Uncertainty in Saturation Flow Predictions

ANDRZEJ P. TARKO
Purdue University, USA
MARIAN TRACZ
Cracow University of Technology, Poland

ABSTRACT

Existing capacity manuals for signalized intersections admit rather considerable standard errors of saturation flow prediction reaching 8–10%. Errors in saturation flow predictions carry over to delay estimates and, consequently, they may lead to erroneous LOS estimates. There are three primary sources of errors in saturation flow predictions related to saturation flow models. First, temporal variance in a saturation flow causes a measurement error. This error increases uncertainty of saturation flow predictions since the measured values are used to develop predictive formulae of saturation flow. Second, omission of some capacity factors in predictive models increases the site-to-site variance of saturation flows not explained by the models. Third, an inadequate functional relationship between model variables and saturation flow rates adds a prediction bias. The discussion of these error components and pertaining countermeasures is presented.

1. INTRODUCTION

The assumption of fixed saturation flow *s* over a saturated green time is a convenient approximation (Figure 1), which makes capacity *c* of lane groups equal to saturation flow *s* times green ratio (*g/C*). The effective green signal is calculated as the displayed green plus the change period minus the lost time. Existing capacity prediction procedures focus on predicting the saturation flow rates that has been proven to depend on a considerable number of geometry and traffic characteristics. The factors of the lost time have not been reported in literature yet.

Predictive equations of saturation flow rates used in the existing capacity manuals (Akcelik 1981; Kimber et al. 1986; Teply 1984; TRB 1997) have been developed from field measurements. The authors of the British and Australian methods admit a rather high standard error of saturation flow predictions reaching 8–10%. It means that in 5% of cases, a committed error can approach or be over 300 veh/h/lane. The effect of this random error on LOS determination is investigated in this paper.

There are several sources of prediction errors. Temporal variance of a saturation flow causes a measurement error. In addition, a measurement technique itself may introduce additional uncertainty. An inadequate functional relationship assumed between model variables and saturation flow rates adds an additional prediction bias. Lastly, capacity factors omitted in the model are responsible for the site-to-site variance of saturation flows not grasped by the predictive model. An additional source, being more of users concern than of researchers, is inaccuracy in the input data that carries over onto the

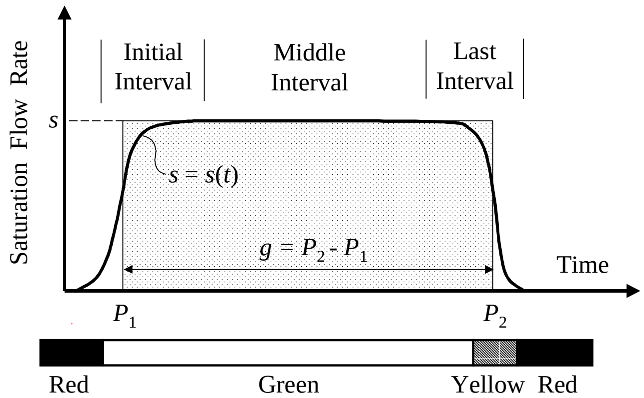

FIGURE 1 The rectangular model of saturation flow rate.

saturation flow predictions. The issues of measuring and modeling capacity are discussed in the Sections 3–5 following Section 2. The final remarks are given in the last section (Section 6) of the presentation.

2. MISCLASSIFICATION OF LOS

In the United States, the performance of highways is evaluated from the perspective of travelers using level of service (LOS). The traffic conditions are evaluated in 15-minute intervals of the analyzed period based on the predicted average control delay. Saturation flow rate is one of the main factors of control delay at signalized intersections.

As mentioned in the introductory part, a standard error of saturation flow prediction can be as high as 8–10%. Apparently, errors in saturation flow rates used to estimate vehicle delays carry over onto delay predictions and LOS predictions. There is no clear indication how frequently the LOS obtained from biased saturation flow rates can be incorrect.

A Monte Carlo experiment was performed to investigate the frequency of incorrectly determined LOS as an effect of inaccurate saturation flow predictions. Lane groups at isolated and signalized intersections were considered. The predicted saturation flow was 1,800 veh/h, the effective green was 50 seconds, and the signal cycle was 100 seconds. It was assumed that the actual saturation flow rates represented by the predicted value varied around the predicted value with a standard deviation of 180 veh/h. Unbiased prediction was assumed. The actual saturation flow rates where generated 1,000 times for each assumed traffic volume. The delays were calculated for the predicted and actual saturation flow rates. The determination of the actual and predicted LOS followed the calculations of the delays.

Figure 2 summarizes the simulation results. The vertical lines mark the ranges of degrees of saturation with the same predicted level of service indicated on the top of the graph box. For example, LOS D was claimed for the considered lane group if the predicted degree of saturation was higher than 0.88 but lower than 1.00. Specifically, if the predicted degree of saturation was 0.95 then the LOS was claimed to be D. The sloped curves (cumulative normal distributions) indicate the frequencies of actual delays being higher than the critical delay for a specific LOS. Each sloped curve corresponds with a particular LOS. For the example of predicted degree of saturation 0.95, the actual delay was higher than the critical delay for LOS C in 74% of cases, for LOS D in 32% of cases, and for LOS E in 11% of cases. This result means that if the predicted degree of saturation was 0.95 with claimed LOS D, the actual LOS was C in (100–74) = 26% of cases, E in (32 –11) = 21% of cases, and F in 11% of cases. Only in (74–32) = 42% of cases, LOS D was correctly determined. This situation was worse for LOS E, which was correctly claimed in 26% of cases.

Saturation flow predictions with the standard error lower than 10% would be represented by the cumulative normal curves steeper than in Figure 2. The absence of errors in saturation flow predictions would correspond to vertical cumulative curves matching the LOS boundaries. This is the ideal case of no prediction errors in LOS. On the other hand, a larger error in saturation flow prediction or additional errors (incorrect traffic volumes or signal timings) would make the cumulative normal curves in Figure 2 flatter.

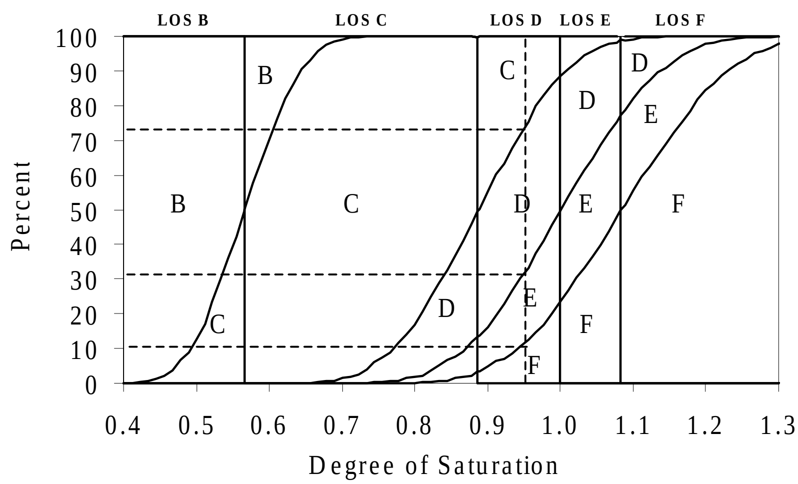

FIGURE 2 Misclassification of LOS caused by errors in saturation flow rate.

The simulation results prompt rather a discouraging conclusion about our limited ability to classify LOS correctly if the prediction error of saturation flows is considerable. It was shown that in most cases, the probability of correctly stating the LOS was lower than 50% and in the worst cases it was as low as 26%.

There is a strong need to improve the predictive methods of saturation flow rates. The following sections discuss the sources of errors in saturation flow predictions. Field data collected several years ago on Polish highways are used for illustration. The conclusions are expected to be transferable to other countries.

3. MEASUREMENT TECHNIQUES

All the existing methods of measuring saturation flows assume that saturation flow rate is fixed during a saturated green signal. Three distinguished measurement methods have been proposed:

a) *Headway method* (Greenshields et al. 1947; TRB 1997) estimates the average time headway between the vehicles discharging from queue as they pass the stop-line. The first several vehicles are skipped to avoid the effect of vehicles' inertia in the initial seconds of green time. The saturation flow rate is calculated as reciprocal of the mean headway.

b) *Regression technique* (Branston and Gipps 1981; Kimber et al. 1985; Stoke et al. 1987) is used to develop an equation involving saturated green time, number of vehicles in various categories, and lost time. A regression analysis yields the saturation flow, the lost times, and the passenger car equivalents for vehicles other than passenger cars.

c) *TRL method* (TRRL 1963), vehicles are counted in three saturated green intervals (Figure 1). The saturation flow is calculated as the number of vehicles in the middle interval divided by the length of this interval.

All these methods are appropriate for actual saturation flow profiles that have the plateau shape (Figure 1). These methods are not practical for cases where the saturated flow profile is more complex as for lane groups with protected and permitted turning movements (Figure 3). Use of a single saturation flow rate to calculate delays introduces a bias to the estimate of deterministic delay d_1.

An optimization technique was proposed by Tracz and Tarko (1992) to fit theoretical profiles similar to the one presented in Figure 3. To improve the estimation of the deterministic delay d_1, the optimization program minimizes the average square error between the observed and theoretical cumulative counts of vehicles departing from queue. The solution yields parameters s_i and P_i shown in Figure 3.

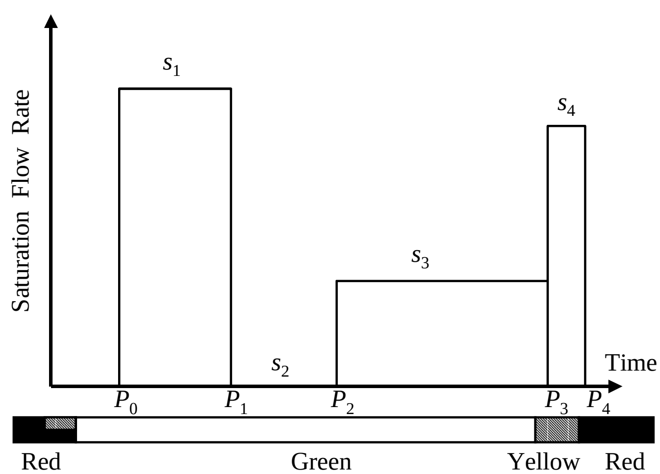

FIGURE 3 A saturation flow profile for a protected and permitted left-turn movement.

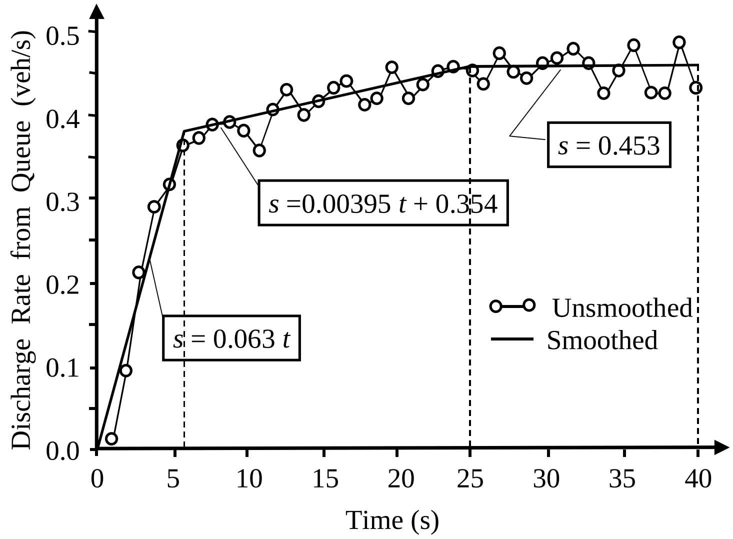

FIGURE 4 The average saturation flow profile for the 21 through lanes.

The plateau shape is commonly assumed for unopposed streams. This assumption was tested using extensive observations of through traffic in 21 traffic lanes located in four largest towns in Poland (Warsaw, Krakow, Gdansk, Katowice). The times when vehicles from queue passed the stop-line were recorded using press-button manual recorders. These times related to the beginning of green time were used to develop an average saturation flow profile representing all the 21 lanes (Figure 4).

Figure 4 shows the mean vehicle departure rates in one-second intervals obtained for the investigated lanes using observations for over 1,100 signal cycles. On average, a saturation flow rate builds up rapidly during the first 6 seconds of green signal. Then, in contradiction to the assumed plateau profile, it does not stabilize but grows further during the next 20 seconds. Typically, a saturation flow rate stabilizes around 25 seconds after the green signal onset. Although the periods with rapid and then gradual buildups of saturation flow were present in all the investigated lanes, the lengths of these periods varied somewhat across lanes.

A non-plateau profile as shown in Figure 4 introduces to the saturation flow measurements an additional site-to-site variance if the counting period varies across sites. To demonstrate this effect, the saturation flow rates have been estimated for various counting periods using the TRL method. The departure rates during the counting period were averaged. The saturation flow values for counting periods varying between 8 and 40 seconds are depicted in Figure 5. Since the first 6 seconds provide too short interval, they were excluded from the estimation. For comparison, the optimization-based method proposed by Tracz and Tarko has been applied too. The results are presented in Figure 5. The dashed line represents the TRL method while the solid line represents the optimization-based method. The results indicate a substantial variation of the saturation flow estimates with the length of counting period.

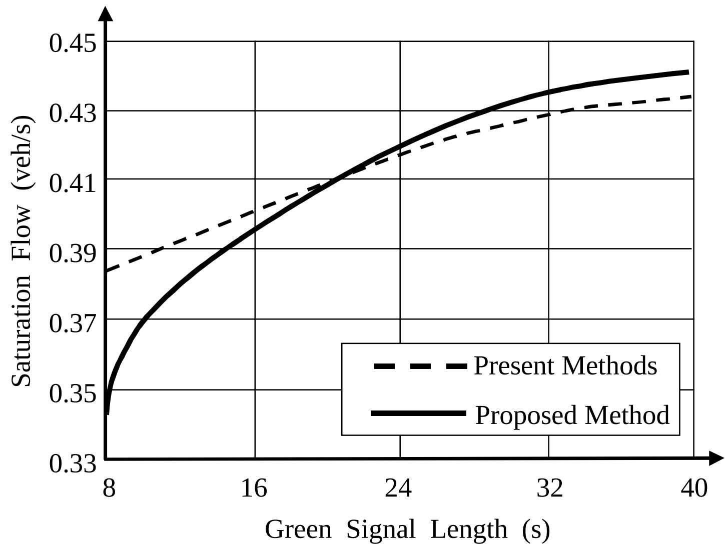

FIGURE 5 Variation of the saturation flow estimates with the length of counting period (green signal).

If the saturation flow rates are estimated with counting periods that vary from site to site, then an additional variance appears in the sample. This effect increases the prediction uncertainty but it does not cause any systematic bias if the counting period is independent of other saturation flow factors. Unfortunately, this may not be the case. For example, it is plausible to expect that the counting periods for straight streams are longer than for turning streams. The difference in the saturation flows between the straight and the turning streams will be a combined effect of the turning maneuver and the counting period. Neglecting the latter effect may lead to an overestimation of the effect of turning maneuvers.

There are two possible methods to avoid the prediction bias. First, the counting period could be included in the saturation flow measurement as an additional explanatory variable. Alternatively, the counting period can have the same length for all investigated traffic lanes. The first solution seems to be more practical since the field data would be used in a more efficient manner. The method of measuring saturation flow rate through measuring the departure times of queues as proposed by the *Highway Capacity Manual* introduces the same problem. The solution is the same — incorporate the effect of departure time to the estimation of the saturation flow rate.

4. HEAVY VEHICLES EFFECT

The saturation flow rate is lower if trucks and buses are present in the stream. Heavy vehicles may influence the saturation flow in two ways: (1) they use time headways longer than the time headways passenger cars use, and (2) they increase time headways of other vehicles.

The effect of heavy vehicles is incorporated into saturation calculations using a passenger car equivalent factor E. The equivalent factor E is the number of passenger cars that replace one heavy vehicle in queue without changing the expected time of queue discharge. The equivalent factor for heavy vehicles that includes only the first effect can be easily determined as h_{hv}/h_{pc}, where h_{hv} and h_{pc} are the average time headways of heavy vehicle and passenger cars, respectively. The average headways are between vehicles that discharge from queue.

TABLE 1 Headway Ratios h_{hv}/h_{pc}

Vehicle category	Equivalent factor	Site-to-site standard deviation
Passenger cars and vans	1.0	0.0
Single trucks and buses	1.6	0.1
Trucks with trailers and articulated buses	2.3	0.2
All heavy vehicles	1.7	0.2

The ratios of time headways of heavy vehicles and cars calculated from the data collected by the authors are presented in Table 1. These values represent the total effect of heavy vehicles only if the second effect of heavy vehicles does not occur. To test the hypothesis that trucks do not affect headway of other vehicles, it is sufficient to show that the average time headway for cars does not depend on heavy vehicle ratio u. The average time headways for cars (including vans) have been regressed against the heavy vehicle ratios observed in the investigated traffic lanes (Figure 6). The regression analysis has indicated a quite significant relationship between these two variables. It is clear that heavy vehicles affect time headways of cars. Consequently, the equivalents in Table 1 do not fully describe the effect of heavy vehicles.

The definition of equivalent factor will be used to determine its correct value. According to this definition

$$n \cdot h_{pc}(u) + m \cdot h_{hv}(u) = n \cdot h_{pc}(u = 0) + E \cdot m \cdot h_{pc}(u = 0) \tag{1}$$

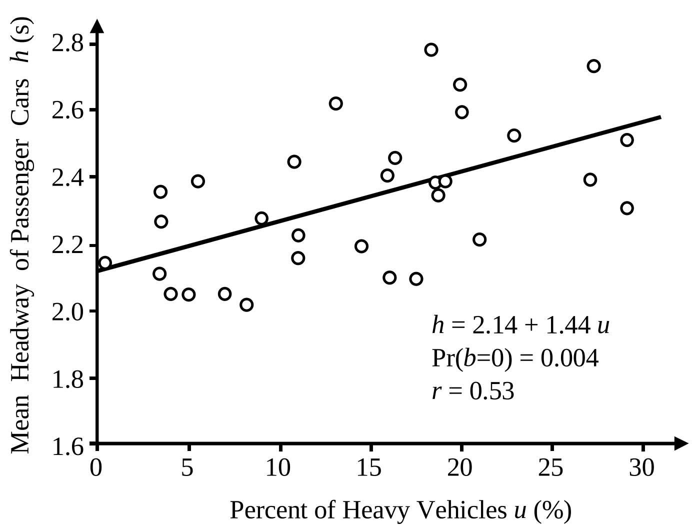

FIGURE 6 The effect of heavy vehicles on non-heavy vehicle time headways.

The left-hand side expresses the discharge time of a queue that is composed of n passenger cars and m heavy vehicles. The proportion of heavy vehicles is $u = n/(n + m)$. The right-hand side equation is the discharge time of the equivalent queue consisting of only passenger cars ($u = 0$). The function $h_{pc}(u)$ describes the relationship between the passenger car headway and the proportion of heavy vehicles in the queue. After dividing the both sides of the equation with $(n + m)$ and taking expectations, the following is obtained:

$$(1 - u) \cdot \mathrm{h}_{pc}(u) + u \cdot \mathrm{h}_{hv}(u) = (1 - u) \cdot \mathrm{h}_{pc}(u = 0) + \mathrm{E} \cdot u \cdot \mathrm{h}_{pc}(u = 0) \qquad (2)$$

where u is the expected proportion of heavy vehicles in queue.

Using sample average values of $u = 0.13$, and $h_{hv}/h_{pc} = 1.7$ (Table 1), and relationship $h_{pc}(u) = 2.14 + 1.44 \cdot u$ (Figure 6), the following is obtained: $h_{pc}(0.13) = 2.33$ s, $h_{hv}(0.13) = 1.7 \cdot h_{pc}(0.13) = 3.96$ s, and $h_{pc}(u = 0) = 2.14$ s. A new value of $E = 2.4$ is obtained by solving the above equation. The new equivalent factor is believed to include both the effects of heavy vehicles. Its value is considerably higher than 1.7 obtained when neglecting the relationship $h_{pc}(u)$.

Another way of evaluating the equivalent factor E is including the u ratio among the explanatory variables of a regression formula for saturation flow. This approach will be demonstrated in the next sections.

5. PREDICTIVE FORMULA

Saturation flow rates are affected by two types of factors. The first type—interruption factor—is an event that interrupts a vehicles stream. The first type includes parking vehicles, stopped buses, opposing traffic for left turns, pedestrians for right turns. These effects are rather well understood, thus they can be simulated or derived analytically. The second type—car following factor—affects spacing between moving vehicles. The second type includes such characteristics as heavy vehicles, lane width, horizontal curvature, and approach grade type. These effects are determined by human behavior and cannot be easily simulated. A statistical analysis of field data is a viable option. Resulting statistical relationships can be embedded in simulation or analytical models. The second type of factors is discussed in this section.

Predictive formulas for saturation flows have three distinctive forms:

multiplicative (TRB 1997),

$$S = S_0 \cdot f_1 \cdot f_2 \cdot f_3 ... \qquad (3)$$

additive (Kimber et al. 1986),

$$S = S_0 + \Delta S_1 + \Delta S_2 + \Delta S_3 ... \qquad (4)$$

and combined (Akcelik 1981; Teply 1984), where f_i and ΔS_i are functions of explanatory variable i.

The main difference between the first two models is the assumption of higher-order interactions between explanatory variables. Since any f_i can be presented as $[1 + (1 - f_i)]$, the multiplicative form can be transformed to the equivalent additive form with all the first and higher-order interaction variables

$$S = S_0 + \Delta S_1 + \Delta S_2 + \Delta S_3 ... + \Delta S_{12} + \Delta S_{23} + \Delta S_{13} ... + \Delta S_{123} ... + \tag{5}$$

where, for example,

$$\begin{aligned} \Delta S_1 &= S_0 \cdot (1 - f_1), \\ \Delta S_{12} &= S_0 \cdot (1 - f_1) \cdot (1 - f_2), \\ \Delta S_{123} &= S_0 \cdot (1 - f_1) \cdot (1 - f_2) \cdot (1 - f_3). \end{aligned}$$

From this perspective, the additive form is more flexible since it may include only significant higher-order variables. Besides, all the components of the additive expression are calibrated separately while the multiplicative form enforces certain dependencies between the components as presented above. On the other hand, the multiplicative formula may be justified by the actual interaction between factors, thus it may correctly imbed all the higher-order components of the additive form with a reduced number of parameters. Let us check the statistical fitness of the two types of models using the field data collected by the authors on Polish urban streets. The collected sample includes 38 traffic lanes with through streams, 21 with left-turning streams, and 10 right-turning streams. The following characteristics were used in the analysis of saturation flow:

- ratio of heavy vehicles,
- lane width,
- type of maneuver,
- turning radius (infinite for straight lanes),
- lane location on the approach (near curb, middle),
- presence of light rail crossing, and
- area type (commercial area, other).

The multiple regression indicated that the following four factors were statistically significant (10% significance level):

- ratio of heavy vehicles, u;
- lane width, w (meters);
- turning radius, r (meters); and
- lane location on the approach ($l = 1$ for near-curb location, 0 for middle location).

No higher-order variables were found significant. The correlation analysis did not show any strong dependence between the explanatory variables.

The resulting additive model is

$$S = 1850 + 80 \cdot (w - 3.5) - 3150/r - 1700 \cdot u - 75 \cdot l \tag{6}$$

The main drawback of this model is unrealistically low saturation flow for streams with high proportion of heavy vehicles ($S = 150$ veh/h for $u = 1.0$, $w = 3.5$, $r = \infty$, and $l = 0$).

This result is the effect of the linear relationship between S and u calibrated for the range of u between 0.0 and 0.30.

The multiplicative model obtained for the same data is

$$S = 1900 \cdot \left(1 + \frac{w - 3.5}{20}\right) \cdot \left(\frac{1}{1 + 2.4/r}\right) \cdot \left(\frac{1}{1 + (2.3 - 1) \cdot u}\right) \cdot \left(0.95 \cdot l + 1 - l\right) \quad (7)$$

The third considered model is a reciprocal model $S = 3600/h$, where h is the average time headway between vehicles departing from queue. The time headway was an additive function of the four significant variables. The obtained model is as follows:

$$S = \frac{3600}{1.88 - 0.14 \cdot (w - 3.5) + 5/r + 2.6 \cdot u + 0.13 \cdot l} \quad (8)$$

The last model regards the fact that the saturation flow rate is the reciprocal of the time headway drivers prefer to maintain in response to traffic and geometry cues. The denominator can be easily used in such simulation models as CORSIM where, so-called, saturation time headway has to be provided by the user.

It should be noticed that the passenger car equivalent for heavy vehicles calibrated for the multiplicative model is E = 2.3, a value close to 2.4 obtained in Section 4. The corresponding E in the reciprocal model can be obtained for ideal conditions ($w = 3.5$, $r = \infty$, $l = 0$) by calculating the average headway for cars (1.88 s when $u = 0$) and for heavy vehicles (1.88 + 2.6 = 4.48 s when $u = 1$). The headway ratio is 4.48/1.88 = 2.38.

The effectiveness of the models is evaluated using standard errors of prediction. Table 2 summarizes the standard errors for the three models already described. Additional three models in Table 2 demonstrate the multiplicative model with improper representation of two factors: heavy vehicles and turning maneuvers. The first additional model uses underestimated passenger car equivalent for heavy vehicles $E = 1.7$. The second model includes the average impact of turns through variable $t = 1$ for turning streams and 0 for through streams. No information about turning radii is used. This treatment yields a single-valued adjustment factor. The third model has both the improper E and the t variable instead of r.

Surprisingly, the effectiveness of the multiplicative, additive, and reciprocal models are is practically the same. Nevertheless, the reciprocal and multiplicative models seem to be more appropriate than the additive model. On the other hand, improper incorporation of heavy vehicles increases the prediction error considerable. The negligence of the turning radius in favor of a single factor for turning maneuver has even a stronger negative effect.

The model with combined deficiency has the prediction error increased from original 90 veh/h to 124 veh/h. These prediction errors correspond to the relative prediction errors 5.8% and 8.1%, respectively.

TABLE 2 Fitness of the Investigated Models

Model	Standard error of prediction (veh/h)
Additive	93
Multiplicative	90
Reciprocal	90
Multiplicative with $E = 1.7$	107
Multiplicative with t instead of r	114
Multiplicative with $E = 1.7$ and t instead of r	124

6. FINAL REMARKS

This presentation focuses on unopposed traffic streams and on saturation flow factors that can be classified as car-following variables. Various types of regression models have been investigated. In addition, effect of some deficiencies in the representation of selected impacts was analyzed.

Prediction errors in saturation flow rates may cause difficulties in correct determination of LOS—the main objective in evaluating signalized intersections. An incorrect LOS may be different from the actual one even by three levels. If the standard error of prediction is 10%, then the odds of correct determination of LOS D or E may be quite low. These results point out to the importance of correct estimation of the saturation flow rates. The reader should keep in mind that there are other sources of incorrect LOS such as inaccurate volumes or inaccurate signal timing.

There are two general sources of errors in saturation flow prediction: measurement and modeling. The actual saturation flow intensity can differ from the plateau shape—widely accepted for unopposed traffic streams. Use of counting periods varying from one site to another may introduce additional variance and uncertainty to the saturation flow measurements and then predictions. A systematic bias is possible if the counting period is correlated with other variables significant for the saturation flow. A practical remedy is to incorporate a counting period (queue discharge time) to the saturation flow predictive formula. This additional factor can be represented in the prediction models through an effective green time. If in some applications, such model is not practical, the counting period can be removed from the model by applying a default value.

The results of analysis have shown a substantial effect of heavy vehicles on time headways of passenger cars. An equivalent factor equal to a simple ratio of average time headways does not grasp the entire effect of heavy vehicles. Such a deficient equivalent factor considerably increases the prediction error of saturation flows. A simplified representation of turning movements also may lead to a considerable increase in the prediction error.

The effect of the model form was found non-consequential to the quality of saturation flow predictions. Nevertheless, the multiplicative and reciprocal forms are recommended. The additive model has slightly worse performance and may produce unrealistic results for inputs outside of the range covered by the field data use in the model calibration.

The strong effect of inaccurate saturation flow rates on the ability to correctly determine the LOS has been found for unopposed streams. Since unopposed streams are usually primary streams in highway networks, this finding indicates an obvious need for frequent updates of predictive formulas for saturation flows and for a careful consideration of local conditions. Where possible, the saturation flow rates should be determined through direct field measurements.

REFERENCES

Akcelik, R. (1981). *Traffic Signals: Capacity and Timing Analysis*. ARRB Research Record 123, Australian Road Research Board.

Branston, D., and Gipps, P. (1981). Some Experiences with a Multiple Regression Method of Estimating Parameters at the Traffic Departure Process. *Transportation Research*, 6A, pp. 445–458.

Greenshields, B.D., Shapiro, D., and Ericksen, E.L. (1947). *Traffic Performance at Urban Intersections*, Technical Report No. 1. Bureau of Highway Traffic, Yale University.

Kimber, R.M., McDonald, H., and Hounsell, N.B. (1985). Passenger Car Units in Saturation Flows: Concept, Definition, Derivation, *Transportation Research*, 1B, pp. 39–61.

Kimber, R.M., McDonald, H., and Hounsell, N.B. (1986). The Prediction of Saturation Flow for Road Junctions Controlled by Traffic Signals. *TRRL Research Report 67*.

Stoke, R.W., Stover, V.G., and Messer, C.J. (1987). Use and Effectiveness of Simple Linear Regression to Estimate Saturation Flows at Signalized Intersections, *Transportation Research Record 1091*, pp. 95–101.

Teply, S. (1984). *Canadian Capacity Guide for Signalized Intersections:* Institute of Transportation Engineers, District 7–Canada and The University of Alberta.

Tracz, M., and Tarko, A. (1992). A Method for Estimating Saturation Flows at Signal-Controlled Junctions. In *Mathematics in Transport Planning and Control*, Griffiths (ed.), *Proceedings of Conference: Mathematics in Transport Planning and Control*, Cardiff, 1989, Clarendon Press, Oxford, pp. 447–463.

Transport and Road Research Laboratory. (1963). *A Method of Measuring Saturation Flow at Traffic Signals*. Road Note No. 34, London.

Transportation Research Board. (1997). *Special Report 209: Highway Capacity Manual*, National Research Council, Washington, D.C.