

Impacts of Sensor Spacing on Accurate Freeway Travel Time Estimation for Traveler Information

ROBERT L. BERTINI

Department of Civil and Environmental Engineering
Nohad A. Toulan School of Urban Studies and Planning
Portland State University
P.O. Box 751
Portland, OR 97207-0751 USA
Email: bertini@pdx.edu
Phone: 503-725-4249
Fax: 503-725-5950

DAVID J. LOVELL

Department of Civil and Environmental Engineering
Institute for Systems Research
University of Maryland
1173 Glenn Martin Hall
College Park, MD 20742 USA
Email: lovell@umd.edu
Phone: 301-405-7995
Fax: 301-405-2585

February 2009

ABSTRACT

Travel time estimation is a critical ingredient for transportation management and traveler information—both infrastructure-based and in-vehicle. Infrastructure managers and operators are interested in estimating optimal freeway sensor density for new construction and retrofits. Focusing on freeway travel time estimation for display on roadside variable message signs, this paper describes a concept developed from first principles of traffic flow for establishing optimal sensor density. The method is based on computing the magnitude of under- and overprediction of travel time during shock passages when using the midpoint method. Two case studies are presented considering representative traffic dynamics situations. Along with other performance metrics, a suggested aggregate measure developed from vehicle hours traveled (VHT) is described for a reasonable range of detector densities. Extensions of the method to account for both recurrent and nonrecurrent congestion are included. Finally some suggestions for future research are described.

Keywords: traveler information, traffic management, freeway sensor

INTRODUCTION

Travel time estimation is used for a wide range of applications, including planning, design, performance measurement, operational analysis, traffic management, incident detection, navigation, traveler information, and more. In the U.S., freeways account for 3% of the national highway mileage, but accommodate more than 30% of the vehicle-miles traveled (VMT). The alleviation of congestion on urban freeways continues to receive substantial attention, and transportation agencies are implementing improved management strategies to reduce congestion and improve travel time reliability. Accurate freeway travel time measurement is important for improving traffic management and for enabling better informed traveler decisions. For example, incident management and traveler information systems can be implemented at relatively low cost. However, such systems rely upon accurate measurement of traffic parameters such as flow, speed, travel time, delay and reliability. Usually these data are measured by fixed sensors (loop detectors, video cameras, microwave detectors, radar sensors, etc.) or by mobile data sources such as automatic vehicle identification (AVI) toll tags or automatic vehicle location (AVL) probe vehicles. In the future, technologies such as vehicular ad hoc networks will likely provide a supplemental data source. It is possible that some applications using travel time are not affected by error; however this paper focuses on a specific application where travel time error is an issue.

Collecting relevant data as input to travel time computations is one ingredient for the successful transmission and display of travel time information for travelers. This can be done via a speed map showing color coded freeway segments or origin-destination listings available via the Internet, handheld devices such as PDAs and mobile phones, or via roadside variable message signs (VMS). While there are many potential uses and display mechanisms for travel time information, the specific focus of this paper is the situation where a driver receives travel time information based on data from fixed sensors upon entering a freeway section via a VMS.

As motivation for this paper, Figure 1 shows a time-space speed plot from an actual freeway corridor (Northbound I-5 in Portland, Oregon on October 31, 2006), with the x -axis as time, the y -axis as distance (in miles), and the “color” indicating speed. The Oregon Department of

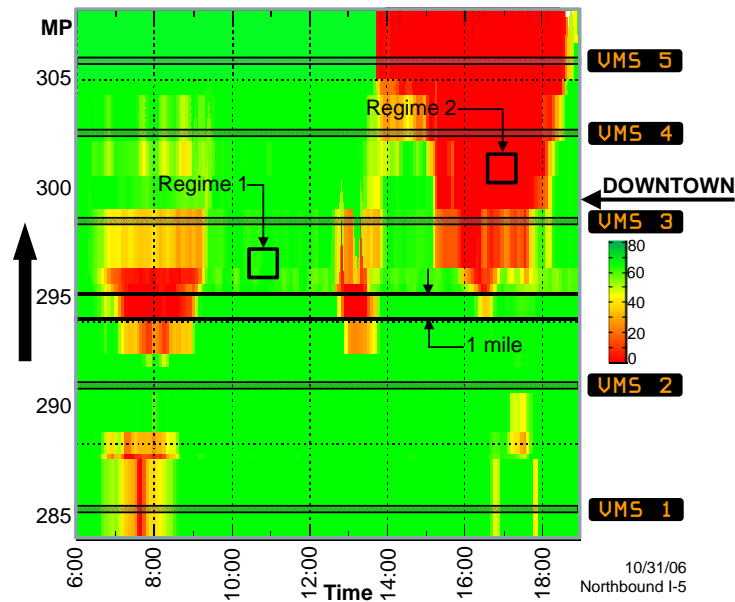


Figure 1 Freeway corridor example.

Transportation (ODOT) has installed loop detectors at an average spacing of 1.2 miles and five VMS along this corridor (locations are shown on the right y-axis of the figure). ODOT displays travel time information to Downtown on VMS 1, 2 and 3 during the AM peak period. ODOT and others are interested in increasing sensor density toward some “optimal” value, in order to improve the accuracy of travel time information in a cost-effective manner. In order to answer the question “how much detection is needed,” this paper focuses on the common use of fixed sensors (such as loop detectors) as a basis for formulating an optimal detector spacing strategy.

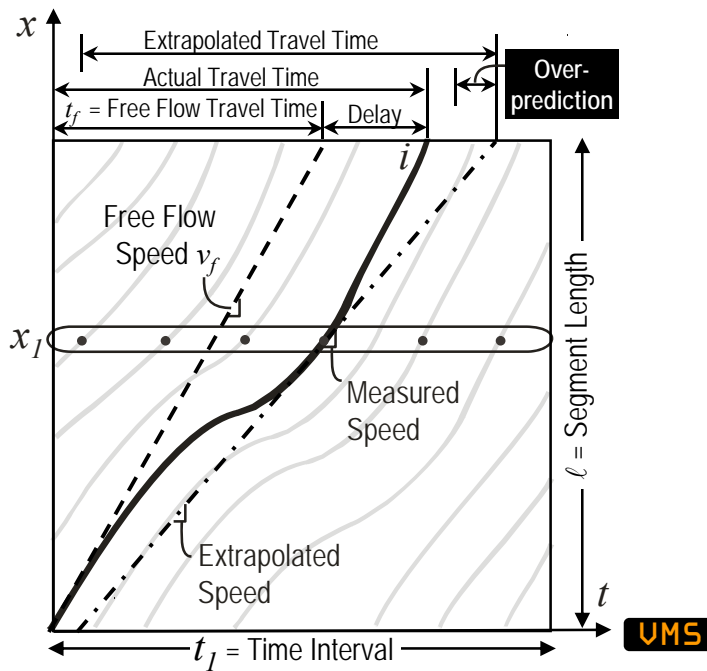


Figure 2 Segment travel time features.

A hypothetical 1-mile segment such as the one between MP 294 and 295 in Figure 1 will be considered as a basis for the analysis. Sensor placement has been the topic of several recent studies (see, for example, Bartin, et al., 2006; Fujito, et al., 2006; Leow et al., 2008).

Turning away from the specific example shown in Figure 1, and in order to simplify and generalize the situation using a consistent analytical framework, Figure 2 illustrates a *hypothetical* time-space (t - x) plane of time interval t_l and length ℓ with a VMS located at the upstream end. Hypothetical vehicle trajectories are shown (in grey) with that of vehicle i highlighted in black. If equipped with AVL or any other data logging system, vehicle i 's trajectory could be constructed and all information necessary to completely describe its path would be known, including its speed (slope of trajectory) at any point and its actual travel time. With an assumed free flow speed, the free flow travel time is known and the delay (actual minus free flow travel time) for vehicle i can be computed. In practice it is more common to use sensors at fixed points (such as point x_l) to measure speed and subsequently extrapolate that speed over a segment. In Figure 2 a speed measured at x_l is extrapolated over a segment of length ℓ resulting in the calculation of an extrapolated travel time. The estimate does not perfectly match the actual travel time (in this case it is overpredicted). The magnitude of this difference, as a function of detector density, is the topic of interest in this paper in the context of providing travel time information to vehicles entering the section via a VMS. The important topic of sensor errors is not considered here, nor are other issues related to the calculation or communication of travel times for other purposes. The passage of a shock represents the “worst case scenario” for travel time prediction, so the methods presented in this paper can be interpreted as a form of robust decision analysis about sensor spacing.

TRAVEL TIME ESTIMATION FRAMEWORK

A simplified framework is assumed as a platform for this analysis. In freeway dynamics there are a variety of transition types as shown in Figure 3A (see May, 1990), some of which move forward and others that move backward. The backward forming shock affects travel time calculations at the onset of congestion since the shock moves against traffic. A forward forming wave travels with traffic and may have less effect. For this analysis, a fundamental traffic flow relation has been assumed, as shown in Figure 3B. In particular, a triangular shape is used; this is common in the literature when broad questions need to be addressed with a minimum of complications. A congested state C (flow q_c and speed v_c) and uncongested states A, B, D, and E (flows q_A , q_B , q_D , q_E , speed v_f) are shown. In order to include the maximum number of representative traffic dynamics situations, two cases are analyzed, shown in Figures 3C and 3D. For Case 1, Figure 3C shows a t - x plane showing a bottleneck (either recurrent or nonrecurrent) at location bn . It is assumed, for the sake of simplicity, that the bottleneck state is binary; it is either active, with some reduced capacity captured by state C, or it is inactive, in which case the nominal capacity can be permitted. Following the rules of standard first-order macroscopic traffic dynamics, and assuming the nominal traffic state was A, there is a transition between uncongested state A and congested state C, marked by a shock of velocity v_{AC} . Figure 3C shows that for an assumed arbitrary highway segment (separated by the two dashed lines with a VMS at the upstream end) transition AC is bounded by a rectangle as the shock passes. If this were an active bottleneck that was deactivated at time t_{deact} , then transition CD would occur, between

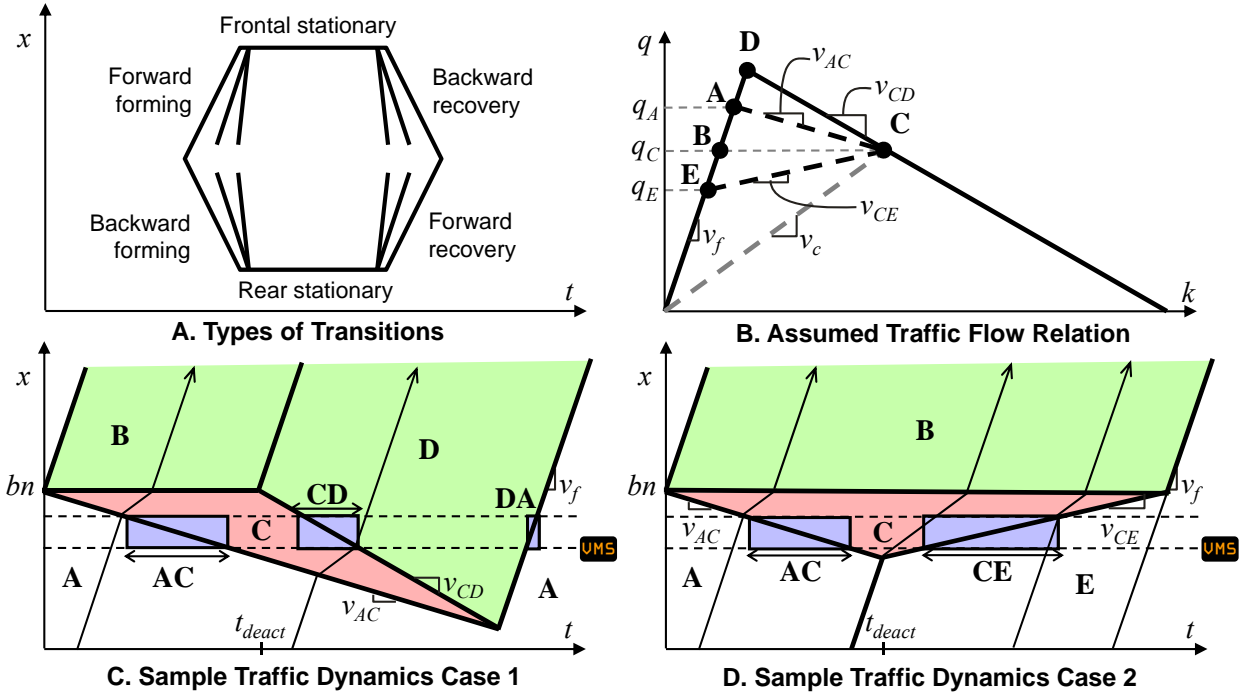


Figure 3 Assumed traffic flow relation and traffic dynamics.

congested state C and uncongested state D, marked by a backward-moving recovery wave of velocity v_{CD} . Transition DA (the return to nominal conditions) is marked by a forward-moving recovery wave of velocity v_f . There is no requirement that these three transitions always occur in this configuration or order, and the computations later in the paper assess each type of state transition separately. To expand the analysis, Figure 3D shows Case 2, which includes a backward forming transition (AC) similar to Case 1 and a forward recovery (CE) marked by a forward-moving recovery wave of velocity v_{CE} . Later computations will examine travel time estimation during transitions AC, CD, DA and CE.

The first-order hydrodynamic model approximates actual vehicle trajectories, such as those in Figure 2, with piecewise-linear trajectories. For the purposes of travel time measurements over an assumed freeway link, this is irrelevant, because one vehicle's travel time is measured as the horizontal difference between the endpoints of its trajectory on the link, and the microscopic details of the shapes of speed transitions in the vicinity of a shock do not matter in this analysis.

CALCULATING TRAVEL TIME DURING STATIONARY CONDITIONS

As noted above, travel time information is used for purposes other than VMS display, and for many of these purposes the error structure is quite different. In particular, for measuring VHT for system evaluation, delays are summed over a rectangular time-space domain and any lag time does not lead to error. For these applications, the errors are a strictly increasing function of the detector spacing, so optimal detector spacing is not an issue. However, the purpose of this paper is to consider travel time information that would be provided to travelers using VMS.

There are situations when travel time can be estimated perfectly well. Referring to Figure 1, travel times during regime 1 (fully uncongested) and regime 2 (fully congested) can be estimated

without error. Viewing the entire t - x plane in Figure 1, it is clear that travel times can be correctly estimated during fully uncongested or congested periods (regimes 1 and 2), but there are notable transitions in and out of congested states—which are of interest for this analysis.

From Figure 2, travel times in states A, B, D and E (uncongested), C (congested) and during transition DA (uncongested) can be estimated without error over an arbitrary section of length ℓ ,

$$\text{by:} \quad tt_f = \ell / v_f, \quad tt_c = \ell / v_c$$

where tt_f is free flow travel time and tt_c is congested travel time. If the sensor where speed is measured is located at the center of the segment, this would be called the midpoint method. Travel time estimation over a segment can be performed accurately if the traffic state within the segment is either fully uncongested or fully congested. In addition, tt_c is an upper bound on the actual travel time through the segment and tt_f is a lower bound on the actual travel time. Empirical studies have shown that travel time estimation during stationary conditions (either uncongested or congested) is usually quite accurate (Cortes, et al., 2002; Monsere, et al., 2006). For simplicity, in this paper, the midpoint method will be used; however, other methods have been developed for improving travel time estimates (Coifman, 2002; Wang and Liu, 2005; Kwon, et al., 2007; Liu and Danczyk, 2008; Ni and Wang, 2008) and can be studied further later.

TRAVEL TIME ESTIMATION DURING TRANSITIONS

As initially illustrated in Figure 3A, there are two basic transition types in freeway traffic flow: uncongested to congested (e.g. AC) and the reverse (e.g. CD or CE). Transitions can move forward or backward and can occur multiple times in a given section as queues propagate and dissipate and sometimes combine with one another. Queues may be caused by recurrent bottlenecks or by nonrecurrent conditions such as incidents. These transient conditions induce errors in travel time estimations based on current measured conditions, and depending on the specifics, can result in either overprediction or underprediction. In the specific travel time information context considered here, underprediction is considered to be more problematic than overprediction, since travelers whose travel times are much longer than predicted at their entry to a segment will be more likely to be dissatisfied. The method involves choosing a sensor spacing as a result of a tradeoff between over- and underprediction and for the reasons cited above, underprediction will be given a higher weight in this tradeoff.

This paper provides an analytical framework and also includes a series of numerical examples using a hypothetical segment of length $\ell = 1$ mile for comparison purposes. For the numerical examples a range of five sensor spacings s (0.1 to 1 mile) will be used and assumed traffic flow parameters will be: $q_A = 2000$ vphpl, $q_C = 1800$ vphpl, $q_E = 1600$ vphpl, $v_f = 60$ mph, $v_c = 30$ mph, $v_{CD} = -17.1$ mph, $v_{AC} = -7.5$ mph, and $v_{CE} = +6.0$ mph. For the calculations that follow, we will consider as the domain of the travel time computations a region of the time-space plane that is bounded by the link endpoints, and includes exactly those vehicles whose trajectories were affected by the passage of the shock. It is also assumed that travel time information can be transmitted to vehicles entering the hypothetical segment via a VMS located at the upstream end of the segment (this concept could be broadened to an in-vehicle or handheld device). Several performance measures are used when comparing results for the range of sensor spacings considered, which will be described below

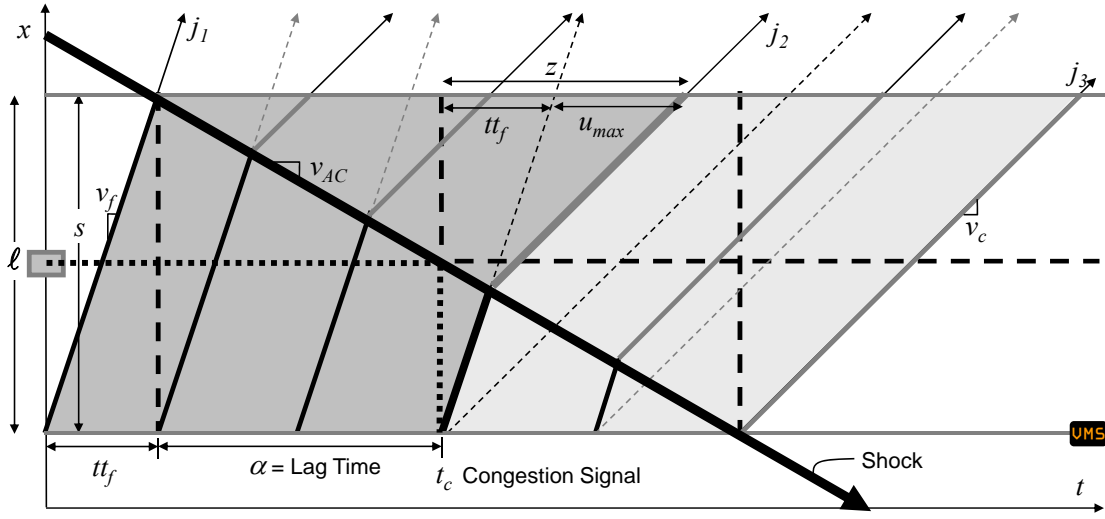


Figure 4 Travel time estimation during regime AC.

Underpredicting Travel Time During Transition AC

Figure 4 illustrates transition AC from uncongested conditions with vehicles traveling at v_f to a congested state with vehicles traveling at v_c . The figure illustrates the backward moving shock passing through the segment of length ℓ with an initial sensor spacing s , at a speed v_{AC} . Trajectory j_1 is the last trajectory to pass through the section at speed v_f and trajectory j_3 is the first to traverse the section at speed v_c . All vehicles between j_1 and j_3 change speeds at some point in the section, and their average speeds are therefore between v_f and v_c . The shading in the figure encapsulates these vehicles. The figure also shows that the sensor continues to record speed v_f until time t_c , a lag time $\alpha = -s/2v_{AC}$ after the shock actually enters the section. After vehicle j_1 and until t_c , based on the VMS message, drivers would expect a free flow trip time through the entire section while their actual trip time will be higher. For example, driver j_2 enters the section an instant before time t_c . Based on the VMS, j_2 expects a speed of v_f through the entire section, but experiences a longer actual travel time z . For vehicles entering the section after j_1 but before time t_c , travel time is *underpredicted* by an amount equal to the difference between the expected travel time (dashed trajectory in the figure) and the actual travel time (solid trajectory). Prior to encountering the shock, vehicle j_2 travels a distance:

$$x_{AC1} = \frac{v_f (\ell - \frac{1}{2}s)}{v_f - v_{AC}} \quad (1)$$

at speed v_f , and drives the remainder of the link at speed v_c . Thus, the travel time experienced by vehicle j_2 is:

$$\begin{aligned} z &= \frac{\left[\frac{v_f (\ell - \frac{1}{2}s)}{v_f - v_{AC}} \right]}{v_f} + \frac{\left[\ell - \frac{v_f (\ell - \frac{1}{2}s)}{v_f - v_{AC}} \right]}{v_c} = \frac{\ell - \frac{1}{2}s}{v_f - v_{AC}} + \frac{\ell (v_f - v_{AC}) - v_f (\ell - \frac{1}{2}s)}{v_c (v_f - v_{AC})} \\ &= \frac{\ell v_c - v_c \frac{1}{2}s + \ell v_f - \ell v_{AC} - \ell v_f + v_f \frac{1}{2}s}{v_c (v_f - v_{AC})} = \frac{\ell (v_c - v_{AC}) + \frac{1}{2}s (v_f - v_c)}{v_c (v_f - v_{AC})} \end{aligned} \quad (2)$$

The amount by which travel time is underpredicted is a maximum for this vehicle, and this maximum error is given by:

$$u_{\max} = z - tt_f = \frac{\ell(v_c - v_{AC}) + \frac{1}{2}s(v_f - v_c)}{v_c(v_f - v_{AC})} - \frac{\ell}{v_f} \quad (3)$$

For example, if $\ell = s = 1$, from the VMS the driver of vehicle j_2 expects a travel time of 1 minute, but actually experiences a 1.56 minute trip, more than 50% longer than expected. If one assigns zero weight to any travel time overprediction (between time t_c and the time the shock reaches the upstream end of the section), for now one can neglect any flow entering the section after time t_c . The remainder of this section only considers vehicles for which travel time is *underpredicted*, since it is assumed that traffic management officials want to avoid giving drivers false expectations of low travel times, when they actually experience longer ones.

In addition to considering as performance measures the values of maximum underprediction (u_{\max}), lag time (α), and percent error, one can quantify the predicted and actual vehicle-hours traveled (VHT_{pred} and VHT_{act}) solely for vehicles experiencing underprediction (superscript u) over the hypothetical segment. Recall that the errors reported would be only for the transitions, since when traffic conditions are homogeneous it is assumed that travel times can be computed without error.

There is an assumption here (as throughout other forms of travel time analysis) that individual vehicles' travel times are additive. In reality, different travelers have different values of time, but this is a common practice in transportation analyses. To compute the VHT_{pred} and VHT_{act} for comparison, the number of vehicles is counted in each situation, and is multiplied by the average travel time (either predicted or actual). Since the trajectories and shocks are all assumed to be linear, the average travel time for a group of vehicles with a similar disposition will always be the midpoint of the best and worst travel times amongst that set.

During transition AC, the trajectories for which travel time is underpredicted cross the upstream end of the link over a time span of:

$$t_{AC_1} = tt_f + \alpha = \frac{\ell}{v_f} - \frac{s}{2v_{AC}} \quad (4)$$

At this location, the flow is q_A ; hence the *number* of vehicles in this condition is:

$$n_{AC_1} = q_A \left(\frac{\ell}{v_f} - \frac{s}{2v_{AC}} \right) \quad (5)$$

Based on the VMS, each of these vehicles expected to travel at free flow speed and therefore have free flow travel time across the link, so the predicted VHT for trajectories for which travel time was underestimated in the AC transition is:

$$VHT_{pred}^u = \frac{q_A \ell}{v_f} \left(\frac{\ell}{v_f} - \frac{s}{2v_{AC}} \right) \quad (6)$$

The actual total travel time can be found by multiplying the same number of vehicles by their expected travel time, which is midway between the highest and lowest travel times for this group of vehicles. The lowest travel time is the free-flow travel time, $tt_f = \ell/v_f$. The vehicle with the

highest travel time is trajectory j_2 of Figure 4, whose travel time was computed in equation (2). The average travel time for all vehicles is then the average between these lowest and highest values, and the total actual travel time for these vehicles, whose travel time was underpredicted, can be found by multiplying this average by the number of vehicles in equation (5), resulting in the following:

$$VHT_{act}^u = \frac{q_A}{2} \left(\frac{\ell}{v_f} - \frac{s}{2v_{AC}} \right) \left(\frac{\ell}{v_f} + \frac{\ell(v_c - v_{AC}) + \frac{1}{2}s(v_f - v_c)}{v_c(v_f - v_{AC})} \right) \quad (7)$$

For $\ell = s = 1$, the predicted VHT/mile is 2.78 veh-hr/mile, yet the actual VHT/mile is 3.55 veh-hr/mile, a 22% error over the collection of vehicles entering transition AC before t_c . Vehicles experience actual underpredictions between 0 and 0.56 minutes.

In order to extend these calculations to arrangements with greater sensor density, Figure 5 illustrates how predicted and actual VHT will change with additional sensors. The increased sensor placement reduces the lag time α so vehicles entering the segment receive the congestion message sooner via the VMS. The speed from each detector will be applied only to a portion of ℓ rather than to the whole section, which reduces the magnitude of the VHT composed of travel time underprediction. The figure indicates this using the darker shaded areas. As the lag time decreases, the VHT of traffic impacted by travel time underprediction will decrease in this section. For now the issue of overprediction is set aside.

Table 1 shows the values of u_{max} and lag time α for transition AC as a function of detector spacing s . The table also shows the predicted and actual VHT/mile for a range of s . To reiterate, for this hypothetical 1 mile segment, vehicles entering the section prior to t_c will expect free flow conditions based on the VMS, but will instead experience progressively longer travel times (the lag time gets shorter with increased detector density). For this set of vehicles, the actual VHT is higher than the predicted VHT. Even with sensors at 0.1 mile spacing, the VHT error is 7% and vehicles expecting a travel time of 1 minute actually experience 1.16 minutes, a 16% underprediction. The gap between the predicted and actual VHT grows with larger sensor spacing. For the range of sensor spacing considered, the VHT error falls between 7% and 22%.

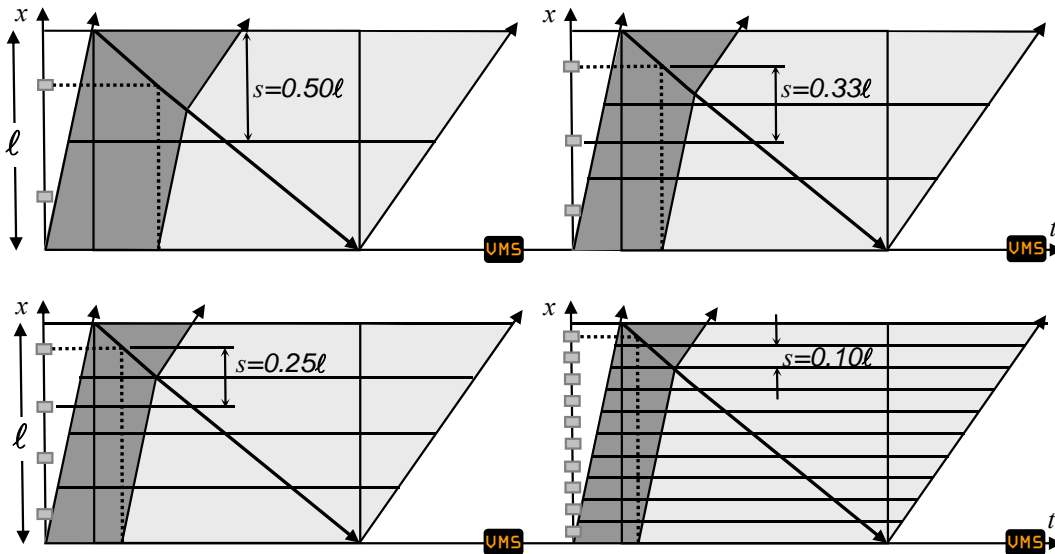


Figure 5 Travel time estimation during regime AC.

For the average ODOT sensor spacing of 1.2 mile, the VHT error would be 24%. In this case, more detection is better if the goal is to minimize underprediction. Spacing between 0.25 and 0.50 mile would keep u_{max} below 33% of t_f . Table 1 also includes a sixth row with $s = 0$, which is equivalent to a form of ubiquitous sensor coverage.

Overpredicting Travel Time During Transition AC

The previous section examined the impact of only the travel time *underprediction* during transition AC. Travel times for vehicles entering the section after time t_c and until the shock crosses the upstream end of the link are *overpredicted*. The duration over which these vehicles cross the upstream end of the link is:

$$t_{AC_2} = \left(\frac{1}{v_{AC}} \right) (\frac{1}{2}s - \ell) \quad (8)$$

and they do so at flow q_A . The *number* of such vehicles is therefore:

$$n_{AC_2} = \frac{q_A}{v_{AC}} (\frac{1}{2}s - \ell) \quad (9)$$

and based on the VMS, each expected the congested travel time ℓ/v_C . For transition AC, including only the overprediction, the predicted VHT (with superscript o) is:

$$VHT_{pred}^o = \frac{q_A \ell}{v_c v_{AC}} (\frac{1}{2}s - \ell) \quad (10)$$

The actual VHT for these vehicles can be computed in a manner identical to equation (7), except that the number of vehicles in this condition is given by equation (9) and the two extreme travel times whose average is the average for all vehicles are ℓ/v_C and the travel time for vehicle j_2 as shown in equation (3). This results in the following total travel time for vehicles in transition AC, whose travel time was overpredicted:

$$VHT_{act}^o = \frac{q_A}{2v_{AC}} (\frac{1}{2}s - \ell) \left(\frac{\ell}{v_C} + \frac{\ell(v_c - v_{AC}) + \frac{1}{2}s(v_f - v_c)}{v_C(v_f - v_{AC})} \right) \quad (11)$$

Table 1 shows that for $\ell = s = 1$, for vehicles entering the section after t_c the predicted VHT is 4.44 veh-hr/mile, and the actual VHT is 3.95 veh-hr/mile, an overprediction of 13%. Many drivers would be pleasantly surprised by the shorter travel time but it is not safe to assume that overprediction is always benign. In a dynamic route choice application, for example, it could disguise a better link and thus lead to suboptimal route recommendations. Travelers could lose their trust in the system. With increased sensor density, the percent error in VHT increases to a maximum of 27% overprediction for $s = 0.10$.

When viewing Figure 6 and Table 1 together, for transition AC it can be seen that when $s \approx 0.5$, the underprediction error is 14%, the overprediction error is 20%, the aggregate error is -11% (overprediction), and the absolute error is 18%. These errors are only for transition AC, recalling that during homogeneous traffic states, travel time estimates can be made without error. This might represent a reasonable compromise for users' acceptance.

Table 1 Segment travel time performance measures during regimes AC and CD

Regime AC										
<i>s</i> (mi)	<i>u_{max}</i> (min)	Lag Time (min)	Underprediction before <i>t_c</i>			Overprediction after <i>t_c</i>			AC Total	
			VHT/mile Pred	VHT/mile Act	Under % Error	VHT/mile Pred	VHT/mile Act	Over % Error	% Error	Abs % Error
1.00	0.56	4.00	2.78	3.55	22%	4.44	3.95	-13%	4%	17%
0.50	0.33	2.00	1.67	1.94	14%	6.66	5.56	-20%	-11%	18%
0.33	0.26	1.33	1.30	1.46	11%	7.40	6.04	-23%	-16%	21%
0.25	0.22	1.00	1.11	1.23	10%	7.77	6.27	-24%	-19%	22%
0.10	0.16	0.40	0.78	0.84	7%	8.44	6.66	-27%	-23%	25%
0	0.11	0.00	0.56	0.59	5%	8.88	6.91	-29%	-26%	27%

Regime CD										
<i>s</i> (mi)	<i>u_{max}</i> (min)	Lag Time (min)	Overprediction before <i>t_r</i>			Underprediction after <i>t_r</i>			CD Total	
			VHT/mile Pred	VHT/mile Act	Over % Error	VHT/mile Pred	VHT/mile Act	Under % Error	% Error	Abs % Error
1.00	0.56	4.00	3.75	3.11	-21%	0.88	1.02	14%	-12%	19%
0.50	0.33	2.00	2.88	2.50	-15%	1.32	1.63	19%	-1%	17%
0.33	0.26	1.33	2.58	2.28	-13%	1.46	1.85	21%	2%	17%
0.25	0.22	1.00	2.44	2.17	-12%	1.54	1.96	22%	4%	17%
0.10	0.16	0.40	2.18	1.96	-11%	1.67	2.17	23%	7%	17%
0	0.11	0.00	2.00	1.82	-10%	1.75	2.31	24%	9%	18%

Regimes AC + CD Total Effect					
<i>s</i> (mi)	Under % Error	Over % Error	Total % Error	Error w/Penalty	Absolute % Error
1.00	20%	-16%	-2%	14%	18%
0.50	17%	-18%	-8%	2%	18%
0.33	17%	-20%	-10%	0%	19%
0.25	17%	-21%	-11%	-1%	20%
0.10	19%	-23%	-12%	-3%	22%
0	20%	-25%	-13%	-3%	24%

Note: Positive error percentages indicate underprediction, while negative error percentages indicate overprediction

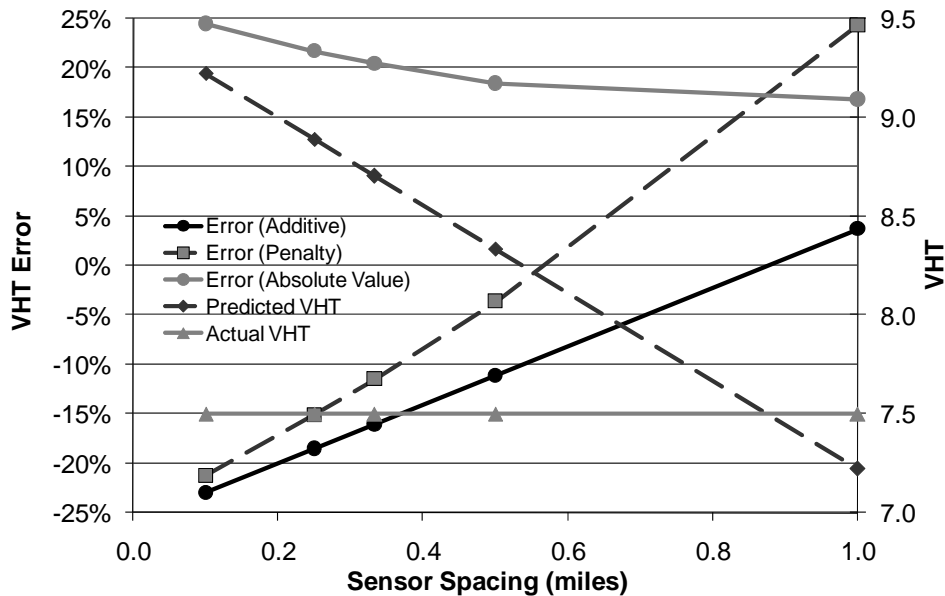


Figure 6 Travel time over- and underprediction in regime AC (VHT/mile).

As an extension of the above, if one assumes that under- and overpredicted VHT can be added together numerically (with one canceling out the other), the total predicted VHT in transition AC would be $VHT_{pred}^u + VHT_{pred}^o$, and the actual VHT in transition AC would be

$VHT_{act}^u + VHT_{act}^o$ (7.50 veh-hr/mile). Here one time unit of underprediction is canceled out by a time unit overprediction. As shown in Table 1, when the two components are added, for $\ell = s = 1$, the effect is still a 4% underprediction in VHT. For decreasing values of s , the aggregate effect is overprediction, up to 23% with $s = 0.10$. Figure 6 illustrates this graphically; here a spacing of $s \approx 0.5$ mile is near where the VHT error line crosses from overall underprediction to overprediction. For the situation where drivers might assign a higher “price” to underprediction than to overprediction, Figure 6 shows, with the line series labeled “penalty,” the percent error in VHT when the underpredicted VHT is weighted at $3\times$ that of overpredicted VHT. From this example, a value of $s \approx 0.5$ mile would be reasonable. Table 1 also includes a column where the total VHT error is considered by adding the absolute value of the differences in actual and predicted VHT (also shown in Figure 6). The absolute error ranges between 17% and 25% for the range of detector spacing considered. Other approaches to the $3\times$ penalty could be explored in the future.

Predicting Travel Time During Transition CD

Case 1 in Figure 3C also includes transition CD from congested state C to uncongested state D, shown in Figure 7. Vehicles from the left are traveling at v_c , and a backward-moving recovery wave passes through the section at speed v_{CD} . The sensor receives the “uncongested” signal at time t_r , which occurs a lag time α' after the wave enters the section. For the sake of brevity, the intermediate computational steps for transition CD are not shown here, since the sequence is identical to that followed for transition AC, only with different parameters. As shown in the figure, travel time is overpredicted for vehicles entering the section after trajectory j_1 and before time t_r . The VHT for these vehicles can be calculated as:

$$VHT_{pred}^o = \frac{q_C \ell}{v_c} \left(\frac{\ell}{v_c} - \frac{s}{2v_{CD}} \right) \tag{12}$$

$$VHT_{act}^o = \frac{q_C}{2} \left(\frac{\ell}{v_c} - \frac{s}{2v_{CD}} \right) \left(\frac{\ell}{v_c} + \frac{\ell(v_f - v_{CD}) + \frac{1}{2}s(v_c - v_f)}{v_f(v_c - v_{CD})} \right) \tag{13}$$

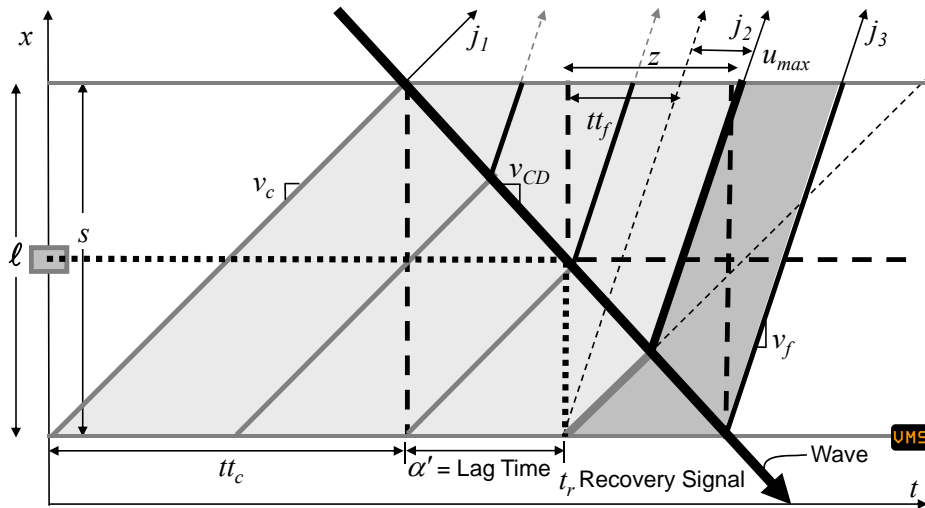


Figure 7 Travel time estimation during regime CD.

Vehicles entering after vehicle j_2 expect free flow travel times but experience higher travel times (underprediction). Their predicted and actual VHT are:

$$VHT_{pred}^u = \frac{q_c \ell}{v_{CD} v_f} (\frac{1}{2}s - \ell) \tag{14}$$

$$VHT_{act}^u = \frac{q_c}{2v_{CD} v_f} (\frac{1}{2}s - \ell) \left(\ell + \frac{\ell(v_f - v_{CD}) + \frac{1}{2}s(v_c - v_f)}{(v_c - v_{CD})} \right) \tag{15}$$

Table 1 shows the under- and overprediction results for transition CD. In this case, the percent VHT error for underprediction increases with increased detection. Figure 8 shows that when the errors are simply added (allowing overprediction to cancel out underprediction) an optimal s would be between 0.33–0.5 mile. Applying a 3× penalty to underprediction results in an optimal s between 0.5–1.0 mile. Adding the absolute values of the under- and overprediction results in errors in the 17–19% range.

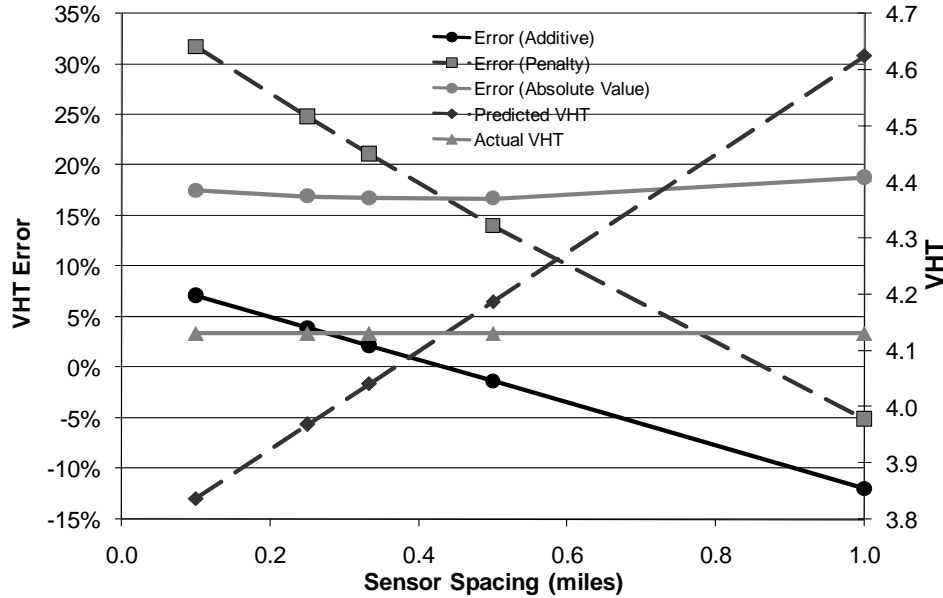


Figure 8 Travel time over- and underprediction in regime CD (VHT/mile).

Combined Effects of Transition Transitions AC and CD

For Case 1, Figure 3C shows a hypothetical freeway section where it is assumed that freeway travel time estimation can be performed accurately using the midpoint method throughout the entire day except during transitions AC and CD. It is possible to consider just the impact of VHT underprediction during the transitions. Figure 9 shows the additive effects of only underpredicted VMT (data from transitions AC and CD in Table 1). The optimal s when considering only underprediction is 0.5 mile (minimum is 16.5% error during the transitions).

Table 1 also shows the overall additive effects of the underprediction and overprediction that occurs in transitions AC and CD. As shown for $\ell = s = 1$, the aggregate effect is a 2% VHT overprediction error. This error increases to 12% VHT overprediction for $s = 0.10$. This is also shown graphically in Figure 10. As before, a 3× penalty is applied to the underprediction error

before it is added to the overprediction error, resulting in a net 14% underprediction for $s = 1$ mile, and a 3% overprediction for $s = 0.1$ mile.

As shown in Figure 10, this reveals an possibly optimal s of 0.33 mile (zero error). It is not

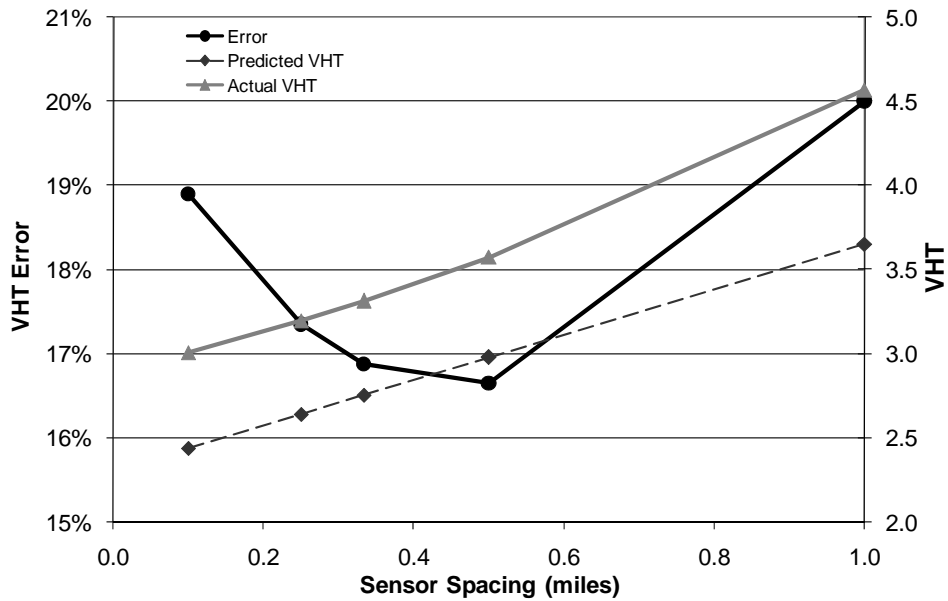


Figure 9 Combined effect of travel time underprediction in regimes AC and CD (VHT/mile).

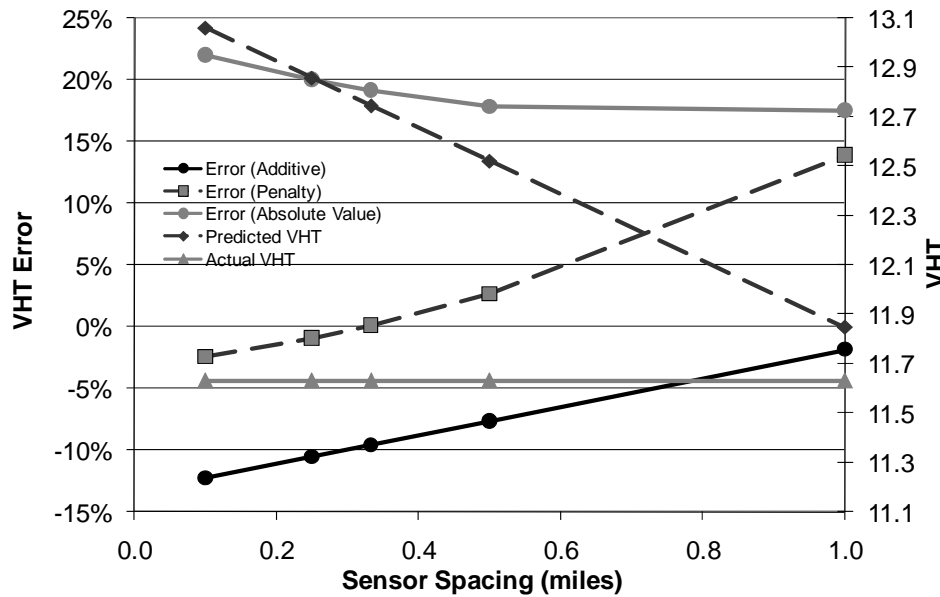


Figure 10 Combined effect of travel time estimation in regimes AC and CD (VHT/mile).

clear how differently drivers actually value underprediction versus overprediction, so this is merely a sample, and could be the topic of further research. The total error applying the sum of the absolute values of the underprediction and overprediction errors is also shown in Table 1 (18–22% error) and in Figure 10.

PREDICTING TRAVEL TIME DURING TRANSITION CE

Case 2 from Figure 3D includes transitions AC and CE. Using the logic for transitions AC and CD, with Figure 3D as a reference, Figure 11 shows transition CE from congested speed v_c to uncongested speed v_f . Transition CE is a forward moving recovery wave passing through a segment of length $\ell = 1$ mile, sensor spacing s , at speed v_{CE} . Trajectory j_1 represents the last vehicle that traveled at speed v_c through the section, and trajectory j_3 represents the first vehicle to traverse the section at speed v_f . Trajectories between j_1 and j_3 (shading in the figure) reduce their speeds in the section, and the average speed is between v_c and v_f . The sensor measures speed v_c until time t_r , a lag time $\alpha = s/2v_{CE}$ after the wave enters the section. After trajectory j_1 and until t_r , a congested trip time is predicted through the section while actual trip times will be lower. For example, trajectory j_2 enters the section just before t_r with an expected speed of v_c through the section, but experiences a shorter travel time z . For trajectories after j_1 entering the section before t_r travel time is overpredicted by an amount equal to the difference between the expected travel time (dashed trajectory in the figure) and the actual travel time (solid trajectory).

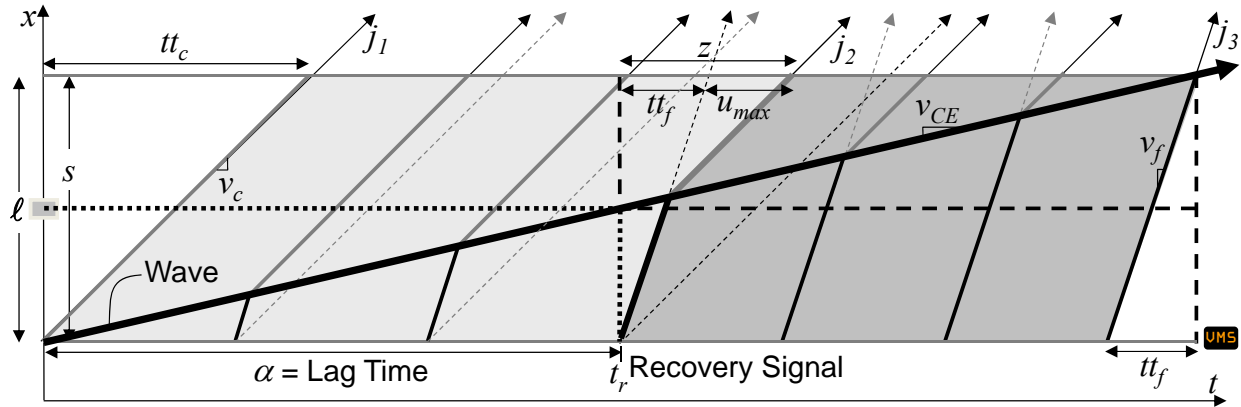


Figure 11 Travel time estimation during regime CE.

Prior to encountering the wave, trajectory j_2 travels a distance:

$$x_{CE} = \frac{sv_f}{2(v_f - v_{CE})} \quad (16)$$

at speed v_f , and traverses the remainder of the link at speed v_c . Thus, the travel time experienced by trajectory j_2 is:

$$z = \frac{\ell(v_f - v_{CE}) + \frac{1}{2}s(v_c - v_f)}{v_c(v_f - v_{CE})} \quad (17)$$

The amount by which travel time is overpredicted is maximal for trajectory j_2 , given by:

$$o_{\max} = z - tt_f = \frac{\ell(v_f - v_{CE}) + \frac{1}{2}s(v_c - v_f)}{v_c(v_f - v_{CE})} - \frac{\ell}{v_f} \quad (18)$$

If $\ell = s = 1$, trajectory j_2 would expect a travel time of 1 minute, but actually experience a 1.44 minute trip, 44% longer than expected based on the VMS. As before, the trajectories for which

travel time is overpredicted cross the upstream end of the link over a time span equal to the lag time α shown in Figure 11:

$$t_{CE_1} = \alpha = \frac{s}{2v_{CE}} \quad (19)$$

At this location, the flow is q_E ; hence the *number* of vehicles in this condition is

$$n_{CE_1} = \frac{q_E s}{2v_{CE}} \quad (20)$$

Each of these vehicles was expected to travel at congested speed so the predicted VHT for vehicles whose travel time was overestimated in transition CE is:

$$VHT_{pred}^o = \frac{q_E \ell s}{2v_c v_{CE}} \quad (21)$$

The actual total travel time can be found by multiplying the same number of vehicles by their expected travel time:

$$VHT_{act}^o = \frac{q_E}{2} \left(\frac{s}{2v_{CE}} \right) \left(\frac{\ell}{v_c} + \frac{\ell(v_f - v_{CE}) + \frac{1}{2}s(v_c - v_f)}{v_c(v_f - v_{CE})} \right) \quad (22)$$

For $\ell = s = 1$, the predicted VHT/mile is 4.44 veh-hr/mile, and the actual VHT/mile is 3.83 veh-hr/mile, reflecting a 16% error over the collection of vehicles entering transition CE before t_r . Vehicles experience actual underpredictions between 0 and 0.44 minute.

Table 2 shows the values of u_{max} , lag time α , and predicted and actual VHT/mile for transition CE as a function of s . For trajectories before t_r , the actual VHT is lower than the predicted VHT. With sensors at 0.1 mile spacing, the VHT error is 1% and vehicles expecting a 2 minute travel time actually experience 1.94 minute, a 3% overprediction. The gap between the predicted and actual VHT grows with larger sensor spacing. For the range of sensor spacing considered, the VHT error falls between 1% and 16%. For the average ODOT sensor spacing of 1.2 miles, the VHT error is 20%. If the goal is to minimize overprediction, more detection is better.

Having examined the impact of travel time overprediction during transition CE, Figure 11 shows that travel time underprediction occurs for trajectories entering the section after time t_r until the wave crosses the downstream end of the link. The duration over which these vehicles cross the upstream end of the link at flow q_E is:

$$t_{CE_2} = \left(\frac{2\ell - s}{2v_{CE}} - \frac{\ell}{v_f} \right), \quad (23)$$

The number of such vehicles is:

$$n_{CE_2} = q_E \left(\frac{2\ell - s}{2v_{CE}} - \frac{\ell}{v_f} \right) \quad (24)$$

each expecting the free flow travel time ℓ/v_f . For transition CE, including only the underprediction, the predicted VHT (with superscript u) is:

Table 2 Segment travel time performance measures during regime CE and combined AC and CE

		Regime CE								
s (mi)	u_{max} (min)	Lag Time (min)	Overprediction (before t_r)			Underprediction (after t_r)			CE Total	
			VHT/mile Pred	VHT/mile Act	Under % Error	VHT/mile Pred	VHT/mile Act	Over % Error	% Error	Abs % Error
1.00	0.44	4.00	4.44	3.83	-16%	1.78	2.17	18%	-4%	17%
0.50	0.72	2.50	2.22	2.07	-7%	2.89	3.93	27%	15%	20%
0.33	0.82	1.67	1.48	1.41	-5%	3.26	4.59	29%	21%	23%
0.25	0.86	1.25	1.11	1.07	-3%	3.44	4.93	30%	24%	25%
0.10	0.94	0.50	0.44	0.44	-1%	3.78	5.56	32%	30%	30%
0	1.00	0.00	0.00	0.00	0%	4.00	6.00	33%	33%	33%

Regimes AC + CE Total Effect					
s (mi)	Under % Error	Over % Error	Total % Error	Error w/Penalty	Absolute %Error
1.00	20%	-14%	-0.44%	18%	7%
0.50	22%	-16%	-0.43%	20%	9%
0.33	25%	-19%	-0.43%	23%	10%
0.25	26%	-21%	-0.42%	24%	11%
0.10	29%	-25%	-0.41%	28%	13%
0	31%	-28%	-0.41%	31%	15%

Positive error percentages indicate underprediction, while negative error percentages indicate overprediction

$$VHT_{pred}^u = \frac{q_E \ell}{v_f} \left(\frac{2\ell - s}{2v_{CE}} - \frac{\ell}{v_f} \right) \quad (25)$$

The actual VHT for vehicles in transition CE whose travel time was underpredicted is:

$$VHT_{act}^u = \frac{q_E}{2} \left(\frac{2\ell - s}{2v_{CE}} - \frac{\ell}{v_f} \right) \left(\frac{\ell}{v_f} + \frac{\ell(v_f - v_{CE}) + \frac{1}{2}s(v_c - v_f)}{v_c(v_f - v_{CE})} \right) \quad (26)$$

Table 2 shows that for $\ell = s = 1$, for vehicles entering the section after t_r , the predicted VHT is 1.78 veh-hr/mile, and the actual VHT is 2.17 veh-hr/mile, an underprediction of 18%. With increased sensor density, the percent error in underpredicted VHT increases to a maximum of 32% underprediction for $s = 0.10$.

When viewing Figure 12 and Table 2 together for transition CE only, it can be seen that when $s = 1$, the total error is a 4% overprediction in VHT. For decreasing values of s when adding numerically the total effect is underprediction, up to 30% with $s = 0.10$. If the absolute values of the errors are added, the effect for $s = 1$ is 17% error (underprediction), increasing to 30% for $s = 0.1$. For $s = 0.5$, the total error is 15% and the absolute error is 20%.

Table 2 also shows the results of adding the effects of travel time estimation errors during transitions AC + CE (as shown in Figure 3D). It is possible to consider just the impact of VHT underprediction during the transitions. The additive effects of only underpredicted VMT (data from transitions AC + CE in Table 2) results in errors between 20–29%. For overprediction only the errors range between -14% and -28%.

Table 2 also shows the additive effects of the underprediction and overprediction that occurs in transitions AC and CE. As shown for $\ell = s = 1$, the aggregate effect is a -0.44% VHT error (very close to zero in the aggregate). The error remains about the same for $s = 0.10$. This is also shown graphically in Figure 13. As before, a 3× penalty is applied to the underprediction error before it is added to the overprediction error, resulting in a net 18% underprediction for $s = 1$ mile, and a 28% underprediction for $s = 0.1$ mile. This is also shown in Figure 13. The total error

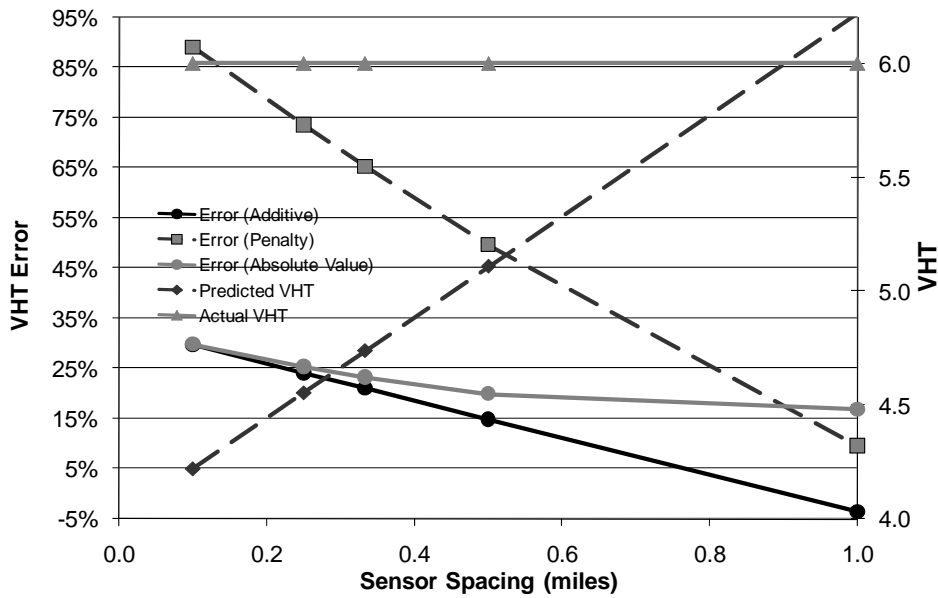


Figure 12 Regime CE under- and overprediction.

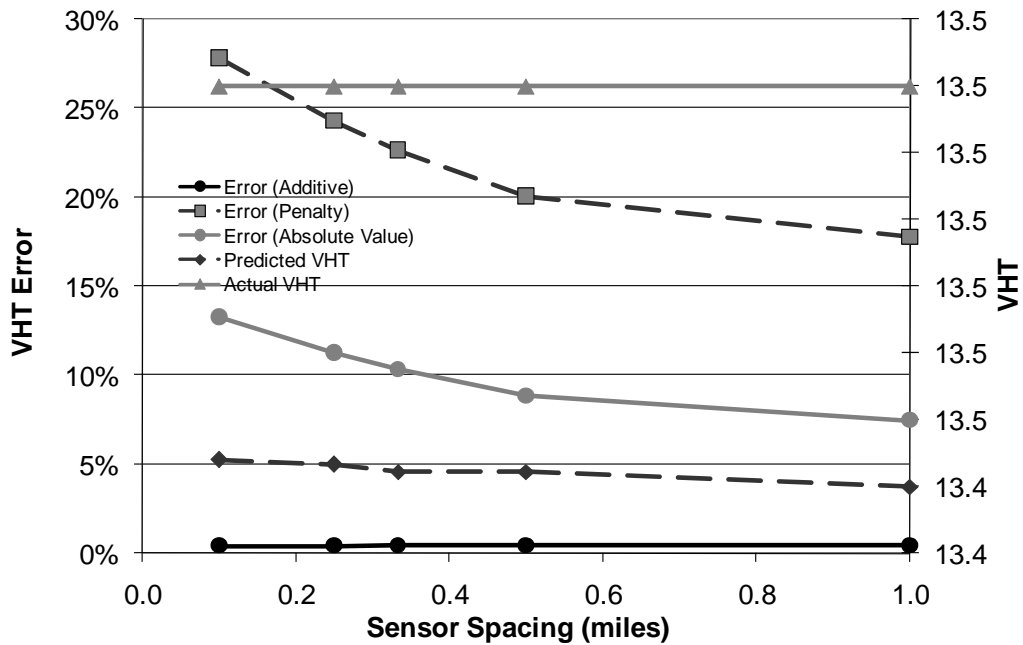


Figure 13 Regimes AC and CE combined under- and overprediction.

applying the sum of the absolute value of the underprediction and overprediction error is also shown in Table 2 (7–15% error) and in Figure 13.

CHANGING DETECTOR LOCATIONS

The results presented so far are dependent on the assumption that a detector is located in the middle of the link (hence the “midpoint” method). We investigate here what would be different if the detectors were located at the upstream and downstream ends of the link. This change

would not be effected by actually moving the detectors, but rather redefining the endpoints of the link. Rather than repeat the analysis of the entire paper, we will repeat here just that part corresponding to the AC transition. It should be clear to the interested reader at that point how to repeat the analysis for other shock transitions, as well as for other detector locations or extrapolation methods not included in this paper.

Figure 14 shows transition AC defined so that single detectors are located at the upstream and downstream ends of the link. This is in contrast with the situation shown in Figure 4. Here we will repeat the analysis that began with equation (2), for travel time underestimation, followed by overestimation. We still focus on the vehicle labeled j_2 , which is the last vehicle to enter the section without the benefit of information from the sensor. In this case, j_2 travels a distance:

$$x_{AC_3} = \frac{v_f \ell}{v_f - v_{AC}} \quad (27)$$

at speed v_f and the remainder of the section at speed v_c . The travel time z for vehicle j_2 is:

$$\begin{aligned} z &= \left[\frac{v_f \ell}{v_f - v_{AC}} \right] \frac{1}{v_f} + \left[\ell - \frac{v_f \ell}{v_f - v_{AC}} \right] \frac{1}{v_c} = \frac{\ell}{v_f - v_{AC}} + \left[\frac{v_f \ell - v_{AC} \ell - v_f \ell}{v_f - v_{AC}} \right] \frac{1}{v_c} \\ &= \frac{v_c \ell}{v_c (v_f - v_{AC})} - \frac{v_{AC} \ell}{v_c (v_f - v_{AC})} = \frac{\ell (v_c - v_{AC})}{v_c (v_f - v_{AC})} \end{aligned} \quad (28)$$

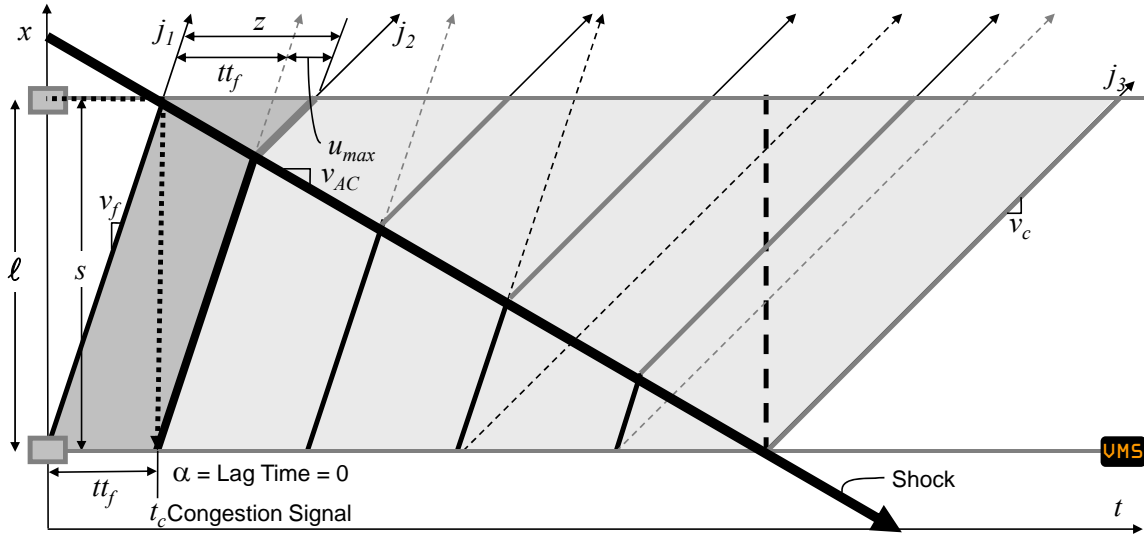


Figure 14 Transition example during transition AC with detector at downstream end.

Because a sensor is located at the upstream end, there is no lag ($\alpha = 0$) and the free-flow travel time is as before. The maximum prediction error u_{\max} is therefore:

$$u_{\max} = z - tt_f = \frac{\ell (v_c - v_{AC})}{v_c (v_f - v_{AC})} - \frac{\ell}{v_f} \quad (29)$$

The time span over which underpredicted vehicles enter the link is the free-flow travel time plus the lag, which is zero. Again, vehicles are counted using the upstream boundary as a reference, and at this location the flow is q_A ; hence the number of vehicles in this condition is:

$$n_{AC_3} = q_A \left(\frac{\ell}{v_f} \right) \quad (30)$$

Each of these vehicles has predicted travel time from the VMS equal to the free-flow travel time, so the total VHT per hour predicted for those vehicles that ultimately end up with underpredicted travel times is:

$$VHT_{pred}^u = q_A \left(\frac{\ell}{v_f} \right) \left(\frac{\ell}{v_f} \right) = q_A \frac{\ell^2}{v_f^2} \quad (31)$$

The actual travel times range from the free-flow travel time to the worst-case scenario given in (28); hence the actual VHT per hour is given by:

$$VHT_{act}^u = \frac{q_A \ell}{2v_f} \left(\frac{\ell}{v_f} + \frac{\ell(v_c - v_{AC})}{v_c(v_f - v_{AC})} \right) \quad (32)$$

For vehicles in this situation whose travel time is overpredicted, the time period over which they cross the upstream end of the link is:

$$t_{AC_4} = -\frac{\ell}{v_{AC}}$$

The flow at this location is q_A , so the number of vehicles in this condition is:

$$n_{AC_4} = -\frac{q_A \ell}{v_{AC}} \quad (33)$$

These vehicles expect to travel the link at speed v_c , so the total predicted VHT per hour is:

$$VHT_{pred}^o = -\frac{q_A \ell^2}{v_{AC} v_c} \quad (34)$$

The actual VHT per hour is:

$$VHT_{act}^o = -\frac{q_A \ell}{2v_{AC}} \left(\frac{\ell}{v_c} + \frac{\ell(v_c - v_{AC})}{v_c(v_f - v_{AC})} \right) \quad (35)$$

Table 3 shows how these new results compare with those from the midpoint method, for a single detector and for transition AC only. Not surprisingly, situating the detector on the downstream

Table 3 Segment travel time performance measures during regime AC for a downstream detector

		Regime AC						
		Underprediction			Overprediction			
Detector Location (min)	u_{max} Lag Time (min)	VHT/mile Pred	VHT/mile Act	Under % Error	VHT/mile Pred	VHT/mile Act	Under % Error	
Midpoint	0.56 4.00	2.78	3.55	22%	4.44	3.95	-13%	
Endpoint	0.11 0.00	0.56	0.59	5%	8.89	6.91	-29%	

Note: Positive error percentages indicate underprediction, while negative error percentages indicate overprediction

end allows the shock information to be known much earlier, therefore, there is a lower tendency to underpredict travel times but a greater tendency to overpredict. Other detector locations or extrapolation methods (Coifman, 2002) could also be tested for specific situations, using the same process shown here but with different details pertaining to detector location.

CONCLUSIONS

This paper has attempted to begin to address the question: “how much detection do you need?” in the context of accurate estimation of freeway travel time using the midpoint method for traveler information applications via upstream roadside VMS. There is no way that one paper can address every possible issue with regard to travel time estimation. Also, it is understood that detection decisions are not made in isolation from other issues, and in fact, sensors are usually placed to enable operation of ramp metering (e.g. Portland) and traffic monitoring (e.g., counting, speed maps, and incident detection). Freeway travel time estimation is often a useful side benefit that can be leveraged from an existing sensor network. It is also possible that in the future some combination of fixed infrastructure based sensors and vehicle based sensing (e.g. AVL) may provide additional answers and improvements. However, this analysis has taken the question of sensor density in some degree of isolation which has resulted in some helpful outcomes. An issue that has been left for further research is the question of *where* to optimally place sensors, beyond simply a question of spacing. The use of other travel time algorithms beyond the midpoint method should also be explored further. The optimal placement of sensors in relation to known bottlenecks, and high incident locations will be examined in the future.

Even were these results to be applied directly, the method does not culminate in a single “optimal” answer because tradeoffs exist in the space blending drivers’ experiences and expectations. Research might be undertaken to reveal some preference structure and illuminate unknown parameters, and it might also be necessary to apply some level of professional judgment and/or policy imposition to the decision. The proper balance between under- and overprediction, for example, is not known, and may not be the same for all drivers. Ultimately, the goal of this paper has been to highlight these issues and to put the quantitative aspects of the problem on sound footing.

ACKNOWLEDGEMENTS

Galen McGill of the Oregon Department of Transportation (ODOT) posed the sensor density question and supports our freeway travel time research. Dennis Mitchell and Jack Marchant of ODOT provide the data. K. Tufte, S. Kothuri, B. Zielke and R. Fernández-Moctezuma of Portland State University assisted in development of this work. We thank the anonymous reviewers of the previous version of this paper for their very extensive and helpful comments and criticisms.

REFERENCES

- Bartin, B., Ozbay, K., Iyigun, C. (2006). A clustering based methodology for determining the optimal roadway configuration of detectors for travel time estimation. *Proceedings of the IEEE Intelligent Transportation Systems Conference*, Toronto, Canada.

- Coifman, B. (2002). Estimating travel times and vehicle trajectories on freeways using dual loop detectors. *Transportation Research*, **36A** 351–364.
- Cortes, C.E., Lavanya, R., Oh, J.S., and Jayakrishnan, R. (2002). General purpose methodology for link travel time estimation using multiple point detection of traffic. *Transportation Research Record*, **1802**, 181–189.
- Fujito, I., Margiotta, R., Huang, W., and Perez, W. (2006). Effect of sensor spacing on performance measure calculations. *Transportation Research Record*, **1945**, 1–11.
- Kwon, J., Petty, K., and Varaiya, P. (2007). Probe vehicle runs or loop detectors? Effect of detector spacing and sample size on the accuracy of freeway congestion monitoring. *Proceedings of the 86th Annual Meeting of the Transportation Research Board*, Washington, D.C.
- Leow, W., Ni, D., and Pishro-Nik, H. (2008). A sampling theorem approach to traffic sensor optimization. *IEEE Transactions on Intelligent Transportation Systems*, **9(2)**, 369-374.
- Liu, H., and Danczyk, A. (2008). Optimal detector placement for freeway bottleneck identification. *Proceedings of the 87th Annual Meeting of the Transportation Research Board*, Washington, D.C.
- May, A.D. (1990). Traffic flow fundamentals. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Monsere, C.M., Breakstone, A., Bertini, R.L., Deeter, D. and McGill, G. (2006). Validating dynamic message sign freeway travel time messages using ground truth geospatial data. *Transportation Research Record*, **1959**, 19–27.
- Ni, D., and Wang, H. (2008). Trajectory reconstruction for travel time estimation. *Journal of Intelligent Transportation Systems*, **12(3)** 113–125.
- Wang, Z., and Liu, C. (2005). An empirical evaluation of the loop detector method for travel time delay estimation. *Journal of Intelligent Transportation Systems*, **9(4)** 161–174.