Reliable design and evaluation method for multilane signalised intersections

Francesco Viti
Researcher, Delft University of Technology, Delft, The Netherlands

Henk J. van Zuylen
Professor, Delft University of Technology, Delft, The Netherlands

Synopsis

Signal controls are designed for increasing safety and regulating conflicting flows. On the other hand they may highly reduce road capacities if not properly set. It is always a challenge for the road authority to find the best design, both from the operational and from the infrastructural point of view. From both sides a good estimation of the traffic to be served and the network costs is needed.

Good design of the road geometry at signalized intersections helps at better managing the different streams approaching the intersection, avoiding conflicts between flows diverging to different directions. Often in practice accumulation lanes are constrained to be smaller than required for spatial reasons, especially in urban areas. In other cases the error comes from bad design of the intersection.

The Highway Capacity Manual (HCM) provides analytical formulae for computing the mean value of the queue but does not give any suggestion on how variable this value can be. An underestimation of how many vehicles to be served at an intersection can raise several problems, i.e. spillback with consequent increase of waiting times and lane blocking together with serious decrease of safety. Cars may not be served within a green phase because blocked by a spillback of another lane; moreover they can try risky maneuvers to be served and increase the chance of accidents.

This paper proposes a probabilistic method based on the Markov chain renewal process, especially suited for practitioners, to compute the dynamics of queues and their variability to design and evaluate the optimal length of accumulation lanes with particular regard to spillback avoidance. The method allows the analyst to estimate the control delay for an individual vehicle in time and the uncertainty of this delay to occur.

We propose also a heuristic model, which computes the average and the standard deviation of queues in time for variable demand and variable signal settings. The model has been implemented in a test scenario and compared with the Markov Chain simulation results. To assess the consequence of spillback we also compute the queue dynamics in a multilane and multiphase signalized intersection, showing how the queue on a restricted turning lane can influence the formation of queues and the behavior of users on the other lanes.

The model is shown to fit better simulation data generated with Markov Chain processes than the HCM especially when the intersection is slightly under- or oversaturated. Simulation data suggests assuming in a design and evaluation study for accumulation lanes a queue distributed as Normal constrained to be non-negative. A risk-averse road administrator can then set a certain probability threshold and compute the maximum amount of passenger car units an accumulation lane needs to be designed for. This method is expected to better estimate this length than using the HCM especially when a responsive control signal is implemented.

Reliable design and evaluation method for multilane signalised intersections

Urban traffic is increasingly keeping the attention of researchers, since congestion characterizes most metropolitan areas and demand is recurrently too large for the actual capacity of the roads. Many urban networks have been, or need to be, designed to satisfy a large amount of flows. In this context intersections represent critical points for assessing the efficiency and reliability of the network. Dynamic traffic control strategies, like adaptive and actuated control strategies, are, among the Dynamic Traffic Management measures, strategies able to modify the capacity of the infrastructure and to adapt it to the demand without physically modify the roads. Aim of these strategies is to both reduce as much as allowed the delays, queue lengths and stops, and guarantee safety to the travelers' maneuvers at such nodes. An important feature of such strategies is also to keep the queue lengths under a certain value in order to avoid the spillback effect (occurring when the length of the queue is larger than the length of the section), which creates problems and extra delay also in other parts of the network. A good infrastructure design should then take into account the intersection demand and offer enough space to efficiently use the signal plan.

Traffic control is one of the determinants of travel times and costs in an urban area. In fact, delays due to waiting times at intersections represent nearly 50% of the total time a traveler spends during his journey. Taking into account the extra costs a traveler commonly feels and associates to these nodes respect to link costs, like stress of waiting at the queues and discomfort produced by frequent stops, it is straightforward to understand the importance both from the road manager point of view and from the travelers' point of view of having a network capacity able to serve the different flows approaching the nodes within acceptable times.

On the other hand, the interactions between multiple vehicle-classes, different streams, heterogeneity of speeds etc. make the assessment of such strategies really difficult, especially when congestion is involved.

Aim of the road manager is then to find the best set-up for the traffic signals in order to guarantee acceptable waiting times for all vehicles. Due to the dynamic and the stochastic nature of the variables involved and the complexity of the system, it is really difficult to compute with a sufficient accuracy the intersection costs.

A model, which estimates as good as possible the dynamics of queues, and also the variability of these queues, is one of the main challenges in the recent research.

A large uncertainty characterizes the queue estimation and prediction at signalized intersections (van Zuylen and Viti 2003). We noticed that, under some quite general assumptions, and considering the inflows and the outflows as stochastic variables, the standard deviation is in some cases comparable and even larger than the average. We showed also that this phenomenon is especially observable when the degree of saturation floats around 1. Since dynamic traffic control strategies (i.e. traffic actuated signals, or adaptive control) are usually set in order to give to the average inflows the minimum amount of green time to clear the intersection (i.e. Smith 1980, Bell 1990). This condition is often met in real life.

If then a road manager has to evaluate or design the proper geometry of an intersection, such as how many lanes to reserve to a certain direction flow, or how long an accumulation lane (here intended as an exclusive turning lane ending up at the intersection) should be, he should primarily have at hand valuable queuing and delay models, which should reflect as good as possible the queues and delays vehicles experience in reality and also be able to provide consistent results when changes in the scenario make a comparison with reality impossible.

Most available methods to evaluate and predict queues and delays at intersections (i.e. the Highway Capacity Manual 2000, Akcelik 1981) provide simply an average value of the queue. A standard design problem implies then the analysis of the daily demand that approaches the intersection and the way these flows split among all available directions.

If the maximum demand among the analyzed is considered for designing the geometry of the intersection the road authority might get the risk to build up such long lanes or too many dedicated lanes that they are nearly always unused, producing enormous costs for the construction and a large impact to the environment.

If on the other hand if the road manager decides to use an average value, there could be a high chance that the space is not sufficient to store the vehicles, and spillback may often occur for some times of the day.

In this paper we aim at providing a method for estimating the queue length dynamics and its distribution. The proposed method allows the road manager a better evaluation and design of the intersection geometry, in terms of number of lanes to dedicate to some streams and in terms of length of the accumulation lanes.

To do so, we first propose a statistical approach for the estimation and prediction of queues based on the Markov Chain process, which allows one to compute not only average values but also a complete probability distribution of queues in time. Later we compute with this method the queue dynamics in a multilane intersection, showing the consequences of spillback of an accumulation lane on the other lanes and on the travelers' behavior.

We present later a heuristic model, which accurately mimics the results of the Markov Chain simulation and that reduces the computational effort required to compute such technique.

We finally solve the problem of finding the optimal length of an accumulation lane given the probability for a spillback to occur a road manager is willing to accept.
We first present the design problem and we explain the need for a model, which is able to compute the probability distribution of the queues. The next section introduces the concept of control delay and of random queues. A description of the Markov Chain mesoscopic simulation model follows together with the presentation of the heuristic analytical expression. Later, an example of application of the reliable design and estimation method is presented. Finally we draw our conclusions.

DESCRIPTION OF THE DESIGN PROBLEM AND MOTIVATION OF THE APPROACH

The scope of this paper is to provide the practitioners a method to evaluate and design the geometry of an intersection, with particular regard to the length of the accumulation lanes, which is also able to evaluate the risk of spillback and the costs this phenomenon produces to the other streams.
In this section we build up the scenario where the studied problem takes place. Figure 1 shows an example of a road section placed upstream of a signalized intersection. In this example the section ends with three accumulation lanes, two dedicated to straight-through vehicles and one is an exclusive right-turn lane.
A signal control may be set in order to give different green times according to the required clearance time for each stream. The flow that approaches the intersection is split randomly between the two lanes before the starting of the accumulation lanes.
In general, the distribution among lanes is not equal. If the left lane is dedicated to overtaking operations usually a larger percentage of vehicles will be observed on the right lane.
If the road geometry allows the driver to check the queue in front of him and on the other alternative lanes, he (or she) will try and move to the lane with the shortest queue observed at the time of his (her) decision.
Another factor to take into account in this problem is the gap acceptance of the users. A user might be willing to change lane but this intention can be impeded by the presence of other vehicles on the target lane at the moment the traveler wants to change lane. Thus, not all vehicles that want to change lane have actually the chance to do it, contributing to the asymmetric evolution of queues on the different lanes.
Furthermore, if a queue builds up on the exclusive right-turn lane and exceeds the length of this lane, it will block the right straight-through lane as well, producing hindrance to the vehicles arriving on that lane.
This phenomenon will push the flows to increase the lane-changing maneuvers towards the left lane, reducing the total intersection capacity.

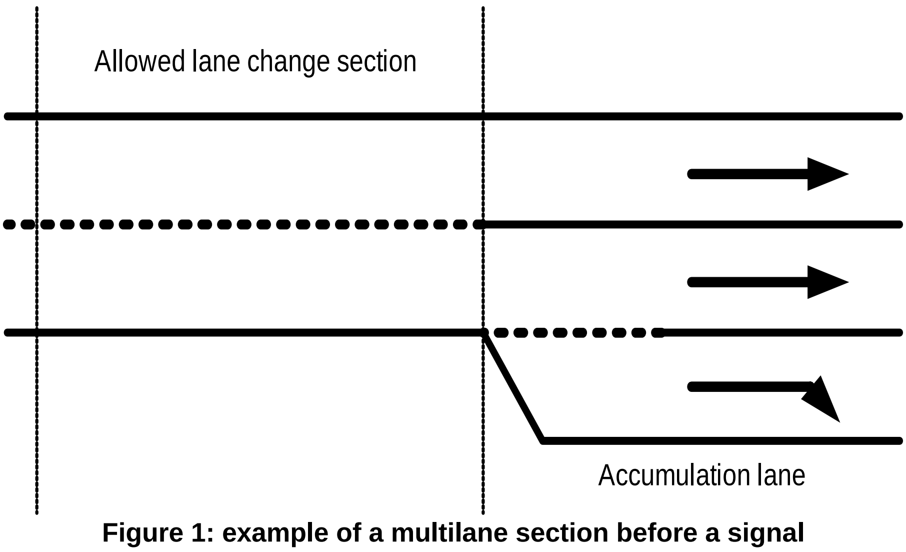

Figure 1: example of a multilane section before a signal

The described problem involves several different operational and psychological factors, which make the estimation of the queues at each lane and of the control delay very difficult.
We aim at finding a method to evaluate and predict such queues. This method should be also able to provide the probability distribution of these queues to occur.
In literature estimation and prediction models are frequently classified in three methods: macroscopic, mesoscopic and microscopic.
Microscopic simulation models are able to provide very accurate results, since they simulate traffic at the vehicle level. These models can provide estimates of delays close to reality and catch the variability of such estimates since they can consider traffic heterogeneity due to i.e. different driving behavior, different traffic conditions, road geometry and so forth. This technique is often preferred for evaluation studies, where the road geometry is already fixed. On the other hand this detail level increases the computational effort and speed of these models, limiting their use for optimization and design purposes. If several different scenarios need to be evaluated several simulations should be made, and this is unfeasible with such simulation

programs. A last drawback of such programs is that the output is highly dependent on the random nature of the input variables. To generate a sufficient number of outputs to estimate an accurate statistical distribution of queues and delays several simulations should be made as well.

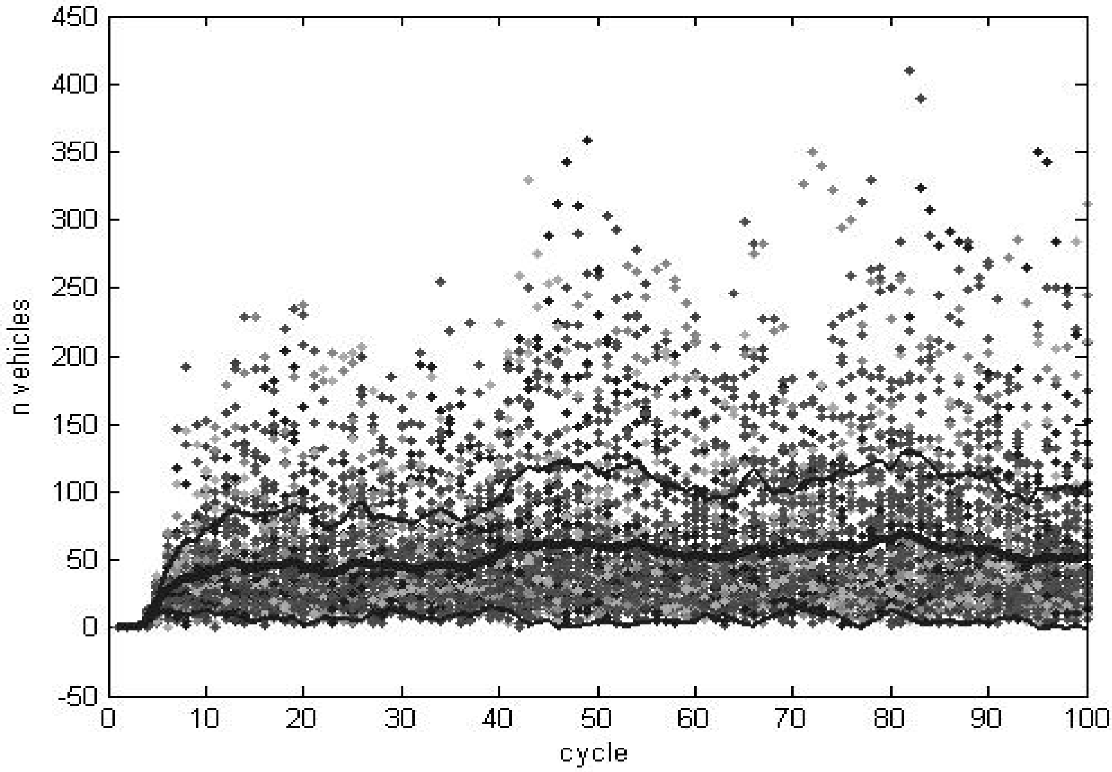

Figure 2: simulation of a queue using the program VISSIM

Figure 2 shows the aggregation of 100 simulations of queues detected for 100 cycles of 60 seconds each at a single lane signalized intersection when the degree of saturation is nearly 1. The used simulation program is VISSIM (PTV 2003). After these iterations the program gives a measure of the average length of the queue but also its variability with sufficient preciseness. As we can see, although the average value tends to stabilize around a certain value, in some cases the queue may have very a large value. In one occasion (the top point displayed on the picture) the queue detected was about 8 times the average value.

Sometimes practitioners are more interested in having results in a faster way. Macroscopic and mesoscopic models are certainly faster methods but they are characterized by a lower level of detail. Macroscopic models are able to compute only average conditions based on average inputs.

Mesoscopic models cannot compute each vehicle delay with the same accuracy of microscopic models but, despite of the macroscopic ones, they can still provide an accurate statistical distribution of network costs. This distribution is computed without the need of several simulations, keeping the computing time low. This makes this method suitable for our research.

In summary the objective of the paper is threefold: 1) we propose a mesoscopic methodology to analyze the performance of a traffic network by taking the example of a simple isolated intersection, 2) we want to provide the practitioners with an easy-to-use analytical formula to evaluate the length of the queue and its uncertainty and 3) we investigate in detail the effects of spillback occurred at an accumulation lane in terms of queue length and travelers' lane-changing behavior at the adjacent lanes.

In the following sections we describe the mesoscopic method we propose for the evaluation of reliability of a signalized intersection. We also propose an analytical expression for the average and the standard deviation of the queue, which accurately follows the trend displayed by the Markov mesoscopic model.

Next section introduces the relationship between the queue length and the control delay and explains the influence of random queues in such delay.

MODELING QUEUES AND DELAYS AT INTERSECTIONS

Control delay is usually defined as the difference between the travel time experienced by a vehicle passing an intersection and the hypothetical travel time experienced if the intersection is not controlled. According to the Highway Capacity Manual (HCM 2000), if the intersection is under saturated, delay is characterized by a deterministic and a stochastic component.

The deterministic component is determined by the average degree of saturation. The stochastic component computes the waiting time due to the random arrival of vehicles during the cycle time at the intersection.
Another component it can be observed in practice is due to the fluctuation of the demand in between cycles. It is common in practice to assume the arrivals as constant at least for a limited period of time, but in reality these arrivals are characterized by a random fluctuation along this average value.
If the green phase is not long enough to handle all the cars, the queue at the end of the green phase is not empty. If the number of arrivals is structurally higher than the number of cars that can depart during the green phase, this term will grow every cycle and it will depend on the duration of the period over which the delay is calculated. If the number of arrivals is, on the average, less than the number that can depart within the cycle, the green phase is under saturated on the average. Due to stochastic variations in the number of arrivals, there is a finite probability that still one or more cars will have to stop twice because the green phase could not handle the whole queue. This component of the delay is called the overflow delay. The queue that evolves in this case is defined accordingly as overflow queue.
The presence of this non-zero initial queue produces an additional delay introduced in the Highway Capacity Manual only in his latest version (HCM 2000). The HCM 2000 specifies then three terms for the estimation of the delay:

$$d = d_1 \cdot PF + d_2 + d_3 \qquad (1)$$

with:

- *d*= total experienced control delay
- *PF* =progression factor, which accounts for signal coordination
- d_1 =uniform delay
- d_2 =random delay and
- d_3 =initial queue delay

The term d_3 depends on the assumed initial queue value. The manual provides a formula for d_3 as function of time *T*, the capacity *c* and the initial queue length *Q*:

$$d_3 = \frac{1800 \cdot Q \cdot (1+u) \cdot t}{c \cdot T} \qquad (2)$$

This formula requires the specification of the type of arrival profile, which determines the value of the parameters *u* and *t* and gives the delay in minutes per vehicle caused by the non-zero initial queue at the starting of the cycle.
A good estimation of the overflow delay is then subject to a good estimation of the overflow queue. The manual does not provide an accurate model concerning the evolution of the queue in time, but suggests the use of a linear deterministic model. When the value of the degree of saturation floats around 1 the initial queue delay is enormously larger than the uniform and the random components and this method tends to underestimate the control delay.
Taking into account that often dynamic signal control systems are set to vary along this range the need for an accurate model especially in these conditions is straightforward.
In past research (Viti and van Zuylen, 2003, 2004a and 2004b) we applied the Markov Chain renewal process to analyze the dynamics of overflow queues at signalized intersections, which contribute to delay propagation in time and also to its uncertainty. The model is valid for single lane intersections, where overtaking operations are not allowed and the FIFO condition strictly holds.
Some authors used already this technique to generate realistic queue lengths, before the authors did, in an isolated intersection context (i.e. Brilon and Wu 1990, Olszewski 1990, and Fu and Hellinga 2000).
An analytical model has been derived for single lane intersections, which mimics accurately the results of the Markov model for variable demand and control settings. The queue model has been compared with the Highway Capacity Manual delay formula and it has been proved to solve the underestimation problem the HCM formula suffers when the demand is slightly under- or oversaturated when compared to the Markov model (Viti and van Zuylen, 2004b).
Even if an analytical expression for this simple case has been found, it is difficult to do the same if a more complex problem is tackled. The Markov Chain process can be still easily extended, as it will be shown in the next section.
Since the paper focuses on the evaluation of queues and delays with particular regard to the occurrence of spillbacks we extend here the Markov model to a multilane section like the one in figure 1.

MODEL DESCRIPTION

This section introduces the evaluation and design model proposed. The meso-simulation model consists of two main parts: the queuing process at each lane, which works similarly to the single-lane case, and the lane-changing process, which at each cycle updates the arrival distribution among lanes depending on the queue length observed at the end of the prior cycle. These two models combined lead to the extension of the single-lane Markov model to multilane intersections.

Furthermore, we show here that, under some assumptions later specified, the queue evolution on the accumulation lanes is not dependent on the lane changing behavior before the intersection, but influences it and the evolution of queues at the adjacent lanes. Thus, the accumulation lane design problem can be independent on this model and treated as an isolated single-lane case. This allows one to use the simpler model presented in Viti and van Zuylen (2003) and the heuristic expression previously proposed in Viti and van Zuylen (2004a, 2005).

Markov chain meso-simulation model of stochastic queues

The average length of the queue E[Q(0)] at the end of a (fixed time) green phase on a controlled intersection is one of the terms in the formula that determine the expected value of the delay, *E*[*W*] (Miller, 1968):

$$E[W] = \frac{q \cdot r}{2(1-y)} \left\{ r + \frac{2}{q} E[Q(0)] + \frac{1}{s} \left(1 + \frac{1}{1-y} \right) \right\} \quad (3)$$

with

- r = duration of the red phase (s)
- q = flow of an intersection arm (veh/h)
- s = saturation flow (veh/h)
- y = load ratio q/s.

For the value of *E*[*Q*(0)] several expressions have been developed (i.e. Catling 1977, Kimber and Hollis 1979, Akcelik 1980). In the case of an oversaturated intersection this quantity will grow from cycle to cycle and it follows a linear behavior.

As previously said, the average queue is non-zero also for intersections that are still undersaturated, but where the queue length is lower than the equilibrium value. If the queue length is longer than the equilibrium value, the queue will instead decrease from cycle to cycle towards the equilibrium value.

Olzewski (1990) studied the queue length dynamics with a model in which the probability distribution of the queue length is calculated from cycle to cycle:

$$P(n,j) = \int_{-\infty}^{+\infty} ds' \frac{1}{\sigma\sqrt{2\pi}} e^{-(s'-s)^2/2\sigma^2} \sum_{l=0}^{j+s't_g} p_l P\left(n-1, j-l+s't_g\right) \quad (4)$$

with

- *P(n.j)* = probability of a queue of j vehicles at the end of the nth green phase
- σ = the variance of the saturation flow and
- p_l = probability of *l* arrivals in the cycle.

To compute this probability distribution in discrete time steps we apply the Markov Chain renewal process formula, where the product of the distribution at time *t*-1 and the transition matrix P_{ij} determines the queue probability distribution at time *t*:

$$P_t(j) = \sum_{i=1}^{N} P_{t-1}(i) \cdot P_{ij} \quad (5)$$

The probability distributions are bounded to be obviously non-negative and to be not larger than the unsignalized road capacity *N*.

The transition matrix keeps track of the joint probability of arrivals A and the service S and is determined by the following conditions:

$$\begin{cases} P_{i0} = P(A-S \le -i) \\ P_{ij} = P(A-S = j-i) \\ P_{iN} = P(A-S \ge N) \end{cases} \qquad (6)$$

The method is valid under assumptions of variable, step-wise average demand, and also time-varying signal settings. Figure 3 displays the evolution of a queue for an assigned fictitious demand, which floats around the capacity of the signal control, set with fixed green time and cycle time. The average queue grows (bold line, figure below) with the demand but the clearance time is longer than the one estimated by the Highway Capacity Manual (which would predict a non-zero length only when the flow is larger than 720 vehicles per hour, thus only for 30 minutes of the entire evaluation period). The Markov model allows one to compute also the standard deviation of the queue (dotted line), giving a measure of the variability observed in a daily scenario.

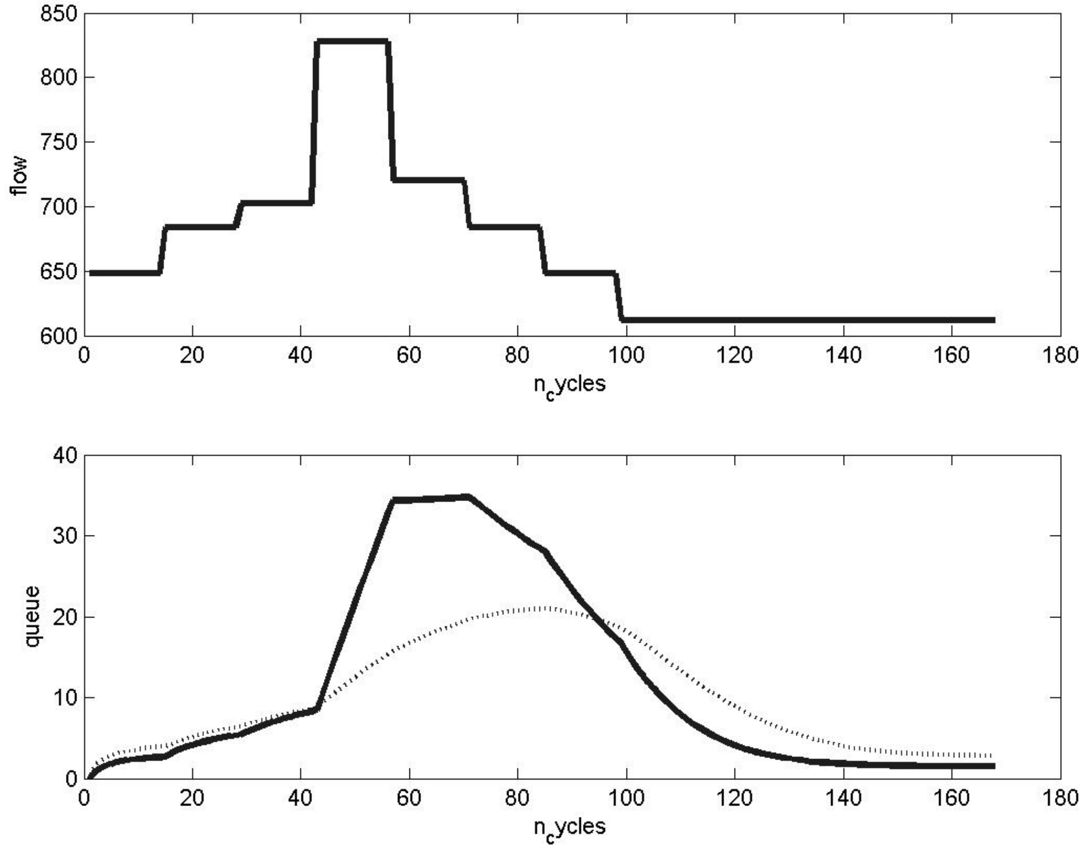

Figure 3: mesoscopic simulation of a queue at an isolated intersection

Under heavy traffic conditions on more than one arm of the intersection dynamic signal plans reduce the average and variance of the queue but they cannot always serve the vehicles within a green time (Viti and van Zuylen 2004b).

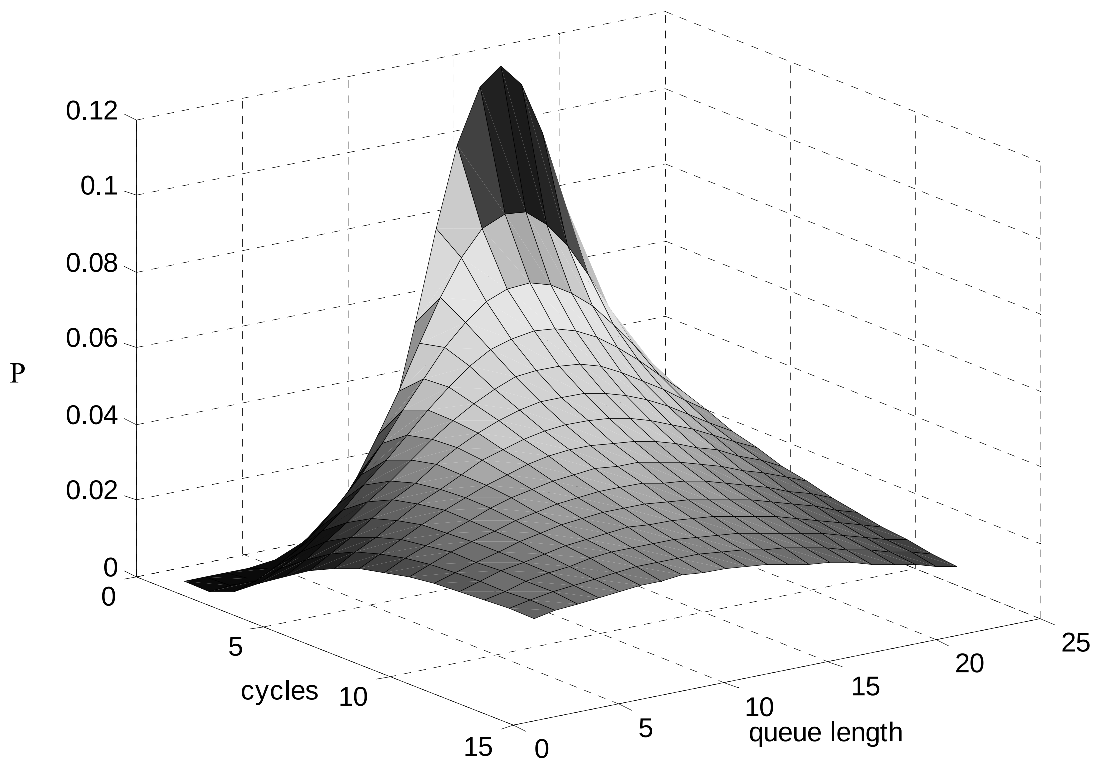

Figure 4: evolution of the queue length distribution in time

The Markov model offers to the road manager the opportunity to estimate the average trend of the queue evolution in time once the input characteristics have been fixed together with their statistical distribution. Moreover, at every cycle, the model computes the total distribution of the queue, as printed in figure 4. Here the queue distribution has been computed for a fixed degree of saturation of 0.95 and with a starting known value of 18 vehicles. We see that already after 15 cycles the distribution is very flat, thus uncertainty and the length of the queue could be very high and with a non-negligible probability.

Queuing model extended to multilane intersections

The model above described has been shown to compute the distribution of the queue for a single-lane isolated intersection. The evolution of the queue at intersections, which belong to an arterial corridor, should be different from the one above described and the reduction of average and standard deviation of the queue is consistent, since the aggregation of vehicles in platoons and the filtering effect due to upstream signals reduces especially the random component and especially when signals are well coordinated (Viti and van Zuylen, 2005). A Markov approach to the signalized arterial corridor problem was recently proposed in Geroliminis and Skabardonis (2005).

In this paper we do not consider this issue, since it does not represent the worst case scenario for the analyzed problem. Thus the above-described model can be still used in an arterial corridor only if signals are not properly coordinated.

The queuing process depends on the randomness of arrivals and departures, but the demand at the following cycles can be influenced by the queue distribution itself because of lane-changing behavior. We assume here that the user increases his “intention” to change lane the larger the difference between the queue at his lane and the target one is. Since queue length is characterized by a probability distribution also this intention can be described by a probability distribution. Once this probability is computed we have to compute whether it is possible for the vehicle to do the maneuver by applying a gap-acceptance model. We apply the Bayesian updating rule to compute this chance, thus we compute the conditional probability that a vehicle changes his current lane. This probability is simply computed by the product of the two probabilities, since we assume they are independent stochastic variables.

Let i be a lane of a road section just before a signalized intersection and j be an adjacent lane. Let us suppose that the user has time to evaluate the queue lengths in front and (eventually) change lane. Suppose also to know the number of vehicles served within a cycle C by knowing the saturation flow S_i and the length of the green phase g.

Knowing the average and distribution of the split rates we can compute for each lane the flow distribution d_i in a cycle by using the relationship $d_i = \alpha_i \cdot d$, where *d* is the flow approaching the intersection and α_i is the percentage of vehicles entering from lane *i*. This component is also a random variable.
Let $q_i(t), q_j(t)$ be the queues at the two lanes at time t and let us suppose to know the queue distribution at the starting of the simulation for each lane $q_i^0(t)$. We assume that the travelers' probability of intention to change lane is:

$$\begin{cases} P_{i\to j}(t) = h(q_i(t) - q_j(t)) & \text{if } q_i(t) - q_j(t) > 0 \\ P_{i\to j}(t) = 0 & \text{otherwise} \end{cases} \tag{7}$$

We assume the probability as a known function *h*, which increases and gets closer to 1 the larger the difference between queues is. Inversely, if the queue at the traversing lane is smaller there is no reason to consider a possible lane change to the adjacent lane.
A user can change lane only if there is enough gap for the maneuver. Known the number of arrivals we can deduce the distribution of distances among cars and consequently the probability of having enough space to change lane. For example if the arrival distribution is poissonian the headway distribution can be approximated as exponential.
The chance of having a sufficient gap to change lane is then equal to the probability $P_j(l \geq \bar{l})$ that the distance between cars observed is higher than a predefined threshold $\bar{l}$.
Once computed these probabilities the number of vehicles $d_{i\to j}$ changing from lane *i* to lane *j* will be given by:

$$d_{i\to j}(t) = d_i \cdot P_{i\to j}(t) \cdot P_j(l \geq \bar{l}) \tag{8}$$

The total number of arrivals f_i at lane *i* will then be:

$$f_i = d_i - d_{i\to j}(t) + d_{j\to i}(t) \tag{9}$$

Known the distribution of departures s_i within a cycle, the queue distribution at the following cycle can be computed like we did for single lane isolated intersections:

$$q_i(t+1) = q_i(t) + s_i - f_i \tag{10}$$

Note that in the single lane problem the arrivals were independent on the present queue, and the transition matrix was invariant with time, while in this case it is time-dependent.
This part completes the modeling of queues with lane changing behavior but not considering the spillback from another lane.
Suppose now that there is an accumulation lane aside of one of the straight through lanes. For simplicity we assume that the vehicles arriving at the intersection and that have to turn are already at the closest lane before entering in the accumulation lane, thus no lane changing to reach the accumulation lane is considered. Since it happens frequently in real conditions, we assume the green time for this extra lane different from the straight through lanes, g_{acc}.
Under these assumptions, the queue at the accumulation lane is computed exactly as a single lane intersection using the standard single lane Markov Chain.
Let β_i be then the fraction of the total demand *d* representing turning flows and let $q_{spillback}$ be the maximum number of vehicles, which can be placed in the accumulation lane without creating a spillback effect. Due to randomness of arrivals there is a non-zero chance of having spillback. The probability that spillback occurs can be computed with the Markov model.
In this condition the adjacent lane will be influenced by this phenomenon. We suppose here that if the length of the adjacent lane queue is smaller than the accumulation lane queue the user will consider for his lane changing behavior the latter.
Thus, all equations above do not change apart conditions (7) that become:

$$\begin{cases} P_{i\to j}(t) = h(\max(q_i, q_{acc}) - q_j) & \text{if } q_{acc}(t) > q_{spillback} \text{ and } \max(q_i, q_{acc}) > q_j \\ P_{i\to j}(t) = h(q_i(t) - q_j(t)) & \text{if } q_i(t) - q_j(t) > 0 \text{ and } q_{acc}(t) < q_{spillback} \\ P_{i\to j}(t) = 0 & \text{otherwise} \end{cases} \quad (11)$$

As long as the probability of spillback is small also the extra delay given by this phenomenon will be small. If on the other hand there is a non-negligible chance that the green time is not sufficient to clear the queue at the accumulation lane, the adjacent lane will also reduce in some cases its capacity, creating extra demand on the other lanes, as it will be shown in the case study section.

Evaluation and design of intersections using the proposed method

The previous section described in detail a method to estimate the queue length at each lane and a way to evaluate the queue length distribution especially taking into account the spillback effect.
If the problem is to evaluate an existent infrastructure, the road manager can use this method to estimate the delay at each lane, since the geometry is already fixed. This method can be also used to find the optimal signal settings, especially when the signal is multiphase.
If on the other hand the road manager has to design a new intersection or has the chance to modify the geometry, he can use the method to calculate the most convenient scenario. Since we showed that under the declared assumptions the lane changing behavior at the section upstream does not influence the queue evolution at the accumulation lanes, the design problem can be restricted to the use of the single-lane queuing model. In the following we present a faster method to solve the design problem.

Proposed analytical model of overflow queues

Traffic practitioners might be interested in having a simpler and easy-to-use formula to have a quick answer about the optimal length of the accumulation lanes instead of using the mesoscopic method.
In Viti and van Zuylen (2004b) we showed that under the assumption of arrivals distributed as poissonian the queue distribution can be well approximated by a Normal distribution. Known the average and standard deviation in time, the road manager can easily compute the value of the queue and consequently the length of the accumulation lane by simply fixing the threshold probability beforehand.
The design problem needs only a method to compute the average and the standard deviation, without the computation of the total probability distribution.
We assumed earlier in the paper that the queue building up at the exclusive lanes is not influenced by the queues at the straight through lanes, and that it can be computed by using the single-lane Markov Chain.
A heuristic model has been shown to accurately mimic the results of the Markov Chain process in terms of average value and standard deviation (Viti and van Zuylen 2004a), also under the assumption of variable demand (Viti and van Zuylen 2004b) and variable signal settings (Viti and van Zuylen 2005).
Here we reprint the simplified formula presented in Viti and van Zuylen (2005) together with an expression for the standard deviation.
The average queue behavior follows in general the deterministic behavior if the degree of saturation is smaller than about 0.70 and when it is larger than about 1.2, while in the intermediate cases it is characterized by two phases, an initial linear trend, Q_{linear}, and an exponential one, Q_{exp}:

$$\begin{aligned} Q_{linear}(T^*) &= Q_0 + (x-1)cT^* \\ Q_{\exp}(T^*) &= Q_e + (Q^* - Q_e)\cdot e^{\beta T^*} \end{aligned} \quad (12)$$

with

- Q_0 = initial queue at the starting of the evaluation period
- Q_e = equilibrium queue, computed with the Miller formula (Miller 1968) $Q_e = \dfrac{1.5(x - x_0)}{1-x}$
- $x_0 = 0,67 + \dfrac{S\cdot g}{600}$ is the limit value of the degree of saturation above which the stochastic effects are relevant (Akcelik 1980)
- c = capacity of the arm (veh/cycle) equal to the saturation flow S multiplied by the green ratio g / C
- x = degree of saturation
- Q^* = conjunction point between the two phases

The parameter β is dependent on the degree of saturation and regulates the time needed for the queue to pass from the initial state to the equilibrium state. An expression of the parameter has been given by:

$$\beta = -\frac{(1-x)^2}{0.15} \tag{13}$$

The conjunction point is computed in such a way that the final function is continuous together with its first derivative and it is computed by the formula (for the derivation of the formula see Viti and van Zuylen 2005):

$$T^* = \frac{lambertW(e^k) - k}{\beta} \tag{14}$$

with $k = \frac{\beta(Q_0 - Q_e)}{(x-1)c}$ and $lambertW$ is the Lambert or Omega function (Corless et al. (1996)).

The standard deviation model is computed accordingly by just replacing the initial linear trend with the following quadratic expression:

$$\mathrm{Q}_{\mathrm{quad}} = \sqrt{\mathrm{c} \cdot \mathrm{x} \cdot \mathrm{T} + Q_0 + 0.01 \cdot Q_0^{\,2}} \tag{15}$$

and by multiplying the equilibrium value by the factor $x + \frac{1-x}{0.15}$.

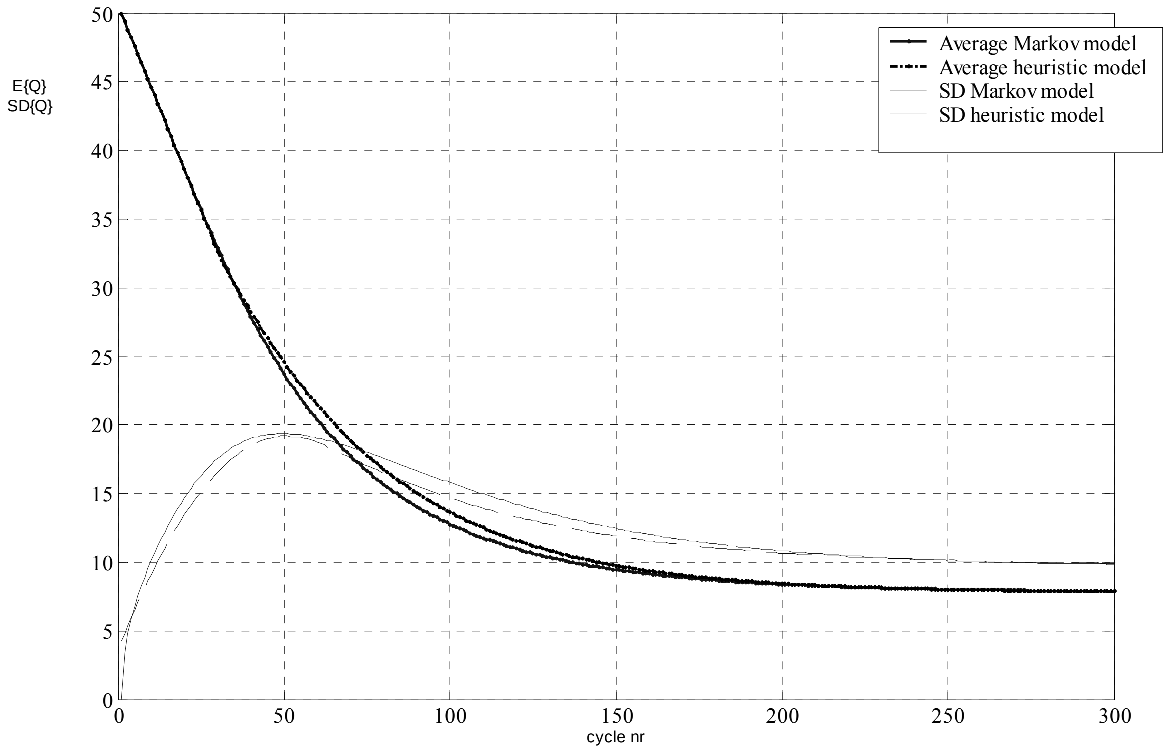

Figure 5: Comparison between the Markov model and the heuristic model

Figure 5 compares the results in terms of average and standard deviation of the mesoscopic and the heuristic models. In this example we chose an initial queue known and equal to 50 vehicles. The models predict the queue length and its variability for 300 cycles at hypothetically the average demand constant for the whole evaluation period. The average degree of saturation has been fixed to 0.95.
As we can see from the picture the difference between the results both in terms of average and of standard deviation is very small, in the order of a single vehicle error.

NUMERICAL EXAMPLE

We solve here the evaluation problem and the design problem for the road section depicted in figure 1. To show the dynamic evolution of lane changing behavior and the influence of spillback on the accumulation lane we assume constant demand for the whole evaluation period. The computation of the reliable design of the accumulation lane length follows.
For the first problem we are especially interested in knowing to what extent the dynamics of the queue can influence lane-changing behavior in slight undersaturated conditions, where queue dynamics do not have linear behavior. First simulation considers only the 2 straight through lanes, initially with 60% of vehicles on the right lane. Second simulation adds up the accumulation lane placed aside the right through lane.
Saturation flow is set to 1800 vehicle per hour, while cycle and green time and set respectively to 60 and 24 seconds, while green time for the accumulation lane is set to 10 seconds and maximum number of vehicles allowed is set to 5 vehicles.
Figure 6 shows, on the top pictures, the evolution of the average and the standard deviation of the queues at the straight through lanes, while the bottom pictures show the lane-changing behavior expressed by the evolution of the split rates in time. Left figures show the condition when the degree of saturation of the signal is set to 0.95. Because of lane changing possibility both lanes reduce their average queue with respect to the single lane case shown in Viti and van Zuylen (2004a). Standard deviation changes accordingly. Due to gap acceptance limitation, the two lanes are still not having the same demand.
When demand increases, here computed for a degree of around 0.97, lane changing also increases, since intention to change lane increases and flows tend to be equally distributed (center). When an equal distribution is met, queues tend asymptotically to increase with the same behavior. The two right pictures show instead the problem when also the right accumulation lane is considered. The chance to have spillback on the right lane due to the right accumulation lane increases the demand on the left lane further (right) and it becomes even larger than the right straight-through lane demand. The more the demand on the accumulation lane increases the larger the chance that an amount of flows moves to the left lane.

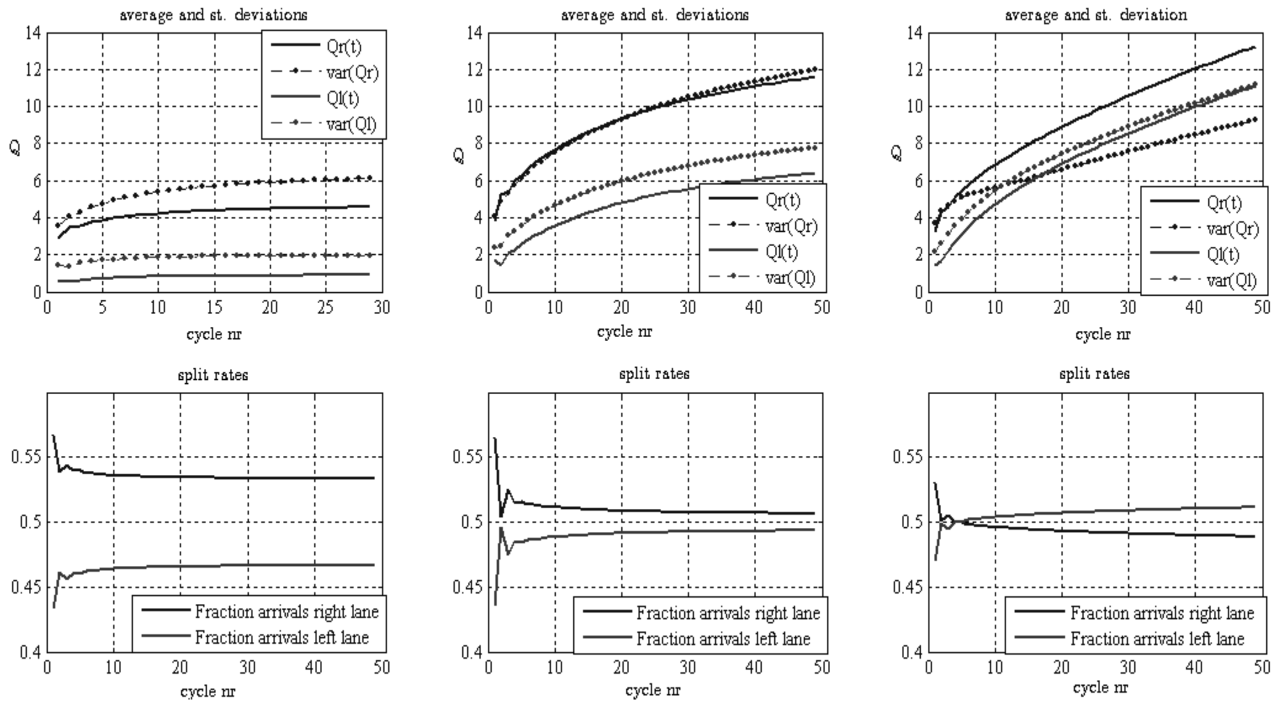

Figure 6: numerical examples of queue evolution and flow distribution in time

We saw from the above picture how the interrelationship between queues, lane-changing behavior and the limitation due to gap acceptance can change the queue evolution and the distribution of flows and queues between the available lanes.
The proposed lane changing model results depend on the choice of the parameters for the lane-changing aspiration model, the gap acceptance and the dynamic queuing model. The calibration of these parameters is needed for its practical use. The parameters chosen for the example are only for illustration purposes and do not pretend to be realistic.
Figure 7 shows the evolution of the distribution of the queue for the first 10 cycles of the queue computed for x=0.95 and starting from a zero initial queue. After these cycles the probability of a queue being for example longer than 10 vehicles is about 17%, thus non-negligible. If an accumulation lane can store at maximum this amount of vehicles there will be the above chance for a spillback to occur.

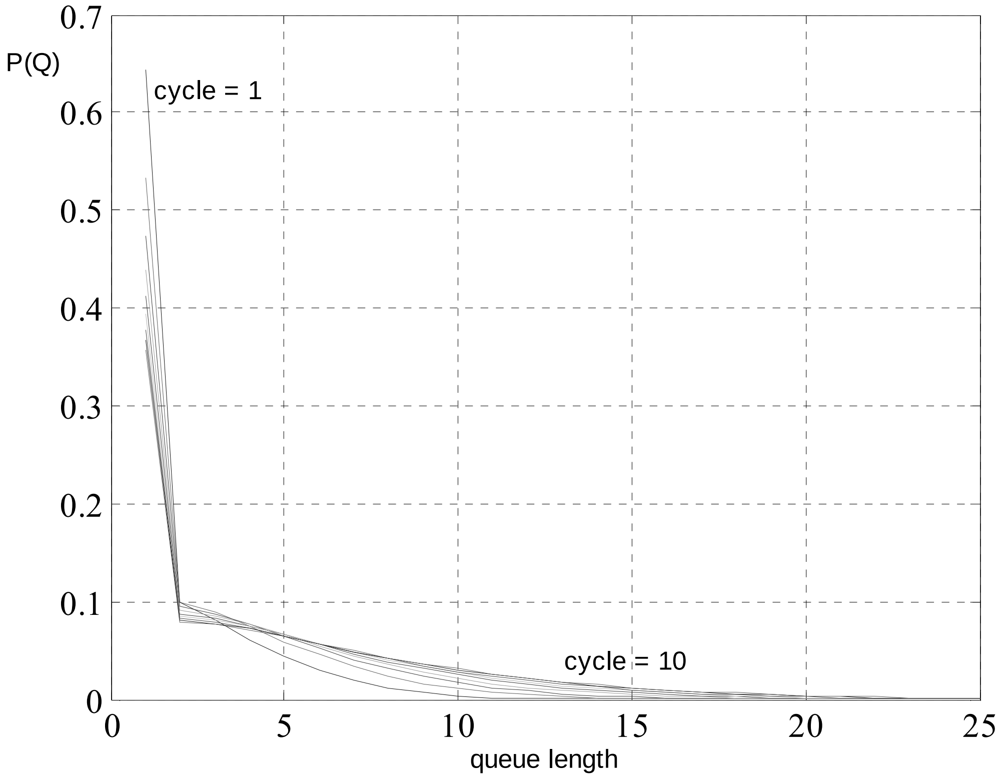

Figure 7: prediction of the queue distribution for 10 cycles

Inversely, if the road manager aims at keeping this percentage under a certain threshold, he can calculate, by analyzing the total daily demand and generating these probability distributions, which is the value of the queue that corresponds to this percentage.
Accordingly, stated the equivalence between average and standard deviation computed with the Mesoscopic method and the heuristic method, this procedure can be done with the use of the simpler analytical expression of the queue and assuming the queue as normally distributed.

CONCLUSIONS

This paper presented a method to estimate and predict the queue distribution at multilane isolated signalized intersections. We used a mesoscopic simulation technique to model the queue dynamics and its variability in time and a probabilistic method to simulate lane preference and gap acceptance of the users to model the interaction between lanes at the intersection.
We show in the paper that queuing models should consider travelers' lane changing due to unequal queues at the intersection combined with a gap acceptance model to accurately estimate the delays that a traveler may incur. Moreover we showed that extra delay could be observed if an accumulation lane is badly designed, since the spillback effect can influence the dynamics of queues at other lanes.
We proposed then to use this method also to design the optimal, in the reliability sense, length of the accumulation lanes in order to efficiently use the intersection with an acceptable chance for a spillback to occur.
To help practitioners at having a handy formula, we proposed also a heuristic analytical expression, which well substitutes the more complex mesoscopic model and makes the design problem easy and fast to be solved.
The method is expected to help road authorities at designing more efficiently the intersection geometry, in terms of length of exclusive turning lanes and required number of lanes for each flow stream. Moreover the method gives an accurate estimation of the costs when the road geometry is fixed and the authority has to evaluate the operational efficiency of the intersection, and it represents a valuable method for solving in a similar way several Dynamic Traffic Management and network design problems.

ENDNOTES

The Dutch Transport Research Center of Rijkswaterstaat (AVV) financially supports the work of both authors. It is a part of the research program Dynamic Traffic Management of the Research School TRAIL and it is framed within the Towards Reliable Mobility spearheads of the Delft University of Technology.

REFERENCES

Akcelik R. (1980). “Time-Dependent Expressions for Delay, Stop Rate and Queue Length at Traffic Signals”. Australian Road Research Board, *Internal Report*, AIR 367-1.

Bell M.G.H. (1990). A probabilistic approach to the optimization of traffic signal settings in discrete time. *Transportation and Traffic Theory*, ed M.Koshi, Elsevier, Amsterdam, pp. 619-632.

Brilon W., Wu N., (1990). “Delays at fixed time signals under time-dependent traffic conditions”. *Traffic Engineering and Control* 31 (12), 623-631.

Catling, I. (1977). “A Time Dependent Approach to Junction Delays”. *Traffic Engineering and Control*, 18, pp. 520-526.

Corless R.M., Gonnet G. H., Hare D. E. G., Jeffrey D. J., Knuth D. E. (1996). "On the Lambert W Function", *Advances in Computational Mathematics*, volume 5, 1996, pp. 329-359.

Fu L., Hellinga B. (2000). Delay variability at signalized intersections. In Proceedings of the 79th Annual Meeting of the Transportation Research Board, January 2000, Washington D.C.

Geroliminis N. and Skabardonis A. (2005). “Prediction of arrival profiles and queue lengths along signalized arterials using Markov decision process”. In *Proceedings of the 84th TRB Meeting*, 11-15 January, Washington D.C., CD-Rom.

Kimber and Hollis (1979). “Traffic queues and delay at road junctions”, *TRL LR* 909.

Miller, A.S. (1968). “The capacity of signalised intersections in Australia”, *Australian Res. Board bulletin* 3

Olszewski P. (1990). “Traffic Signal Delay for Non-uniform Arrivals”. *Transportation Research Rec.* 1287. TRB, National Research Council, Washington D.C..

PTV Planung Transport Verkehr AG (2003). “VISSIM version 3.70 user manual”, January 2003, Germany.

Smith M.J. (1980). A local traffic control policy, which automatically maximizes the overall travel capacity of an urban road network. Paper presented at the International Conference on Urban Traffic Control Systems, Berkeley, California (August, 1979). In *Conference proceedings* and in *Traffic Engineering and Control*, 21, 298-302.

Transportation Research Board (2000). “Highway Capacity Manual”, *special report* 209, National Research Council, Washington D.C., TRB.

Van Zuylen H.J. and Viti F (2003). “Uncertainty and the Dynamics of Queues at Signalized Intersections”. *In Proceedings of the CTS-IFAC conference*, 6-8 August, Tokyo. Elsevier, Amsterdam, 2003.

Viti F. and Van Zuylen H.J (2004a). “Modeling queues at signalized intersections”. *Transportation Research Record* 1883, 68-77. National Research Council, Washington D.C..

Viti F. and Van Zuylen H.J (2004b). “Modeling overflow queues on urban signalized intersections”. *In Proceedings of the 10th World Conference of Transport Research*, 7-8 July, Istanbul, Turkey, 2004.

Viti F. and van Zuylen H.J. (2005). “Modeling Overflow Queues at Signalized Arterial Corridors and with Responsive Control”. In *Proceedings of the 84th TRB Meeting*, 11-15 January, Washington D.C., CD-Rom.