Comparative Analysis of Models for Estimating Delay for Oversaturated Conditions at Fixed-Time Traffic Signals

W. B. CRONJÉ

To optimize a fixed-time traffic signal, a model is required to estimate with sufficient accuracy the measure of effectiveness necessary for the optimization process. Suitable models have been developed for the degree of saturation (x) in the range $0 < x < 0.9$. Reliable models have also been developed for the zone of the degree of saturation $x > 1.1$. In this zone, traffic can be treated deterministically. However in the range $0.9 \leq x$ 1.10, the deterministic approach falls and a model should be based on the probabilistic approach to traffic flow. The ideal model should be applicable over the entire range of the degree of saturation. Only two such models have been encountered in the literature. Because of shortcomings of these models and the lack of a reliable model in the transition zone from undersaturation to oversaturation, an alternative model was developed by Cronjé (*Transportation Research Record 905*, 1983). This model is based on a Markov process and the geometric probability distribution, and is referred to in this paper as the M Geom Model. In this paper, the M Geom Model is compared with the models developed by Mayne and Catling on a cost basis. Monetary rates are assigned to the measures of effectiveness, namely, total delay and number of stops, for a wide range of cycle lengths, flows, and degree of saturation. The results indicate that the M Geom Model estimates cost more accurately and is consequently recommended for optimizing fixed-time traffic signals.

The predominant equations for estimating delay for undersaturated conditions have been developed by Webster, Miller, and Newell. However, these equations are only reliable for the degree of saturation (x) in the range $0 < x <$ about 0.9. However, because of oversaturated conditions existing for periods during the peak hours in urban conditions, there is a need for a model applicable over the entire range of the degree of saturation. Three such models have been developed by Mayne *(1)*, Catling *(2)*, and Cronjé *(3)*.

Catling adapted equations of classical queuing theory to oversaturated traffic conditions and developed comprehensive queue length and delay formulas that can represent the effects of junctions in a time-dependent model.

Mayne used some results on transient queuing theory to derive comprehensive queue and delay formulas for use in a procedure to optimize traffic signal settings during peak periods. He transformed existing formulas into formulas using the Poisson distribution, then applied appropriate statistical distribution approximations and other techniques of numerical analysis.

Cronjé treated traffic flow through a fixed-time signal as a Markov process and developed equations for estimating delay, queue length, and stops. The properties of the geometric probability distribution were then applied to these equations to obtain simpler equations, reducing computing time. The equations were then modified to further reduce computing time without sacrificing too much accuracy.

The measures of effectiveness used in this paper to compare the three models are delay and stops, but Catling and Mayne did not derive equations for estimating stops. However, the number of stops for the Mayne and Catling models can be obtained from their queue length equations by using the definition that the number of stops is the number of vehicles arriving while there is a queue plus the overflow of vehicles at the beginning of the cycle.

The probabilistic model used to describe the traffic arrivals at the intersection is the Poisson distribution because the Mayne and Catling models are both based on the Poisson distribution.

The basis of comparison of the models is simulation. Traffic arrivals at a signalized intersection is a stochastic process. If a traffic count is taken on a specific day for a particular period at an intersection and the delay and stops calculated, and if a count is then repeated for the same time period the next day, it is possible that the values obtained for delay and stops may differ appreciably. Consider the following set of arrivals generated that were obtained over 15 cycles for the following prescribed data:

- Cycle length: $c = 50$ sec;
- Degree of saturation: $x = 1.05$;
- Flow: $q = 500$ vehicles per hour; and
- Ratio of variance to mean of arrivals per cycle: $I = 0.5$.

Arrivals (a): 8, 8, 6, 6, 8, 8, 7, 8, 8, 5, 9, 8, 9, 8, 10 Σ 116. Initial queue length is 2 vehicles; average number of departures per cycle is 7.4 vehicles.

The equation for obtaining the overflow at the end of a cycle is as follows:

$$K = B + Z - V \tag{1}$$

where

K = overflow at the end of the cycle,
B = overflow at the beginning of the cycle,
Z = arrivals per cycle, and
V = departures per cycle.

Department of Civil Engineering, University of Stellenbosch, Stellenbosch 7600, Republic of South Africa.

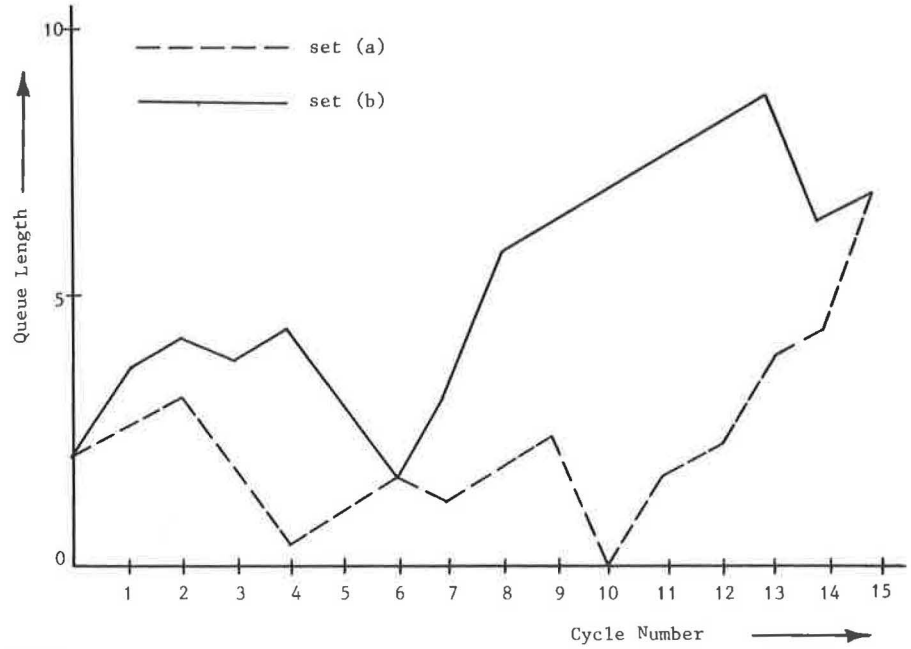

FIGURE 1 Queue length at end of 15 cycles.

Applying Equation 1 to the arrivals generated, the following values for K are obtained at the end of 15 cycles:

2.6, 3.2, 1.8, 0.4, 1.0, 1.6, 1.2, 1.8, 2.4, 0.0, 1.6, 2.2, 3.8, 4.4, 7.0.

Suppose the order of arrivals is changed to the following:

Arrivals (b): 9, 8, 7, 8, 6, 6, 9, 10, 8, 8, 8, 8, 8, 5, 8 $\Sigma = 116$.

The following values for K are now obtained:

3.6, 4.2, 3.8, 4.4, 3.0, 1.6, 3.2, 5.8, 6.4, 7.0, 7.6, 8.2, 8.8, 6.4, 7.0.

A plot of the two sets of overflow is shown in Figure 1. Although the area under the curves is not the delay, it is directly related to delay and obviously the delay differs appreciably for the two sets.

Calculating delay as the area under the queue length curve and calculating stops from the definition previously defined, the values for the two sets of arrivals obtained are given in Table 1 alongside those obtained by the M Geom Model.

From the values indicated in Table 1 it is therefore clear that the order of the arrivals affects delay and stops considerably. If actual counts have to be used, a fairly large sample will be required to properly estimate the population measures of effectiveness. On the other hand, the larger the sample, the longer is the calendar period over which the counts are taken; during this period, changes in the traffic flow conditions may occur, biasing the sample. This problem is eliminated with computer simulation, which shall therefore be applied.

Traffic arrivals are generated by random numbers over a wide range of cycle lengths, flows, and degrees of saturation. To introduce as large a degree of variation as possible into the sample, arrivals are generated for the binomial and negative binomial distributions for values of I varying from 0.5 to 1.5, excluding the case of $I = 1$, which gives the Poisson distribution. For each model the cost, obtained from assigning monetary rates to the measures of effectiveness—namely, delay and stops—is compared with the actual cost obtained for the arrivals generated. The results are analyzed and it is found that in the case of the M Geom and Catling models the cost difference follows the normal distribution at the 5 percent level. The models are compared extensively and the M Geom Model is found to estimate cost more accurately than do the Mayne and Catling models.

TABLE 1 DELAY AND STOPS FOR TWO SETS OF GENERATED DATA

	Delay		Stops			
Set	Actual	M Geom	Actual	M Geom	ΔD (%)	ΔN (%)
a	3,770	6,797	150	210	80.3	40.0
b	6,053	6,797	192	210	12.3	9.4

Note: ΔD = difference in delay and ΔN = difference in stops.

ANALYSIS

The equations used in the Mayne, Catling, and M Geom models for estimating delay, stops, and overflow shall now be given.

Mayne Model

Mayne used some results on transient queuing theory to derive comprehensive queue and delay formulas for use in a pro-

cedure to optimize traffic signal settings during peak periods. He transformed existing formulas into formulas using the Poisson distribution, and then applied appropriate statistical distribution approximations and other techniques of numerical analysis.

Queue Length Formulas

After applying further approximations, Mayne obtained the following queue length formulas.

Rising Queue The rising occurs under the following conditions:

$x \geq 1$; $x < 1$, and $Q < Q_e$.

$$Q(t) = (w^2 \cdot u + k_1)\,(w^2 \cdot u + k_2)^{-1/2}\,u^{1/2} - w \cdot u \qquad (2)$$

$$Q_e \doteqdot x/[2(1 - x)] = 1/4 \cdot w \qquad (3)$$

where

x = degree of saturation,
Q = queue length at start of period under consideration,
Q_e = equilibrium queue length,
$Q(t)$ = queue length at time t,
w = $(1 - x)/(2 \cdot x)$,
u = $q \cdot t$,
k_1 $\doteqdot$ 0.3417, and
k_2 $\doteqdot$ 0.1834.

Falling Queue The falling queue occurs under the following conditions:

$x < 1$ and $Q > Q_e$

$$Q(t) = (w^2 \cdot u^2 + 0.5 \cdot u + k_5 \cdot w^{-2})^{1/2} - w \cdot u \qquad (4)$$

where k_5 is 0.1202.

For rising queues, if $Q \neq 0$, then the origin, where the queue length is zero, has to be determined. This condition is shown in Figure 2. In the case of the falling queue, the zero origin is determined similarly except that the queue length at the origin is larger than Q.

Formulas for Obtaining Zero Origin

The formulas for obtaining zero origin are as follows.

Rising Queue The rising queue occurs under the following conditions:

$0 < x < 1$ and $Q < Q_e$, or $x > 1$.

$$R(Q) = (C - B)/(2 \cdot A \cdot q) \qquad (5)$$

where

$R(Q)$ = inverse function of $Q(t)$, which gives the approximate value of the time parameter for which the mean queue has value Q;
C = $(B^2 + 4 \cdot k_2 \cdot A \cdot Q^2)^{1/2}$;
B = $k_1^2 - 2k_2 \cdot w \cdot Q - w^2Q^2$;
A = $w^2(0.5 - 2 \cdot w \cdot Q)$; and
q = average arrival rate of traffic.

Falling Queue The falling queue occurs under the following conditions:

$0 < x < 1$ and $Q > Q_e$.

$$R(Q) = (k_5/w^2 - Q^2)/q(2 \cdot w \cdot Q - 0.5) \qquad (6)$$

Having determined $R(Q)$ from Equations 5 and 6, $Q(t)$ is then obtained from Equations 2 and 4 such that

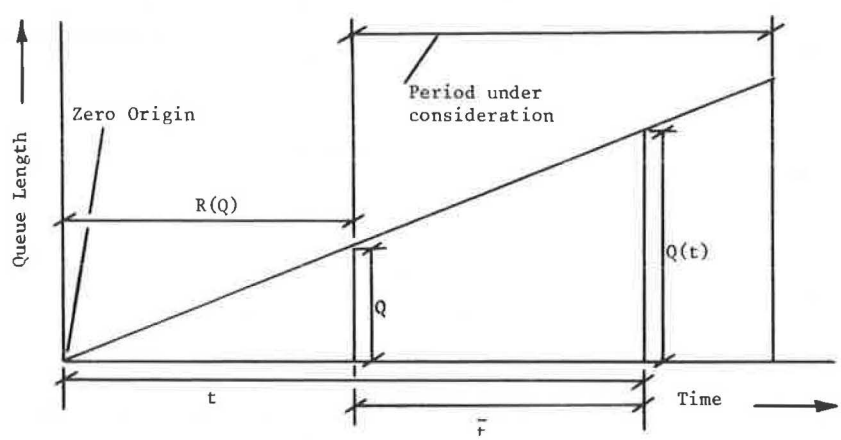

FIGURE 2 Determination of zero time origin for a rising queue.

$Q(t) = Q[R(Q) + \bar{t}]$

where $\bar{t}$ is the time from the start of the period under consideration to where the queue length is required.

Delay Formulas

The delay formulas are as follows:

Oversaturation For oversaturation, the following conditions are implied:

$x > 1$, $x < 1$, and $Q > Q_e$.

Delay during the peak period $[T, T + L]$, that is, of duration L, is given by

$$R(T,T + L) = S(T,T + L) + 0.5 \cdot c\,(1 - \lambda)\,[W + Q(T + L) - Q(T)] \quad (7)$$

and

$$S(T,T + L) = \int_T^{T+L} Q(t) \cdot dt = c\,[0.5(Q_o + Q_n) + Q_1 + Q_2 + \ldots + Q_{n-1}]$$

where

W = $q \cdot L$ = total approach flow during period under consideration,
L = length of period under consideration,
Q_o = $Q(T)$ = queue length at start of period under consideration,
Q_n = $Q(T + L)$ = queue length at end of period under consideration,
c = cycle length,
λ = g/c = proportion of cycle that is effectively green, and
g = effective green time.

Undersaturation For undersaturation, the following conditions are implied:

$x < 1$ and $Q < Q_e$.

Mayne, however, suggests the use of Webster's approximate Equation 8 for average delay for $0 < x <$ approximately 0.9 and $Q < Q_e$.

$$d = 0.9\,[c(1 - \lambda)^2/2(1 - \lambda \cdot x) + x^2/2 \cdot q(1 - x)] \quad (8)$$

where d is average delay per cycle.

It follows that

$$D = d \cdot q \cdot c$$

where D is the total delay per cycle.

Catling Model

Catling adapted existing equations of classical queuing theory to oversaturated traffic conditions and developed comprehensive queue and delay formulas that can represent the effects of junctions in a time-dependent model.

Queue Length Equations

The queue length equations are as follows.

Rising Queue The same conditions for a rising queue apply as for the Mayne model. As in the case of the Mayne model, if there is a queue at the start of the period under consideration, the origin at which the queue length is zero has to be determined. For the Mayne model the origin is calculated directly but in the Catling model it has to be obtained by trial and error.

$$G(x, t) = [(\beta^2 + 2 \cdot x^2 \cdot Q^2 \cdot t^2 \cdot \alpha \cdot C)^{1/2} - \beta]/\alpha \quad (9)$$

where

$G(x, t)$ = queue length at time t with degree of saturation x,
t = time from zero origin to point where queue length is required,
α = $2(Q \cdot t - C)$,
β = $Q \cdot t\,[Q \cdot t\,(1 - x) + 2 \cdot C \cdot x]$,
Q = $s \cdot g/c$, and
C = 0.55.

The equilibrium queue length, $Gl(x)$, with degree of saturation x, derived by Catling, is given by

$$Gl(x) = C \cdot x^2/(1 - x) \quad (10)$$

Falling Queue The same conditions for a falling queue apply as for the Mayne Model. In this case, a zero origin is not determined. The queue length formula is as follows:

$$G(x,t,G_o) = G_o - Q \cdot t(1 - x) \quad (11)$$

where t is the time measured from the start of the period under consideration, and G_o is the queue length at the start of the period under consideration.

Delay Formulas

Catling gives average delay formulas as follows.

Rising Queue The following conditions are implied: $x \geq 1$, $x < 1$, and $G_o \leq Gl(x)$.

$$d(x,t,G_o) = [(t + t') \cdot D(x,t + t') - t' \cdot D(x,t')]/t \quad (12)$$

where

t	=	length of time period under consideration;
t'	=	time from the zero origin to start of time period under consideration;
$d(x,t,G_o)$	=	average delay over period under consideration with degree of saturation x and initial queue length G_o;
$D(x,t + t')$	=	average delay over period $(t + t')$, that is, from the zero origin to the end of the period under consideration; and
$D(x,t')$	=	average delay over period t', that is, from the zero origin to the start of the period under consideration.

In general,

$$D(x,t) = A + (N - M)/4 \cdot Q \quad (13)$$

where

M	=	$2 \cdot C + t \cdot Q(1 - x)$,
N	=	$(M^2 + 8 \cdot C \cdot Q \cdot t \cdot x)^{1/2}$,
$D(x,t)$	=	average delay over time period t with degree of saturation x, and
t	=	time period from zero origin.

Falling Queue The following conditions are implied: $x < 1$ and $G_o > Gl(x)$.

In addition, the following conditions are treated:

$$t \le [G_o - Gl(x)]/Q(1 - x) \text{ and } t > [G_o - Gl(x)]/Q(1 - x)$$

$$t \le [G_o - Gl(x)]/Q(1 - x)$$

The average delay equation is as follows:

$$d(x,t,G_o) = A - 0.5 \cdot t(1 - x) + G_o/Q \quad (14)$$

where A is $0.5 \cdot c\ (1 - \lambda)^2$ and t is length of time period under consideration.

$$t < [G_o - Gl(x)]/Q(1 - x)$$

$$d(x,t,G_o) = A + Gl(x)/x \cdot Q + [G_o - Gl(x)]\{x[G_o + Gl(x)] - 2 \cdot Gl(x)\}/2 \cdot x \cdot Q^2 \cdot t(1 - x) \quad (15)$$

For use in a comparative analysis, Equations 8, 12, 14, and 15 have to be transformed to total delay such that

$$D(x,t,G_o) \doteq d(x,t,G_o) \cdot q \cdot t$$

where $D(x,t,G_o)$ is the total delay over time period t under consideration with degree of saturation x and initial queue length G_o, and q is the average arrival rate of traffic over time period t.

M Geom Model

The equations of the M Geom Model are as follows:

$$E(K) = E(B) + E(Z) - E(V) - \sum_v p(v) \sum_{z=o}^{v-1} p_Z(z)[E(B)(1 - F^{v-z}) - v + z] \quad (16)$$

$$E(N) = E(B) + E(Z) + \sum_v p_V(v) \sum_{z=o}^{v-1} p_Z(z)\ (z/c) + (v/g - z/c)[E(B)(1 - F^{v-z}) + z - v] \quad (17)$$

$$E(D) = E(B)c + [E(Z)c - E(V)g]/2 + \sum_v p_V(v) \sum_{z=o}^{v-1} p_Z(z)\{E(B)[2 - F^{v-z}(1 + F)] + (v - z)^2 - (v - z + 1)^2 F\}/2(1 - F)(v/g - z/c) \quad (18)$$

where

$E(K)$	=	expected value of the overflow at the end of the cycle,
$E(N)$	=	expected value of the number of stops per cycle,
$E(D)$	=	expected value of the total delay per cycle,
$E(B)$	=	expected value of the overflow at the beginning of the cycle,
$E(Z)$	=	expected value of the arrivals per cycle,
$E(V)$	=	expected value of the maximum number of departures per cycle, and
F	=	$E(B)/[1 + E(B)]$.

The range of the degree of saturation is divided into three zones as follows:

- Zone 1: $0 < x < 0.9$,
- Zone 2: $0.9 \le x < 1.009$, and
- Zone 3: $x \ge 1.009$.

In some zones, Equations 16, 17, and 18 shall be applied in a modified form.

Zone 1: $0 < x < 0.9$

In this zone there are two possibilities, namely, $B < Q_o$ and $B > Q_o$, where B is this queue length at the beginning of the period under consideration, and Q_o is the equilibrium overflow, that is, the stabilizing value of the queue length after some time.

Cronjé found that the following Newell equations are the most accurate for estimating delay and stops for undersaturated conditions *(4)*:

$$d = c(1 - \lambda)^2/2(1 - \lambda \cdot x) + Q_o/q$$

$$D = d \cdot q \cdot c \quad (19)$$

$$N = q[(q \cdot r + Q_o)/(s - q) + r] + Q_o \quad (20)$$

where

d = average delay (sec per vehicle);
D = total delay per cycle (vehicle seconds);
N = number of stops per cycle;
Q_o = $I \cdot H(\mu)x/2\ (1 - x)$;
c = cycle length (sec);
g = effective green time (sec);
λ = g/c = proportion of the cycle that is effectively green;
I = ratio of the variance to the average of the arrivals per cycle;
x = degree of saturation, the ratio of the average number of arrivals per cycle to the maximum number of departures per cycle;
q = average arrival rate (vehicles per second);
s = saturation flow (vehicles per second);
$H(\mu)$ = $e^{-(\mu+\mu^2/2)}$;
μ = $(1 - x)(s \cdot g)^{1/2}$; and
r = effective red time (sec).

The method of obtaining delay and stops in Zone 1 shall now be given.

For $B < Q_o$: When there is an increase in flow, delay does not increase instantaneously to the new value but fluctuates until it stabilizes at the end of the nonstationary zone, as indicated in Figure 3.

Determine delay and stops from Equations 19 and 20. Now also determine delay and stops by using Equations 16–18 for the initial number of cycles in the nonstationary zone. Check the delay obtained for each cycle against 0.9 (delay obtained by using Equation 19). As soon as the calculated delay exceeds this value, the calculation of delay and stops ceases and the sums of delay and stops for the period under consideration are as follows:

Σ Delay = Σ delay over the first number of cycles + $M \cdot$ (delay obtained by using Equation 19), and

Σ Stops = Stops over first number of cycles + $M \cdot$ (stops obtained by using Equation 20).

where

M = $N - Y$,
N = number of cycles in this period under consideration, and
Y = number of cycles until calculated delay: ≥ 0.9 (delay obtained by using Equation 19).

For $B > Q_o$: This case is treated exactly as when $B < Q_o$ except that after the delay for each cycle has been obtained, it is checked against the following value:

$1.1 \cdot$ (delay obtained by using Equation 19)

and Y is the number of cycles until the calculated delay is $\leq 1.1 \cdot$ (delay obtained by Equation 19).

Zone 2: $0.9 \leq x < 1.0009$

Also in this case, delay and stops are determined from Equation 16–18 for the period under consideration.

Zone 3: $x \geq 1.0009$

In this zone, Equation 16 is modified to the following:

$$E(K) = E(B) + E(Z) - E(V) + H/E(B) \quad (21)$$

where

$$H = E(B)_1 \sum_v p_V(v) \sum_{z=o}^{v-1} p_Z^{(z)}[E(B)_1 (1 - F^{v-z}) + z - v]$$

$E(B)_1$ = expected queue length at the beginning of the first cycle.

In this zone, stops and delay are evaluated from Equations 17 and 18 but the summation terms are omitted.

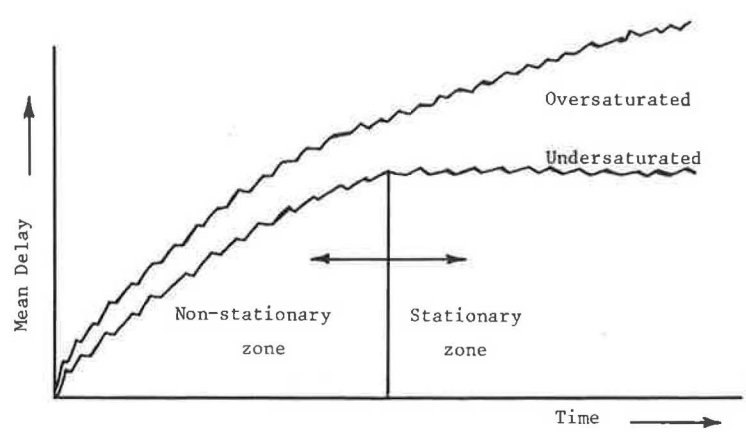

FIGURE 3 Transition from nonstationary to stationary zone as flow increases.

TABLE 2 DATA FOR GENERATING ARRIVALS

Input				Output			
c	x	q	I	c	x	q	g
50	1.05	500	0.5	50	1.05	542.4	14.4
			0.6			547.2	14.5
			0.7			556.8	14.7
			0.8			499.2	13.2
			0.9			552.0	14.6
			1.1			499.2	13.2
			1.2			537.6	14.2
			1.3			432.0	11.4
			1.4			480.0	12.7
			1.5			465.6	12.3
50	1.10	650	0.5	50	1.10	633.6	16.0
			(and so forth)				

Note: c = cycle length; x = degree of saturation; q = average arrival rate; I = ratio of variance to mean of arrivals per cycle; and g = effective green time (sec).

TABLE 3 FITTING DATA WITH NORMAL DISTRIBUTION FOR M GEOM

Group	Observed Frequency (O)	Theoretical Frequency (T)	$(O-T)^2/T$
< − 40	6 } 12	1.92 } 10.30	0.28
−40 to −30	6	8.38	
−30 to −20	23	27.14	0.63
−20 to −10	54	57.95	0.27
−10 to 0	81	78.62	0.07
0 to 10	80	76.32	0.18
10 to 20	51	45.70	0.61
20 to 30	17	17.95	0.05
30+	2	6.02	2.68
Σ =	320	320	4.77

Note: $\bar{x} = -1.906$ and $\bar{s} = 15.173$.

DATA GENERATED

The data generated are based on the following values for c, I, x, and q:

c (sec) = 50, 60, 70, 80, 90, 100, 110, 120
I = 0.5, 0.6, 0.7, 0.8, 0.9, 1.1, 1.2, 1.3, 1.4, 1.5

The degree of saturation and the flow are assigned the following paired values:

x	q (veh/hr)
1.05	500
1.10	650
1.15	800
1.20	900

Arrivals are generated over 15 cycles. The arrivals are converted to flow. The cycle length and degree of saturation are kept the same and the corresponding effective green time is calculated. To limit the length of the paper, only input and output data obtained for c = 50 sec are given in Table 2. A total of 320 sets of data are obtained.

The equations used in the M Geom, Mayne, and Catling models are applied to the data generated. The cost difference is then checked against the normal distribution.

M Geom

The histogram of a sample from arrivals generated for the M Geom Model is shown in Figure 4 and the processed data are given in Table 3.

The number of groups = n = 8. The number of degrees of freedom lost = 3. Therefore, the number of degrees of freedom $f = n - 3 = 8 - 3 = 5$.

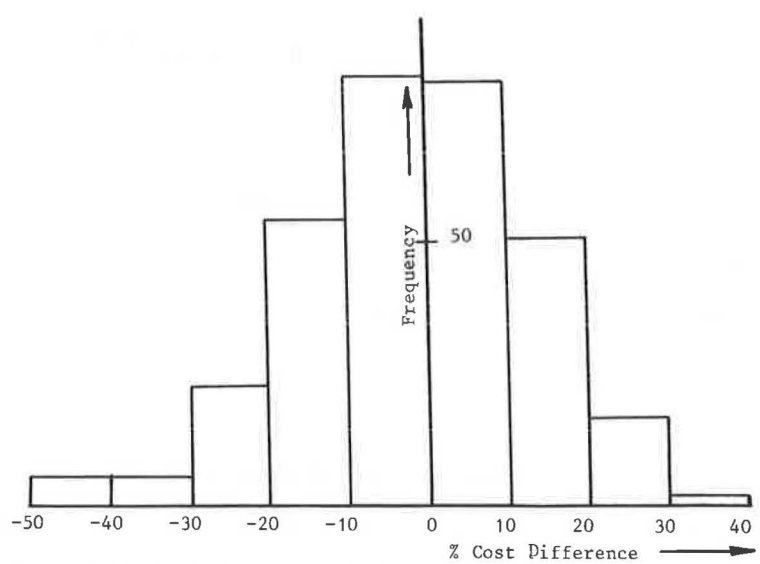

FIGURE 4 Histogram of sample from arrivals generated for M Geom Model.

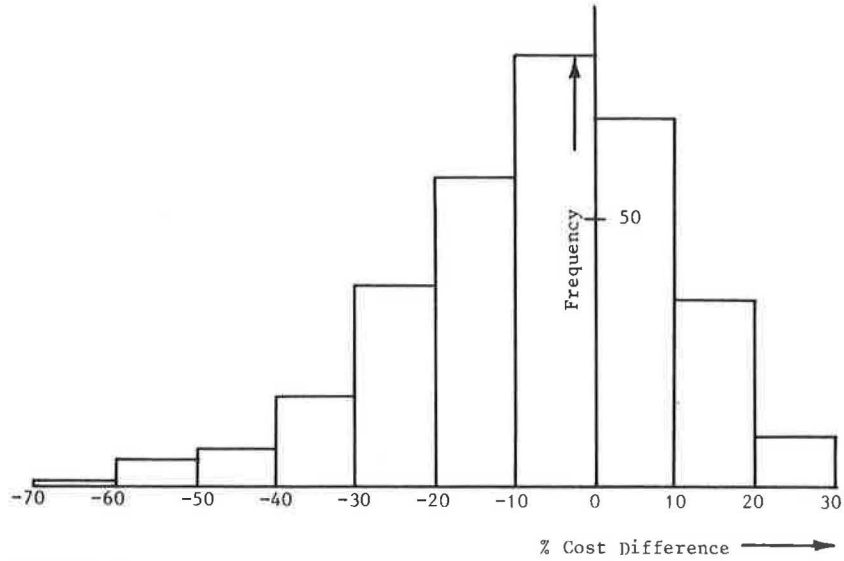

FIGURE 5 Histogram of sample from arrivals generated for Mayne Model.

At the 5 percent level $\chi^2_{5;0.05} = 11.07$

$\Sigma(0 - T)^2/T = 4.77 < 11.07$

Therefore, there is insufficient evidence to reject the hypothesis that the percentage cost difference follows the normal distribution. The 95 percent confidence interval is given by

$x \pm 1.96 \cdot \bar{s}$

Therefore, the confidence interval is

$[-1.906 \pm 1.96 \cdot 15.173]$

$= [-31.64\%; 27.83\%]$

The following monetary rates are used: $3.1 \cdot 10^{-2}$ rand per stop; $1.74 \cdot 10^{-4}$ rand per vehicle second for total delay. The rand is the unit of currency in the Republic of South Africa [1 rand = \$0.62 (1984)]. These rates are from research to be published by the National Institute for Road and Transport Research, Republic of South Africa.

Mayne Model

The histogram of a sample from arrivals generated for the Mayne Model is shown in Figure 5 and the processed data are given in Table 4.

The number of groups $n = 8$. The number of degrees of freedom lost = 3. Therefore, the number of degrees of freedom = 8 − 3 = 5.

At the 5 percent level $\chi^2_{5;0.05} = 11.07$.

$\Sigma(0 - T)^2/T = 12.07 > 11.07$

Therefore, there is insufficient evidence to accept the hypothesis that the percentage cost difference follows the normal distribution.

TABLE 4 FITTING DATA WITH NORMAL DISTRIBUTION FOR MAYNE MODEL

Group	Observed Frequency (O)		Theoretical Frequency (T)		$(O - T)^2/T$
< −60	1		0.29		
−60 to −50	5	13	1.60	8.58	2.28
−50 to −40	7		6.69		
−40 to −30	17		20.26		0.52
−30 to −20	38		44.64		0.99
−20 to −10	58		67.46		1.33
−10 to 0	81		73.47		0.77
0 to 10	69		57.12		2.47
10 to 20	35		31.97		0.29
20+	9		16.51		3.42
Σ =	320		320.01		12.07

Note: $\bar{x} = -7.438$ and $\bar{s} = 16.862$.

Because the Mayne Model does not follow the normal distribution at the 5 percent level, the confidence interval shall be obtained from the histogram in Figure 5 by locating the value of the variant at each tail for 2.5 percent of the area. The number of observations given by 2.5 percent is given by

$2.5 \cdot 320/100 = 8$

The 95 percent confidence interval is given by

$[-47.5\%; 20.0\%]$

Catling Model

The histogram of a sample from arrivals generated for the Catling Model is shown in Figure 6, and the processed data are given in Table 5.

The number of groups = $n = 8$. The number of degrees of

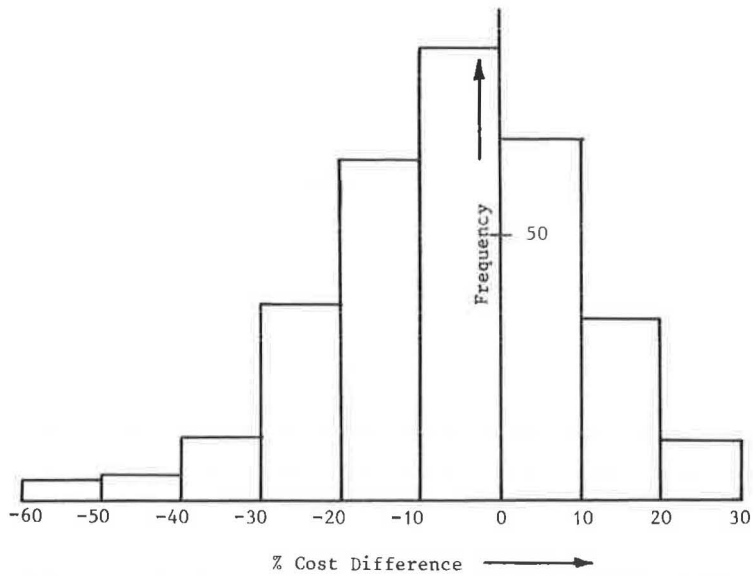

FIGURE 6 Histogram of sample from arrivals generated for Catling Model.

freedom lost = 3. Therefore, the number of degrees of freedom = 8 − 3 = 5.

At the 5 percent level $\chi^2_{5;0.05} = 11.07$.

$$\Sigma(0 - T)^2/T = 6.59 < 11.07$$

Therefore, there is insufficient evidence to reject the hypothesis that the percentage cost difference follows the normal distribution.

The 95 percent confidence interval is given by

$$\bar{x} \pm 1.96 \cdot \bar{s}$$

Therefore, the confidence interval is

$$[-6.406 \pm 1.96 \cdot 15.803]$$

$$= -37.4\%;\ 24.6\%$$

TABLE 5 FITTING DATA WITH NORMAL DISTRIBUTION FOR CATLING MODEL

Group	Observed Frequency (O)	Theoretical Frequency (T)	$(O - T)^2/T$
< − 50	4 } 9	0.93 } 5.31	
−50 to −40	5	4.38	2.56
−40 to −30	12	16.48	1.22
−30 to −20	37	40.58	0.32
−20 to −10	64	68.51	0.30
−10 to 0	85	80.03	0.31
0 to 10	68	61.34	0.65
10 to 20	34	32.54	0.07
20+	11	15.20	1.16
Σ =	320	319.99	6.59

Note: $\bar{x} = -6.406$ and $\bar{s} = 15.803$.

The confidence intervals for the three models have been developed for the 320 sets of generated data. However, it is possible that on the average one model may appear preferable but that for specific ranges of the data one of the other models may estimate cost more accurately. To investigate this possibility, the variation in percentage cost difference is examined against the following:

1. Variation in cycle length (c),
2. Variation in paired sets of flow (Q) and degree of saturation (x), and
3. Variation in the variance to mean ratio of the arrivals per cycle (I).

The values obtained are given in Tables 6–8 and plotted in Figures 7–9. Because the negative values for percentage cost difference actually indicate an overestimation of cost, these values shall be plotted on the positive axis in Figures 7–9.

First, the variation of percentage cost difference against cycle length variation is examined for the M Geom, Mayne, and Catling models. The values obtained are given in Table 6.

TABLE 6 PERCENTAGE COST DIFFERENCE AGAINST CYCLE LENGTH VARIATION FOR THE THREE MODELS

Cycle Length (c)	M Geom	Mayne	Catling
50	−0.72	−6.27	−5.02
60	−6.93	−12.96	−11.78
70	−6.42	−12.57	−11.47
80	0.82	−4.93	−4.02
90	−1.17	−7.15	−6.19
100	1.30	−4.47	−3.67
110	1.19	−4.54	−3.78
120	−2.04	−8.12	−7.23

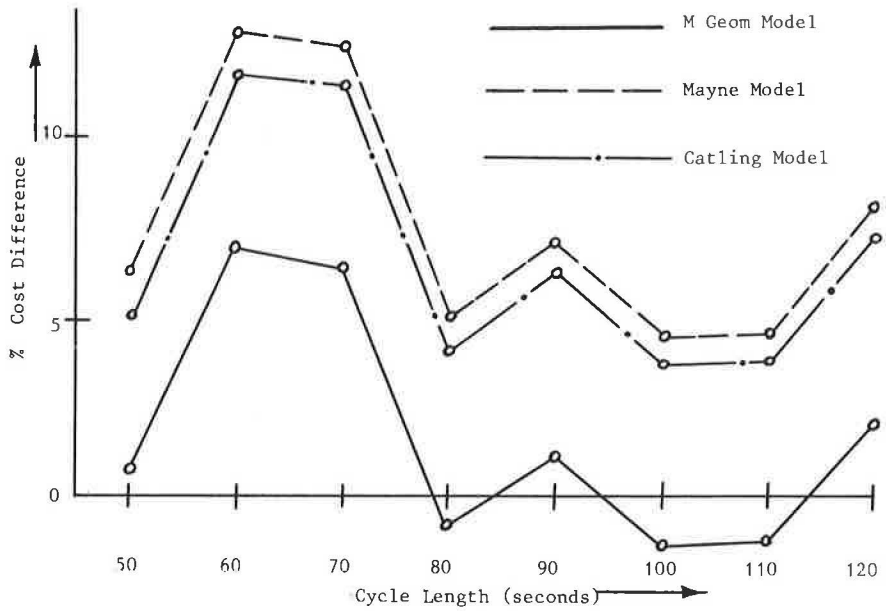

FIGURE 7 Percentage cost difference against cycle length variation for the three models.

TABLE 7 PERCENTAGE COST DIFFERENCE AGAINST FLOW AND DEGREE OF SATURATION FOR THE THREE MODELS

Flow (Q)	Degree of Saturation (x)	M Geom	Mayne	Catling
500	1.05	−6.64	−16.94	−12.63
650	1.10	1.34	−4.42	−2.89
800	1.15	−0.35	−4.46	−4.65
950	1.20	−1.34	−4.68	−6.41

The percentage cost difference given in Table 6 for the three models is plotted in Figure 7.

Second, the variation of percentage cost difference against paired values of flow and degree of saturation is examined for the M Geom, Mayne, and Catling models. The values obtained are given in Table 7. The percentage cost difference indicated in Table 7 for the three models is plotted in Figure 8.

Last, the variation of percentage cost difference against the ratio of the variance to the mean of the arrivals per cycle is examined for the M Geom, Mayne, and Catling models. The values obtained are given in Table 8. The percentage cost difference given in Table 8 for the three models is plotted in Figure 9.

DISCUSSION OF M GEOM, MAYNE, AND CATLING MODELS

The 95 percent confidence intervals and means for the three models for the data generated are as follows:

- M Geom Model: [−31.64; 27.83]; $\bar{x} = -1.906$.
- Mayne Model: [−47.5; 20.0]; $\bar{x} = -7.438$.
- Catling Model: [−37.4; 24.6]; $\bar{x} = -6.406$.

Although the range in the confidence interval is approximately the same, the M Geom Model estimates cost on the average considerably closer than do the Mayne and Catling models. From an inspection of Figures 7–9 it is also clear that throughout the range of varying cycle length, flow and degree of saturation, and variance to mean ratio of arrivals per cycle, the M Geom Model estimates cost more accurately. The only exception is when $I > 1.45$, which is negligible. Computing time for the three models for the generation of the arrivals and for the computation of cost is as follows:

M Geom Model: 1 min 54.5 sec
Mayne Model: 1 min 52.8 sec
Catling Model: 1 min 53.8 sec

It is clear that in computing time for the three models, there is so little difference that there is no reason to prefer one model in particular. It is therefore concluded that the M Geom Model

TABLE 8 PERCENTAGE COST DIFFERENCE AGAINST I FOR THE THREE MODELS

I	M Geom	Mayne	Catling
0.5	−8.10	−14.46	−13.32
0.6	0.75	−5.02	−4.05
0.7	−6.27	−12.48	−11.36
0.8	−3.40	−9.32	−8.32
0.9	−4.82	−10.88	−9.87
1.1	−0.62	−6.44	−5.47
1.2	−0.20	−5.88	−5.05
1.3	0.41	−5.34	−4.38
1.4	0.52	−5.27	−4.29
1.5	4.28	−1.21	−0.32

Note: I = ratio of variance to mean of arrivals per cycle.

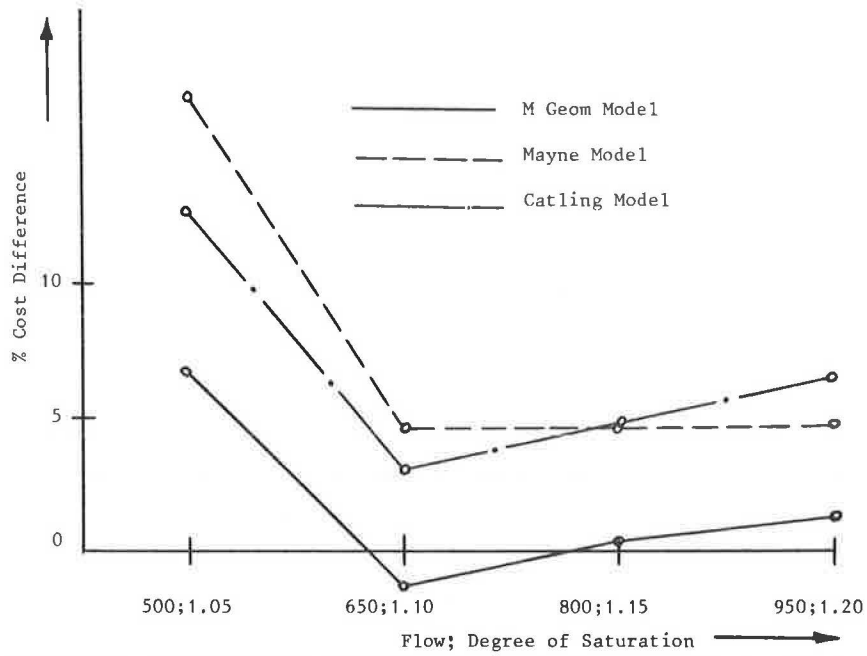

FIGURE 8 Percentage cost difference against flow and degree of saturation for the three models.

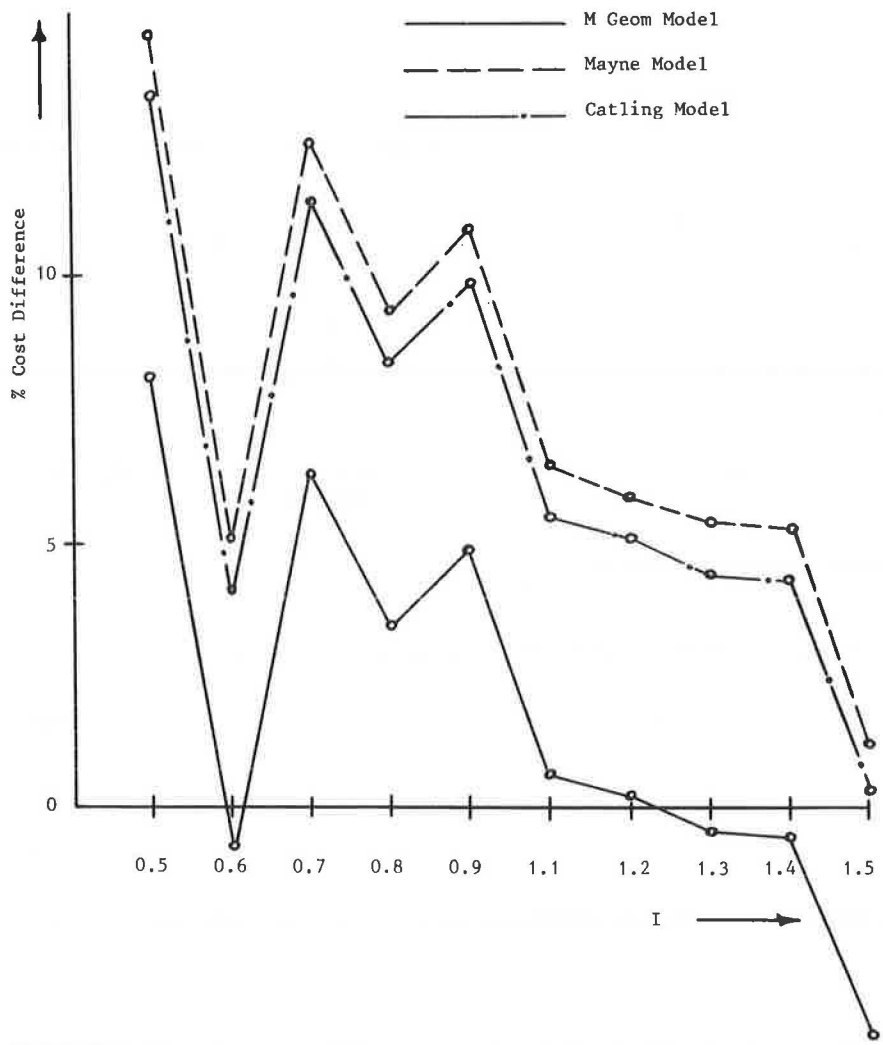

FIGURE 9 Percentage cost difference against variation in *I* for the three models.

is the preferred model for estimating cost in the optimization procedure for fixed-time traffic signalized intersections for undersaturated and saturated conditions.

REFERENCES

1. A. J. Mayne. "Traffic Signal Settings for Congested Traffic." *Proc., Engineering Foundation Conference on Research Directions in Computer Control of Urban Traffic Systems*, Calif., Feb. 1979.
2. I. Catling. "A Time-Dependent Approach to Junction Delays." *Traffic Engineering and Control*, Nov. 1977.
3. W. B. Cronjé. "Optimization Model for Isolated Signalized Traffic Intersections." In *Transportation Research Record 905*, TRB, National Research Council, Washington, D.C., 1983, pp. 80–83.
4. W. B. Cronjé. "Analysis of Existing Formulas for Delay, Overflow, and Stops." In *Transportation Research Record 905*, TRB, National Research Council, Washington, D.C., 1983, pp. 89–93.

Publication of this paper sponsored by Committee on Traffic Flow Theory and Characteristics.

Peaking Characteristics of Rural Road Traffic

PIETER W. JORDAAN AND CHRISTO J. BESTER

A methodology for estimating the traffic volume on rural roads for any of the highest thousand hours of the year is presented. The methodology requires estimates of the annual average daily traffic and a peaking characteristic. Also presented is the derivation of the latter from the data obtained from permanent traffic counters. It is found that the peaking characteristic is related to the fraction obtained by dividing the thirtieth highest hourly volume by the annual average daily traffic as well as by the average length of through trips on the road link to which it is applicable. It is hoped that this methodology will create a sound base for further research into traffic volume variations and design hour volume.

Peaking of rural traffic is a well-known phenomenon. How to take such peaking into account is less well established. Two approaches are in use in highway engineering practice, namely, the design hour approach and the flow regime approach.

The design hour approach was established in the United States by Peabody and Norman in 1941 *(1)*. They found that if the hourly traffic counts from a year's records were sorted from high to low and plotted against rank number, there was a "knee" at about the thirtieth highest hour. Since 1941 this concept of using some hourly volume to represent the peaking characteristic of the traffic on a road has received much support. It is embodied in many design textbooks *(2–4)*, usually of U.S origin.

In South Africa this concept is also well-known and applied by most road authorities *(5)*. Some local research has been done into the relationship between K [the fraction obtained by dividing the thirtieth highest hourly volume by the annual average daily traffic (AADT)] and the AADT *(6)*. To date, no real causal parameter for estimating K has been found. K is generally believed to vary as follows:

- Roads with a high proportion of recreational traffic: $K > 0.25$,
- Average roads: $0.15 < K \leq 0.25$, and
- Roads with little peaking: $K \leq 0.15$.

It has been observed that as AADT increases, there is a tendency for K to decline *(6)*.

The flow regime approach was established by Dawson in the United Kingdom in the late 1960s *(7)*. In this approach, the 8,760 hr of the year are divided into groups with constant flow characteristics. Dawson's original flow regimes for rural roads are given in Table 1. Over the years, this approach has been refined by the Transport and Road Research Laboratory in the United Kingdom *(8)*. It is also used in the COBA procedures for project evaluation *(9)*.

In South Africa, this approach is also used in the RODES2 program developed by Bester *(10)*, in which the flow regimes are differentiated by K, the fraction of the thirtieth highest hourly volume.

From the foregoing it is clear that a knowledge of hourly vehicular flow is important to the roads engineer. Not only is a knowledge of the flow during the thirtieth highest hour required, but the use of sophisticated project evaluation pro-

P. W. Jordaan, University of Pretoria, Lynwood Rd., Pretoria, Republic of South Africa, 0002. C. J. Bester, National Institute for Transport and Road Research, CSIR, Pretoria, Republic of South Africa.