data from the actual arterial corridor. These data included signal timing information, cycle-by-cycle traffic volumes, and GPS data from probe vehicles.

Pseudo "ground-truth" trajectories were generated by the simulation for comparison purposes. Data used to estimate performance were limited to what could be gleaned from inductive loop detectors in the field, including vehicle flows, occupancy (percentage of time a vehicle is detected), and speed. Data for estimates were aggregated at different intervals, including event based, per cycle, and per 5 minutes. Finally, for determining the effects of detector density, detectors were placed 1 per segment (at stop bars), 1 per 2 segments (at stop bars), and 3 per segment (at the stop bar, just past the intersection, and at the back of the signal queue, 200’ from the stop bar). Data were collected for 1 simulated afternoon peak hour.

Simulation Configuration

A simulation was constructed with intersections and lane configurations identical to this section of roadway. Access points aside from intersections were ignored and all turning movements were assumed
to take place at intersections. Data collection points ("detectors") were placed every 5 meters near
intersections, and every 10 meters between detectors (see Figure 1). Signal timing was taken directly
from the SCATS strategic monitor records. SCATS is an adaptive system; to simplify the experiment,
timings from the peak hour were averaged in the simulation. To get input volumes, the readings from stop
bar detectors at Burnside and Eastman (eastbound) and Burnside and Kelly (westbound) were totaled for
the peak hour. Input volumes for the cross streets were taken from stop bar detectors on Eastman,
Fairview, and Kelly. Right turning volumes were estimated based on all other counts. More detail can be
found in (11).

To calibrate the model, signal offsets were adjusted to minimize total delay for vehicles traveling
on Burnside in either direction. Then simulated speeds were compared to both the posted speed limit and
the average speed of probe vehicles. Simulated vehicle speeds were found to be consistent with probe
vehicle speeds and the speed limit – the average speed of eastbound probe vehicles was 28.5 mph versus
29.6 mph for simulated eastbound vehicles, and westbound probe vehicles traveled at 28.3 mph compared
to 30.7 mph for simulated westbound vehicles. The posted speed limit for this segment of East Burnside
is 35 mph, and the maximum speed in the simulation is 38 mph.

Simulation Data

Data were collected from the simulation by way of data collection points placed every 5 to 10 meters on
the roadway, which functioned as idealized inductive loop detectors. Data available from collection points
included arrival time, departure time, and speed, which were easily translated into count, occupancy, and
speed. In addition, collection points also provided data such as acceleration, vehicle length, vehicle type,
and a unique vehicle ID. This was used to create "ground truth" plots to compare estimates to.

Estimates were made based on subsets of the available data. Analysis of the quality of
information available from loop detectors is done by looking only at particular data collection points and
only at the data items that a loop detector could provide.
a) Aerial view of study area

b) Simulation network

FIGURE 1 Aerial Photo and Simulation Network for Study Area.
1 METHODOLOGY

In assessing arterial performance, travel time was considered because it is the most intuitive, broadly comparable, and easily applicable way to measure performance. In order to make a valid comparison between actual vehicle travel time and travel time that is calculated from detector data, it is necessary to have both "ground truth" data and detector data. Therefore, a simulated corridor was designed, based on measurements of an actual arterial. This enabled us to compare complete vehicle trajectories to trajectories generated from some set of data collected at discrete points.

Vehicle delay was considered in addition to travel time because it is easily derivable from travel time and because it can be depicted visually in a very effective way. Delay can be shown by an input-output diagram, in which one curve shows the cumulative count of vehicles that enter a given road segment, and a second curve shows the cumulative count of vehicles that leave the road segment. The first curve minus the second curve yields travel time, and that minus the free-flow travel time for the segment yields delay. Points where the two curves overlap indicate zero delay, while any gap is proportional to the system delay. This illustration can be created for both the "ground truth" trajectories and the estimated trajectories, giving us the actual delay and the estimated delay, and a clear way to compare the two. Quantifying this comparison lets us evaluate the accuracy of estimation methods and the impacts of detector density and data aggregation.

Six methods for estimating travel time on an arterial were tested against actual travel times and each other. In order to make these comparisons, two types of diagrams were employed: time-space diagrams and input/output diagrams. These diagrams allow visual comparison of estimation methods to ground truth and to each other.

2 Travel Time Estimation Methods

As mentioned above, the six methods for estimating travel time under consideration include:

- The **midpoint method** is essentially the same method currently used to measure freeway performance with loop detectors, applied to an arterial environment. To estimate a vehicle’s travel time over a roadway segment at a given point in time, the most recent speed readings are taken at each detector in the segment. The vehicle is assumed to travel over the section of road from half the distance to the previous detector to half the distance to the next detector at the measured speed. In this way, a predicted trajectory can be constructed over any instrumented section of road. It is hypothesized that this method will perform poorly on arterials, as arterial operations are characterized by frequent changes in speed.

- The **vehicle matching method** (and subsequent methods) takes advantage of two facts: first, detectors on arterials are most commonly located near signals, and second, signals aim to group vehicles into platoons. To estimate travel time, the vehicle matching method starts with the departure time for each detector in the roadway segment. The segment is broken down into sub-segments between each consecutive detector pair. For each vehicle departure at the beginning of a sub-segment, its arrival time at the second detector at the 90th percentile speed is calculated. This vehicle is then matched to the next departure from the second detector after this predicted arrival. Total travel time is then estimated to be the difference between the departure time from the first detector and the departure time from the second detector. It is hypothesized that this method of estimation will give a lower bound on vehicle delay, and to give the best estimates in highly under-saturated conditions.

- The **saturation matching method** is a refinement of the vehicle matching method. The saturation matching method seeks to address the fact that the vehicle matching method often matches multiple departures from the first detector to a single departure at the second detector. The saturation matching method is the same as the vehicle matching method, but instead of matching to the next departure at the second detector, it matches to the greater of the next departure time or the last matched departure time plus some delta equal to the saturation headway.

- The **phase matching method** seeks to account for platoon dispersion by distributing vehicle arrivals evenly throughout the green phase. The phase matching method once again estimates travel times by breaking segments down into sub-segments between each consecutive detector pair. This method...
assumes that each detector is located at the stop bar upstream of a traffic signal. Once again, for each vehicle departure at the beginning of a sub-segment, its arrival time at the second detector at the 90th percentile speed is calculated. This indicates which green phase the vehicle is estimated to arrive in. Then, the total number of vehicles arriving during each green phase is evenly distributed throughout that green phase. Travel time is calculated by subtracting the measured departure time at the first detector from the estimated departure time at the second detector. The phase matching method is hypothesized to give an upper bound on vehicle delay, and to perform best in saturated conditions.

The vehicle synchronization method takes advantage of closely-spaced detectors to assume that vehicle conservation more-or-less holds. Detectors are not necessarily assumed to be placed at signals. As before, though, for each vehicle departure at the beginning of a sub-segment, find its arrival time at the next detector near a signal at the 90th percentile speed. This indicates which green phase the vehicle is estimated to arrive in. Then enumerate vehicles at the arriving detector and vehicles at the departing detector for each green phase. Vehicle 1 at the entry detector is matched to vehicle 1 at the egress detector, vehicle 2 to vehicle 2, and so on. Excess vehicles are ignored (they are assumed to have turned off at some point in the segment).

Finally, the average departure method takes advantage of some mathematical properties of the average travel time equation to offer a way to compute average travel time and delay without computing any individual travel times. Consider a road segment with a detector on each end. The entry detector detects two entering vehicles, at times A1 and A2, while the egress detector detects two departing vehicles, at times B1 and B2. If A1 and B1 are the same vehicle, and A2 and B2 are the same vehicle, then we can make the following observation about the equation for calculating the average travel time:

$$\frac{(B_1 - A_1) + (B_2 - A_2)}{2} = \frac{(B_1 + B_2)}{2} - \frac{(A_1 + A_2)}{2}$$

Thus, the average travel time is simply the average time of entry subtracted from the average time of egress. In applying this method, it is assumed that each detector is located at the stop bar before a traffic signal. The arrival time at the second detector assuming 90th percentile speed is used to determine which green phase cohort each vehicle belongs to.

**ANALYSIS AND RESULTS**

To assess the effectiveness of estimation methods and the effects of data aggregations, comparisons to "ground truth" were made both visually, with input/output diagrams, and quantitatively, by examining the distribution of the ratio of estimated travel times to actual travel times. Input/output diagrams are examined to test whether an estimation method seems to capture the general nature of traffic behavior. Methods that appear to reflect the same trends that govern "ground truth" are then quantitatively evaluated. First, the analysis is conducted assuming access to vehicle actuations and detailed signal timing information. Following this, we aggregate detection in time and space to demonstrate the loss of information. Detailed views of trajectory plots and input-output diagrams showing each method applied to the Kelly-Fairview segment are provided below. In the left plot, the simulated “pseudo ground truth” trajectories are shown as black lines, the trajectories extrapolated by one of the methods are shown in light gray.

Figure 2 shows a detailed view of the input/output curves for the midpoint method overlaid onto the ground truth curves. The midpoint estimation method shows itself immediately as being unsuitable for application to arterials – the output curve (purple) for this method does not show similar trends to the actual output curve. The midpoint method estimates an average delay per vehicle of 51.0 seconds, far different from the actual value of 20.1 seconds.
Figure 3 shows a detailed view of the input and output curves for the vehicle matching method. This method for estimating performance is a clear improvement over the midpoint method. The relationship between the two curves resembles and is proportional to that of the cumulative curves for ground truth. However, there is room for improvement, as the queue discharge portion of the output curve is clearly truncated. But note that the vehicle matching method tends to under-estimate delay, which is consistent with the hypothesis that estimates made by this method represent a lower bound for delay. The vehicle matching method estimates an average delay of 8.6 seconds per vehicle.

The saturation matching method represents a refinement of the vehicle matching method, as Figure 4 shows. The estimated input/output curves are an even better fit to the actual input/output curves than those of the vehicle matching method. The assumption that vehicles would discharge at saturation
rate if they were too closely spaced for the vehicle matching method to distinguish them appears to be a sound one. The saturation matching method estimates an average delay of 13.6 seconds per vehicle.

FIGURE 4 Saturation matching method trajectory and input-output diagrams.

Figure 5 shows the detail of the input/output curves for the phase matching method. The estimated output curve appears to track the actual output curve reasonably well. However, compared to the saturation matching method, the phase matching method loses quite a bit of detail, due to the fact that it does not use some information about actual departure times at the egress detector. The input/output diagram appears to support the hypothesis that aggregated phase matching represents an upper bound on actual delay; however, it does appear to under-predict delay in certain cases.

FIGURE 5: The phase matching method trajectory and input-output diagrams
Aggregated Data

This analysis is repeated for data where only the total number of detected vehicles per signal cycle is known, and not their arrival times. In this case, vehicle “detections” are assumed to be evenly spaced throughout the green phase (when phase information is available), both at the entry and egress detectors. When signal timing data are not available, there is too little information to apply the estimation methods under consideration in this paper.

The vehicle matching, saturation matching, and phase matching methods were applied to both event-based and cycle aggregated data (with phasing information) between Eastman and Fairview, and between Fairview and Kelly, in both directions (4 road segments). Table 1 summarizes the actual delay and the estimated delay for each method and each road segment.

Looking at Table 1, several things are clear. First, the vehicle matching estimation method sometimes over-estimates and sometimes under-estimates delay. Similarly, the phase matching method sometimes over-estimates and sometimes under-estimates delay. Thus, neither can be said to give a strict upper or lower bound on delay.

The vehicle matching method performs well in three of the four road segments, but substantially under-estimates delay on the Kelly-Fairview segment. This segment has the most signal delay of the four, and the vehicle matching method assigns many incoming vehicles to the first departing vehicle reading at the beginning of the second detector’s green phase. This accounts for much of the difference.

The saturation matching method improves on the vehicle matching method for both the Kelly-Fairview and the Fairview-Kelly segments, and does slightly worse for the other two. We would expect vehicles to be spaced at saturation density as they leave a signal queue, which explains the saturation matching method’s improvement on the Kelly-Fairview segment. As for the two segments where the saturation matching method did worse than the vehicle matching method, in those cases, the vehicle

<table>
<thead>
<tr>
<th></th>
<th>Kelly-Fairview</th>
<th>Eastman-Fairview</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average delay</td>
<td>Total delay</td>
</tr>
<tr>
<td></td>
<td>(seconds)</td>
<td>(vehicle hours)</td>
</tr>
<tr>
<td>Actual</td>
<td>20.1</td>
<td>4.4</td>
</tr>
<tr>
<td>Event based, vehicle matching</td>
<td>8.6</td>
<td>2.1</td>
</tr>
<tr>
<td>Event based, saturation matching</td>
<td>13.6</td>
<td>3.3</td>
</tr>
<tr>
<td>Event based, phase matching</td>
<td>17.8</td>
<td>4.3</td>
</tr>
<tr>
<td>Cycle by cycle, vehicle matching</td>
<td>5.7</td>
<td>1.4</td>
</tr>
<tr>
<td>Cycle by cycle, saturation matching</td>
<td>10.9</td>
<td>2.6</td>
</tr>
<tr>
<td>Cycle by cycle, phase matching</td>
<td>21.4</td>
<td>5.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Fairview-Eastman</th>
<th>Fairview-Kelly</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average delay</td>
<td>Total delay</td>
</tr>
<tr>
<td></td>
<td>(seconds)</td>
<td>(vehicle hours)</td>
</tr>
<tr>
<td>Actual</td>
<td>7.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Event based, vehicle matching</td>
<td>7.7</td>
<td>1.8</td>
</tr>
<tr>
<td>Event based, saturation matching</td>
<td>9.7</td>
<td>2.3</td>
</tr>
<tr>
<td>Event based, phase matching</td>
<td>8.7</td>
<td>2.1</td>
</tr>
<tr>
<td>Cycle by cycle, vehicle matching</td>
<td>20.3</td>
<td>4.8</td>
</tr>
<tr>
<td>Cycle by cycle, saturation matching</td>
<td>22.1</td>
<td>5.2</td>
</tr>
<tr>
<td>Cycle by cycle, phase matching</td>
<td>15.2</td>
<td>3.6</td>
</tr>
</tbody>
</table>
matching method over-estimated delay, and the saturation matching method will, by definition, predict
more delay than vehicle matching. The differences in those cases are marginal.

The phase matching method performs very well in the westbound direction, but very poorly in the
eastbound direction. The reason for this is that the signals at the egress detectors in the eastbound
direction (Fairview, Kelly) have very long green times and lower traffic volumes. Since the phase
matching method evenly spaces output, the last few departing vehicles are estimated to leave at the end of
the green phase. In a less saturated environment, the assumptions made by the phase matching method are
clearly unwarranted.

Finally, all three methods display similar characteristics when they are applied to aggregated data
as they do with event based data. The vehicle matching method does well for the same three segments,
but poorly on the Kelly-Fairview segment, the saturation matching method improves on the Kelly-
Fairview estimate and is essentially the same as the vehicle matching method for the other three, and the
phase matching method does poorly for both eastbound segments. The results appear to be good enough,
particularly for the vehicle matching and saturation matching methods, to be encouraging about the
usefulness of estimates made from aggregated data.

Detector Density

To evaluate the effects of higher detector density, two detectors were added to the simulated Kelly-
Fairview westbound segment. The detectors were placed just past the Kelly intersection and behind the
furthest extent of the Fairview signal queue. This broke the road segment down into three sub-segments.
The vehicle matching method and the vehicle synchronization method were applied to all three sub-
segments, and the phase matching method was applied to the third sub-segment.

Table 2 shows the results of this analysis. The vehicle matching method performs very well, with
errors in segments 2 and 3 canceling each other out. The synchronization method consistently
overestimated delay – this was the result of errors from imperfect vehicle conservation that compounded.
It was thought that resetting the vehicle count with the start of every green phase would ameliorate this,
but this was complicated for data taken from the intermediate detectors, as it was unclear where the green
phase should be thought to begin. From this, it appears that increasing detector density beyond 1 per stop
bar does not significantly improve the quality of delay estimates. The value in adding detectors would be
in improving the quality of vehicle counts, and in correspondingly improving total delay calculations.

To evaluate the effects of lower detector density, estimates were made with the vehicle matching,
saturation matching, and phase matching methods from Kelly to Eastman and from Eastman to Kelly. The
ground truth that the estimates were compared to consisted of the set of vehicles that traveled the whole
distance from the entry detector to the exit detector. Vehicles that turned on or off of Burnside at any
point were excluded for this analysis. Each estimation method assumes that every vehicle entering the
road segment exits at the egress detector. Therefore, ground truth consisted of a much smaller vehicle
total than any of the estimates provided, and overlaying the input/output diagrams held no value.
However, the estimates still seemed to provide useful average delay and average speed information, even
if the total vehicle delay numbers weren’t comparable (see Table 3 for a summary of estimated average
delay). Consistent with the previous analysis, the phase matching method performed poorly in the
eastbound direction, while the saturation matching method appeared to be the most robust. This suggests
that these estimation methods can still be useful in low detector densities and in the absence of vehicle
conservation.
TABLE 2 Delay Summary for Increased Detector Density

<table>
<thead>
<tr>
<th>Segment</th>
<th>Average Delay (seconds)</th>
<th>Total delay (vehicle hours)</th>
<th>Average Speed (miles per hour)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td>2.5</td>
<td>0.6</td>
<td>15.6</td>
</tr>
<tr>
<td>Vehicle Matching</td>
<td>3.2</td>
<td>0.8</td>
<td>13.6</td>
</tr>
<tr>
<td>Synchronization</td>
<td>3.5</td>
<td>0.8</td>
<td>12.9</td>
</tr>
</tbody>
</table>

Segment 2

<table>
<thead>
<tr>
<th>Segment</th>
<th>Average Delay (seconds)</th>
<th>Total delay (vehicle hours)</th>
<th>Average Speed (miles per hour)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td>-4.4</td>
<td>-1.1</td>
<td>38.7</td>
</tr>
<tr>
<td>Vehicle Matching</td>
<td>2.2</td>
<td>0.6</td>
<td>27.8</td>
</tr>
<tr>
<td>Synchronization</td>
<td>4.5</td>
<td>1.1</td>
<td>25.2</td>
</tr>
</tbody>
</table>

Segment 3

<table>
<thead>
<tr>
<th>Segment</th>
<th>Average Delay (seconds)</th>
<th>Total delay (vehicle hours)</th>
<th>Average Speed (miles per hour)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td>17.7</td>
<td>4.1</td>
<td>10.4</td>
</tr>
<tr>
<td>Vehicle Matching</td>
<td>11.4</td>
<td>3.0</td>
<td>13.6</td>
</tr>
<tr>
<td>Phase Matching</td>
<td>19.6</td>
<td>5.1</td>
<td>9.7</td>
</tr>
<tr>
<td>Synchronization</td>
<td>20.4</td>
<td>4.8</td>
<td>9.5</td>
</tr>
</tbody>
</table>

Totals

<table>
<thead>
<tr>
<th>Segment</th>
<th>Average Delay (seconds)</th>
<th>Total delay (vehicle hours)</th>
<th>Average Speed (miles per hour)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td>15.8</td>
<td>3.6</td>
<td>20.7</td>
</tr>
<tr>
<td>Vehicle Matching</td>
<td>16.8</td>
<td>4.3</td>
<td>20.3</td>
</tr>
<tr>
<td>Synchronization</td>
<td>28.5</td>
<td>6.7</td>
<td>16.4</td>
</tr>
</tbody>
</table>

TABLE 3 Estimates with Missing Detector Travel Time Estimation

<table>
<thead>
<tr>
<th>Westbound</th>
<th>Average Delay (seconds)</th>
<th>Total Delay (vehicle hours)</th>
<th>Average Speed (miles per hour)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td>24.6</td>
<td>3.9</td>
<td>21.1</td>
</tr>
<tr>
<td>Vehicle Matching</td>
<td>13.2</td>
<td>3.1</td>
<td>24.6</td>
</tr>
<tr>
<td>Saturation Matching</td>
<td>21.2</td>
<td>5.0</td>
<td>22.0</td>
</tr>
<tr>
<td>Phase Matching</td>
<td>22.0</td>
<td>5.3</td>
<td>21.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Eastbound</th>
<th>Average Delay (seconds)</th>
<th>Total Delay (vehicle hours)</th>
<th>Average Speed (miles per hour)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td>0.8</td>
<td>0.1</td>
<td>30.2</td>
</tr>
<tr>
<td>Vehicle Matching</td>
<td>-2.7</td>
<td>-0.6</td>
<td>32.2</td>
</tr>
<tr>
<td>Saturation Matching</td>
<td>-0.9</td>
<td>-0.2</td>
<td>31.1</td>
</tr>
<tr>
<td>Phase Matching</td>
<td>-26.2</td>
<td>-5.5</td>
<td>58.7</td>
</tr>
</tbody>
</table>

In order to better evaluate the efficacy of the proposed methods it was necessary to look more closely at disaggregated estimates. Two questions in particular were addressed: Was the success in estimating average delay simply a case of regression to the mean? And could we trust these performance estimation techniques to estimate the behavior of vehicles in a single signal cycle or even that of individual vehicles, as well as estimating performance over the whole peak hour?

To answer these questions, the three most promising estimation methods (the vehicle matching method, the saturation matching method, and the phase matching method) were used to make cumulative vehicle diagrams for four road segments (Kelly to Fairview and Fairview to Eastman in the westbound direction).
direction, and Eastman to Fairview and Fairview to Kelly in the eastbound direction). One detector at the stop bar per intersection was assumed.

From the cumulative vehicle diagrams, it was possible to extract a travel time estimate for each vehicle, simply by finding the horizontal distance between the curves for any particular counted vehicle. This was also done with the ground truth data, to use as a reference. Then, for each method, the ratio of the estimated travel time to the ground truth travel time was calculated. The better the estimation method, the closer the mean of this set of values will be to one, and the closer its variance will be to zero.

In addition, for the three estimation methods listed above plus the average departure method, both the estimated and actual average travel time per signal cycle was calculated. The ratios of these two values were also found, and the distribution of these ratios was analyzed as above. Because an individual vehicle’s travel time is heavily influenced by when the vehicle arrives during the signal cycle, and vehicle arrivals are stochastic in general, it was felt that the average travel time on a road segment over the last signal cycle would be a more useful basis for comparison over time.

Both of these evaluations were performed on both event-based data and on cycle aggregated data. Table 4 provides a statistical summary for each of these distributions.

<table>
<thead>
<tr>
<th>TABLE 4 Travel time estimate distribution summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event Based Data</td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>Individual Vehicles</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Cycle Average</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Cycle Aggregated Data</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Individual Vehicles</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

None of the estimation methods perform very well for individual vehicles. This is probably due to the fact that the estimation methods make an estimate for every vehicle that enters the road segment, while ground truth travel time was only calculated for vehicles that crossed both the entry and egress detectors. Thus, it is easy to see how comparisons could get out of sync. When predicting the average travel time in a signal cycle from event-based data the phase matching method performs poorly, but the other three methods (vehicle matching, saturation matching, and average departure) all perform very (and equally) well.

As noted above, the vehicle matching method, the saturation matching method, and the average departure method were all roughly equal in tracking the actual average travel times per signal cycle. All have low variances and high adjusted R-squared values, confirming that they model reality very well. This was the case even where estimates were made from cycle-aggregated data.
CONCLUSION

This research clearly shows that useful arterial performance information can be extracted from data collected by inductive loop detectors in configurations that are already widely implemented. It was shown that one detector per lane per intersection yields useful information, and that the quality of information does not deteriorate substantially even when there is no detector coverage at one intervening intersection. Some arterial performance measures, such as total vehicle delay, rely on accurate counts of vehicles on each segment of roadway, and computation of these measures benefits from increased detector density, but the impact on estimates of travel time and delay appear to be minimal.

Temporal aggregation of detector data was found to impact the quality of travel time estimates, but if cycle-level signal data and phase information was available, accurate estimates could be made. The absence of signal phase information rendered aggregated data useless for estimating travel time, as did data aggregated in larger blocks than one signal cycle.

Six travel time estimation methods were proposed and evaluated: the midpoint method, the vehicle matching method, the saturation matching method, the phase matching method, the vehicle synchronization method, and the average departure method. Three were found to provide superior travel time estimates: the vehicle matching method, the saturation matching method, and the average departure method. Of these, the saturation matching method stood slightly above the other two. The estimates generated by this method were consistently the most accurate or not far from the most accurate. It is robust and easy to implement, and should handle saturated conditions and signal failure in a more realistic manner.

Temporal aggregation of detector data was found to impact the quality of travel time estimates, but if cycle-level signal data and phase information was available, accurate estimates could be made. The absence of signal phase information rendered aggregated data useless for estimating travel time, as did data aggregated in larger blocks than one signal cycle.

Six travel time estimation methods were proposed and evaluated: the midpoint method, the vehicle matching method, the saturation matching method, the phase matching method, the vehicle synchronization method, and the average departure method. Three were found to provide superior travel time estimates: the vehicle matching method, the saturation matching method, and the average departure method. Of these, the saturation matching method stood slightly above the other two. The estimates generated by this method were consistently the most accurate or not far from the most accurate. It is robust and easy to implement, and should handle saturated conditions and signal failure in a more realistic manner.

Interest question that remains open is what contribution to the quality of travel time estimates can probe vehicles make? This is easily investigated in a simulation; one would just need to pick some random subset of actual vehicles in the ground truth sample to represent the probes, and to apply some statistical analysis to see how representative it is of the whole population. From there, the probe data could be examined to see how it could be applied to fill in the gaps in estimates made from loop detector data.

ACKNOWLEDGEMENTS

Funding and support for this project is provided by the National Science Foundation, Oregon Department of Transportation, Federal Highway Administration, City of Portland, and TriMet. The authors particularly thank David Crout and Steve Callas of TriMet. Peter Koonce and Bill Kloos of the City of Portland and Fred Liang of the City of Bellevue also provided valuable advice and input to the project. Dr. Miguel Figliozzi was a member of the committee for the thesis on which this paper is based.

REFERENCES


