Analysis for the Use of Cumulative Plots for Travel Time Estimation on Signalized Network

Ashish Bhaskar · Edward Chung ·
André-Gilles Dumont

Abstract This paper provides fundamental understanding for the use of cumulative plots for travel time estimation on signalized urban networks. Analytical modeling is performed to generate cumulative plots based on the availability of data: *a*) Case-D, for detector data only; *b*) Case-DS, for detector data and signal timings; and *c*) Case-DSS, for detector data, signal timings and saturation flow rate. The empirical study and sensitivity analysis based on simulation experiments have observed the consistency in performance for Case-DS and Case-DSS, whereas, for Case-D the performance is inconsistent. Case-D is sensitive to detection interval and signal timings within the interval. When detection interval is integral multiple of signal cycle then it has low accuracy and low reliability. Whereas, for detection interval around 1.5 times signal cycle both accuracy and reliability are high.

A. Bhaskar (✉) · A.-G. Dumont
Ecole Polytechnique Fédérale de Lausanne,
EPFL-ENAC-ICARE-LAVOC, station 18,
Lausanne CH-1015, Switzerland
e-mail: ashish.bhaskar@epfl.ch

A.-G. Dumont
e-mail: andre-gilles.dumont@epfl.ch

E. Chung
Queensland University of Technology,
QUT Gardens Point Campus, GPO Box 2434, Brisbane,
Queensland 4001, Australia
e-mail: edward.chung@qut.edu.au

1 Introduction

Travel time is the time required to travel between two spatially separated points on the network. It is an important network performance measure and it quantifies congestion in a manner easily perceived by all users.

Literature is abundant with research on travel time estimation and prediction. Number of models ranging from simple naïve regression [1–5], traffic flow theory [6, 7], pattern recognition [8–12] to advance machine learning [13–16] and data fusion models [17–21] are proposed. New models are still being sought by many researchers as there are avenues for improvement especially in terms of transferability, applicability and robustness.

A majority of such models are limited to freeways and cannot be applied as it is on urban network due to different behavior of traffic on the two facilities. The complexities related to urban network includes interrupted traffic due to conflicting areas such as intersections; and significant traffic flow from (to) mid-link sources (sinks), etc.

Cumulative plot is the plot of cumulative counts of values (here, vehicles at a specific location) versus time, starting from an arbitrary initial count e.g., value is zero at time equals zero [22]. In traffic engineering, Newell [23] is pioneer to use cumulative plots for dynamic analysis of deterministic congested systems.

Using cumulative plots to model travel time [6, 7] and/or travel delay [24] is not new to the field. However, its performance with respect to numerous model parameters related to detection interval; signal timing, etc., is not investigated in literature.

The classical analytical procedure for travel time estimation considers cumulative plots $U(t)$ and $D(t)$ at upstream entrance and downstream exit of the link, respectively. Refer to Fig. 1, N is the number of vehicles
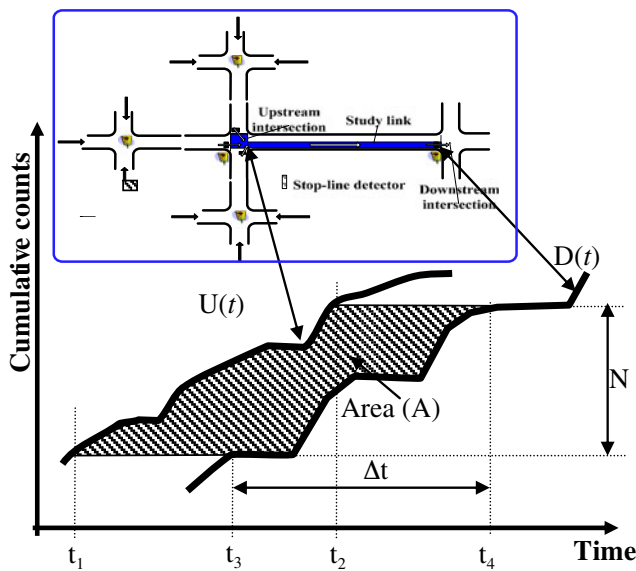
Fig. 1 Illustration of the classical analytical methodology for travel time estimation

that arrives at upstream from time t_1 to time t_2 or departs from downstream from time t_3 to time t_4. The area A, between the plots (considering $U(t)$ from time t_1 to time t_2 and $D(t)$ from time t_3 to time t_4) is the total travel time for N vehicles. Average travel time $\overline{TT}$ is A/N and is mathematically expressed as follows:

$$\overline{TT} = \frac{\sum_{i=1}^{N} D^{-1}(i) - U^{-1}(i)}{N} = \frac{\sum_{i=1}^{N} D^{-1}(i) - \sum_{i=1}^{N} U^{-1}(i)}{N} \tag{1}$$
$$N = D(t_4) - D(t_3) = U(t_2) - U(t_1)$$

Where $D^{-1}(i)$ and $U^{-1}(i)$ are the time corresponding to the ith cumulative count observed at $D(t)$ and $U(t)$, respectively.

For real application there are certain issues to be addressed such as, *a*) relative deviation amongst the cumulative plots (also termed as drift) due to mid-block sources and sinks (e.g. parking), and detector counting error; and *b*) unknown fluctuations in traffic flow due to aggregated detector counts from stop-line detector.

Due to relative deviation the cumulative plots can either diverge from each other or can even cut each other. Bhaskar et al. [25] have addressed the relative deviation issue by integrating cumulative plots with probe vehicle data. Therefore, in this paper no relative deviation amongst the cumulative plots is assumed and we focus on addressing the following aggregated counts issue.

If the available detector data is aggregated counts then due to aggregation the information about the actual traffic fluctuations is lost. For instance, if the detector detection (aggregation) interval is 5 min and 100 vehicles are observed during a 5 min interval, then we do not know how these 100 vehicles are distributed within the interval.

Therefore, the granularity of the cumulative plot depends on the aggregation interval.

The objectives of this paper are to: perform analytical modeling for estimation of cumulative plots; and empirically study the performance of the estimated cumulative plots for their application in travel time estimation using classical analytical procedure. The performance is analyzed with respect to the following parameters: *a*) detection interval; *b*) signal timings; *c*) offset of signal timings within detection interval; and *d*) degree of saturation. Such fundamental analysis is not being performed in literature, although cumulative plots are foundation for numerous analytical models.

The structure of the paper is as follows: First, analytical modeling is performed in Section 2. Thereafter, the developed models are tested using simulation in Section 3 and finally, the results from its sensitivity analysis are presented in Section 4.

2 Analytical modeling

Here, stop-line detector data; signal controller data and saturation flow rate is integrated. Stop-line loop detectors are detectors at the intersection stop-line and generally provide aggregated counts during detector detection interval. Signal controller provides signal timings such as signal phase plan and time corresponding to the start and end of signal green for each phase. Here aggregated counts from the stop-line detectors are known and fluctuations in the traffic flow due to signals are unknown. To generalize the model, cumulative plots at the location of the detector are estimated for the three cases depending on the availability of the data: *a*) *Case-D:* Only detector data is available; *b*) *Case-DS:* Detector data and signal controller data is available; and *c*) *Case-DSS:* Detector data, signal controller data and saturation flow rate is available. The slope of the plot defines the flow pattern at the respective entrance of the intersection.

Here, N_d and q as the counts and flow, respectively during the detection interval of DI seconds. Signal timings considered are the effective signal timings.

2.1 Case-D

The flow pattern (Fig. 2) is assumed to be uniform throughout the detection interval (Eq. 2). The assumption is reasonable for shorter detection intervals and in the absence of any further information can be applied for larger detection intervals.
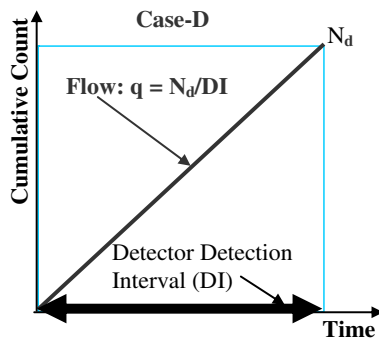
$$q = \frac{N_d}{DI} \tag{2}$$

Fig. 2 Flow pattern for case-D

2.2 Case-DS

A stepwise flow pattern is defined (Eq. 3) where it is uniform only during the signal green period within the detection interval and during signal red period there is no flow (Fig. 3).This captures the fluctuations in the flow pattern even for larger detection intervals. Flow patterns during each green period of the detection interval are parallel to each other.

During *green periods in the detection interval*
$$q = \frac{N_d}{\sum g_{d,i}}$$
During *red periods in the detection interval*
$$q = 0 \tag{3}$$

We define $g_{d,i}$ as the *i-th* green period during the detection interval.

In Fig. 3, two green periods are present during the detection interval and the counts are distributed to each green interval in proportion to the corresponding green time. The count, N_i, during each *i-th* green period ($g_{d,i}$) in the detection interval is assumed to be in proportion to $g_{d,i}$ (Eq. 4).

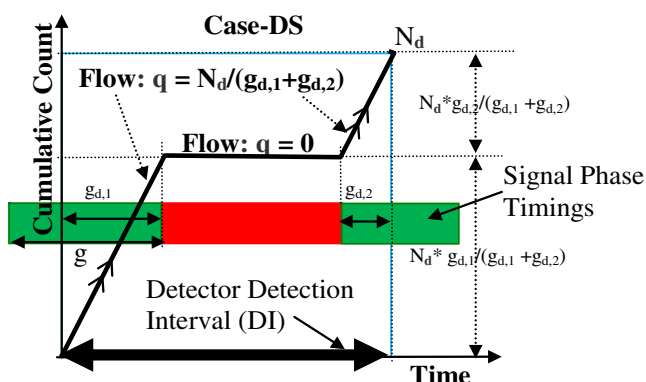$$N_i = \frac{N_d * g_{d,i}}{\sum_i g_{d,i}} \tag{4}$$

Fig. 3 Flow pattern for case-DS

2.3 Case-DSS

For realistic representation of the cumulative plots, saturation flow rate is considered and the counts during the green interval are segregated into counts from the saturation flow pattern and counts from the demand pattern. We define the demand, which is the cumulative plot *(CP_demand)* at the location of the stop-line detector assuming point (vertical) queue at intersection. It can also be defined as the expected cumulative plot at the location of stop-line detector if there is no restriction, at the intersection, on the flow of the vehicles.

At a signalized intersection (during the green phase) the vehicles from the queue are effectively discharged at saturation flow. Thereafter, the flow pattern follows the demand pattern. If demand and saturation flows are known, then accurate and realistic flow pattern considering saturation flow and non-saturation flow can be estimated.

For simplicity, we focus on green (g) for a complete signal cycle instead of $g_{d,i}$ (*i-th* green period during the detection interval). The g can extend in more than one detection interval. For instance, in Fig. 3, the first green g has the component $g_{d,1}$ during the indicated detection interval. The count, N_g, during a g is obtained by respectively adding the counts from all its components, if split in more than one detection interval. Out of N_g vehicles, n_s vehicles enter the intersection at saturation flow pattern and the remaining ($N_g - n_s$) follow the demand pattern. The maximum number of vehicles which can depart during g is $N_{max}(=s*g)$, where s is saturation flow rate (vehicles/second).

For a link between two consecutive intersections as shown in Fig. 1, the demand pattern, for the detector at the downstream end of the link, can be deduced from $U(t)$. However, for a network there can be certain links where the $U(t)$ is unknown such as at the entrance of the network, here demand can be assumed. Therefore, the following two cases of assumed and deduced demand patterns are considered to estimate cumulative plots for case-DSS.

2.3.1 Assumed demand pattern

The detector counts represent demand for under-saturation situations. However, for over-saturation situation, the counts are upper bounded by capacity and that is less than true demand. Therefore, demand estimated in this case is termed as "*assumed demand*".

The demand flow pattern can be assumed to follow a uniform pattern (deterministic) or can be assumed to be distributed according to some probability distribution (stochastic). To simplify the analysis it is assumed that demand is uniform during the signal cycle. As shown in Fig. 4, N_g numbers of vehicles are counted during the green phase that represents the uniform demand for the signal cycle. By superimposing saturation flow pattern (during the
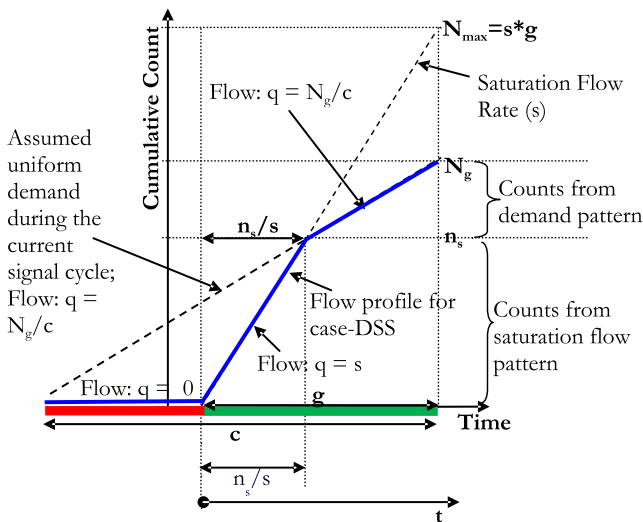
Fig. 4 Assumed demand: Geometrical relationship between ns and Ng assuming uniform demand pattern during the current signal cycle

green phase) on the uniform demand pattern the following relationship can be geometrically obtained:

$$\frac{n_s}{N_g} = \frac{(1-g/c)}{\left(1-X*\frac{g}{c}\right)} \quad for\ X < 1$$
$$= 1 \quad for\ X \geq 1 \quad (5)$$
$$where:$$
$$X = \frac{N_g}{N_{max}} \quad when\ Ng < N_{max}$$

The saturation flow starts at the beginning of the green period and lasts for n_s/s time units. Therefore, the flow pattern is defined as follows:

During Red Period
$$q = 0$$
During Green Period (6)
$$q = \begin{cases} s & 0 < t < n_s/s \\ \frac{Ng}{c} & n_s/s \leq t \leq c \end{cases}$$

where: t is the time since the start of the green within the green period and c is the signal cycle time.

Equation 5 provides the ratio of the counts in saturation flow rate (n_s) to the total counts during a green interval (N_g). For under-saturation situations, the ratio N_g/N_{max} represents degree of saturation (X) and n_s/N_g is the proportion of demand in saturation flow rate. For a given degree of saturation, the higher the green split (g/c) the lower the n_s/N_g ratio; and for near to saturation situations the ratio is close to one. This is as expected, because as the demand approaches capacity almost all the vehicles are at saturation flow rate.

2.3.2 Deduced demand pattern

In this case, we are interested in estimating D(t), given U(t). The demand can be deduced from the upstream cumulative

plot. We name this demand the "*deduced demand*" and it is the horizontal shift of the U(t) by free-flow travel time ($tt_{freeflow}$) of the link i.e., U($t-tt_{freeflow}$). Here for simplicity, no platoon dispersion is assumed. It is found (presented in next section) that the model performs reasonably well with this assumption. Nevertheless, a platoon dispersion model can be adopted to estimate the demand.

The flow is defined as zero for red intervals. For green intervals, if CP_{demand} is greater than the cumulative counts (D(t)) then the flow is at saturation flow rate otherwise the flow pattern is same as demand pattern (see Eq. 7).

During Red Period
$$q = 0$$
During Green Period
$$if\ CP_{demand}(t) > D(t)$$
$$q = s \quad (7)$$
$$else$$
$$q = \frac{\partial CP_{demand}(t)}{\partial t}$$
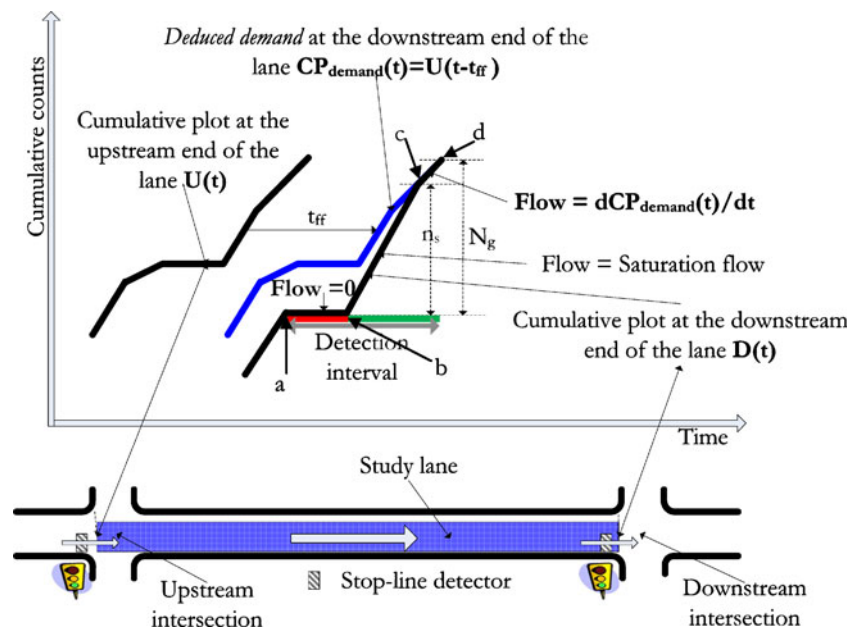$$where: CP_{demand}(t) = U(t - tt_{freeflow})$$

In Fig. 5, the known parameters are: the upstream cumulative plot, U(t); reference position (position *a* in the figure) for the D(t); signal timings at downstream intersection; and counts from the downstream end of the link (N_g). Flow pattern at downstream intersection for the current detection interval is unknown. This is obtained by no flow during red period (*a* to *b*) and during green period, the flow is at saturation flow until $CP_{demand}(t)$ is greater than D(t) (*b* to *c*) and thereafter flow follows the demand pattern (*c* to *d*).

Note: The flow pattern is estimated for each detection interval (case-D and case-DS) or for each signal cycle (case-DSS); and the polyline for the cumulative plots are generated by cumulating the profiles from each estimation interval taking into account the residual queue from the last interval

3 Model testing

The model is tested using a microscopic traffic simulator, AIMSUN [26] on a link between two consecutive signalized intersections. Travel time is defined as the time required travelling from the entrance of the upstream intersection to the entrance of the downstream intersection. The simulation network is similar to one illustrated in Fig. 1, where flow from three different directions at upstream intersection and a through movement at downstream intersection are considered. The flow on the upstream links is metered by signals at the further upstream. Scenarios for different degrees of saturation in the range of 0.5 to 1.2 at downstream intersection are simulated. The performance of the model, defined in terms of accuracy (%)

Fig. 5 Illustration of estimation of $CP_{DS}(t)$ for case-DSS with demand deduced from $CP_{US}(t)$

(Eq. 8), is evaluated for different detection intervals from 10 s to 360 s.

$$\text{Accuracy (\%)} = \left(1 - \frac{\sum_{i=1}^{N} \frac{|actual_i - estimated_i|}{actual_i}}{N}\right) *100 \qquad (8)$$

Where, N is the total number of travel time estimation intervals. $Actual_i$ and $estimated_i$ are the average actual travel time and average estimated travel time for i^{th} estimation interval, respectively.

The results presented here are from simulation with signal cycle time of 120 s and green split of 0.5 at both upstream and downstream intersections. Average travel time for 6 min is estimated from simulation of 1 h for each scenarios mentioned above.

For over-saturation situations if links are short then the queues are likely to extend to the upstream end of the link. Such situation will affect the saturation flow rate at the upstream intersection and therefore for case-DSS, the saturation flow rate has to be appropriately corrected. It should not affect the estimation for case-D and case-DS. The aim of the current analysis for case-DSS is to test the methodology for a given saturation flow rate and therefore a constant saturation flow rate is considered. For higher degree of saturation the queue may extend to the upstream end of the link (spill-back) resulting in drop in the observed flow at the upstream stop-line detector. This should affect the methodology because the methodology does not depend on the flow rate.

Figure 6 represents the graphs for detection intervals versus accuracy for the three cases. Each point on the graph represents the average of the accuracies obtained from different degree of saturation for a given detection interval. As expected, short detection intervals have higher accuracy levels irrespective of the cases and for detection intervals less than 30 s the estimation is very accurate. Detection interval is not critical if signal timings are available (Case-DS). Comparable accuracy can be obtained from a) detector data from larger detection intervals with signal timings; and b) detector data from shorter detection intervals without signal timings. If detection interval is short, then signal timings and saturation flow rate are not required. For case-D, the performance is not consistent for different detection intervals and in this example the accuracy drops significantly to 80% when detection interval is close to integral multiple of signal cycle for instance 120 s, 240 s and 360 s. This inconsistency in the performance for case-D is analyzed in the next section.
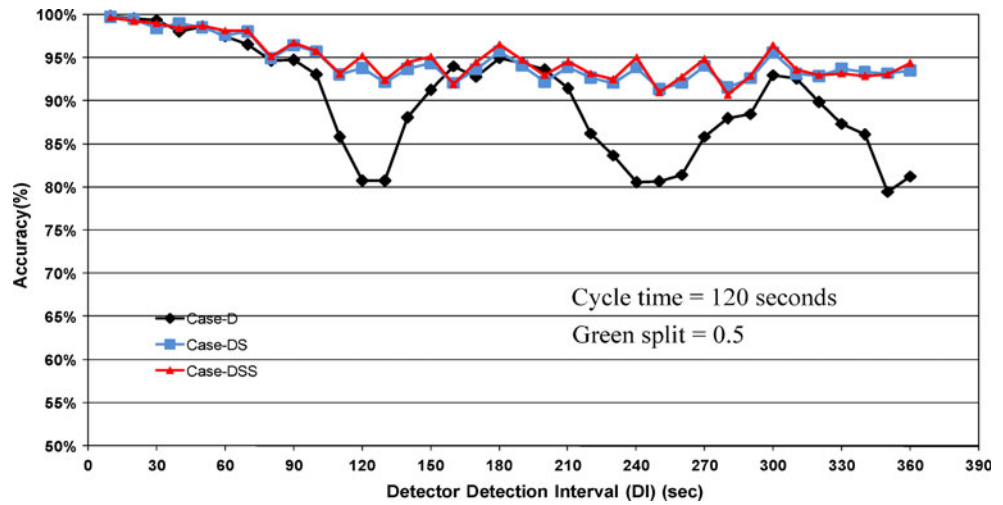
4 Sensitivity analysis

The cumulative plots generated for case-DSS are realistic and accurate (Fig. 7), whereas for case-D they are simplest but with inconsistency in the performance for travel time estimation. Therefore, the sensitivity analysis for case-D and case-DS is performed, considering case-DSS as a reference, and with a goal to determine a) the parameters which contribute to the inconsistency; and b) the values of the parameters for which the model is most accurate and reliable.

Following parameters are considered for the analysis:

I. Detection Interval: defined by variables $\beta*c$, where β is a rational number and c is signal cycle time.

Fig. 6 Comparative overview of the performance of the three cases as detection interval versus accuracy graphs

II. Signal green time: defined by variable g. It can be shown that as the green split (g/c) increases case-D and case-DS approaches case-DSS.

III. Sequence of signal phases in the detection interval defined by variables $\alpha*c$ ($0 \leq \alpha \leq 1$) which is the time from the start of the detection interval to the start of the green period within the detection interval or in other words it is the offset of the green with respect to detection interval.

For a given detection interval and signal timings there can be different patterns of signal timings within the detection interval. Figure 8a illustrated different patterns of signal timings within a detection interval. These patterns determine the shape of the cumulative plot for case-DSS. For consecutive detection intervals these patterns will change from one detection interval to another, except for detection interval which is integral multiple of signal cycle.

However, for fixed signal cycle with rational value of β, the pattern will repeat itself after certain time. For instance, in Fig. 8b where $\beta=1.5$ third pattern is similar to first pattern.

To make the analysis valid for any cycle time the above defined variables are normalized with signal cycle time (c).

IV. Degree of saturation (X): The degree of saturation determines the proportion of counts in the saturation flow rate and hence the shape of plots for case-DSS.

4.1 Sensitivity analysis case-D

In the previous section, the performance was evaluated by comparing the estimated and actual travel time using traffic simulation on a network. In this section, for sensitivity analysis, we evaluate the deviation of cumulative plot for case-D from that of case-DSS and estimate the corresponding

Fig. 7 Illustration of the cumulative plot for different cases and individual vehicle identification
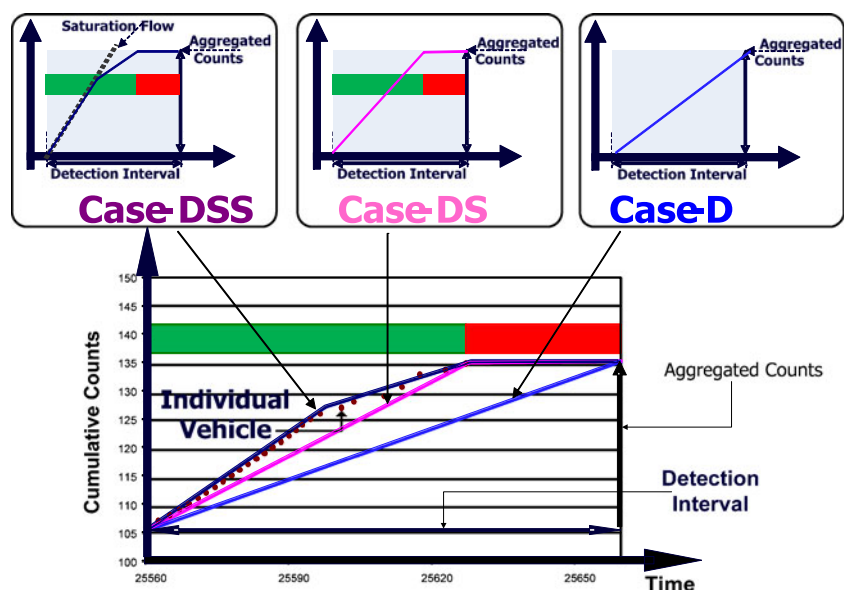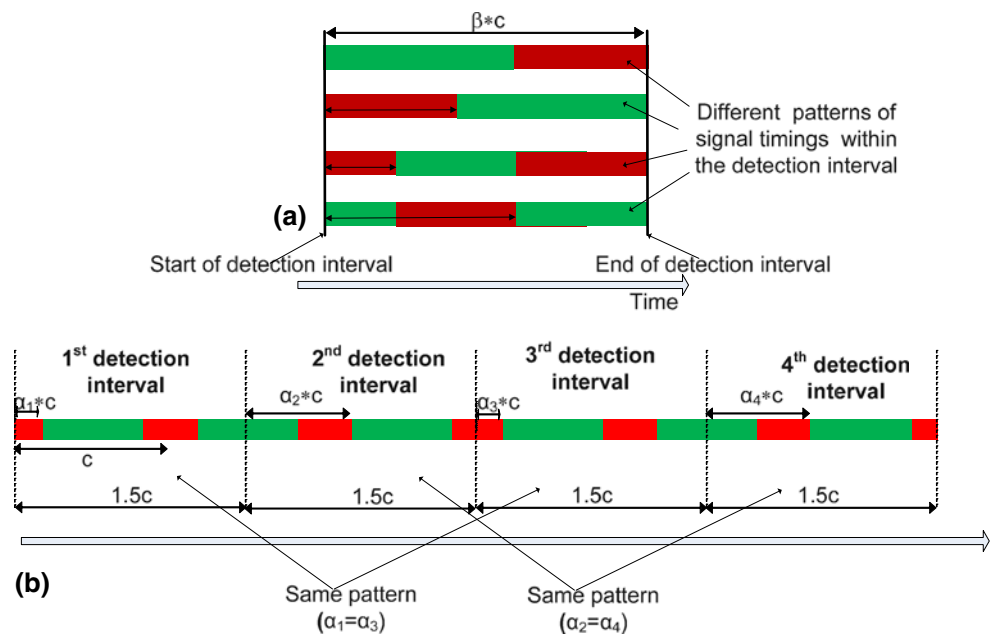
Fig. 8 Illustration of **a** several patterns of signal timings within a detection interval; and **b** patterns for consecutive detection intervals with $\beta=1.5$

accuracy in the travel time estimation by simple geometry. To explain the methodology for sensitivity analysis we consider an example (Fig. 9). For travel time estimation one is interested in the area between the cumulative plot for U(t) and D(t). The shape of the plot is defined by the parameters: β, g/c, α and X. In the figure, the performance for U(t), is evaluated for a scenario where $((\alpha + g/c \leq 1)$ and $(1 + \alpha + g/c < \beta \leq 2)$ and $(X < 1))$.

The accuracy for a combination of the parameters is defined as follows (Eq. 9):

$$\text{Accuracy (\%)} = \left(1 - \frac{|Area_{Case-DSS} - Area_{Case-D}|}{Area_{Case-DSS}}\right)*100 \tag{9}$$

Where: $Area_{Case-DSS}$ and $Area_{Case-D}$ are the areas under the plots for case-DSS and case-D, respectively as presented in the figure.

The analysis is performed for $0 < \beta \leq 2$; $0.5 \leq g/c \leq 0.9$; $0 \leq \alpha \leq 1$ and $0.5 \leq X \leq 1.2$. The performance of the model is defined in terms of accuracy (%) and standard deviation (σ^2). The accuracy and standard deviation presented for a parameter (say β) is the 5th percentile and standard deviation of the accuracies (Eq. 9) obtained from different scenarios considering all possible combinations with other parameters (g/c, α and X), respectively. So, accuracy is the minimum accuracy for 95% of the scenarios and standard deviation is an indicator for the relative reliability of the model. Higher σ^2 indicates lower reliability and vice versa.

4.1.1 Sensitivity with respect to β

The model is highly sensitive to detection interval and β is identified as a critical parameter. Figure 10a and e represent

graphs for accuracy and σ^2 versus β, respectively. The accuracy decreases from more than 95% to less than 85% with increase of β from 0 to 1, respectively and thereafter it increases(> 90%) till $\beta=1.5$ and decreases again to less than 85% for β close to 2. On the contrary, σ^2 monotonically increases for $0 < \beta < 1$ and $1.5 < \beta < 2$ and decreases for $1.5 < \beta < 2$ (Fig. 10e). This indicates that the model is least reliable when β is close to an integer (1 and 2) and for $1 \leq \beta \leq 2$ the model is most accurate and reliable when β is close to 1.5.
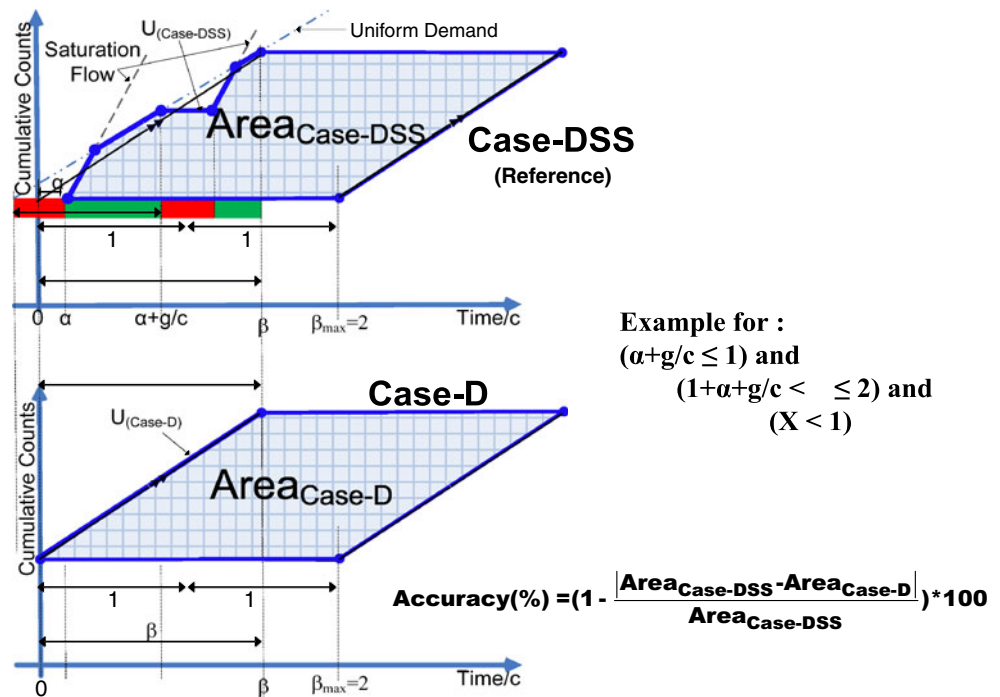
4.1.2 Sensitivity with respect to g/c

Figure 10b and f represent graphs for accuracy and σ^2 versus g/c, respectively, for β equal to 1, 1.5 and 2. The graphs for β equal to 1 and 2 are the same. Accuracy increases and σ^2 decreases (reliability increases) with increase of g/c. For high g/c (> 0.85) the model is relatively insensitive to β and accuracy is more than 95% (Fig. 10b). Whereas, for lower g/c (<0.4) the model is highly sensitive to β. Relatively higher value of σ^2 (Fig. 10f) for β equal 1 and 2 is consistent with the results of the sensitivity analysis for β i.e., the model is least reliable for integer values of β and most reliable for β around 1.5.

4.1.3 Sensitivity with respect to α

Figure 10c and g represent graphs for accuracy and σ^2 versus α, respectively, for β equal to 1, 1.5 and 2. For β equal 1.5 (Fig. 10c), the accuracy is generally more than 90% whereas, for β equal 1 and 2, there is a significant fluctuation in the accuracy from less than 75% (for α

Fig. 9 Example for evaluation of case-D with case-DSS as reference

around 0.85) to more than 90% (for α around 0.3). Aforementioned (Fig. 8), for consecutive detection intervals with fixed signal timings the pattern of signal timings (α) within a detection interval is constant for integral values of β. The sensitivity of the model to α for integral values of β, makes it unreliable for such values of β. However, if tuned properly by choosing an appropriate α (e.g. $\alpha=0.3$) it can give good estimations.

4.1.4 Sensitivity with respect to X

Figure 10d and h represent graphs for accuracy and σ^2 versus X, respectively, for β equal to 1, 1.5 and 2. Relatively the model is less sensitive with respect to X and the accuracy increases by 2% for increase in X from 0.5 to 1. The relatively higher value of σ^2 (Fig. 10h) for β equals to 1 and 2 is due to low reliability of the model for integer β.

4.2 Sensitivity analysis: case-DS

The difference between case-DSS and case-DS is that case-DSS considers saturation flow rate, and its flow profile depends on demand pattern, green split and degree of saturation. Similarly to the sensitivity analysis for case-D, the sensitivity analysis for case-DS is performed considering case-DSS as reference and the results are presented in Fig. 11. It is found that the accuracy for case-DS is generally higher than 94% ($\sigma^2<2\%$) and is slightly sensitive to the parameters.
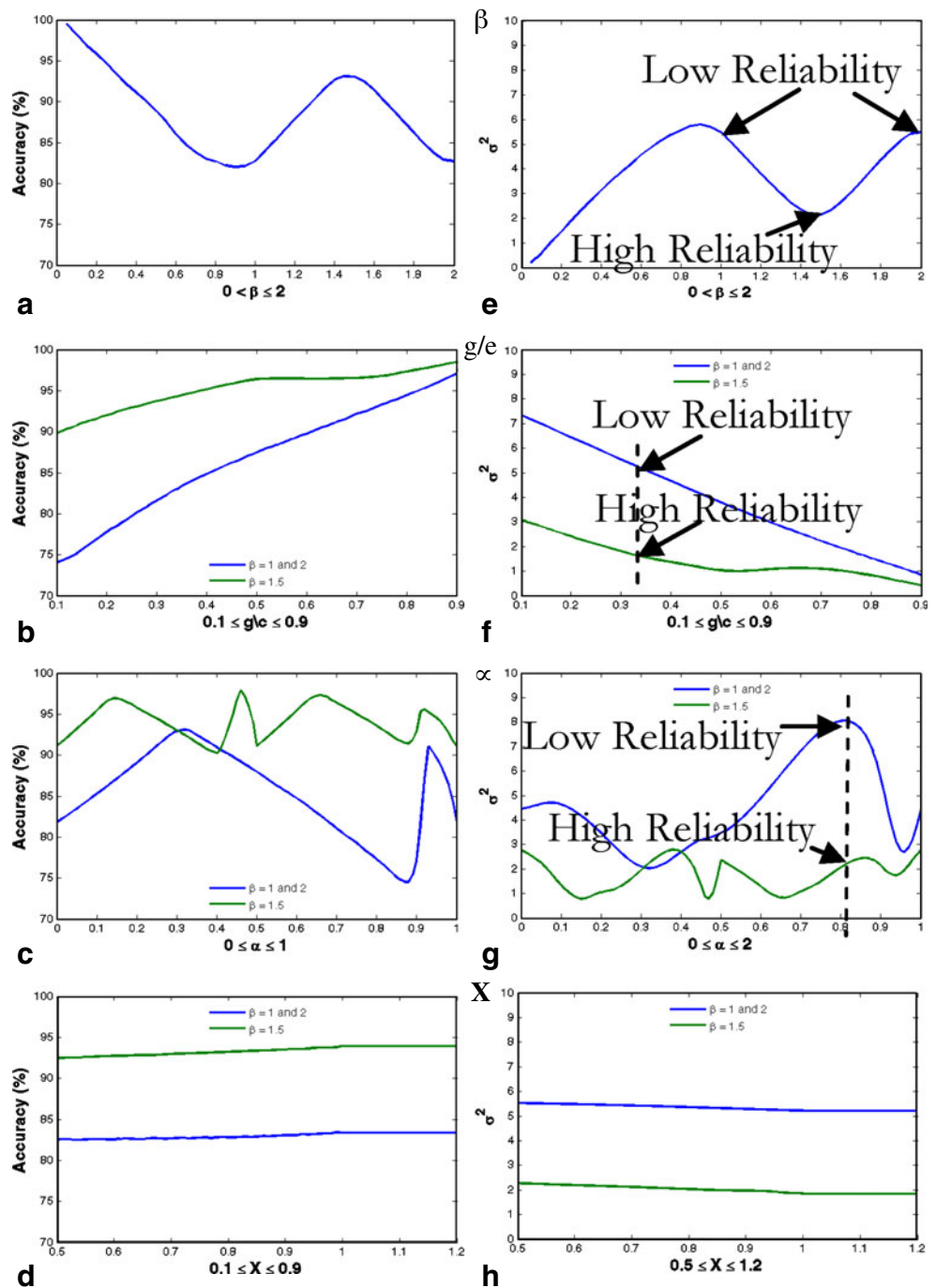
Similarly to case-D, for $1\leq\beta\leq2$ the highest accuracy and reliability are for β close to 1.5 (Fig. 11a and e). For $0.1\leq g/c\leq0.5$ the accuracy drops from more than 99% to 96% and for $0.5\leq g/c\leq0.9$ the accuracy increases back to 99% (Fig. 11b and f). The model is almost insensitive to α and there is slight decrease in accuracy and reliability for $0.75\leq\alpha\leq0.85$ (Fig. 11c and g). Accuracy and reliability of the model actually increases with the increase in the degree of saturation (Fig. 11d and h). For $X \geq 1$, all the counts are in saturation flow and case-DS is same as case-DSS (100% accuracy).

As the accuracies are generally more than 94% with σ^2 less than 2%, it is reasonable to conclude that case-DS is generally consistent with case-DSS and even in the absence of saturation flow rate information one could obtain travel time with reasonable accuracy by integrating aggregated detector data with signal controller data.

5 Explanation of the findings

The reason for low reliability at integral values of β and high accuracy for β around 1.5 can be explained with a help of an example. Let us consider $\beta=1.5$ (Fig. 12) with different patterns of signal timings ($\alpha=0$, g/c, $0.5*(1-g/c)$, $(1-0.5*g/c)$) and compare deviation of the areas for travel time estimation from flow profiles under case-D and case-DS. Then, for case-D, there is always a counter balance for under-estimation or over-estimation of area, which explains the improvement in accuracy. However, for $\beta=1$ (Fig. 13),

Fig. 10 Accuracy and standard deviation obtained from the sensitivity analysis for case-D with case-DSS as reference
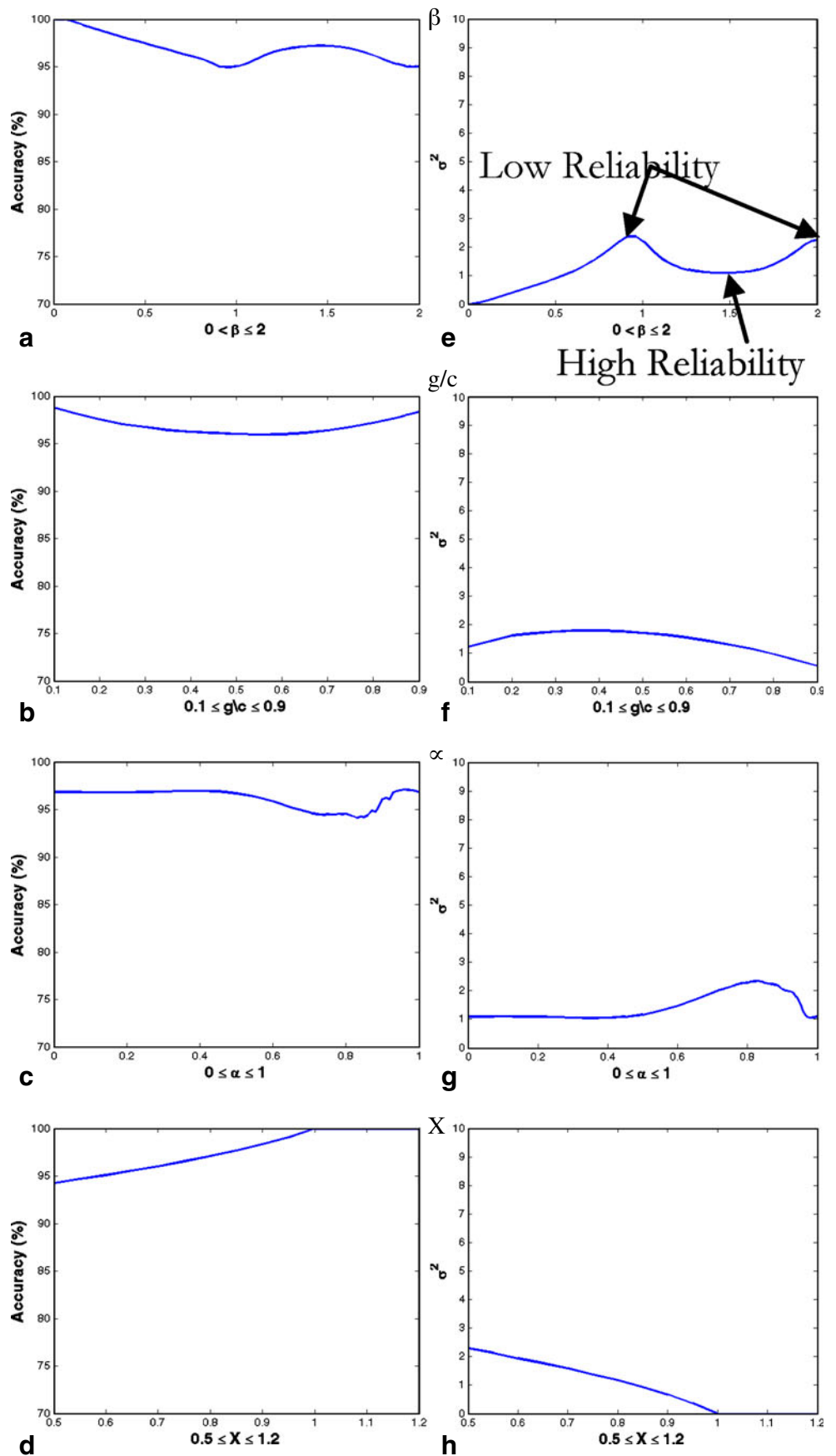

with $\alpha = 0$ and $\alpha = g/c$ there is either underestimation or overestimation with no counter balance area (lowest accuracy) and for $\alpha = 0.5*(1-g/c)$ and $(1-0.5*g/c)$ there is a perfect balance of areas (highest accuracy). Therefore, for integral values of β, the estimation can range from perfect to worst which accounts for its low reliability.

In the above qualitative comparison we have considered case-DS as a reference instead of case-DSS due to simplicity in illustration of flow profiles. Consideration of flow profiles for case-DSS will not affect the above qualitative comparison.

6 Discussion and conclusion

This paper analyses the performance of the classical analytical procedure, utilizing cumulative plots, for travel time estimation on signalized urban network, with respect to various parameters: β, for the detection interval; g/c, for the green split; α, for offset between detection interval and green period; and X, for degree of saturation. Such fundamental analysis is not being performed in literature, although cumulative plots are foundation for numerous analytical models.

Fig. 11 Accuracy and standard deviation obtained from the sensitivity analysis for case-DS with case-DSS as reference
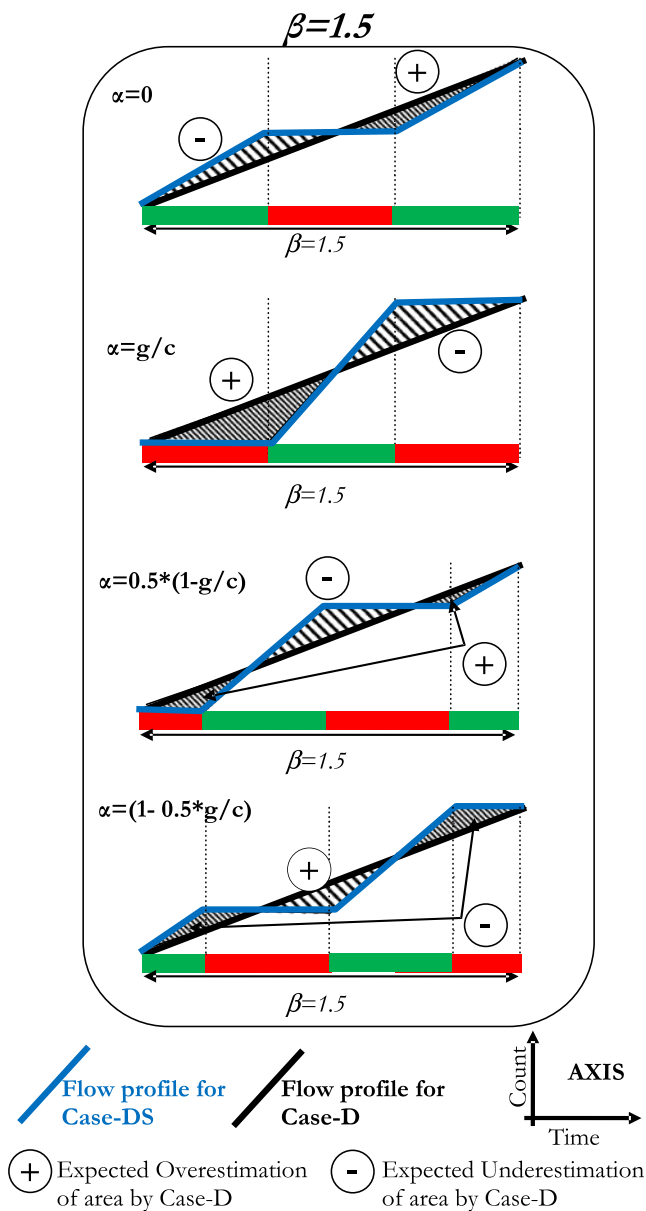
Fig. 12 Deviation in area for travel time estimation of case-D from case-DS under different values of α and for β=1.5 (assuming area to the right of cumulative plot is of interest)

The travel time estimates for case-DSS is very accurate and can be used as a reference. For the other cases, sensitivity in accuracy is tested with respect to the model parameters (β, g/c, α and X).

For case-DS, the accuracies are generally more than 94% with standard deviation less than 2%, so it is reasonable to conclude that even in the absence of saturation flow rate information one should obtain travel time with reasonable accuracy by integrating detector data with signal controller data.

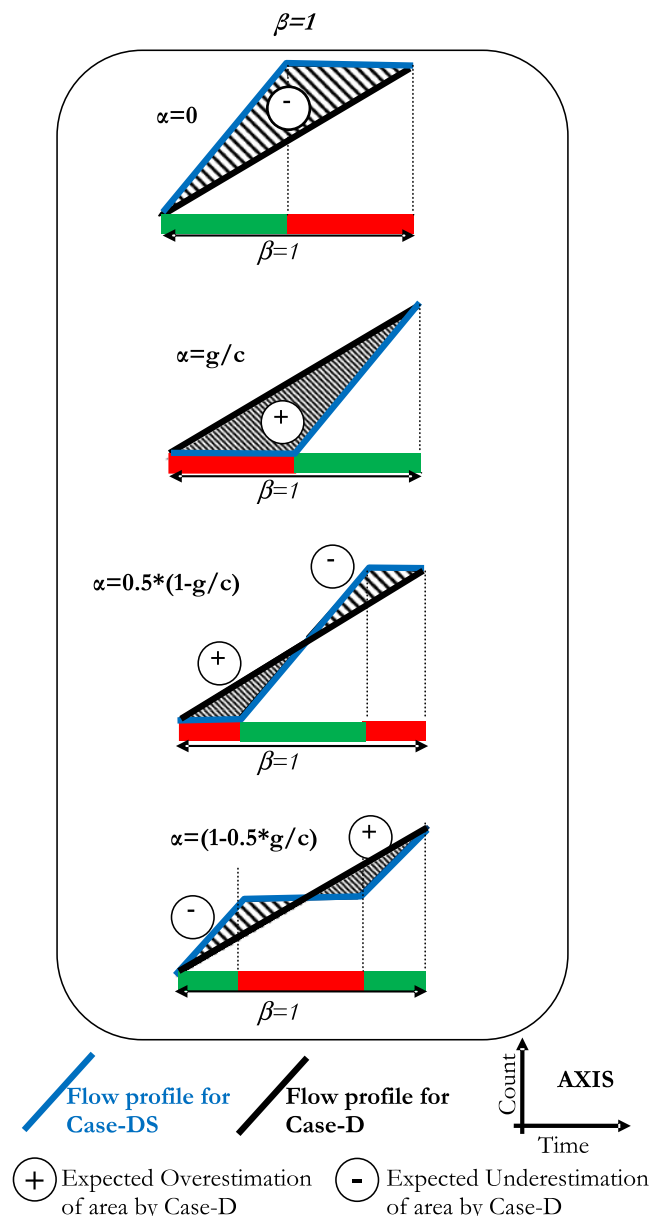In fact, with small values of β, accuracy is close to perfection in case-D also. Yet, the sensitivity analysis for

Fig. 13 Deviation in area for travel time estimation of case-D from case-DS under different values of α and for β=1 (assuming area to the right of cumulative plot is of interest)

The classical procedure is vulnerable to relative deviation ("drift") amongst the plots. Methodology that address the relative deviation issue is presented by Bhaskar et al. [25]. Therefore, in this paper we assume no relative deviation amongst the plots and focus on the issue when the available data from the detectors are aggregated resulting in loss of information for the actual fluctuations in traffic flow due to signals. Analytical modeling is performed to generate three different models based on the availability of data: case-D: only detector data is available; case-DS: detector and signal controller data is available; and case-DSS: detector data, signal controller data and saturation flow rate is available.

$1\leq\beta\leq2$ indicates that case-D is highly sensitive to detection interval. For β around 1.5, the model is most accurate with high reliability, whereas, for β close to 1 and 2, the model is least accurate with low reliability. Therefore, it can be argued, if only aggregated data is available then, for better confidence in travel time estimation, aggregation interval close to integral multiple of signal cycle time should be avoided. In fact, X has relatively little impact on the sensitivity of case-D. As for g/c and α, they are the two secondary most important factors for the sensitivity of case-D; that is when β is close to 1 or 2.

The explanations for these findings (Section 5) are also provided, which enable us to generalize the results when only detector data is available. For $\beta>2$, detection interval should be chosen such that β is close to the half of an odd number (e.g., 2.5, 3.5, 4.5, etc.), because of high accuracy and more stability. On the contrary, integral values of β should be avoided because of its low reliability. If β is close to an integer, then accuracy and reliability can still be improved with high g/c or choosing $\alpha=0.3$. This generalization is consistent with the simulation results presented in model testing section for $1/12\leq\beta\leq3$, $g/c=0.5$ and $\alpha=0$.

References

1. Wardrop, J.G.: Journey speed and flow in central urban areas. Traffic Eng Control **9**, 528–532 (1968)
2. Gault, H.E.: An on-line measure of delay in road traffic computer controlled systems. Traffic Eng Control **22**, 384–389 (1981)
3. Sisiopiku, V.P., Rouphail, N.M., Santiago, A.: Analysis of correlation between arterial travel time and detector data from simulation and field studies. Transp. Res. Rec. 166–173 (1994)
4. Xie, C., Cheu, R.L., Lee, D.H.: Calibration-free arterial link speed estimation model using loop data. J Transp Eng **127**, 507–514 (2001)
5. Rice, J., Van Zwet, E.: A simple and effective method for predicting travel times on freeways. IEEE Trans Intell Transp Syst **5**, 200–207 (2004)
6. Nam, D.H., Drew, D.R.: Automatic measurement of traffic variables for intelligent transportation systems applications. Transp Res Part B Methodol **33**, 437–457 (1999)
7. Oh, J.S., Jayakrishnan, R., Recker, W.: Section travel time estimation from point detection data, in In proceedings of the 82th Annual Meeting of Transportation Research Board, Washington, D.C., U.S.A. (2003)
8. You, J., Kim, T.J.: Development and evaluation of a hybrid travel time forecasting model. Transp Res Part C Emerg Technol **8**, 231–256 (2000)
9. Bajwa, S.I., Chung, E., Kuwahara, M.: A travel time prediction method based on pattern matching technique. in 21st ARRB and 11th REAAA Conference Cairns, Australia (2003)
10. Dailey, D.J.: Travel-time estimation using cross-correlation techniques. Transp. Res. **27B**, 97–107 (1993)
11. Petty, K.F., Bickel, P., Ostland, M., Rice, J., Schoenberg, F., Jiang, J., Ritov, Y.A.: Accurate estimation of travel times from single-loop detectors. Transp Res Part A Policy Pract. **32**, 1–17 (1998)
12. Robinson, S., Polak, J.W.: Modeling urban link travel time with inductive loop detector data by using the k-NN method. Transp Res Rec **1935**, 47–56 (2005)
13. Park, D., Rilett, L.R., Han, G.: Spectral basis neural networks for real-time travel time forecasting. J Transp Eng **125**, 515–523 (1999)
14. Van Lint, J.W.C.: Online learning solutions for freeway travel time prediction. IEEE Trans Intell Transp Syst **9**, 38–47 (2008)
15. Chen, H., Grant-Muller, S., Mussone, L., Montgomery, F.: A study of hybrid neural network approaches and the effects of missing data on traffic forecasting. Neural Comput & Applic **10**, 277–286 (2001)
16. Dia, H.: An object-oriented neural network approach to short-term traffic forecasting. Eur J Oper Res **131**, 253–261 (2001)
17. Westerman, M., Litjens, R., Linnartz, J.-P.: Integration of probe vehicle and induction loop data: estimation of travel times and automatic incident detection. (1996)
18. Berka, S., Tarko, A., Rouphail, N.M., Sisiopiku, V.P., Lee, D.-H.: Data fusion algorithm for ADVANCE Release 2.0. Advance Working Paper Series, No. 48, University of Illinois at Chicago, Chicago, IL (1995)
19. El Faouzi, N.E.: Data fusion in road traffic engineering: an overview. in Proceedings of SPIE—The International Society for Optical Engineering, Orlando, FL, pp. 360–371 (2004)
20. Choi, K., Chung, Y.: A data fusion algorithm for estimating link travel time. ITS J **7**, 235–260 (2002)
21. Xie, C., Cheu, R.L., Lee, D.H.: Improving arterial link travel time estimation by data fusion. 83rd Annual Meeting of the Transportation Research Board (2004)
22. Daganzo, C.F.: Fundamentals of transportation and traffic operations. Pergamon, Oxford (1997)
23. Newell, G.F.: Applications of queueing theory. London New York—N.Y.: Applications of queueing theory (1982)
24. Webster, F.V., Cobbe, B.M.: Traffic signals: road research technical paper no. 56, Her Majesty's Stationery Office, London, England (1966)
25. Bhaskar, A., Chung, E., Dumont, A.-G.: Estimation of travel time on urban networks with midlink sources and sinks. Transp Res Rec **2121**, 41–54 (2009)
26. AIMSUN: www.aimsun.com. accessed December (2009)

Ashish Bhaskar is a post-doctoral scholar at Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. He obtained his PhD degree in Transportation engineering from EPFL; Master of Engineering degree from The University of Tokyo, Japan and Bachelor of Technology degree in Civil Engineering from Indian Institute of Technology, Kanpur, India.

Edward Chung is Professor of ITS research in Queensland University of Technology, Brisbane, Australia. He is also an Adjunct Professor at the Collaborative Research Centre for Advanced Mobility, University of Tokyo, Japan. He received his Bachelor degree in Civil Engineering and PhD in Traffic Engineering from Monash University, Australia.

André-Gilles Dumont is a Professor at EPFL and Director of LAVOC (Laboratory des voies de la Circulation). He teaches and conducts research in telematics, road design, pavement design and road maintenance.