Webster's Delay Formula – revisited.

August 1, 2013

Peter Wagner*
German Aerospace Center (DLR)
Institute of Transportation Systems
Rutherfordstrasse 2, 12489 Berlin, Germany
phone: +49 30 67055 - 237, fax: +49 30 67055 - 291
email: peter.wagner@dlr.de, web: http://www.dlr.de/ts/en/

Nathan H. Gartner
Civil & Environmental Engineering Dept
University of Massachusetts Lowell
Lowell, MA 01854, U.S.A.
phone: +1 978 934 - 2289, fax: +1 978 934 - 3052
email: Nathan_Gartner@uml.edu

Ting Lu, Robert Oertel
German Aerospace Center (DLR)
Institute of Transportation Systems
Rutherfordstrasse 2, 12489 Berlin, Germany
phone: +49 30 67055 - 213, - 650 , fax: +49 30 67055 - 291
email: ting.lu@dlr.de, web: http://www.dlr.de/ts/en/

For Presentation and Publication
93rd Annual Meeting
Transportation Research Board
January 12 – 16, 2014
Washington, D. C.

Submission date: August 1, 2013

words:	3 933
plus 9 figures:	2 250
plus 0 tables:	0
total count:	6 183
word limit:	7 500

*corresponding author

ABSTRACT

The equations developed by Webster in his famous 1958 report[10] are still the basis of traffic signal planning today. They are being used in handbooks like the HCM and similar instruments world-wide. However, the handbook approach typically works with approximations to the original equations which have stood the test of time, but may nevertheless not be the best to be done today. This work analyzes Webster's approach and advocates a more modern use of it which utilizes the tremendous advances in computer hardware and software. This is being done by comparing approximations to exact solutions, and by a comparison between various models and Webster's equations itself. It is shown that there can be significant differences in the calculation of optimal cycle times and consequent delay times.

1 Introduction

In 1958, Webster published his famous report [10], in which the working of a fixed cycle traffic signal is analyzed in depth. Especially the formula for the optimum cycle time of a n-phase intersection is still used in every day's work and is put into handbooks like the Highway Capacity Manual [1] and similar works. A large amount of research has been put into the comparison of Webster's equations with micro-simulation tools and to comparing them to real data (see e.g. chapter 18 of [1]), and into the theoretical description of what happens at a signal controlled intersection [3, 2, 6, 9, 7]. Nevertheless, there are still open questions in this field, a few of them will be highlighted in the current study. In this paper we compare Webster's theory with results obtained from a micro-simulation model. The deterministic model by Webster fits fairly well with a host of different deterministic modeling approaches. For the stochastic part, however, differences between theory and simulation reality have been found. It is shown that there can be significant differences in the calculation of the optimal cycle times and the resulting delay times. The latter differences can be in the range of 4% to 40%.

2 A detailed account on Webster's equation

As given in his report [10], the total delay per vehicle of an n phase intersection controled by a fixed cycle controler reads:

$$D(c) = \frac{1}{2Q}\sum_{r=1}^{n}\left(\frac{c y_r s_r (1-\lambda_r)^2}{1-y_r} + \frac{y_r^2}{\lambda_r(\lambda_r - y_r)}\right) \tag{1}$$

where the demands for each phase q_r, the corresponding saturation flows s_r, the green times g_r, the cycle time c and the short-hand notations $y_r = q_r/s_r$, $\lambda_r = g_r/c$, $Y = \sum_{r=1}^{n} y_r$, and $Q = \sum_{r=1}^{n} q_r$ have been used. Slightly different from the original formula, the normalization has been made correct by dividing the sum with the total demand Q, to get the metric (delay/vehicle) correct. This eases the comparison with simulation results. The first term in the sum accounts for the deterministic delay, while the second term is due to fluctuations in the vehicles' arrivals and therefore named the stochastic term.

Essentially, the first term is just a simple-looking function of the cycle time, $D(c) = k_1 c + k_2 + k_3/c$. However, the coefficients k_i turn out to be very complicated functions of the various parameters. The second term of equation (1) adds $k_4/(c-L) + k_5/(c-L/(1-Y))$, which will be shown in a moment. Furthermore, strictly speaking $D(c)$ is only then a valid equation if $c > L/(1-Y)$ holds, however there are some funny exemptions from this rule. Definitely, for $c \leq L$ no green time is left and so the intersection does not work anymore.

Of course, other delay formulas can be used in the same manner as will be discussed here, in section 2.3 this is done in a rudimentary fashion for the HCM 2010 function.

The assumptions that have been used in the derivation of this formula are that the incoming traffic is generated according to a Poisson process, and that the outgoing traffic follows a deterministic process: the time-interval between two vehicles leaving the signal is constant as long as there are queued vehicles in front of the signal. Also, there is no detailed vehicle dynamics involved: vehicles arrive and stop immediately at the end of the queue, and they leave and accelerate in zero time. All the missing delay times like reaction and acceleration times are subsumed into the total loss time L, together with the times needed to switch between the phases. In a certain sense, L is therefore not exactly measurable but can be understood as a parameter.

To make progress, the variables λ_r which are the ratio of the green times to the cycle time, have been specified by assuming that the green times are proportional to the saturation values y_r, more specifically to set $\lambda_r = \left(1-\frac{L}{c}\right)\frac{y_r}{Y}$. Note, that by inserting this into equation (1), the second term in the sum simplifies considerably and can even be eliminated from the summation, because it reads:

$$\frac{y_r^2}{2\lambda_r(\lambda_r - y_r)} = \frac{Y^2}{2\left(1-\frac{L}{c}\right)\left(1-\frac{L}{c}-Y\right)} = \frac{Y}{1-\frac{L}{c}-Y} - \frac{Y}{1-\frac{L}{c}}$$

From here, it can be readily seen that this stochastic term, and therefore D, has two poles at $c = L$ and at $c = L/(1-Y)$. This simplification is a bit surprising and will be picked up later on. Note, that this does not happen if the λ_r are not assumed to be proportional to the saturation degrees y_r. Right now, it will be ignored and equation (1) will be used. In [10], equation (1) is then differentiated with respect to the cycle time c and the resulting expression is set to zero, in order to find the optimal cycle time c_{opt}. However, the resulting expression is so complicated as to be useless, or even not solvable at all, therefore [10] used a long and complicated chain of reasoning after which the famous formula results for the optimal cycle time:

$$c_{hcm} = \frac{1.5L+5}{1-Y} \tag{2}$$

It is clear, that this is an approximation, therefore one might ask the question how good is it. This will be investigated in the following.

2.1 Quality of the handbook formula

To assess the quality of equation (2) note that it is very simple to compute the optimum cycle time numerically. This can be done either by finding the minimum in equation (1) or by computing the root of $dD(c)/dc = 0$, the latter approach being numerically more efficient. In fact, this can be put into a spreadsheet, and compute a table of values c versus $D(c)$ and find the minimum in this table which is a good approximation to the true minimum.

To test the quality of the handbook equation, a large number of scenarios have been generated. A scenario is defined as a set of numbers $\{q_r\}_{r=1,\ldots,n}$, $\{s_r\}_{r=1,\ldots,n}$, and L. In the following, this has been done for $n = 2,3,4$, $s_r = s = 1800$ veh/h, while the q_r have been drawn randomly from the interval [36, 1800] veh/h and with the additional condition that the resulting Y-values should be smaller than 0.9. Several 10,000 scenarios can be analysed with ease and give already a fairly complete picture. The most direct comparison is the frequency distribution of the difference between c_{opt} and c_{hcm}which can be seen in Figure 1 for the case $n = 2$. The mean value of this difference is around 14 s, i.e. the true optimal cycle times are on average 14 s larger than c_{hcm} which corresponds to a relative difference of 11%. Consequently, the delay times resulting from the handbook formula are always larger, however the difference is not very big, on average 0.8 s or 4%. The maximum difference ("worst case") turns out to be 14 s or 41%.

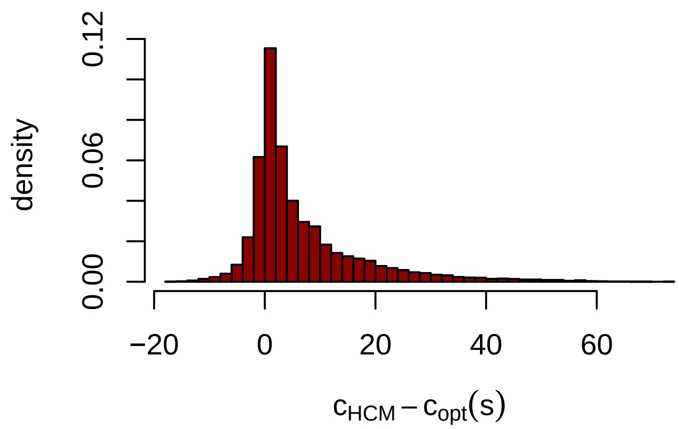

Figure 1 Frequency distribution of the difference between the true minimal cycle time and the handbook minimal cycle time.

To generalize these results for $n > 2$, a different analysis have been performed. The following has a certain similarity to [4]. For each scenario, the optimum cycle time c_{opt}, the loss time L, and the total saturation Y is known. Therefore, by plotting $(1 - Y)c_{opt}$ as function of L, the difference between the true function and the approximation that led to equation (2) can be demonstrated. This is shown in Figure 2.

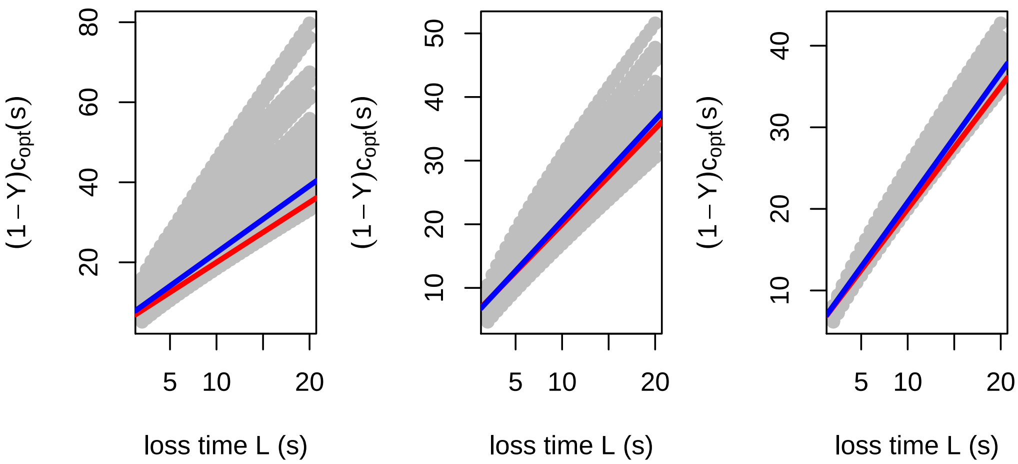

Figure 2 **For $n = 2,3,4$ phases (from left to right) plotted is $(1-Y)c_{opt}$ as function of L for all the scenarios investigated, together with the corresponding handbook linear function (in red) and a robust fit (in blue) to the data-points. In general, the data follow the handbook equation quite well, the largest deviations are found for large loss times and for a small number of phases.**

Obviously, the handbook formula does a fairly good job. E.g., for the worst case $n = 2$, the fit to the data-points yields $c_{opt} = (1.68L + 5.9)/(1-Y)$, for $n = 4$ it yields $c_{opt} = (1.59L + 4.94)/(1-Y)$. The only bad thing is that there are a certain number of cases, where this simple linear approach gives completely wrong results, as the gray shaded area in Figure 2 demonstrate.

2.2 A more general approach

So far, the approach has followed [10] completely. However, albeit perfectly sensible, it may not be true in any case that $g_r \propto y_r$, especially if there is a large difference between the different y_r-values. Therefore, equation (1) can be rewritten as a function of n green times g_r:

$$D(g_1,\ldots,g_n) = \frac{1}{2Q}\sum_{r=1}^{n}\left(\frac{y_r s_r (S-g_r)^2}{(1-y_r)S} + \frac{y_r^2 S^2}{(g_r - S)(g_r - (1-y_r)S)}\right) \tag{3}$$

Here, the term $S = L + \sum_i g_i$ has been introduced to shorten the equation and to make it easier to read. Now, the symmetry in the second (stochastic) term of the Webster equation is broken, so different results than before can be expected. By pursuing the same approach as in the case of the $D(c)$ equation, again a large number of different scenarios can be analyzed, e.g. to test the assumption $g_r \propto y_r$. This times, the numerical approach is more involved, since now for each of the scenarios the minimum of an n-dimensional function must be found, with the risk that the optimization does not find any result at all. So, in roughly one third of the cases, the optimization does not find a sensible solution. This has not been further explored, since only the validity of the assumption $g_r \propto y_r$ was of interest.

From Figure 3 it turns out, that this assumption is almost always true, so there is no need to look deeper into the remaining differences.

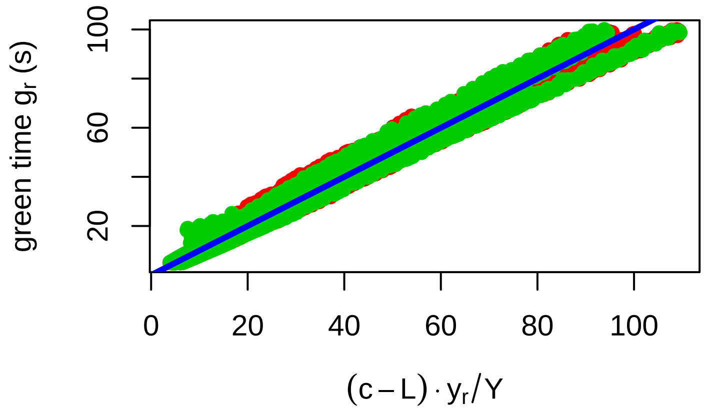

Figure 3 The results of the different scenarios for $n = 2$. Shown is the plot of g_r versus $(c - L)y_r/Y$ for the two green times, and a line with slope 1. The data-points follow Webster's assumption $g_r = (c - L)y_r/Y$ well, the slope of the two curves relating g_r to $(c - L)y_r/Y$ is 0.998 and 0.988, respectively.

2.3 Comparison to the HCM equation

The HCM uses a different delay formula that can handle over-saturation. The version that will be used here is given as:

$$d = \frac{c}{2}\frac{\left(1 - \frac{g}{c}\right)^2}{1 - \min\left\{1, \frac{q}{s}\right\}\frac{g}{c}} + 900T\left(\left(\frac{q}{s} - 1\right) + \sqrt{\left(\frac{q}{s} - 1\right)^2 + 4\frac{q/s}{ST}}\right) \quad (4)$$

where, in addition to the variables already introduced, we have the flow period T measured in hours, and the lane group capacity S, measured in vehicles/hour. The original version contains factors k and l which have been set to 0.5 and 1, respectively, as needed for an isolated intersection. To find the total delay from this delay equation for a single leg, the summation over all legs divided by the total demand has to be performed:

$$D = \frac{1}{Q}\sum_{r=1}^{n} q_r d_r$$

where the d_r is given by equation (4). In this case, nothing has been tried to solve this equation for the optimum cycle time analytically. Furthermore, for this equation the whole approach displayed so far does not seem to work at all: when computing $D(c)$ for a set of fixed demand and saturation values, the second term just adds a constant. This can be seen if one carefully inspects the second term in equation (4): there is no c visible in this part of the equation, not even something hidden in one of the constants. Therefore, the minimum is essentially that of Webster's deterministic approach, since the first term is close to the deterministic term in Webster's equation. Therefore, the application of the approach as in section 2.1 yields nothing new despite the fact that the delay equation itself is more general than the one of Webster.

3 Comparison to models

3.1 Deterministic approach

Consider the deterministic queueing model first. For a single leg, the delay formula can be computed in a very simplistic manner to yield:

$$d = \frac{c}{2} \frac{\left(1 - \frac{g}{c}\right)^2}{1 - \frac{q}{s}} \tag{5}$$

From this, the question may be asked, what is actually modeled by this equation? The derivation of equation (5) uses a continuous flow model. I.e., traffic arrives at a constant rate q, and it leaves the leg as long as the traffic light is green with the constant rate s. If the signal is red, the arriving traffic flow is stored in a single “cell”. It can be simulated with a finite time-step size Δt and describing each phase by the continuous “number” of jammed vehicles $\{n_r\}_{r=1,\ldots,n}$. In each time-step the variable n_r is increased by $q_r\Delta t$, while it may in addition being decreased by $s_r\Delta t$ during the green period of phase r. This model does not care for any spatial effects, and it does not care for any discrete vehicles. Therefore, it is very easy to simulate, and a simulation of a two-leg intersection reproduces the first term of equation (1) almost exactly, see Figure 4 for a comparison.

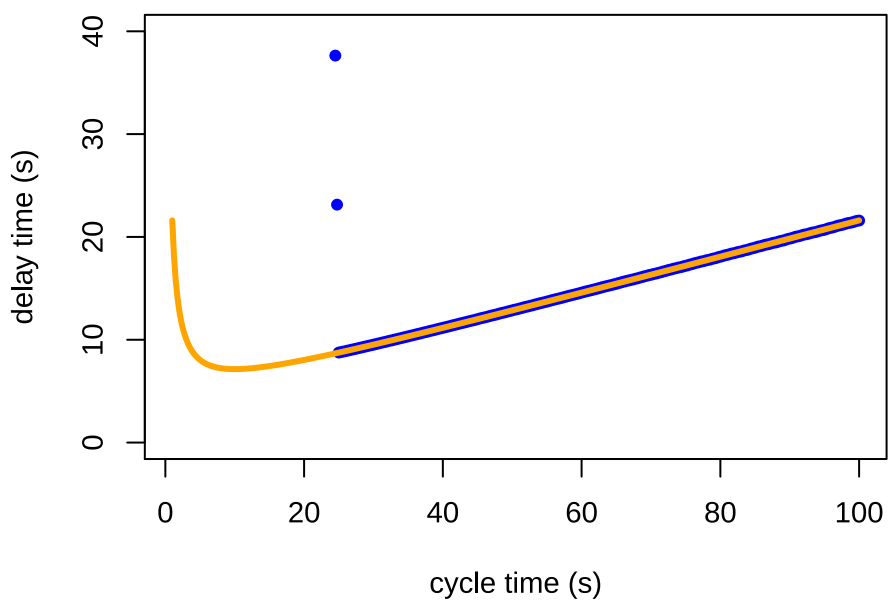

Figure 4 Comparison of the continuous model to the first term in equation (1). For cycle times larger than $c_{min} = L/(1-Y)$ simulation and theory are in exact agreement, differences are of the order of 10^{-4}s in the delay time. For the differences below c_{min} see the text. The parameters chosen are the same for all the remaining simulation results: $s_r = 1800$ veh/h, $q_r = 540$ veh/h, and $L = 10$ s.

Below c_{min}, the two results diverge. For $c \leq L$, no vehicle can obviously pass at all. This has not been stated explicitely in the derivation of equation (1), but is clear from construction. Between L and c_{min}, the simulated delay diverges with simulation time and proportional to the distance to c_{min}, since there not all vehicles can leave during green.

There are several possible steps to make such an intersection model more realistic. The first one is to leave the assumption that traffic arrives in continuous packets and assume discrete vehicles that arrive any $1/q$ seconds and leave the intersection any $1/s$ seconds (if the light is green). Obviously, the results now depend on the fact, whether or not the number of leaving vehicles times fits exactly within a green time or not. This can be seen in the following Figure 5. The queueing model behind is exactly what has been used by Webster itself and which is described in the appendix of his book – except for the fact, that the input flow in this example is deterministic. In Kendall's notation [5], such a queue is named a D/D/1 queue, where "D" is for deterministic inflow/outflow and the 1 stands for one server. Later on, also the exponential distribution (Kendall's notation is "M") and the Erlang distribution ("E") will be needed.

Again, there is a fairly good agreement between the theory and such a D/D/1 queueing model, see Figure 5.

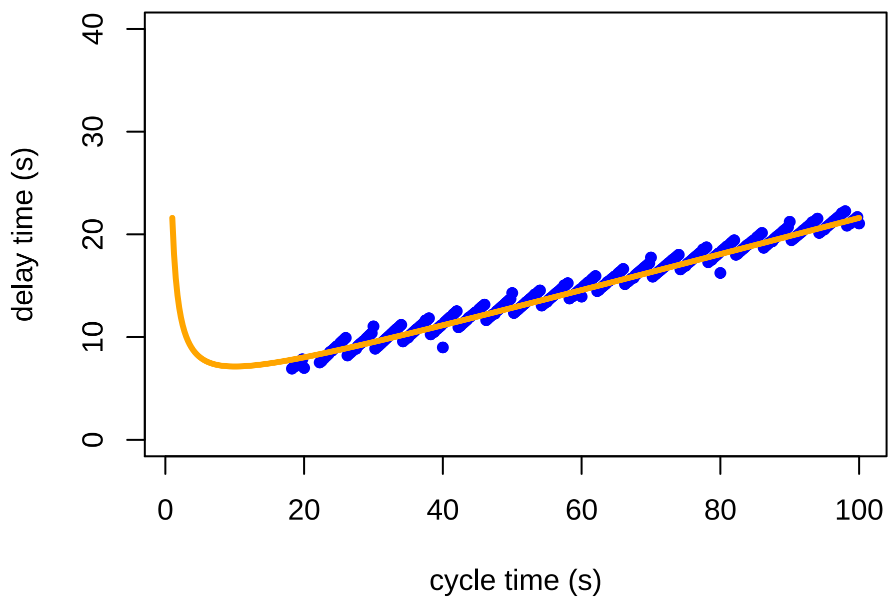

Figure 5 Comparison with a simple queue-model which is essentially a D/D/1 model in Kendall's notation.

In a final step, a micro-simulation has been used. Again, nothing complicated is needed here, since different micro-simulation models do not show dramatic differences due to the fact that the dynamics is strongly controlled by the traffic light. Therefore, a discrete variant of Newell's "lower-order model" [8] has been used. However, especially the saturation flow of such a model is usually difficult to control, so one may expect deviations from theory. Of course, the most important aspect is the spatial dimension that is now explicitly modeled. And a bit surprisingly, as long as we stick to the deterministic simulation, again the simulation fits fairly well to the theoretical result, as is shown in Figure 6.

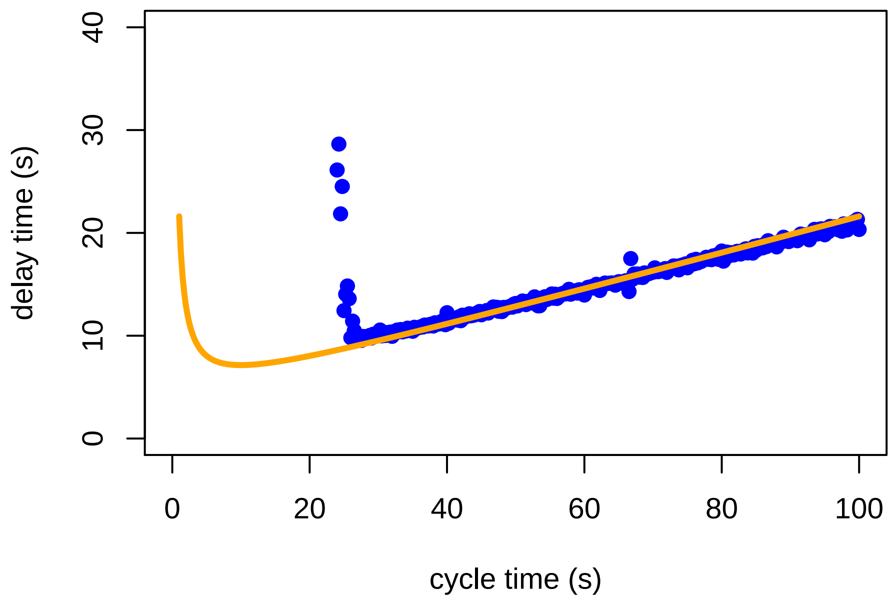

Figure 6 Comparison with a simple micro-simulation model.

3.2 Stochastic approach

It has been demonstrated so far, that there are small and understandable differences between Webster's deterministic theory and a corresponding modelling with dynamical models. In this section, the focus is on stochastic models. There is no clear idea how to make the continuous model stochastic, so the discussion focuses here on the queue-models and on the microscopic traffic flow model.

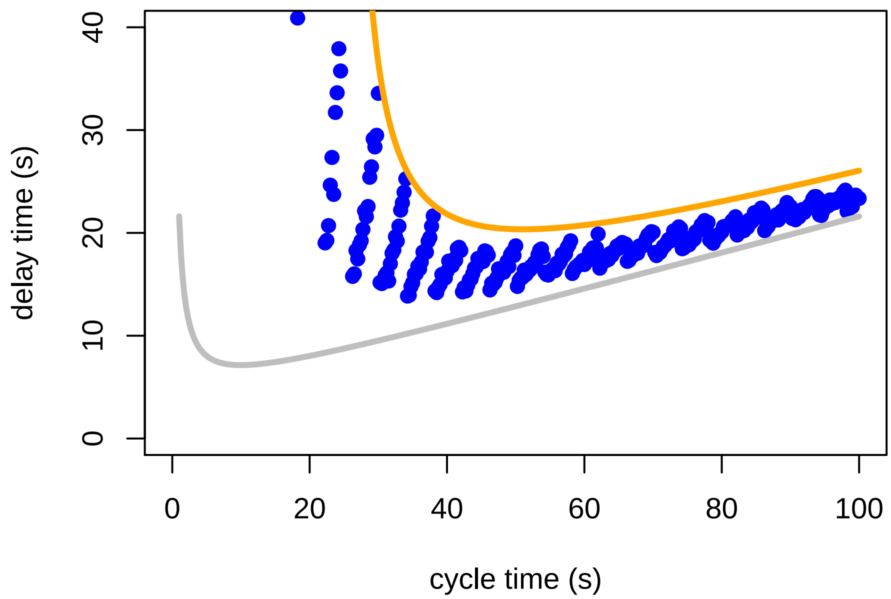

Figure 7 Comparison between an M/D/1 queueing model (blue dots) and Webster's theory. The gray line is the first term, while the orange line is the full expression in equation (1).

For the queueing model, it is quite simple to make it stochastic. Webster's stochastic theory assumes that the vehicles stream is modelled as a Poisson process. That is easy to simulate and yields the result shown in Figure 7. Due to the deterministic outflow, there is still a lot of structure to be seen that is similar to the results presented in the Figure 4. Note, that the data-points fall into the interval between Webster's deterministic and stochastic models. So, the stochastic model is not wrong, but it is not completely in line with the simulation results either. This is important since it determines where the optimum cycle time will finally end up.

Assuming deterministic outflow does not look like a realistic assumption. Therefore, this can be either changed into Poisson outflow, too, or, more realistically, into a stochastic process that leads to a realistic headway distribution such as an Erlang distribution. This is displayed in Figure 8. The blue points belong to a queueing model with a Poisson process both on the inflow and on the outflow, while the green points are from a simulation of a queueing model with Poisson inflow, but with an outflow drawn from an Erlang distribution. It seems that this combination reproduces Webster's model best, while the Poisson approach is overdoing it: the fluctuations are too large.

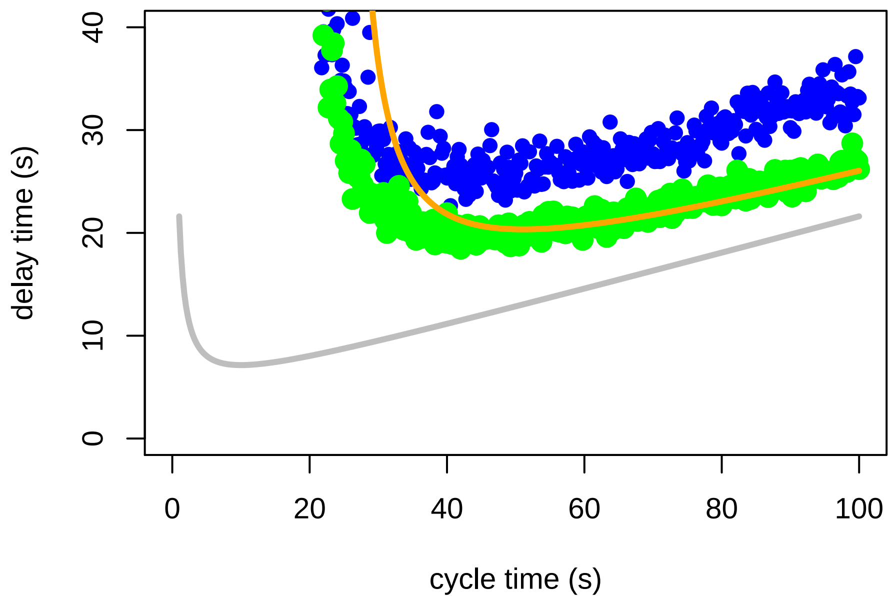

Figure 8 Comparison between an M/M/1 queueing model (blue points) and an M/E/1 queueing model (green points) to Webster's theory. The lines are as in Figure 7.

So far, all these queueing approaches have ignored spatial effects. This is taken into account by the final simulation. Here, again Newell's lower-order model has been used as in the previous section. To make this model stochastic, acceleration noise has been added to the car-following equations. The results (displayed in Figure 9) seem to interpolate between the stochastic part of Webster's model and, especially for larger cycle times, the deterministic part.

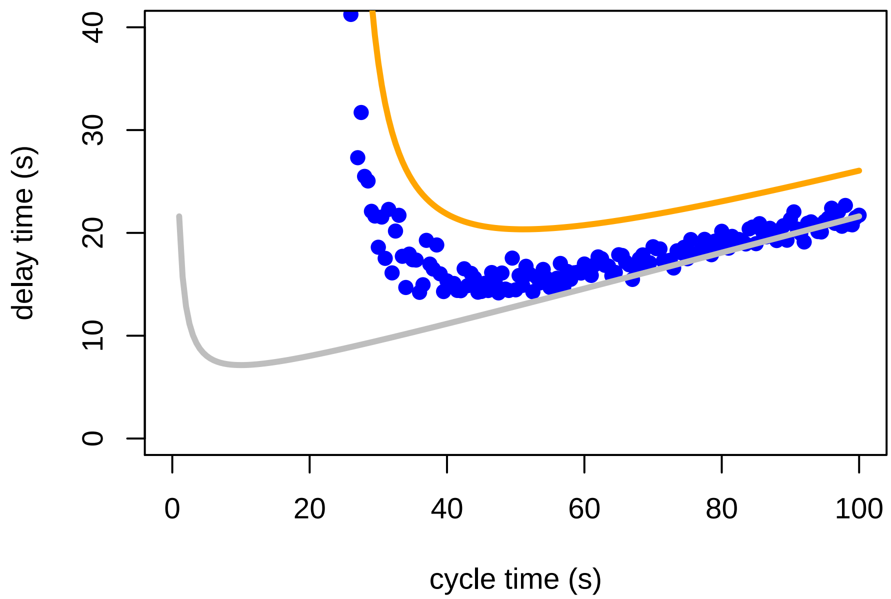

Figure 9 Comparison between a microscopic simulation model (blue points) and Webster's theory. The lines are as in Figure 7.

3.3 Consequences

Several interesting results can be drawn from these comparisons. First of all, if traffic would behave in a deterministic manner, Webster's deterministic theory would be a really good match to a wide range of different models. A bit surprisingly, there is little reason to worry about spatial effects. This will surely change if more than one intersection is considered, since then the propagation of congestion becomes important. There is a difference between the deterministic flow model that is at the heart of Webster's theory and a discrete analogue, such as a queueing model. This is mainly due to the fact that it is important whether integer multiples of the deterministic time headway between two vehicles just fit precisely into the green-interval or not.

For the stochastic part, the results are more mixed. Whereas Webster's theory is explicitly constructed around the assumption that the queueing process at an intersection is modelled as an M/D/1 process, the simulation results presented here display clear differences. The best agreement so far has been found if the "D" in M/D/1 is changed into a distribution that has a certain width, but is not as wide as the Poisson distribution. The microsimulation displays a more mixed result: for small cycle times, the stochasticity clearly shows up, while for larger cycle times the process gets more ordered which may be due to the fact that once a queue of vehicles gets moving, they will drive for a certain time quite ordered. It may be very interesting to compare this to data from real traffic, which however might be very difficult since one usually cannot change the cycle time at will to see how the delay changes, and this for unchanged conditions in demand and supply.

What is worse with this result is that the optimum cycle time is not as clearly defined as what one would like to have. To remedy this, more details about the stochastic processes that shape traffic have to be collected and analysed.

4 Conclusions

A number of different tests have been used to evaluate the quality of Webster's approximations. The basic conclusions that can be drawn from this research are as follows:

- The deterministic approach by Webster fits fairly well with a host of different deterministic modelling approaches.
- For the stochastic part, however, differences between theory and simulation reality have been found.
- The consequences to be drawn for every day planning are a bit more complicated. First of all, since there are differences especially in the stochastic description, there will be also differences in the quality that can be reached from the equations that have been derived from Webster's theory.
- In addition to that, the approximation formulas that are used today, especially the one for the optimum cycle time, is not in any case the real optimum. This should not be too surprising, since it definitely is an approximation, and it seems unlikely that somebody can find a better analytical solution.
- However, this is simple to deal with: there is no longer the need for handbook formulas. Instead, it might be much more flexible to use numerical tools to actually solve these equations. In the case of the optimum cycle time, and under the additional assumption that the green times g_r are proportional to the saturation values y_r, the finding of a good optimal cycle time can be coded in a spread-sheet program with very little effort. The analysis in this paper has shown that this yields on average a 4% improvement, which can be as large as 40%.
- It is argued here that this numerical approach will be the one of the future, there is no need to use the handbook approach any longer. Although the direct numerical estimation is more complicated, there is no need to worry about the quality of the results, and, as has been shown, may even lead to small improvement in performance.
- Finally, as the discussion of stochasticity has shown, there is an urgent need to account for it correctly, especially when it comes to simulation tools and their results. Here, a certain amount of empirics is needed, and, eventually, some improvement of the micro-simulation models as well may be important in accordance of the empirical results.

References

[1] *Highway Capacity Manual 2010*, volume Volumes 1 - 4. Transportation Research Board., 2010.

[2] R. Akcelik. Time-dependent expression for delay, stop rate and queue lengh at traffic signals. Technical Report AIR 367-1, Australian Road Research Board, Vermont South, Australia, 1980.

[3] R. Akcelik. The highway capacity manual delay formula for signalized intersections. *ITE Journal*, pages 23 – 27, March 1988.

[4] D Cheng, C. J. Messer, Z. Z. Tian, and J. Liu. Modification of webster's minimum delay cycle length equation based on hcm 2000. In *TRB 2003 Annual Meeting CD-ROM*, 2003.

[5] D. G Kendall. Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded markov chain. *The Annals of Mathematical Statistics*, 24(3):338, 1953.

[6] R.M. Kimber and E.M. Hollis. *Traffic Queues and Delays at Road Junctions*. TRRL laboratory report. Traffic Systems Division, Transport and Road Research Laboratory, 1979.

[7] G. F. Newell. *The Theory of Highway Traffic Signals*. Number COURSE NOTES UCB-ITS-CN-89-1. Institute of Transportation Studies, University of California at Berkeley, 1989.

[8] G.F. Newell. A simplified car-following theory: a lower order model. *Transportation Research Part B*, 36(3):195–205, 2002.

[9] H.J. Van Zuylen and F. Viti. Delay at controlled intersections: The old theory revised. In *Intelligent Transportation Systems Conference, 2006. ITSC '06. IEEE*, pages 68–73, 2006.

[10] F. V. Webster. Traffic signal settings. Technical Report Road Research Technical Paper No. 39, Her Majesty's Stationery Office, Department of Scientific And Industrial Research Road Research Laboratory, 1958.