ARTICLE IN PRESS

JOURNAL OF TRAFFIC AND TRANSPORTATION ENGINEERING (ENGLISH EDITION) XXXX; XXX (XXX): XXX

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.keaipublishing.com/jtte

Original Research Paper

A multi-class time-dependent model for the analysis of waiting phenomena at a motorway tollgate

*Andrea Pompigna**, *Raffaele Mauro*

Department of Civil, Environmental and Mechanical Engineering, University of Trento, Trento 38123, Italy

HIGHLIGHTS

- A time-dependent multiclass queue model, with compact closed-form equations.
- An agile way to work beyond probabilistic queues, without resorting to simulations.
- A formulation with parameters that can be calibrated on real data, directly usable in practice.
- A good approximation level, compared with discrete-state simulation model reiterations.
- Satisfactory flexibility requirements for current and future uses in traffic engineering.

ARTICLE INFO

Article history:
Received 16 September 2019
Received in revised form
20 March 2020
Accepted 10 April 2020
Available online xxx

Keywords:
Traffic engineering
Time-dependent queue model
Shared queue
Coordinate transformation method
Tollgate

ABSTRACT

The planning, design and operational management of motorway toll booths are of great interest in traffic engineering, as these facilities directly influence the quality of the service offered to users. This paper focuses on a time-dependent queue model based on the coordinates transformation criterion for operations assessment at a motorway tollgate. This model allows to face the whole spectrum of situations that may characterize a toll booth, some of which often fall outside the boundaries of the probabilistic theory for stationary queues.

The paper proposes an $M/\Gamma/1$ multi-class queue model for the evaluation of evolutionary profiles of waiting times and queue lengths by closed-form equations. The results obtained for three numerical test cases show a good approximation level, compared with the mean values of queue parameters obtained reiterating a discrete-state simulation model.

The proposed time-dependent equations will be useful in technical cases, allowing to operate quickly and compactly even when probabilistic queue theory is not applicable or produce unrealistic results, and the burden of complexity of the simulation approach is not conveniently absorbable. The discussion highlights a significant flexibility of the model proposed in addressing situations with conventional vehicles, i.e., with total human control and proposes some considerations for application in future scenarios with the presence of connected vehicles (CVs).

* *Corresponding author*. Tel.: +39 339 2686976.
E-mail address: andrea.pompigna3@gmail.com (A. Pompigna).

Peer review under responsibility of Periodical Offices of Chang'an University.
https://doi.org/10.1016/j.jtte.2020.09.001

1. Introduction

Planning, designing and managing toll collection facilities is of great interest in motorway engineering. Motorway toll stations and toll barriers, in fact, can represent a crucial problem both for the concessionaire in relation to reductions in motorway capacity that can result from their congestion during the daily operations, and for the users who increase their travel time because of the queues and delays in the toll collection system. Operating conditions analysis of such infrastructures, therefore, represent a significant matter supporting decisions and influencing both the service quality offered by the management and the comfort and satisfaction for the users.

The Highway Capacity Manual (HCM), published by the US TRB since the 1950s of the last century, is a widely used reference for the analysis of road transport systems. While addressing several aspects that affect the conditions of the motorway service, the manual in its various editions up to the 6th (TRB, 2015) (i.e., the most updated edition) does not include guidelines or operating procedures for motorway toll stations or barriers. From this point of view, therefore, there are no standardized references and analytical procedures that can serve in the various phases of planning, design, operation and management.

Several studies, however, have dealt with this theme by proposing methods and models for performance and service levels assessment, differentiating themselves for the analytical or simulative approach and for the performance indicators choice.

As regard the first topic, various researches consider the analytic application of the queue theory (Boronico and Siegel, 1998; Chakroborty et al., 2016; Cherng et al., 2005; Edie, 1954; Haight, 1958; Kim, 2009; Mehri and Djemel, 2011; Wang, 2017; Zarrillo et al., 1997) or the simulation analysis to evaluate waiting times and queue lengths, sometimes combining the two approaches and comparing the results (Aksoy et al., 2014; Astarita et al., 2001; Ceballos and Curtis, 2004; Kim, 2011; Lin and Su, 1994; Obelheiro et al., 2011; Sadoun, 2005; Shanmugasundaram and Punitha, 2014; Van Dijk et al., 1999; Woo and Hoel, 1991).

Regarding the second topic, international researches propose the delays associated with the waiting time in the queue (Al-Deek and Radwan, 1995; Edie, 1954; Klodzinski and Al-Deek, 2002; Lin and Su, 1994; Zarrillo et al., 1997), the queue lengths (Van Dijk et al., 1999; Zarrillo et al., 1997) or the demand-capacity ratio (Al-Deek and Radwan, 1995; Woo and Hoel, 1991) as main performance indicators, i.e., the so-called measures of effectiveness (MOEs). Using these MOEs or their combination, several studies have proposed service level ranges, resorting to the HCM style (i.e., 6 levels of service, LOS, in increasing order of criticality from A to F) (Klodzinski and Al-Deek, 2002; Lin and Su, 1994; Obelheiro et al., 2011; Woo and Hoel, 1991). Of great interest is therefore the search for simple and compact models that allow to estimate these parameters reliably, allowing to deal with the different operating situations and with the best possible level of approximation.

Moreover, as will be discussed in the following literature review, traffic engineers often assess waiting times and queue lengths at a tollgate using the results of the probabilistic theory for stationary queue. These results allow, in some recurring cases of direct application, to identify closed-form expressions for queue lengths and waiting times calculations. It is also well known, however, that the probabilistic queue theory does not allow to fully deal with all the situations in which a motorway toll facility can be found, generating realistic results only in stationary situations sufficiently far from saturation levels (Ceballos and Curtis, 2004; Lin and Su, 1994). Discrete-state simulations (based on arrivals and services probability distribution hypotheses) or micro-simulation models (based on the formalization of vehicles interactions, according to the traffic flow theory), still allow to treat such situations. It should be emphasized that simulation techniques may have some drawbacks, regarding for example: the burden of implementation, verification, calibration and validation of the models; the computational effort to evaluate the possible variations in time profiles of arrival and the service characterization for vehicles; the difficulty of generalizing, often operating with case-by-case models; the need to reiterate the simulation for a sufficiently large number of times in order to produce statistically significant results.

In light of the above, this paper aims to deepen the application of the queue theory for a motorway tollgate in such situations that cannot be tackled using the results of the stationary queue theory, due to the presence of statistical non-equilibrium conditions, in whatever way they may occur.

In highway engineering the so-called time-dependent models are of great interest. These models have been developed starting from the works of Catling (1977) and Kimber and Hollis (1979) and allow to treat all the saturation cases of a generic service counter through an unified approach. As explained further below in the paper, a time-dependent model identifies some closed-form expressions that allow to conveniently combine the solutions deriving from the stationary probabilistic queue theory with those from the deterministic theory of the waiting phenomena (Pompigna, 2020).

The purpose of this work is, therefore, the definition of a time-dependent model for a motorway tollgate that allows to fill the gap highlighted by literature review. Actually, as clarified below, the reviewed researches usually overcome the limits of probabilistic theory using simulations with a few exceptions (Levinson and Chang, 2003). The results of the time-dependent model proposed in this paper are compared with what the authors get from a large number of trials of a stochastic discrete event simulation (SDES) model to verify the reliability in practical uses.

The results showed in this paper can be of great interest for technical applications such as: the design of a toll collection system (e.g., number and type of tollgates); the assessment of service levels, which can be related to the current situation or for expected scenarios, due to changes in traffic demand and/or service supply; the quantitative characterization of transport networks for traffic demand assignment in transport planning models.

The paper is structured as follows. Section 2 reviews and discusses the main researches appeared in the international literature since the 1960s of the last century, going through the different approaches related to the performance analysis of toll stations and barriers. Section 3 shows an example of a motorway tollgate characterization, examining the nature of the arrival and service processes, and presents the closed-form equations for a multi-class $M/\Gamma/1$ stationary queue model. Section 4 deals with the time-dependent modeling for the non-stationary queue. Starting from an introduction to the coordinate transformation method, the time-dependent solutions for the multi-class $M/\Gamma/1$ queue model are deducted for a tollgate. The closed-form equations for the average waiting time and the average number of vehicles in the system during a certain observation interval are provided, which are also applicable sequentially allowing to evaluate any time profile for demand and service. Section 5 presents the SDES model used to compare time-dependent model results. A stochastic simulation model with Poisson arrivals and multi-class Gamma-distributed service times, corresponding to the already mentioned multi-class $M/\Gamma/1$ queue model, is presented. Section 6 discusses some numerical results, comparing time-dependent model outputs with a sufficiently high number of iterations of the SDES model by applying the Monte Carlo method. Section 7 offers an insight for a future scenario, which considers an increase in automatic payment technologies and a growing popularity for connected vehicles (CVs). Section 8 contains the concluding remarks.

2. Literature background

Edie (1954) was the first to deal with the analysis of a toll booth using the analytical theory of the queues, analyzing its performance through the probabilistic $M/M/s$ and $M/D/s$ queue models with multiple serving channels, assuming Poisson arrivals (i.e., $M/\cdots/\cdots$) and exponentially (i.e., $\cdots/M/\cdots$) or deterministically (i.e., $\cdots/D/\cdots$) distributed service times. The model was used to estimate the time taken by a vehicle to carry out the operations and to clear the gate, in the hypothesis of homogeneity in vehicle types and gates characteristics. Compared to the discussion of Edie (1954), which appears not properly in accordance with the actual operations of a station with a sufficiently sized toll plaza, Haight (1958) proposed a model applicable to a double-entry toll station. With respect to this model, the incoming vehicle chooses time to time the shortest between the two parallel queues. The model considers Poisson arrivals and exponential service times with possibly different average values for the two gates, in the double hypothesis that each vehicle must necessarily keep the gate chosen on the arrival or that it can change queue while waiting.

To address queue lengths and waiting times for a toll station with a variable number of gates and different types of toll collection systems (e.g., manual, automatic, electronic, eventually mixed and with distinction for vehicular class), Zarrillo et al. (1997) used a deterministic model based on the cumulative curves of arrivals and service processes. In presence of multiple gates with different types of service, with the aim of studying the costs minimization, Boronico and Siegel (1998) applied an approach derived from the probabilistic theory of queues. The arrivals are determined under the assumption of uniformly distributed demand on gates of the same type, with different service times depending on whether the collection system is automatic or manual. These authors used exponentially distributed service times, while noting that the variance in real-life data is lower than that of the chosen distribution. In light of these considerations, and focusing the analysis on a New Jersey Parkway toll station, an $M/M/1$ queue model was developed in order to obtain the higher limits of the investigated MOEs.

Also Cherng et al. (2005) used the probabilistic theory for steady-state queues to evaluate the waiting times at the gates of a motorway toll station and at the convergence points downstream. These authors formulated some basic assumptions to define the model: a constant flow uniformly distributed on the gates; exponential arrival and service times; total waiting time obtained as the sum of time to receive and complete the service plus the lost time at the confluence to reach the exit segment. An $M/M/1$ model for waiting at the gate and an $M/G/1$ model for the downstream confluence were used, in order to determine the optimal number of gates for minimizing the total time spent. Also Wang (2017) presented an $M/M/1$ model for the estimation of the waiting time spent at a motorway tollgate, considering different types of payment services. The queue model was used within an optimization method for the identification of the number and the configuration of the gates, according to a non-dominated sorting genetic algorithm II (NSGA-II).

As an alternative to the $M/M/1$ model, Kim (2009) proposed an $M/G/1$ model with service times generically distributed with known mean and variance, for estimating the average waiting time required by an integer and non-linear programming algorithm for a toll station optimization. The model considers the hypothesis that the arrivals by payment type are uniformly distributed on the gates. The assumption regards a real-life evidence: an incoming vehicle tries to reach the shorter queue among those that offer the requested service and, once in the waiting lane, the vehicle does not abandon it for another.

Mehri and Djemel (2011) provided further specifications for probabilistic models regarding service times and gates number. The authors identify a Gamma (Γ) distribution for service times starting from the analysis of experimental data collected for a Tunisian motorway and using a stationary $M/\Gamma/s$ for performance analysis. The analysis considers an approximation of the Pollaczek-Kintchine (PK) formula for the $M/\Gamma/s$ model. Using a multi-agent simulation model, they confirm the lower adequacy of an $M/M/s$ model. Karsaman et al. (2014) used an $M/M/s$ model for the stationary analysis of the differences in capacity between cash and electronic payment systems on Indonesian motorways. Also Liu et al. (2017) proposed an $M/M/s$ model for cost analysis and optimization of a toll station, including a cellular automata model to evaluate some traffic indicators.

Regarding the aforementioned single or multiple channel models, Chakroborty et al. (2016) proposed a model called coupled multiple-queue queuing system (CMQ2S). The model is valid under stationary conditions of under-saturation and incorporates some of the considerations in Haight (1958). As for the parallel queue model, but with multiple gates, the CMQ2S model allows to manage the arrivals at the single gate queue as dependent by the queue lengths for all the gates.

An important part of the literature recurs to the simulation techniques, which allow to simulate the state of the tollgate and of the vehicles, and to evaluate the main indicators for the analysis of the waiting phenomenon. Customized simulation models can be implemented as discrete-state models (based on hypotheses in the distribution of arrivals and services) or as micro-simulation models (based on the formalization of the interactions between vehicles using the traffic flow theory), and in turn considering ad-hoc implementations or commercial software.

Woo and Hoel (1991) developed a custom implementation, using an event simulation model to correlate average density and flow-to-capacity ratio for a tollgate starting from real-life data. The model considers an exponential distribution of arrivals and a Gamma distribution for service times, with parameters differentiated by vehicle type. Also Sadoun (2005) used an event simulation model, considering homogeneous and non-homogeneous gates regarding the payment type, to test the facility performance and the optimal gates number under specific time-dependent traffic profiles.

Lin and Su (1994) developed an ad-hoc event simulation model, structured in eight modules that consider arrivals according to a shifted negative exponential distribution, a probabilistic model for gate selection and different service times depending on the gate type. The outputs concern each gate (with the analysis of the average time in the system, the average and maximum length of the queue) or refer to the entire toll station (with the analysis of the total number of queued users and the average vehicle speed). They test the model also against analytical formulations, based on the queue theory in terms of average queue lengths and average times in the system. The authors do not find significant differences for saturations lower than 0.95, revealing, on the contrary, a substantial over-estimation in the analytical models results for higher saturation values. Shanmugasundaram and Punitha (2014) developed a discrete-state model for the analysis of the entrances at an Indian toll station, considering Poisson arrivals and differentiated exponential service times (cars, light commercial, heavy commercial, multi-axle vehicles).

Ito (2005) used the software ARENA (systems modeling corporation, Rockwell Automation, Inc.) to analyze motorway gates in Japan and to identify solutions that reduce congestion and increase the efficiency of electronically controlled gates. Also Van Dijk et al. (1999) used ARENA, proposing a combined approach of queue theory and micro-simulation to determine the number of gates with different payment types. The hybrid analytical-simulative approach combines the conceptual framework of the queue theory with the computational potential of the simulations, allowing to limit the number of options and scenarios and to compare them in testing the complexity of real-life situations, as also in Van Dijk (2000). Kim (2011) also used a discrete state simulation model implemented with the ARENA software to analyze queue lengths and waiting times with variable traffic demand. The author highlights some gaps in the queues theory models, related to the fact that in real conditions many situations distant from the their base assumptions can occur, because of variations in arriving flows or non-homogeneous characteristics of the gates. Duhan et al. (2014) proposed an application of WinQSB software (Chang, 1998) for the analysis of a toll station in Northern India using a multiple channel model with Poisson arrivals and exponential service times. This software allows analyzing the performances with both the closed-form expressions from the queue theory and the results of discrete-state simulations. Magsino and Ho (2016) developed an intelligent highway tollgate queue selector using a stochastic event traffic model in MATLAB/Simulink for minimizing the queuing time of vehicles at the tollgates. They use a fuzzy logic-controlled queue (FLCQ) system, which takes into consideration the current lane density of the server and its service time to make the queue selection.

About micro-simulations with specifically implemented models, Astarita et al. (2001) proposed a micro-simulation model that represents the interactions between supply and demand. It uses car-following, lane-changing and response-to-traffic-control models, to show the performances of tollgates with mixed types of payment systems. The model allows to keep track of the behaviors variability among vehicles and drivers with different characteristics and of the effects related to the size and layout of the toll station. The latter possibility allows to represent spill-back situations and mutual interference of queues because of the limited capacity of manual gates, which can nullify the benefits deriving from introducing gate with high automation capacity levels.

Finally, the literature highlights some applications of commercial traffic micro-simulation software for the analysis of a motorway toll station. Ceballos et al. (2004) compared the results of the analytical queue models M/M/1 and M/M/s with what obtained using the VISSIM software (PTV). The authors highlight how the use of simulation models for toll collection facilities allows to resolve some shortcomings of the analytical approaches. These problems concern the significant deviations of the estimated queue parameters from the real-life situations. The authors advise to avoid single-channel or multiple serving channels analytical models with high levels of traffic demand, due to their asymptotic behavior. Also Obelheiro et al. (2011) used a micro-simulation model implemented in VISSIM to evaluate the performance of a toll station with variable gate number. They propose and test a method for level of service analysis, based on the user's quality perception. Aksoy et al. (2014) used VISSIM for the analysis of the toll facilities on the Turkish Fatih Sultan Mehmet Bridge in order to evaluate some management strategies and to calculate the optimal number of gates, with the aim of minimizing delays and maximizing the capacity of the system.

At the end of this literature review, it should be noted how the referenced researches generally consider, sometimes in comparison with simulations, the applications of the stationary queue theory for under-saturated gates, highlighting the non-reliability of these models for situations close to saturation. These limits are addressed and removed only by using simulative applications. An exception exists in Levinson and Chang (2003), in which the authors propose to calculate the waiting time at the gate using the Akçelik and Troutbeck (1991) formula. This formula is reported in the HCM (TRB, 2000) for the non-signalized intersections and, even if it makes possible to estimate queue parameters in non-stationary and over-saturated situations, is not rigorous and totally suitable for the case under discussion.

3. The stationary queue model

3.1. *Gate characterization for a motorway toll station*

As discussed in the previous sections, a motorway toll station can be modeled using mathematical tools derived from the probabilistic queue theory. All the gates of a motorway toll station can be modeled as a set of single serving channels, one for each gate (e.g., in the studies of Boronico and Siegel, 1998; Ceballos and Curtis, 2004; Cherng et al., 2005; Kim, 2009; Mehri and Djemel, 2011; Shanmugasundaram and Punitha, 2014; Wang, 2017; Woo and Hoel, 1991). In the space just upstream of the gate, the accumulation of vehicles occurs; vehicles will have to wait for ticket collection, toll payment through traditional methods (manual or automatic) or vehicle recognition and tracking (electronic collection).

A generic user arriving at the toll station chooses the lane which meets the requirements for the transaction type he/her wants to carry out. From this point of view, it is necessary to clarify some issues about the transaction types and the toll collection technologies available at a toll station. Kim (2009) classifies toll gates in provided with: manual operator (MO), automatic device (AD), electronic exaction (EE), a mix of technologies (MT), for example with a mix of MO/EE, AD/MO, AD/EE. Besides this, AD gates are further differentiable, because of automatic cash payment machines (CM) or differences in payment transactions by point of sale (POS) or credit card terminal (CC) (with or without PIN, with or without receipt). Without loss of generality, a typical layout on the Italian motorway network in Fig. 1 can represent a concrete example with dedicated EE gates, EE and AD mixed gates, mixed gates with AD for payment with CM/CC or MO.

At this point it is necessary to consider the relationship between the total traffic arriving at the toll station and the relative share arriving at each gate. As pointed out by Kim (2009), if the aim is to evaluate operations based on the collected data, then the actual transit data for each gate are available for the analysis. If the aim concerns, for example, a predictive analysis for the design of a new facility or a scenario analysis to test optimizing solutions, several studies have proposed different criteria and solutions for vehicles distribution, as discussed in the literature background.

This paper, however, does not propose specific models for the arrivals distribution between the individual gates—a further analysis that will concern the research continuation, but starting from the considerations of Chakroborty et al. (2016), Kim (2009), Boronico and Siegel (1998) and Zarrillo et al. (1997) the authors assume that an user who arrives at the toll plaza chooses the minimum-queue gate, compared to those with the required technologies. The authors also assume that the user, after choosing the gate, maintains the selection as conditioned by vehicles already queued in the adjoining gates and those that arrive in the following instants. Also according to Kim (2009), Boronico and Siegel (1998) and Zarrillo et al. (1997), under these conditions the percentage of arrivals by type of transaction is assumed as uniformly distributed over the gates that propose the same. If the analyst does not have the actual arrival figures for the gate, the arrival rate can be estimated by virtue of the presence of various technologies and the users' segmentation, assuming that equal gates see arriving uniformly distributed users' portions (Boronico and Siegel, 1998; Kim, 2009).

Fig. 2 shows the basic model for this type of problem. For the generic gate j of the station, the first space a vehicle can occupy immediately near the stop line is the service point (i.e., the service counter) where the tolling operations take place. In the further waiting positions behind the service counter, the actual queue line forms.

In a non-empty queue system, the number of vehicles in the system, $L_s^{(j)}$, is equal to the number of vehicles in the queue, $L_c^{(j)}$, plus the vehicle in the service counter. The waiting time for a vehicle at the gate j, $w^{(j)}$, is equal to the time in the queue, $d^{(j)}$, plus the service time, $s^{(j)}$ (i.e., the time that the vehicle spends in the first position to complete the payment operation).

If $\lambda^{(j)}$ is the average rate of vehicles arrivals and $\mu^{(j)}$ is the average rate of services (i.e., the reciprocal of the average service time $s^{(j)}$), the quotient $\lambda^{(j)}/\mu^{(j)}$ is the saturation degree, $\rho^{(j)}$. The value of $\rho^{(j)}$ indicates the degree of saturation of the gate: under-saturation, if $\rho^{(j)} < 1$; saturation, if $\rho^{(j)} = 1$; over-saturation, if $\rho^{(j)} > 1$.

If T is a certain time period sufficiently large and if $\lambda^{(j)}$ and $\mu^{(j)}$ are constant during the same period with $\rho^{(j)} < 1$, for the applications we can use the results of the probabilistic queue theory models for statistical equilibrium. These results are even more realistic the more the gate is far from saturation ($\rho^{(j)} \ll 1$). The probabilistic queue theory provides queues length and waiting times determinations rapidly tending to

Fig. 1 — Gates with different types and mix of payment options (Italian motorway system).

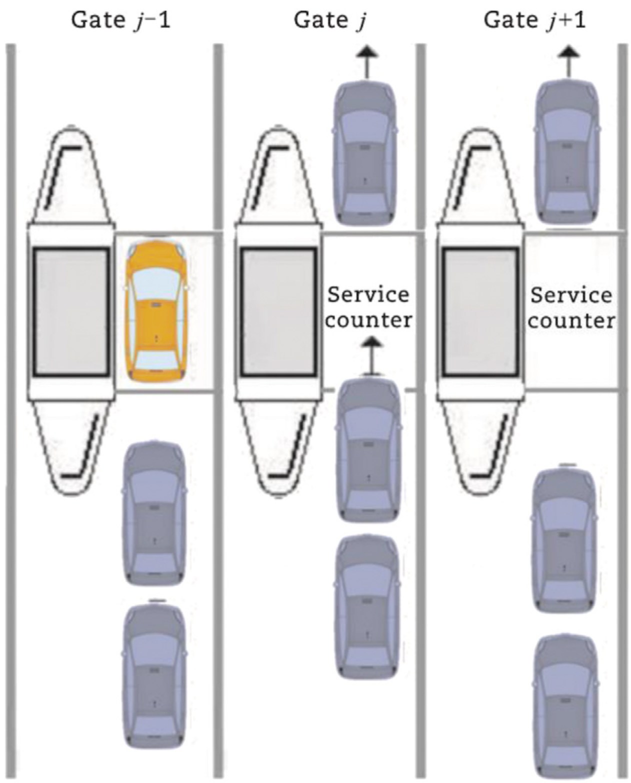


Fig. 2 – Gates layout at a motorway toll station.

infinity with $\rho^{(j)} \rightarrow 1$ starting from values of $\rho^{(j)}$ generally higher than 0.6–0.8, and therefore they appear unrealistic. So overall, the probabilistic theory does not apply if the gate is close to saturation, saturated or over-saturated, or with variable $\lambda^{(j)}$ and $\mu^{(j)}$ (Mauro, 2010; Mauro and Pompigna, 2020; Pompigna, 2020).

For a whole toll station with many gates, the available entry data may concern the total number of arrivals at the toll plaza and any segmentation of users based on the characteristics of the motorway traffic demand (e.g., percentage of EE users, percentage of propensity to use an AD rather than an MO toll system, percentage of Cash/Credit Card transactions, etc.). If F is the total flow rate arriving at the station, the generic gate j will be characterized by a flow rate $\lambda^{(j)}$ such that $F = \sum_j \lambda^{(j)}$. It appears that $\lambda^{(j)} = Fp_j$, with p_j the probability of using gate j. It is clear that p_j can be estimated on the basis of the users segmentation with respect to the operations that can be performed at the gate.

In light of this, the authors focus on the generic gate j assuming that the flow rate $\lambda^{(j)}$ and the relative segmentation, with respect to the vehicle classes that use the toll collection types available on the same, are known or estimated. For simplicity of notation, the parameters and status variables indexes related to the generic gate j are dropped and considered as implied.

3.2. *Arrivals and services processes*

As highlighted in Section 2, the models used to study the waiting phenomena at a gate of a motorway station are substantially M/M/1 or M/G/1 with first-in-queue-first-out-of-queue (FIFO) service discipline. In both cases, according to Kendal's notation, the arrivals process is Poisson with inter-arrival times distributed exponentially. Regarding service times, the literature presents various types of probability distributions, since the service rate depends on the operation carried out and on the vehicles type involved in each operation (Sangavi et al., 2017). As evidenced by empirical experience, EE systems have faster service times than non-EE systems, while MO modes express higher service times than any AD. From this point of view, the design of a motorway toll station or barrier involves optimization choices in the gates number and in mixing toll collection options (Kim, 2009).

For service time modeling related to different transaction types (with differences in payment methods and/or vehicle types) and in relation to mixed situations at each gate, different assumptions regarding the probability distribution of service time have been used in literature, such as: average (deterministic) service times for each type of service (Astarita et al., 2001; Sadoun, 2005); an average time for each type of service and vehicle class (Van Dijk et al., 1999); log-normal distribution for queuing and non-queuing vehicles for ticketing/toll-collection operations with MO/AD (Pratelli et al., 2001); triangular distributions for EE and non-EE systems (Ito, 2005); exponential distributions for cash and EE payments (Karsaman et al., 2014) and for parallel servers (Magsino and Ho, 2016); general probability distribution for the single gate with different and eventually mixed technologies (Kim, 2009); Gamma distribution with parameters dependent on the vehicle type (Woo and Hoel, 1991).

In the applications that consider the probabilistic theory of the queues, in general the probability distributions for the service times are exponential (e.g., in the studies of Boronico and Siegel, 1998; Chakroborty et al., 2016; Cherng et al., 2005; Karsaman et al., 2014; Liu et al., 2017; Wang, 2017) or general (e.g., in the studies of Kim, 2009; Mehri and Djemel, 2011). As reported by Kim (2009), the exponential distribution allows more compact expressions in the probabilistic treatment of the queue, but it has the drawback of having a high variance (equal to the square of service time, which could generate an overestimation of expectations as noted by Boronico and Siegel (1998)) and the absence of memory (which could not describe realistically some real-life situations, such as the cumulated working hours for a manual toll operator). The generic distribution G allows to operate in more general conditions, with a bit more complex expressions (PK equation).

3.3. *A multi-class queue model*

In relation to the inter-arrival and service times distributions, this paper proposes a multi-class model. According to the probabilistic methods for the queues, this model allows to treat the case of a tollgate serving users belonging to a defined number of classes. Each class follows its own inter-arrival and service time probability distribution (Gautam, 2012; Ravindran, 2007). This situation actually represents the most general case of a gate at a motorway toll station.

The users' disaggregation is by the type of transaction performed at the tollgate, considering the technology installed and the vehicles segmentation. As already specified, a tollgate can allow different types of operations, present also as a mix, which orient users' choice to wait for the service at the corresponding waiting lane. In these terms, service times depend on the operations that occur at the service counter, related with users' and technology heterogeneity and regarding the toll operation (ticket issuance, vehicle recognition, fixed toll payment, kilometric toll payment, etc.), the vehicle type (e.g., light or heavy vehicle), the presence or absence of the manual operator (MO, AD) and the allowed payment systems (e.g., EE, cash, cards, etc.). Depending on this characterization, the multi-class model is adequately flexible as it allows to consider several classes of users, each characterized by its specific parameters in the probability distributions that regulate the processes of arrivals and services.

In this paper the authors consider users belonging to R classes, variously composed taking into account the segmentation, with i being the generic class belonging to the set $\{1, 2, \cdots, R\}$. In the multi-class model the users belonging to the generic class i ($i \in \{1, 2, \cdots, R\}$) arrive at the tollgate with independent and identically distributed arrival times according to a class-specific probability distribution. In addition, users belonging to the generic class i ($i \in \{1, 2, \cdots, R\}$) are served respecting the FIFO discipline, which require independent and identically distributed service times according to a class-specific probability distribution to perform the operations at the service counter and to leave the system upon completion of the service.

3.4. The multi-class $M/\Gamma/1$ model for the single gate

The multi-class model (Gautam, 2012; Ravindran, 2007) proposed in this research is an M/Γ/1 model, i.e., with a Poisson distribution for arrivals (e.g., in the studies of Boronico and Siegel, 1998; Chakroborty et al., 2016; Cherng et al., 2005; Karsaman et al., 2014; Liu et al., 2017; Wang, 2017) and a Gamma distribution for service times (Mehri and Djemel, 2011).

For the R classes, the authors assume a Poisson process with parameter λ_i for the class i ($i \in \{1, 2, \cdots, R\}$). For the service times s_i of each class i, to be considered as independent and identically distributed variables, it is convenient to hypothesize a Gamma distribution with a shape parameter α_i, a scale parameter β_i and a shift parameter δ_i. It turns out that $S_i \sim \Gamma_i(\alpha_i, \beta_i, \delta_i)$, i.e., a shifted Gamma distribution, having a probability density function expressed by

$$f(s_i) = \frac{1}{\Gamma(\alpha_i)\beta_i^{\alpha_i}}(s_i - \delta_i)^{\alpha_i - 1} e^{-\frac{s_i - \delta_i}{\beta_i}} \tag{1}$$

where $\Gamma(\alpha_i)$ is the Euler's Gamma function.

The convenience of using a Γ distribution with known mean and variance consists in the possibility to express its probability function and to calculate, at least numerically, the relative quantile distribution, making possible simulation scenarios as will be clarified in the following sections of the paper.

The shifted Gamma distribution is obtained by translating a Gamma distribution of shape parameter α_i and scale parameter β_i. This shift allows to consider a minimum time of service δ_i for each class i. The expected value for $\Gamma_i(\alpha_i, \beta_i, \delta_i)$ is

$$E[S_i] = 1/\mu_i = \delta_i + \alpha_i\beta_i \tag{2}$$

which is shifted of δ_i if compared with the expected value for $\Gamma_i(\alpha_i, \beta_i)$. The variance for $\Gamma_i(\alpha_i, \beta_i, \delta_i)$ is

$$VAR[S_i] = \alpha_i\beta_i^2 = \frac{(1 - \mu_i\delta_i)^2}{\mu_i^2\alpha_i} \tag{3}$$

which is coincident with the variance for $\Gamma_i(\alpha_i, \beta_i)$.

It is necessary to remember that for an M/G/1 model (with Poisson arrivals at rate λ and service times with a general probability distribution with mean $1/\mu$ and variance σ^2), the probabilistic theory of waiting phenomena allows to determine the average waiting time in the system for the queue in statistical equilibrium (considering a sufficiently long period of time T and a condition of under-saturation of the gate with $\rho = \lambda/\mu < 1$) through the well known PK relationship.

$$w = \frac{1}{\mu} + C\frac{\lambda/\mu^2}{1 - \lambda/\mu} \tag{4}$$

$$C = \frac{1}{2}\left(1 + \mu^2\sigma^2\right) \tag{5}$$

In addition, according to Little's law, for the average number of users in the system it results that

$$L_s = \lambda w = \frac{\lambda}{\mu} + C\frac{\lambda^2/\mu^2}{1 - \lambda/\mu} \tag{6}$$

The Eqs. (4)–(6) are valid for a general service time distribution G. If the distribution is a shifted Gamma $\Gamma(\alpha, \beta, \delta)$, the average waiting time in the system and the average number of vehicles in the queue system in statistical equilibrium must consider the specific value of C.

$$C = \frac{1}{2}\left[1 + \frac{(1 - \mu\delta)^2}{\alpha}\right] \tag{7}$$

If we consider a multi-class M/Γ/1 model with a FIFO discipline (i.e., no class receives preferential treatment), Little's law is valid in relation to each class i($i \in \{1, 2, \cdots, R\}$). Being $L_{s,i}$ the average number of users in the system and w_i the average waiting time in the system for the single class $i \in [1, R]$, it results that $L_{s,i} = \lambda_i w_i$. Similar results can also be obtained for the average number of queued users, $L_{c,i}$, and for the average waiting time in the queue, d_i. In this way, each class can be treated as a separate system with a rate of arrivals λ_i, a rate of departures μ_i and a degree of saturation ρ_i.

The R classes can be aggregated into a unique Poisson process of arrivals with rate $\lambda = \lambda_1 + \lambda_2 + \cdots + \lambda_R$. If L_s and w are respectively the average number of users in the system and the average waiting time in the system considering all the classes $i \in [1, R]$, it results that $L_s = L_{s,1} + L_{s,2} + \cdots + L_{s,R}$ and $w = L_s/\lambda$ (Gautam, 2012; Ravindran, 2007).

Whereas for each class of users i the random variable service time S_i is distributed according to $\Gamma_i(\alpha_i, \beta_i, \delta_i)$, the actual service time of an arbitrary user without class specification is a random variable S with a mixture Gamma probability distribution. This distribution is obtained as the weighted average of the individual distributions per class, considering the probability that the arbitrary user can belong to each class:

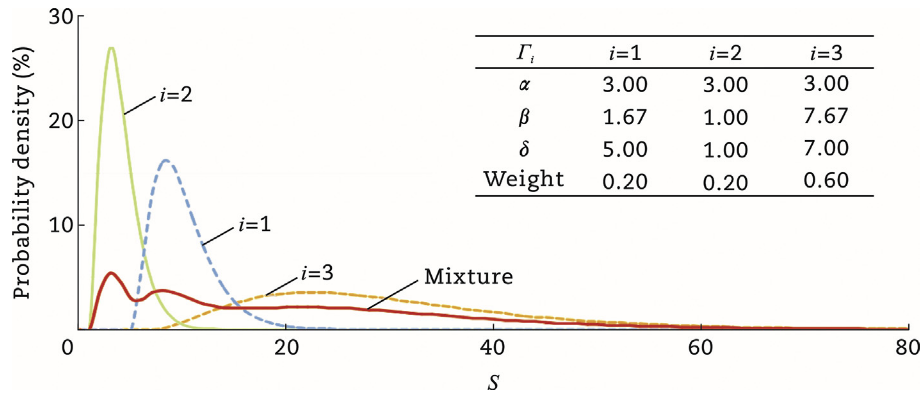


Fig. 3 – An example of the probability density for a Gamma mixture distribution deriving from three Γ_i distributions.

$$\Gamma(T)=\mathrm{Prob}(S\le T)=\frac{1}{\lambda}\sum_{i=1}^{R}\lambda_i\Gamma_i(T) \tag{8}$$

With purely illustrative purposes, Fig. 3 shows an example of mixture Gamma probability density distribution, considering $\Gamma_i(\alpha_i,\ \beta_i,\ \delta_i)$ for $i=1,\ 2,\ 3$, with relative weights and parameters in the embedded table.

The PK relationship for a multi-class M/Γ/1 can be obtained by identifying the value for C, based on the expectation of S

$$\mathrm{E}[S]=\frac{1}{\lambda}\sum_{i=1}^{R}\lambda_i\mathrm{E}[S_i]=\frac{1}{\lambda}\sum_{i=1}^{R}\lambda_i\frac{1}{\mu_i}=\frac{1}{\lambda}\sum_{i=1}^{R}\lambda_i(\delta_i+\alpha_i\beta_i) \tag{9}$$

and the second moment for S

$$\mathrm{E}\left[S^2\right]=(\mathrm{E}[S])^2+\mathrm{VAR}[S]=\frac{1}{\lambda}\sum_{i=1}^{R}\lambda_i\mathrm{E}\left[S_i^2\right] \tag{10}$$

and replacing in Eq. (5).

Being $\mathrm{E}[S_i^2] = (\mathrm{E}[S_i])^2 + \mathrm{VAR}[S_i] = (\delta_i+\alpha_i\beta_i)^2 + \alpha_i\beta_i^2$, and knowing the vector $(\alpha_i,\ \beta_i,\ \delta_i)$ for each class i $(i\in\{1,\ 2,\ \cdots,\ R\})$, the second moment for S_i is known and then, by Eq. (10), also the second moment for S. Being $\mathrm{VAR}[S] = \mathrm{E}[S^2] - (\mathrm{E}[s])^2$, the value for C results as

$$C=\frac{1}{2}\left(1+\frac{\mathrm{VAR}[S]}{(\mathrm{E}[S])^2}\right)=\frac{1}{2}\frac{\mathrm{E}\left[S^2\right]}{(\mathrm{E}[S])^2} \tag{11}$$

According to Eq. (11), for the queue in statistical equilibrium, for $\rho = \lambda\mathrm{E}[S]<1$, it results as

$$w=\mathrm{E}[S]+\frac{1}{2}\frac{\lambda\mathrm{E}\left[S^2\right]}{1-\lambda\mathrm{E}[S]} \tag{12}$$

$$L_s=\lambda\mathrm{E}[S]+\frac{1}{2}\frac{\lambda^2\mathrm{E}\left[S^2\right]}{1-\lambda\mathrm{E}[S]} \tag{13}$$

4. The time-dependent model for the non-stationary queue

4.1. *The coordinate transformation method*

The solutions provided by Eqs. (12) and (13) according to the probabilistic approach are valid under statistical equilibrium conditions, which however are only partially achievable for a system operating under real-life conditions (Mauro, 2010). The two equations produce acceptable approximations only if T is sufficiently large and if the gate is under-saturated (i.e., $\rho = \lambda\mathrm{E}[S]<1$). Moreover, the results are realistic if the gate is far from saturation ($\rho\ll 1$). As already highlighted, the probabilistic queue theory provides queue lengths and waiting times determinations tending rapidly to infinity for $\rho\to 1$, starting from $\rho>[0.6, 0.8]$ (Pompigna, 2020).

With a saturated or over-saturated gate, respectively with $\rho = \lambda\mathrm{E}[S]=1$ or $\rho = \lambda\mathrm{E}[S]>1$, the results of the probabilistic queue theory in conditions of statistical equilibrium are not used. On the other hand, in these two cases, the use of probabilistic queue theory for conditions other than statistical equilibrium leads to excessively complex results for practical applications. In relation to the case of over-saturated gate, traffic engineers can resort, however, to the deterministic theory of waiting phenomena, which treats traffic as a continuous fluid (May and Keller, 1967), with results that are more reliable the more the gate is over-saturated ($\rho\gg 1$).

Ultimately, it can be summarized that for a gate that is sub-saturated but close to saturation, saturated or over-saturated with saturation degree ρ not significantly greater than 1, the probabilistic theory of the queues is not used in the first case, nor the deterministic one in the other two cases.

Furthermore, if λ and E[S] vary in time, whatever the state of the gate (i.e., values for ρ), the results of the equilibrium of the queues in the non-equilibrium conditions are not useful in the technical practice (Mauro, 2010). In conditions of over-saturation and for any type of variation of λ and E[S] in T, the deterministic theory of the queues can be applied providing more reliable results as the gate becomes over-saturated ($\rho\gg 1$).

To have a unified approach in dealing with all the saturation cases, the so-called "time-dependent" queue models have long been developed in traffic engineering. These models allow us to combine the solutions provided by the stationary probabilistic theory with those of the deterministic theory. With a "time-dependent" model, the probabilistic model for $\rho\ll 1$ is best approximated, tending asymptotically to the deterministic model for $\rho\gg 1$. As highlighted in Section 2, among the reviewed studies there are no time-dependent applications of the waiting phenomena for motorway toll gates, except in Levinson and Chang (2003). These authors, in fact, use the Akçelik and Troutbeck (1991) formula, as

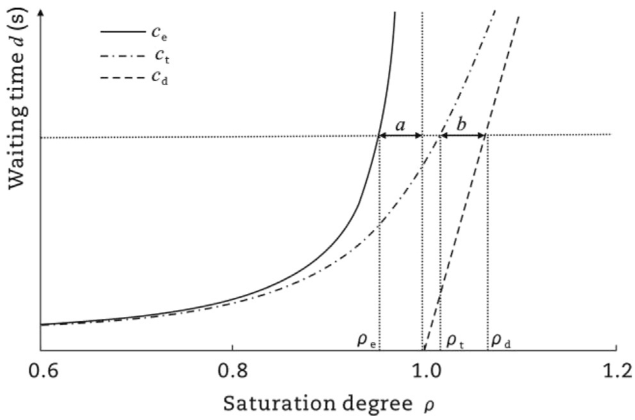


Fig. 4 – Transitional curve for waiting time inferred with coordinate transformation with respect to ρ between deterministic and probabilistic solid cases.

reported in the HCM (TRB, 2000) for un-signalized intersections.

The method behind time-dependent models considered in this paper is based on the coordinate transformation criterion, developed at the Transportation Research Laboratory for the TRANSYT program (Robertson, 1979). This is subsequently extended by Kimber and Hollis (1979). A similar approach is due to Doherty (1977), extended and generalized by Catling (1977). Over the years, the analysis of waiting phenomena by means of the coordinate transformation method has found numerous applications, especially in the queue assessment at intersections (e.g., Akçelik and Troutbeck, 1991; Brilon, 1995, 2007, 2008, 2015, Brilon et al., 1997; Cvitanić et al., 2007; Heidemann, 2002; Heydecker and Verlander, 1998; Mauro, 2010; Mauro and Pompigna, 2020; Troutbeck and Brilon, 2000; Wong et al., 2003; Wu, 1998). Among the studies reviewed, however, there are not applications of the method in analyzing the waiting phenomena at the gates of a motorway toll station.

It should be noted that the coordinate transformation method is able to provide closed-form equations for the state variables of a time-dependent queue system. Other methods for the analysis of time-dependent queuing problems, such as the PSSFA method (Hu et al., 2018; Wang et al., 1996), come to the solution of the problem in a less direct way. The same PSSFA, in fact, involves the integration of differential equations for the queue length and the waiting time, which must necessarily be performed by applying numerical methods.

In general, the coordinate transformation method makes it possible to determine, for the length of the queue or the waiting time (dependent variables), the asymptotic transition curve c_t from the probabilistic model c_e and the deterministic one c_d by operating a coordinate transformation with respect to the saturation degree ρ (independent variable). The Kimber and Hollis (1979) model is the first in temporal order among the time-dependent models obtained with the criterion of the coordinates transformation.

With reference to Fig. 4, the transition curve c_t can be obtained by imposing the conditions $a = b$ in relation to the curve c_e that provides the solutions for the statistical equilibrium queue system (steady-state queue) and to the asymptote c_d that describes the system states in a deterministic way.

With reference to Fig. 4, the transition curve c_t can be obtained by imposing the conditions $a = b$ in relation to the curve c_e that provides the solutions for the statistical equilibrium queue system (steady-state queue) and to the asymptote c_d that describes the system states in a deterministic way. Fixing a generic value of the state variable (in Fig. 4 the waiting time d in the queue), a is the distance between the fixed value on c_e and the vertical asymptote for $\rho = 1$, and b is the distance between the curve c_t (that is to be determined) and the half-line representing the oblique asymptote of the curve c_d. If ρ_e, ρ_d and ρ_t are the values of the independent variable ρ uniquely identifiable by the fixed value of the state parameter on c_e, c_d and c_t, it results that $a = 1 - \rho_e$ and $b = \rho_d - \rho_t$. The equation for the coordinate transformation regarding ρ is

$$\rho_e = \rho_t - (\rho_d - 1) \tag{14}$$

The coordinate transformation allows to manage all the saturation conditions for the gate, providing the status of the system in terms of number of vehicles and waiting times in the system experienced during a specified observation period with constant average demand and capacity. The method, if appropriately applied in succession over several consecutive time slices T_k, with the constraint that the final queue $L_{cT,k}$ of each slice T_k represents the initial queue $L_{c0,k+1}$ of the following T_{k+1}, allows to approximate and process any demand and capacity profile, evaluating the evolving characteristics of queues and waiting times accordingly (Mauro and Pompigna, 2020).

4.2. *Time-dependent solutions of the multi-class M/Γ/1 queue with coordinate transformation*

The model can be inferred by imposing the conditions $a = b$ for a system in statistical equilibrium with a multi-class M/Γ/1 queue and for a deterministic curve, to which the transformed curve tends, of a multi-class D/D/1 model.

As regards the M/Γ/1 multi-class model, the arrivals distribution probability can be assumed according to a single Poisson process with a rate $\lambda = \lambda_1 + \lambda_2 + \cdots + \lambda_R$ and the distribution of service times for the arbitrary user of the mixture Γ type, obtained as the average of the single class-specific distributions, weighed considering the probability that the same arbitrary user may belong to each class, with E[S] and E[S^2] first and second moments, respectively. For this model, the PK formulas are valid for the expected value of the number of vehicles in the system L_s and for the expected value of the waiting time in the system w (Eqs. (12) and (13)).

For the deterministic model multi-class D/D/1 model, it results that

$$L_{sT} = \max\left\{0, L_{s0} + \left(\lambda - \frac{1}{\mathrm{E[S]}}\right)T\right\} \tag{15}$$

$$w_T = \max\left\{0,\ \frac{1}{\mu}\left[L_{s0} + 1 + \left(\lambda - \frac{1}{\mathrm{E[S]}}\right)\right]\ \frac{T}{2}\right\} \tag{16}$$

Considering that λ_i and $(\alpha_i, \beta_i, \delta_i)$ are constant during T, and that at the beginning of the same period an initial queue L_{s0} is observed, the application of the coordinate transformation method to Eq. (13) and to Eq. (15) using the condition in Eq. (14) gives the following approximate formula for the number of users in the system at the end of the period T.

$$L_{sT} = \frac{1}{2}\left(\sqrt{A^2 + B} - A\right) \tag{17}$$

$$A = \frac{\left(\frac{1}{E[S]} - \lambda\right)\frac{1}{E[S]}T^2 + (1 - L_{s0})\frac{1}{E[S]}T - 2(1 - C)(L_{s0} + \lambda T)}{\frac{1}{E[S]}T + 1 - C} \tag{18}$$

$$B = \frac{4(L_{s0} + \lambda T)\left[\frac{1}{E[S]}T - (1 - C)(L_{s0} + \lambda T)\right]}{\frac{1}{E[S]}T + 1 - C} \tag{19}$$

$$C = \frac{1}{2}\frac{E\left[S^2\right]}{(E[S])^2} \tag{20}$$

Similarly, the application of the coordinate transformation method to Eqs. (12) and (16) obtained by imposing the Eq. (14) gives the following approximate formula for the average waiting time in the system during the period T.

$$w_T = \frac{1}{2}\left(\sqrt{J^2 + K} - J\right) \tag{21}$$

$$J = \frac{T}{2}(1 - \lambda E[S]) - E[S](L_{s0} - C + 2) \tag{22}$$

$$K = 4E[S]\left[\frac{T}{2}(1 - \lambda E[S]) + \frac{1}{2}E[S]\lambda TC - E[S](L_{s0} + 1)(1 + C)\right] \tag{23}$$

where C can be obtained using the Eq. (20).

It can be observed that, if Γ_i is replaced by a single distribution G with expected value $1/\mu$ and variance σ^2, the time-dependent Eqs. (16)–(21) coincide with what is reported in (Heidemann, 2002). It is worth pointing out that the Γ_i functions can be calibrated in the relevant parameters $(\alpha_i, \beta_i, \delta_i)$ based on real-life data, directly monitored at a motorway tollgate. A calibration of these parameters with real data samples can be obtained by resorting to maximum likelihood estimates (MLE) (Minka, 2002; Shanker et al., 2016). The appendix of this paper contains methodological indications for the calibration of the parameters and a concrete example of analysis on real data at a tollgate.

5. Stochastic discrete event simulation

Stochastic discrete event simulations (SDES) offer useful tools for testing the results of waiting phenomena at a motorway tollgate, as shown in Section 2. SDES approach allows to simulate the evolution of the state variables in a waiting system by randomly generating the arrivals at the tollgate and the departures from the service counter by means of appropriate probability distribution functions.

In this specific case SDES represents a control model, that is a verification and validation tool for the time-dependent equations obtained with the coordinates transformation.

In the simulation model prepared for this research, the discrete states coincide with the arrivals at the gate. For the instant at which an arrival occurs, the model estimates the instant in which each arrived vehicle leaves the system, taking into account the other waiting vehicles already queued and the service time distribution for all vehicles in system. In this case the queue model is a multi-class $M/\Gamma/1$ with users divided into R classes and served according to the FIFO discipline. The users belonging to the generic class $i\ (i \in \{1, 2, \cdots, R\})$ arrive at the gate according to a Poisson process with inter-arrival intervals with mean $1/\lambda_i$, resulting R independent and identically distributed arrival processes. The service times s_i are also independent and identically distributed according to R shifted Gamma distributions S_i with parameters $(\alpha_i, \beta_i, \delta_i)$, mean $E[S_i] = 1/\mu_i = \delta_i + \alpha_i\beta_i$ and variance $VAR[S_i] = \alpha_i\beta_i^2$.

The state-to-state evolution of the system is represented by the arrival of a new vehicle v at the gate. For each new vehicle v, the inter-arrival interval time with respect to the previously arrived vehicle at the gate is generated randomly. Within a given period of observation T, the model simulates the random extraction of an exponentially distributed variable with a parameter $\lambda = \sum_i \lambda_i$ with constant λ_i. In this way, the model assigns the class $i\ (i \in \{1, 2, \cdots, R\})$ at the arriving vehicle, according with the probability λ_i/λ. Referring to the assigned class i, the corresponding service time is identified by simulating the extraction of a random variable with a $\Gamma_i(\alpha_i, \beta_i, \delta_i)$ probability distribution.

By evaluating the service end-time in the previous state of the system (i.e., the instant in which the previous vehicle disengages the service counter) and taking into account both the arrival instant and the service time (as a function of the class i) for the current vehicle, the instant of the service end for the same vehicle (i.e., the instant in which it leaves the system) can be obtained. Operating in this way, given the values of λ_i and $(\alpha_i, \beta_i, \delta_i)$ constant during T, the arrival and departure times for each vehicle can be simulated. These allow to obtain the trend of the average number of vehicles in the queue and in the system and of the average waiting time in the system. Extending the analysis at several intervals in consideration of time-dependent profiles, the evolution of the waiting phenomenon in terms of vehicles and times can be simulated.

It should be emphasized that, in order to obtain statistically significant results, the simulation process must be reiterated a sufficiently large number of times, so that in every system state (and therefore for each vehicle arrived) all the probability distributions are respected asymptotically (i.e., the class membership, the inter-arrival intervals and service times). Each alternative evolution of the waiting system, represented by a different set of arrival and departure times, constitutes a single Monte Carlo iteration or a trial of the simulation. The values of λ_i and $(\alpha_i, \beta_i, \delta_i)$, constant during T but possibly variables on sequential intervals, are considered common to all the trials of the simulation.

For an $M/\Gamma/1$ multi-class system, the inter-arrival time and the service time are simulated considering the quantile

Table 1 – Exponential and shifted Gamma distribution parameters depending on users' classes for the three test cases.

Class i	Vehicle type	Payment type	λ_i/λ (%)	$E[S_i]$	δ_i	α_i	β_i	$VAR[S_i]$	$E[S_i^2]$
1	Light	Coin	7	23	5	3	6.00	108.00	637.00
2	Light	Card	10	20	5	3	5.00	75.00	475.00
3	Light	Electronic	53	4	2	3	0.67	1.33	17.33
4	Heavy	Coin	3	30	5	3	8.33	208.33	1108.33
5	Heavy	Card	7	28	5	3	7.67	176.33	960.33
6	Heavy	Electronic	20	4	2	3	0.67	1.33	17.33

functions (inverse of the cumulative probability functions) respectively of the exponential distribution and of the shifted Γ_i distribution for the class i. For the generic vehicle v, given the properties of the exponential distribution, the inter-arrival time with respect to the previous vehicle can be calculated in closed form with

$$x_v = -1/\lambda \ln(1-u) \tag{24}$$

where u is a uniformly distributed variable with values between 0 and 1 ($u \in U(0, 1)$).

The vehicle v can therefore be assigned to a specific class based on λ_i/λ probabilities with $i \in \{1, 2, \cdots, R\}$, representing the proportion of vehicles belonging to each class. The same class can be assigned using the sampled values for $u \in U(0, 1)$.

Regarding the simulation of the service time related to each class i, which are distributed according to Γ_i (α_i, β_i, δ_i), it should be noted that for the Gamma probability distribution it is not possible to express the quantile function by closed-form expressions and therefore it is necessary to operate using numerical approximation methods (Devroye, 1986; Hörmann et al., 2004). The most common calculation software with statistical applications contain specific functions that allow for the function inversion, using support variables distributed uniformly between 0 and 1. Alternatively, if it is necessary to proceed without a specific function for the Gamma probability distribution inversion, the algorithm of Marsaglia and Tsang (2000) works well for the purpose.

Regarding the model validation, the average values of a high number of Monte Carlo iterations (averages on 1800 simulations) of the SDES model were compared with the results of the probabilistic theory for stationary queue in under-saturated situations. For this validation task, an M/M/1 model was used as a reference, assuming parameters for Γ_i as to replicate a single exponential distribution of service time S with $E[S] = 1/\mu$ (i.e., $\alpha_i = 1$, $\beta_i = \frac{1}{\mu}$, $\delta_i = 0$). A further validation task involved the comparison with the results of an M/G/1 model obtained by applying the PK formulas (once the average and variance of a single generic distribution of service times has been set).

It is worth pointing out that SDES model calibration can be obtained by calibrating the values of λ_i and of the parameters (α_i, β_i, δ_i) for each Γ_i on the basis of the real data. For these distributions, as already mentioned, we can use MLE (Minka, 2002; Shanker et al., 2016). However, in the numerical applications of the following Section 6, these parameters are supposed to be known and differentiated for each users' classes. As said previously, the appendix of this paper contains methodological indications for the calibration of the parameters and a concrete example of analysis on real service time data at a tollgate.

6. Test case: results and discussion

Three test cases have been identified to verify the results of the time-dependent model, making possible to compare the output of the Eqs. (16)–(21) and what is obtained with the SDES model. As already mentioned, the SDES model must be reiterated for a sufficiently high number of times (i.e., trials) using a Monte Carlo procedure to obtain statistically significant results. Actually, only when we have the queue results for a large number of SDES iterations, their average values can be compared with the results of time-dependent equations. For the analysis of the test cases, a tollgate that includes electronic and automatic toll collection system (in turn with card or coin) for light (namely, cars) and heavy vehicles (namely, trucks or buses) is considered. A specific class i is associated to each of the 6 combinations of collection and vehicle types in Table 1.

Table 1 also shows the λ_i/λ probability of classes occurrence with respect to the flow rate λ arriving at the gate, and the parameters of the Γ_i probability distribution with the relative expected value, second moment and variance. Considering values in Table 1 as constants, Figs. 5 and 6 show the trends of L_{sT} and w varying λ and ρ, which are obtained by considering the time-dependent Eqs. (17)–(23) with a zero initial queue. These trends are compared with PK results (exclusively in their validity range given by $\rho < 100\%$), showing a substantial coincidence for values of $\rho < 70\%$. For higher values of ρ, the PK determinations are rapidly and unrealistically tending to infinity for $\rho \to 1$, with substantial differences with respect to the time-dependent model results.

Furthermore, in Figs. 5 and 6, the results of time-dependent model are compared with the output of the SDES model obtained for 1800 Monte Carlo iterations. From the analysis, always considering a zero initial queue, an overlap of the results of the time-dependent model emerges with respect to what can be obtained using the SDES model iterations, synthesized by the empirical average values on 0.01 – width binning intervals for ρ. Further numerical evaluations are carried out considering an evolutionary traffic demand. Table 2 describes traffic demand composition related to the 6 toll collection classes (i.e., 2 vehicle types and 3 types of payment) during eight intervals, each lasting 15 min. Table 1 shows again the parameters of the probability distributions functions Γ_i.

Taking into account the demand distribution during each time slice in Table 2, as percentages of traffic for each of the 6 vehicle and payment combinations with respect to the characterization of the entire vehicle fleet, three test cases

Please cite this article as: Pompigna, A., Mauro, R., A multi-class time-dependent model for the analysis of waiting phenomena at a motorway tollgate, Journal of Traffic and Transportation Engineering (English Edition), https://doi.org/10.1016/j.jtte.2020.09.001

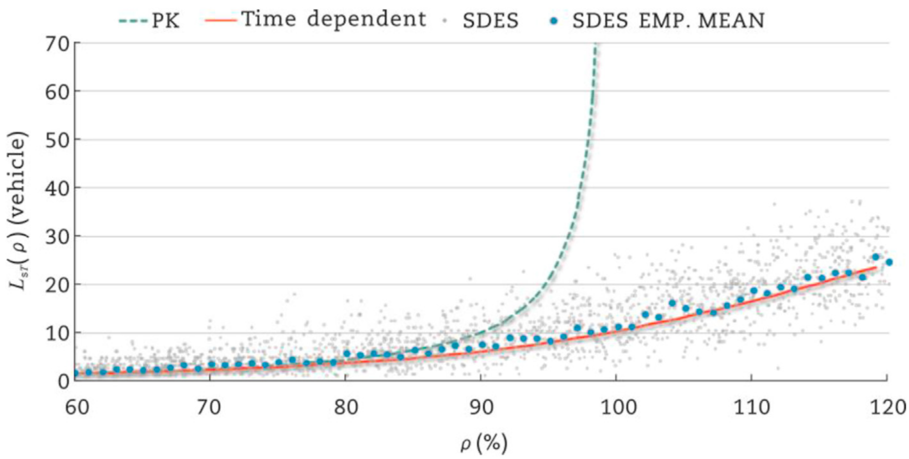


Fig. 5 – L_{sT} at the end of each interval according to the PK formulas, the time-dependent model and the SDES trials ($L_{s0} = 0$, varying ρ).

A, B and C were considered. These three test cases differ each other for the trend of arrivals during the 8 time intervals. Case A and case B have a peak demand in the initial portion of the period (i.e., third period). In particular, case A has a more pronounced traffic peak than case B. In case C, the peak demand is comparable with case A, but this is located in a more central position with respect to the analyzed time intervals (fourth and fifth periods).

Table 3 shows the total arrivals at the gate during each time slice for the three test cases. The demand composition percentage in Table 2 are applied to these total arrivals. Table 3 also shows the trend of the parameters of the multi-class queue for each of the three test cases A, B and C. The evolutionary profiles of ρ, L_{sT} and w are obtained using the multi-class queue formulas, the respective traffic demand shown in the same table and the percentage distribution of the arrivals over the 6 toll collection classes in Table 2, given the relative distributions $\Gamma_i(\alpha_i, \beta_i, \delta_i)$ for service times in Table 1.

The values obtained for ρ in each time slice of each test case are shown graphically in Fig. 7. The ρ trends show periods of over-saturation in all the three cases, corresponding to traffic peaks. For the test case A there is an over-saturation of 122% in the third interval; in the same interval also case B reaches saturation, with ρ equals to 100%. For case C, instead, a maximum over-saturation (with ρ equals to 109%, intermediate between cases A and B) is placed in the fourth and fifth interval.

Each of the three test case was simulated using the SDES model in Section 4 considering 600 trials. Starting from the results obtained for each trial, the mean and the 95% confidence interval were evaluated for L_{sT} and w for each 15-min interval, allowing to evaluate the time profiles of the averages queue length and waiting time and their confidence intervals reported in Table 4.

The evolutionary profiles for L_{sT} and w obtained separately with the time-dependent model and the SDESs for each scenario can be compared, in order to evaluate the capacity of the time-dependent multi-class queue model to fit the SDES results. The results are shown graphically in Figs. 8–10.

The comparisons for all the three numerical test cases highlight the ability of the multi-class time-dependent queue model to fit the trend of the average queue and waiting time values obtained by SDES. Table 5 shows the determination coefficient R^2 and the mean absolute percentage error (MAPE) as indicators of the goodness of fit. The values of the determination coefficient R^2 obtained for each scenario are

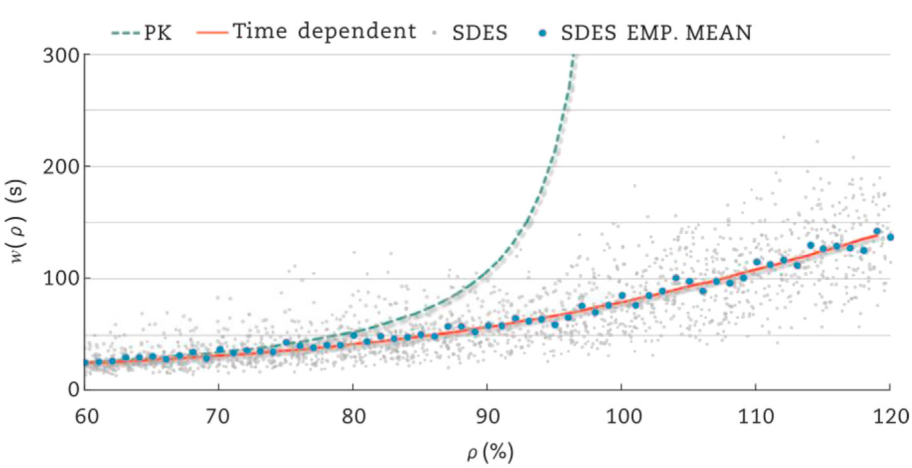


Fig. 6 – Average w in each interval according to the PK formulas, the time-dependent model and the SDES trials ($L_{s0} = 0$, varying ρ).

Table 2 – Demand distribution with respect to the users' classes for each time slices (15-min intervals).

T-slice k	Class i	Vehicle type	Payment type	λ_i/λ(%)
1	1	Light	Coin	7
	2	Light	Card	10
	3	Light	Electronic	53
	4	Heavy	Coin	3
	5	Heavy	Card	7
	6	Heavy	Electronic	20
2	1	Light	Coin	8
	2	Light	Card	9
	3	Light	Electronic	50
	4	Heavy	Coin	6
	5	Heavy	Card	7
	6	Heavy	Electronic	20
3	1	Light	Coin	7
	2	Light	Card	12
	3	Light	Electronic	48
	4	Heavy	Coin	3
	5	Heavy	Card	8
	6	Heavy	Electronic	22
4	1	Light	Coin	4
	2	Light	Card	15
	3	Light	Electronic	58
	4	Heavy	Coin	2
	5	Heavy	Card	5
	6	Heavy	Electronic	16
5	1	Light	Coin	5
	2	Light	Card	14
	3	Light	Electronic	58
	4	Heavy	Coin	1
	5	Heavy	Card	6
	6	Heavy	Electronic	16
6	1	Light	Coin	1
	2	Light	Card	6
	3	Light	Electronic	60
	4	Heavy	Coin	3
	5	Heavy	Card	10
	6	Heavy	Electronic	20
7	1	Light	Coin	5
	2	Light	Card	7
	3	Light	Electronic	58
	4	Heavy	Coin	2
	5	Heavy	Card	11
	6	Heavy	Electronic	17
8	1	Light	Coin	7
	2	Light	Card	10
	3	Light	Electronic	53
	4	Heavy	Coin	3
	5	Heavy	Card	7
	6	Heavy	Electronic	20

extremely close to 100%, confirming the capacity of the time-dependent model to reproduce optimally the variability of the values of L_{sT} and w obtained with simulations. Also MAPE shows good level of approximation, ranging levels between 7% and 12%.

7. An insight for a future scenario

In recent years the rise of new technologies has made great progress in intelligent transportation systems (ITS) directing more and more attention towards connected vehicles (CVs).

CVs are equipped with advanced communication technologies that allow an exchange of information between the various elements of the transport system, configuring what is generically identified with vehicle to everything, or V2X. V2X network connection is actually specified with respect to the nature of the relationship between the vehicle and the outside world, including vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), vehicle-to-people (V2P), vehicle-to-network (V2N). The new V2X technologies, even mixed with non-CVs, will awaken a growing interest as usher in new operation models, change in traffic flow fundamentals and management, and redesign safety and mobility management (Mahmassani, 2016; Mostafizi et al., 2017). To this technological revolution, that will be increasingly dominant in future years, we must surely add a progressive increase in card/electronic payments. This section of the paper offers a glance to these aspects, to be specifically treated in a more thorough way in the continuation of the research. It is useful to anticipate here some considerations regarding the use of the formulas presented in this paper for the treatment of these aspects, which will characterize future scenarios.

Regarding the second topic, the proposed time-dependent model presents an adequate flexibility to consider the aspects related to the increase of automatic technologies users (card or electronic toll) or to the introduction of new payment technologies (e.g., license plate recognition (LPR)). The proposed time-dependent multi-class queue model makes it possible to represent every type of payment and any variation in its use, once the parameters of the relative probability distributions of arrivals and service times have been calibrated.

Going back to the first topic, it is appropriate to distinguish between two policies for a mixed fleet of CVs and non-CVs that can be implemented by the motorway concessionaire: I) the "designated policy" under which the gate is designated to CVs; II) the "integrated policy" under which CVs and non-CVs can use the same gate (Mirzaeian et al., 2018).

In the first case, with only CVs at a tollgate in fully V2X environment, the queue model appears as totally deterministic. In this case, in fact, the arrivals at the gate are deterministic, as they are temporally predetermined according to a scheduling program within a certain time interval. For these non-human controlled vehicles, even the service times are totally deterministic as the toll operations are conducted with fully automated procedures. In this case, the queue model is D/D/1, with arrivals and departures managed by the CVs Control System. In this case, i.e., under a "designated policy", the queue parameters may be calculated according with the Eqs. (13) and (14).

In the second case, i.e., under an "integrated policy", traffic demand is a mix of non-CVs and CVs. Human and non-human controlled vehicles alternate with each other in arriving at the toll plaza by composing into batches or platoons. Using multi-class queue model, we can consider each platoon of CVs as a single vehicle belonging to the CVs class. For reason of simplicity, the hypothesis of a fixed length of the CVs platoon can be assumed that is a single class of CVs is introduced in the model. Without this assumption the concept can be generalized to platoons with variable length, according to a

Table 3 – Parameters and results (L_{sT} and w) of the time-dependent multi-class queue model.

Test case	T-slice k	Arrival (veh/min)	E[S]	E[S^2]	VAR[S]	C	λ	1/E[S]	ρ	L_{sT}	w
A	1	4.00	9.39	205.22	117.04	1.16	0.067	0.106	0.63	1.85	27.54
	2	6.00	10.20	239.57	135.53	1.15	0.100	0.098	1.02	11.53	93.58
	3	7.33	9.95	223.80	124.80	1.13	0.122	0.101	1.22	33.98	242.69
	4	5.33	8.88	179.74	100.89	1.14	0.089	0.113	0.79	18.48	235.11
	5	4.67	8.89	179.88	100.85	1.14	0.078	0.112	0.69	5.17	86.32
	6	4.00	8.33	178.02	108.63	1.28	0.067	0.120	0.56	1.66	25.29
	7	3.33	9.23	205.90	120.71	1.21	0.056	0.108	0.51	1.19	21.32
	8	3.33	9.39	205.22	117.04	1.16	0.056	0.106	0.52	1.18	21.24
B	1	2.67	9.39	205.22	117.04	1.16	0.044	0.106	0.42	0.77	17.17
	2	5.33	10.20	239.57	135.53	1.15	0.089	0.098	0.91	6.17	60.79
	3	6.00	9.95	223.80	124.80	1.13	0.100	0.101	1.00	12.93	113.26
	4	4.00	8.88	179.74	100.89	1.14	0.067	0.113	0.59	2.50	43.01
	5	3.33	8.89	179.88	100.85	1.14	0.056	0.112	0.49	1.10	19.81
	6	2.67	8.33	178.02	108.63	1.28	0.044	0.120	0.37	0.66	14.78
	7	2.67	9.23	205.90	120.71	1.21	0.044	0.108	0.41	0.75	16.88
	8	2.67	9.39	205.22	117.04	1.16	0.044	0.106	0.42	0.76	17.17
C	1	2.67	9.39	205.22	117.04	1.16	0.044	0.106	0.42	0.77	17.17
	2	3.33	10.20	239.57	135.53	1.15	0.056	0.098	0.57	1.38	24.66
	3	4.67	9.95	223.80	124.80	1.13	0.078	0.101	0.77	3.33	41.15
	4	7.33	8.88	179.74	100.89	1.14	0.122	0.113	1.09	17.98	115.80
	5	7.33	8.89	179.88	100.85	1.14	0.122	0.112	1.09	30.38	227.60
	6	4.00	8.33	178.02	108.63	1.28	0.078	0.120	0.65	7.57	137.39
	7	6.00	9.23	205.90	120.71	1.21	0.056	0.108	0.51	1.51	27.94
	8	7.33	9.39	205.22	117.04	1.16	0.044	0.106	0.42	0.79	17.65

certain probability distribution (i.e., one class for each fixed length of the platoons, between 1 and the maximum possible length).

Having to mix with non-CV vehicles due to the "integrated policy", we expect the CVs platoons can no longer be described with deterministic arrivals. In relation to the conditioning of non-CV vehicles, with stochastic arrivals, also the CVs will have stochastic arrivals. In these terms, we can still consider CVs platoons arrivals according to Poisson with parameter λ_{CV}. At this point, the CVs class will have its own deterministic service time S_{CVs} ~ constant $= s_{CV}$, which is equal to the service time of each CV multiplied by the number of vehicles in the CVs platoon.

From this point of view, the service time distribution of the CVs class can be considered as a degenerate probability distribution with $E[S_{CVs}] = s_{CV}$ and $VAR[S_{CVs}] = 0$, and then with $E[S^2_{CVs}] = s^2_{CV}$. This degenerate probability distribution can be approximated with a shifted Gamma function, assuming that $\alpha_{CVs} = 1$, $\delta_{CVs} = s_{CV}$ and considering a small enough value for β_{CVs}.The value for C to be used in time-dependent model (Eqs. 17–23) can be obtained by considering Eq. (11) and using the first and second moments of the Gamma mixture variable as in Eqs. (9) and (10). The values identified above for the CVs class are used in these last equations for class i = CVs (i.e., $E[S_{CVs}] = s_{CV}$ and $E[S^2_{CVs}] = s^2_{CV}$).

The considerations above highlight the flexibility of the proposed model and its compatibility with some future issue for motorway traffic planning and assessment. Further in-depth analysis and validation of the model will be necessary to test and confirm the use of the dependent time model also in this case, extremely important for future application.

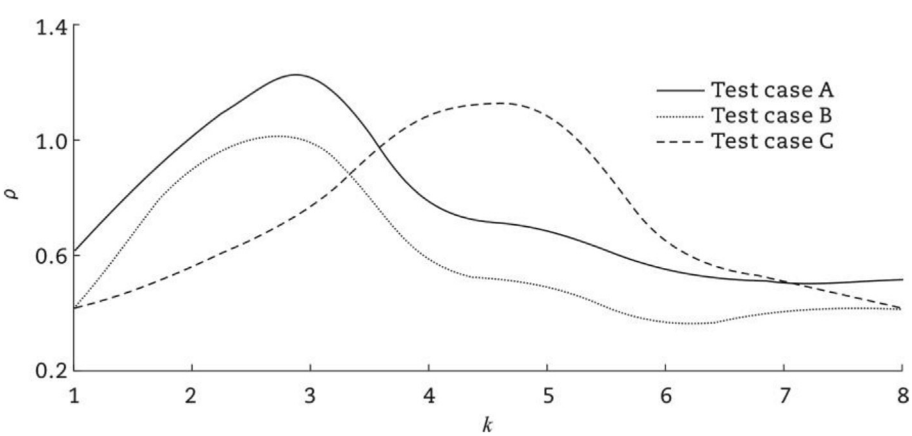


Fig. 7 – Evolution during time slices for the three test cases.

Please cite this article as: Pompigna, A., Mauro, R., A multi-class time-dependent model for the analysis of waiting phenomena at a motorway tollgate, Journal of Traffic and Transportation Engineering (English Edition), https://doi.org/10.1016/j.jtte.2020.09.001

Table 4 – SDES simulation results for L_{sT} and w (average values and 95% confidence interval with 600 trials).

Test case	T-slice k	L_{sT}				w			
		Mean	Standard deviation	95% confidence interval		Mean	Standard deviation	95% confidence interval	
A	1	1.92	1.58	1.79	2.05	26.97	8.89	26.26	27.68
	2	12.47	4.97	12.07	12.87	94.59	30.78	92.12	97.05
	3	33.05	8.13	32.40	33.70	238.21	66.11	232.92	243.50
	4	20.60	10.02	19.80	21.40	258.02	88.73	250.93	265.12
	5	7.56	7.40	6.97	8.15	123.74	78.68	117.45	130.04
	6	1.94	2.90	1.71	2.17	39.35	40.73	36.09	42.60
	7	1.16	1.14	1.07	1.26	21.60	10.34	20.78	22.43
	8	1.26	1.24	1.17	1.36	21.69	6.96	21.13	22.25
B	1	0.79	0.84	0.72	0.85	16.99	4.33	16.64	17.34
	2	7.43	3.69	7.13	7.72	62.88	22.54	61.08	64.69
	3	13.10	5.88	13.62	14.57	117.10	44.07	115.57	122.62
	4	2.72	3.06	2.48	2.96	71.55	43.81	68.04	75.05
	5	1.16	1.18	1.06	1.25	21.94	10.96	21.06	22.82
	6	0.64	0.82	0.57	0.70	14.55	4.05	14.23	14.87
	7	0.74	0.85	0.67	0.81	16.81	4.47	16.45	17.17
	8	0.86	0.96	0.78	0.93	17.67	5.12	17.26	18.08
C	1	0.82	0.87	0.75	0.89	17.36	4.46	17.00	17.72
	2	1.42	1.30	1.32	1.53	24.99	8.10	24.34	25.64
	3	3.80	2.78	3.58	4.02	42.73	16.30	41.42	44.03
	4	18.50	6.53	17.98	19.03	115.41	37.41	112.41	118.40
	5	29.47	8.99	28.75	30.18	223.39	70.60	217.74	229.04
	6	8.99	8.51	8.31	9.68	162.56	79.38	156.21	168.91
	7	1.57	2.31	1.38	1.75	41.98	41.06	38.69	45.26
	8	0.79	0.87	0.72	0.86	18.52	7.82	17.89	19.14

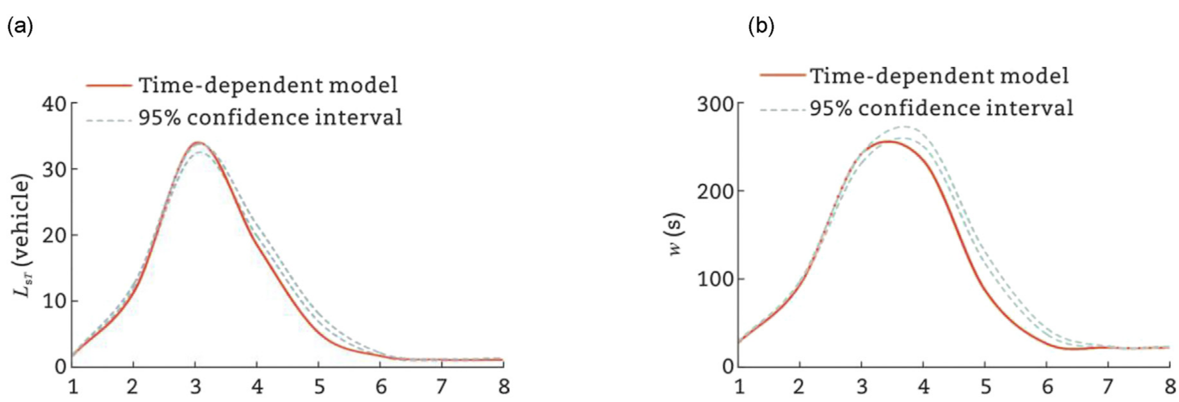


Fig. 8 – Test case A time-dependent model and 95% confidence interval of the 600 SDES trials. (a) L_{sT}. (b) w.

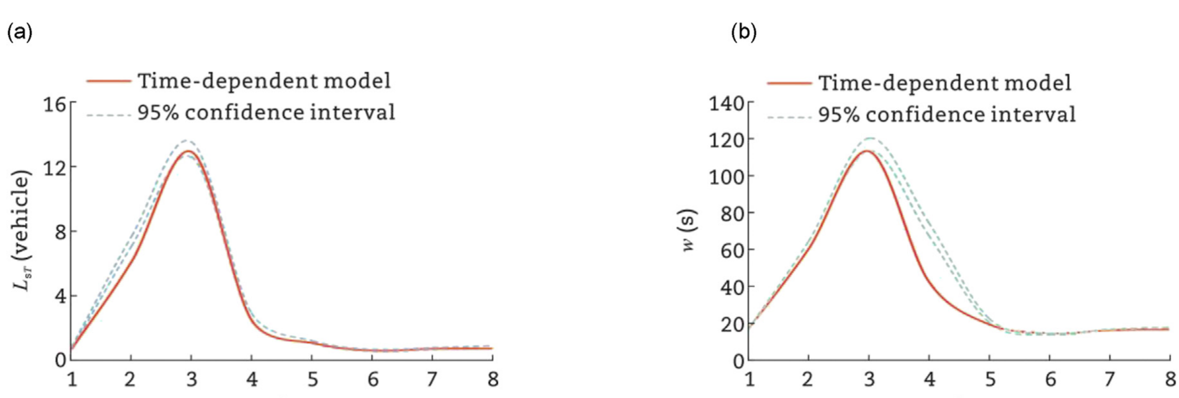


Fig. 9 – Test case B time-dependent model and 95% confidence interval of the 600 SDES trials. (a) L_{sT}. (b) w.

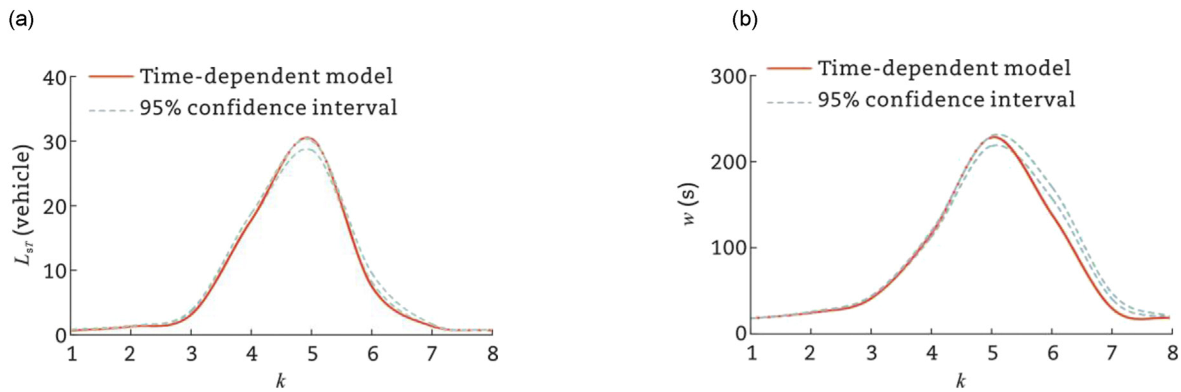


Fig. 10 – Test case C time-dependent model and 95% confidence interval of the 600 SDES trials. (a) L_{sT}. (b) w.

Table 5 – Goodness of fit (GOF) indicators for L_{sT} and w evolution (time-dependent model versus 600 SDES trials averages).

Queue parameter	GOF indicator	Test case A (%)	Test case B (%)	Test case C (%)
L_{sT}	R^2	99.10	99.20	99.70
	MAPE	12	7	7
w	R^2	97.70	93.50	98.40
	MAPE	10	8	8

8. Conclusions

The paper directs reader's attention to the application of time-dependent queue models that allows to face the whole spectrum of situations that may characterize the relationship between demand and supply service at a motorway tollgate. These situations include also over-saturation or any case distant from stationary that render unrealistic the application of the probabilistic queues theory. Moreover, these models solve the problems related to the simulation approach due to the possibility of compact formulas for an easy use in real applications.

The current article examines the time-dependent modeling for the non-stationary queue using the coordinate transformation criterion. In particular, a multi-class $M/\Gamma/1$ model has been proposed and developed, which allows a segmentation of the probability distributions of arrivals and service times in consideration of a clustering of users (i.e., vehicle and payment types) at a tollgate.

Starting from the deduction of the closed-form equations for the queue in statistical equilibrium, the time-dependent solutions of the multi-class $M/\Gamma/1$ queue are obtained for a motorway tollgate by the coordinates transformation for a certain time period with constant arrivals and calibrated service time distributions, and with a non-null initial queue. Probability distributions parameters can be calibrated using real-life data, in order to reproduce the actual behaviors at a tollgate. The equations for the average waiting time and the average number of vehicles in the system in a certain observation interval are therefore detailed. These equations, identified for a “basic case”, are also applicable for a time slices sequence and therefore they enable the assessment of any demand and services evolutionary profiles.

The paper presents also SDES model that can be used to simulate the operating of a motorway tollgate. The simulative model, adequately verified and validated, considers the presence of multiple class of users stratified by vehicle type and toll transaction type. In this specific case SDES represents a control model, that is a verification and validation tool for the results obtained using the coordinates transformation formulas.

Three different numerical cases are proposed to test the application of the time-dependent equations against the results of a sufficiently high number of simulation with the SDES model, in term of the average waiting time and the length of the queue at intervals of 15 min. A Monte Carlo stochastic simulation model has been used to reiterate SDES model for 600 times, in order to obtain statistically significant results.

The comparison shows a good capacity of the time-dependent multi-class queue model to approximate the mean values of the SDES model reiterations. In this way, the coordinate transformation method is confirmed as fully capable of resolving cases that fall outside the boundaries of the probabilistic theory of waiting phenomena. One of the objectives of the continuation of the research, from the point of view of comparison with alternative methods, may be to evaluate the convenience of the proposed formulas (i.e., effectiveness in producing acceptable approximations and efficiency in use in technical applications) with respect to the application of the PSSFA method (Hu et al., 2018; Wang et al., 1996), adequately specified for the multiclass case.

The proposed formulas, therefore, are easy to use in practical cases for the evaluation of the main MOEs for LOS assessment of a motorway toll station, allowing to operate quickly, compactly and with a good level of input segmentations and results approximation. Furthermore, they are easier to apply from a computational point of view, if compared with the level of complexity of the preparation and reiteration of the SDES model to infer on the average values of queue length and waiting times. Moreover, the same formulas will serve in the specification of the cost functions for the characterization of the supply system within the traffic demand assignment models. This aspect, in particular, can be deepened in the continuation of the research, together with the implications concerning the distribution of the demand on several gates depending on the operations types, the toll plaza geometry and the vehicles trajectories.

Please cite this article as: Pompigna, A., Mauro, R., A multi-class time-dependent model for the analysis of waiting phenomena at a motorway tollgate, Journal of Traffic and Transportation Engineering (English Edition), https://doi.org/10.1016/j.jtte.2020.09.001

It is worth pointing out that in the examples proposed in this paper the time-dependent equations are used to analyze situations with conventional vehicles, i.e., with total human control. As studies and applications that take into account the presence of vehicles without human control are increasingly in demand in transport engineering, in this paper a glance on CVs is presented. The proposed discussion highlights a significant flexibility of the model proposed in addressing these emerging topics, which needs however further investigations in the continuation of the research.

Conflict of competing interest

The authors do not have any conflict of interest with other entities or researchers.

Appendix

The Gamma shifted probability distribution, having a probability density function expressed by

$$f(x)=\frac{1}{\Gamma(\alpha)\beta^{\alpha}}(x-\delta)^{\alpha-1}e^{-\frac{x-\delta}{\beta}} \tag{A1}$$

can be calibrated in its parameters α (shape), β (scale) and δ (shift) resorting to maximum likelihood estimates (MLE) (Minka, 2002; Shanker et al., 2016).

Calibration can be obtained from an observed set of n independent data points $X=\{x_1,x_2,\cdots,x_n\}$. First of all, starting from the available data, we can estimate the shift parameters as

$$\hat{\delta}=\min\{x_1,x_2,\cdots,x_n\} \tag{A2}$$

At this point we can translate the whole data set taking into account $\hat{\delta}$, obtaining $Y=\{y_1,y_2,\cdots,y_n\}$, where $y_i=x_i-\hat{\delta}$. The Gamma probability density function expressed by

$$f(x)=\frac{1}{\Gamma(\alpha)\beta^{\alpha}}(y)^{\alpha-1}e^{-\frac{y}{\beta}} \tag{A3}$$

can be calibrated in its parameters α (shape) and β (scale) using $Y=\{y_1,y_2,\cdots,y_n\}$.

The likelihood function for n independent and identically distributed observations $Y=\{y_1,y_2,\cdots,y_n\}$ is

$$L(\alpha,\beta)=\prod_{i=1}^{n}f(y_1,\alpha,\beta) \tag{A4}$$

and the log-likelihood is

$$l(\alpha,\beta)=(\alpha-1)\sum_{i=1}^{n}\ln(y_i)-\sum_{i=1}^{n}\frac{y_i}{\beta}-n\alpha\ln(\beta)-n\ln(\Gamma(\alpha)) \tag{A5}$$

The maximum likelihood estimator $\hat{\beta}$ for β is obtained by deriving $l(\alpha,\beta)$ with respect to β and setting the derivative equal to zero. The maximum for β is easily found to be

$$\hat{\beta}=\frac{1}{\alpha n}\sum_{i=1}^{n}y_i \tag{A6}$$

Substituting this into Eq. (A5) gives

$$l(\alpha)=(\alpha-1)\sum_{i=1}^{n}\ln(y_i)-n\alpha-n\alpha\ln\frac{\sum_{i=1}^{n}y_i}{n\alpha}-n\ln(\Gamma(\alpha)) \tag{A7}$$

The maximum likelihood estimator $\hat{\alpha}$ for α is obtained by deriving $l(\alpha)$ with respect to α and setting the derivative equal to zero, that yields

$$l(\alpha)-\psi(\alpha)=\ln\left(\frac{1}{n}\sum_{i=1}^{n}y_i\right)-\frac{1}{n}\sum_{i=1}^{n}\ln(y_i) \tag{A8}$$

where $\psi(\alpha)$ is the Digamma function.

As the Eq. (A8) has not a closed-form solution for α, a numerical method can be used, for example the Newton–Raphson method in Choi and Wette (1969). In this way the value of $\hat{\alpha}$ can be obtained by successive approximations.

We can now consider a concrete example, taking the data set of the cash tollgate of Cikupa (Jakarta, Indonesia) reported in (Karsaman et al., 2014). The service time values for 80 monitored vehicles are shown in Fig. A1.

According to the data, the shift parameters $\hat{\delta}$ can be set equal to 3 s. Gamma distribution parameters $\hat{\alpha}$ and $\hat{\beta}$ can be estimated using the service time data translated by a value equal to $\hat{\delta}$. The MLE can be obtained using the statistical functions available in commonly used software. The analysis carried out with Matlab probability distribution fitting tools shows the following estimates: $\hat{\alpha}=1.298$, $\hat{\beta}=2.181$. Fig. A2 shows the trend of the estimated Gamma probability density function, together with the histogram representing the relative frequencies of the data at the tollgate, realigned with respect to $\hat{\delta}$.

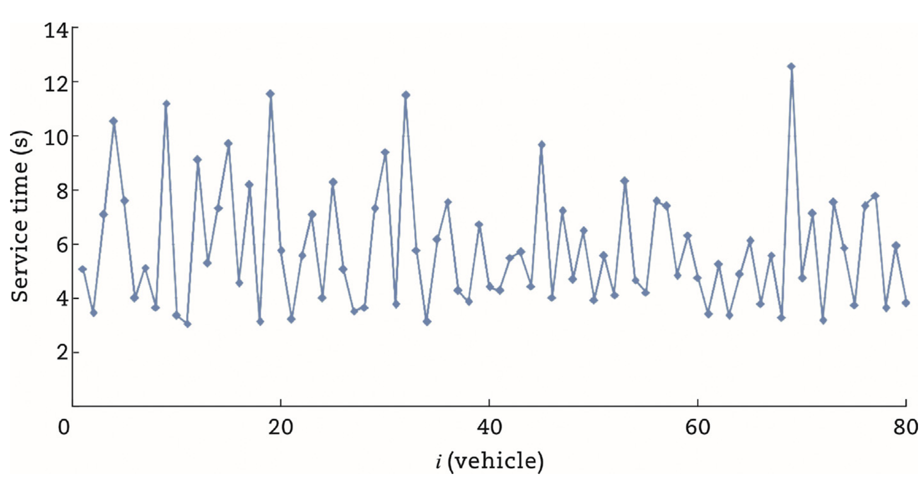


Fig. A1 – Service time values at the cash tollgate of Cikupa (Jakarta, Indonesia) (Karsaman et al., 2014).

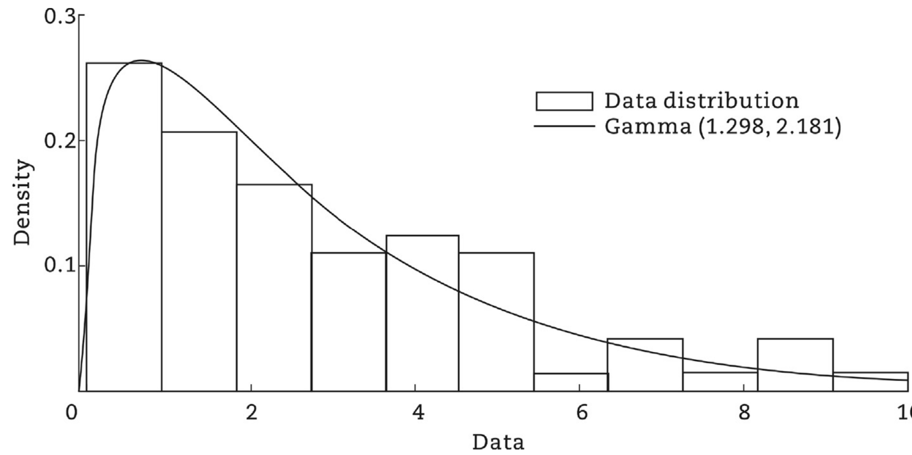


Fig. A2 – Data distribution and fitted Gamma distribution for service times at the cash tollgate of Cikupa (Jakarta, Indonesia) (Karsaman et al., 2014).

The Kolmogorov–Smirnov (Kolmogorov, 1933; Smirnov, 1948) and the Anderson-Darling (Anderson and Darling, 1954) tests can be used to determine the goodness of fit (Karadağ and Aktaş, 2015), i.e., how well the data fit to the underlying Gamma distribution. The tests consider the null hypothesis that the data follow the gamma distribution with parameters: $\widehat{\alpha} = 1.298$, $\widehat{\beta} = 2.181$. The calculated value of the Kolmogorov–Smirnov test statistic is $D_n = 0.062$, which is less than the critical value of $D_{\text{crit}} = 0.152$ (significance level 0.05). Since $D_n < D_{\text{crit}}$, we cannot reject the null hypothesis concluding that there is no significant difference between the real-life data and data coming from a Gamma distribution with parameters: $\widehat{\alpha} = 1.298$, $\widehat{\beta} = 2.181$.

The calculated value of the Anderson-Darling test statistic is $\text{AD} = 0.262$, which is less than the value critical $\text{AD}_{\text{crit}} = 0.780$ for Gamma distributions (significance level 0.05). Since $\text{AD} < \text{AD}_{\text{crit}}$, we cannot reject the null hypothesis concluding that there is no significant difference between the real-life data and data coming from a Gamma distribution with parameters: $\widehat{\alpha} = 1.298$, $\widehat{\beta} = 2.181$.

Both tests show a high p-value that is certainly much greater than 0.25, as the D and AD values are still lower than the tabulated values with significance equal to 0.25.

REFERENCES

Akçelik, R., Troutbeck, R., 1991. Implementation of the Australian roundabout analysis method in SIDRA. In: International Symposium on Highway Capacity, Karlsruhe, 1991.

Aksoy, G., Celikoglu, H.B., Gedizlioglu, E., 2014. Analysis of toll queues by micro-simulation: results from a case study in Istanbul. Procedia-Social and Behavioral Sciences 111 (5), 614–623.

Al-Deek, H.M., Radwan, A.E., 1995. A Framework for evaluating level of service at electronic toll collection plazas. In: Transportation Congress: Civil Engineers Key to the World Infrastructure, Orlando, 1995.

Anderson, T.W., Darling, D.A., 1954. A test of goodness of fit. Journal of the American Statistical Association 49, 765–769.

Astarita, V., Florian, M., Musolino, G., 2001. A microscopic traffic simulation model for the evaluation of toll station systems. In: Intelligent Transportation Systems, Oakland, 2001.

Boronico, J.S., Siegel, P.H., 1998. Capacity planning for toll roadways incorporating consumer wait time costs. Transportation Research Part A: Policy and Practice 32 (4), 297–310.

Brilon, W., 1995. Delays at oversaturated intersections based on reserve capacities. Transportation Research Record 1484, 1–8.

Brilon, W., Troutbeck, R., Tracz, M., 1997. Review of International Practices Used to Evaluate Unsignalized Intersections. Transportation Research Circular, 468. Transportation Research Board, Washington DC.

Brilon, W., 2007. Time dependent delay at unsignalized intersections. In: The 17th International Symposium on Transportation and Traffic Theory (ISTTT17), London, 2007.

Brilon, W., 2008. Delay at unsignalized intersections. Transportation Research Record 2071, 98–108.

Brilon, W., 2015. Average delay at unsignalized intersections for periods with variable traffic demand. Transportation Research Record 2483, 57–65.

Catling, I., 1977. A time-dependent approach to junction delays. Traffic Engineering & Control 18 (11), 520–526.

Ceballos, G., Curtis, O., 2004. Queue analysis at toll and parking exit plazas: a comparison between multi-server queuing models and traffic simulation. In: ITE 2004 Annual Meeting and Exhibit, Lake Buena Vista, 2004.

Chakroborty, P., Gill, R., Chakraborty, P., 2016. Analysing queueing at toll plazas using a coupled, multiple-queue, queueing system model: application to toll plaza design. Transportation Planning and Technology 39 (7), 675–692.

Chang, Y.L., 1998. WinQSB: Decision Support Software for Ms/Om. John Wiley, New York.

Cherng, C., Lewis, A., Pai, S., 2005. Modelling Toll Plaza Behaviour Using Queuing Theory. University of Washington, Washington DC.

Choi, S.C., Wette, R., 1969. Maximum likelihood estimation of the parameters of the Gamma distribution and their bias. Technometrics 11 (4), 683–690.

Cvitanić, D., Breški, D., Vidak, B., 2007. Review, testing and validation of capacity and delay models at unsignalized intersections. Promet Traffic-Traffico 19 (2), 71–82.

Devroye, L., 1986. Non-uniform Random Variate Generation. Springer, New York.
Doherty, A.R., 1977. A Comprehensive Junction Delay Formula. LTR1 Working Paper. Department of Transport, London.
Duhan, D., Arya, N., Dhanda, P., et al., 2014. Application of queuing theory to address traffic problems at a highway toll plaza. Applied Mechanics and Materials 592–594, 2583–2587.
Edie, L.C., 1954. Traffic delays at toll booths. Journal of the Operations Research Society of America 2 (2), 107–138.
Gautam, N., 2012. Analysis of Queues: Methods and Applications, first ed. CRC Press, Boca Raton.
Haight, F.A., 1958. Two queues in parallel. Biometrika 45, 401–410.
Heidemann, D., 2002. Mathematical analysis of non-stationary queues in traffic flow with particular consideration of the coordinate transformation technique. In: The 15th International Symposium on Transportation and Traffic Theory, Adelaide, 2002.
Heydecker, B.G., Verlander, N.Q., 1998. Transient delay in oversaturated queues. In: The 3rd IMA International Conference on Mathematics in Transport Planning and Control, Cardiff, 1998.
Hörmann, W., Leydold, J., Derflinger, G., 2004. Automatic Nonuniform Random Variate Generation. Springer, Berlin.
Hu, L., Zhao, B., Zhu, J., et al., 2018. Two time-varying and state-dependent fluid queuing models for traffic circulation systems. European Journal of Operational Research 275 (3), 997–1019.
Ito, T., 2005. Process simulation approach to design and evaluation of toll plaza with ETC gates. International Journal of Simulation 6 (5), 14–21.
Karadağ, Ö., Aktaş, S., 2015. Goodness of fit tests for generalized Gamma distribution. In: Internation Conference of Numerical Analysis and Applied Mathematics 2015 (ICNAAM 2015), Rhodes, 2015.
Karsaman, R.H., Mahendra, Y., Rahman, H., et al., 2014. Measuring the capacity and transaction time of cash and electronic toll collection systems. Journal of Engineering and Technological Sciences 46 (2), 180–194.
Kim, B.J., 2011. Conceptualization of traffic flow for designing toll plaza configuration: a case study using simulation with estimated traffic volume. International Journal of Industrial Engineering 18 (1), 51–57.
Kim, S., 2009. The toll plaza optimization problem: design, operations, and strategies. Transportation Research Part E: Logistics and Transportation Review 45 (1), 125–137.
Kimber, R.M., Hollis, E.M., 1979. Traffic Queues and Delays at Road Junctions. Transport and Road Research Laboratory (TRRL), Wokingham.
Klodzinski, J., Al-Deek, H.M., 2002. New methodology for defining level of service at toll plazas. Journal of Transportation Engineering 128 (2), 173–181.
Kolmogorov, A., 1933. Sulla determinazione empirica di una legge di distribuzione. Giornale dell'Istituto Italiano degli Attuari 4, 83–91.
Levinson, D., Chang, E., 2003. A model for optimizing electronic toll collection systems. Transportation Research Part A: Policy and Practice 37 (4), 293–314.
Lin, F.B., Su, C.W., 1994. Level-of-service analysis of toll plazas on freeway main lines. Journal of Transportation Engineering 120 (2), 246–263.
Liu, Y.J., Cao, J., Cao, X.Y., et al., 2017. Optimization of design scheme for toll plaza based on M/M/C queuing theory and cellular automata simulation algorithm. Modern Applied Science 11 (7), 1–13.
Magsino, E.R., Ho, I.W.H., 2016. An intelligent highway tollgate queue selector for improving server utilization and vehicle waiting time. In: 2016 IEEE Region 10 Symposium (TENSYMP), Bali, 2016.
Mahmassani, H.S., 2016. Autonomous vehicles and connected vehicle systems: flow and operations considerations. Transportation Science 50 (4), 1140–1162.
May Jr., A.D., Keller, H.E., 1967. A deterministic queueing model. Transportation Research 1 (2), 117–128.
Minka, T.P., 2002. Estimating a Gamma Distribution. Microsoft Research, Cambridge.
Marsaglia, G., Tsang, W., 2000. A simple method for generating gamma variables. ACM Transactions on Mathematical Software 26 (3), 363–372.
Mauro, R., 2010. Calculation of Roundabouts: Capacity, Waiting Phenomena and Reliability. Springer, Berlin.
Mauro, R., Pompigna, A., 2020. State of the art and computational aspects of time-dependent waiting models for non-signalised intersections. Journal of Traffic and Transportation Engineering (English Edition), https://doi.org/10.1016/j.jtte.2019.09.007.
Mehri, H., Djemel, T., 2011. Study and simulation of queuing theory in the toll motorway. Studia Informatica Universalis 9 (2), 96–130.
Mirzaeian, N., Cho, S.H., Scheller-Wolf, A., 2018. A Queueing Model and Analysis for Autonomous Vehicles on Highways. Available at: https://ssrn.com/abstract=3278330 (Accessed 15 July 2019).
Mostafizi, A., Dong, S., Wang, H., 2017. Percolation phenomenon in connected vehicle network through a multi-agent approach: mobility benefits and market penetration. Transportation Research Part C: Emerging Technologies 85, 312–333.
Obelheiro, M.R., Cybis, H.B.B., Ribeiro, J.L.D., 2011. Level of service method for Brazilian toll plazas. Procedia-Social and Behavioral Sciences 16, 120–130.
Pompigna, A., 2020. Queue formation: general models. In: Guerrieri, M., Mauro, M. (Eds.), A Concise Introduction to Traffic Engineering-Theoretical Fundamentals and Case Studies. Springer, Berlin, https://doi.org/10.1007/978-3-030-60723-4.
Pratelli, A., De Sio, O., Palavisini, A., 2001. Indagine Sperimentale e Tecnologie per L'esazione del Pedaggio Sulle Autostrade. Firenze University Press, Florence.
Ravindran, A.R., 2007. Operations Research and Management Science Handbook, first ed. CRC Press, Boca Raton.
Robertson, D.I., 1979. Traffic models and optimum strategies of control: a review. In: International Symposium on Traffic Control Systems, Berkeley, 1979.
Sadoun, B., 2005. Optimizing the operation of a toll plaza system using simulation: a methodology. Simulation Transactions of the Society for Modeling and Simulation International 81 (9), 657–664.
Sangavi, G.V., Megha, G.C., Prajendra, H.R., et al., 2017. Application of queuing theory of a toll plaza-a-case study. International Journal of Engineering Research & Technology 6 (6), 541–554.
Shanker, R., Shukla, K.K., Shanker, R., et al., 2016. On modeling of lifetime data using two-parameter gamma and Weibull distributions. Biometrics & Biostatistics International Journal 4 (5), 00107.
Shanmugasundaram, S., Punitha, S., 2014. A simulation study on toll gate system in M/M/1 queueing models. IOSR Journal of Mathematics 10 (3), 1–9.
Smirnov, N.V., 1948. Table for estimating the goodness of fit of empirical distributions. The Annals of Mathematical Statistics 19 (2), 279–281.
TRB, 2000. Highway Capacity Manual, forth ed. Transportation Research Board, Washington DC.

TRB, 2015. Highway Capacity Manual, sixth ed. Transportation Research Board, Washington DC.

Troutbeck, R.J., Brilon, W., 2000. Unsignalized intersection theory. In: Gartner, N., Messner, C.J., Rathi, A.K. (Eds.), Traffic Flow Theory. Transportation Research Board, Washington DC, pp. 1–47.

Van Dijk, N.M., Hermans, M.D., Teunisse, M.J.G., et al., 1999. Designing the westerscheldetunnel toll plaza using a combination of queueing and simulation. In: The 31st Conference on Winter Simulation, Phoenix, 1999.

Van Dijk, N.M., 2000. On hybrid combination of queueing and simulation. In: The 32nd Conference on Winter Simulation, Orlando, 2000.

Wang, W.P., Tipper, D., Banerjee, S., 1996. A simple approximation for modeling nonstationary queues. In: The 15th Joint Conference of the IEEE Computer and Communications Societies Conference, New York, 1996.

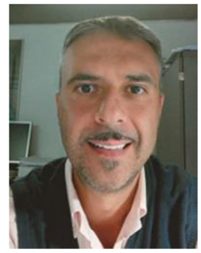
Wang, C., 2017. Study on toll plaza design based on M/M/1 queue theory. In: The 7th International Conference on Education, Management, Computer and Society (EMCS 2017), Shenyang, 2017.

Woo, T.S.H., Hoel, L.A., 1991. An Investigation of Toll Plaza Capacity and Level of Service. Virginia Transportation Research Council, Charlattesville.

Wong, W.T., Wong, S.C., Tong, C.O., 2003. Sheared delay formulae for the TRANSYT traffic model: a review and calibration. Transport Reviews 23 (1), 1–20.

Wu, N., 1998. Estimation of queue lengths and their percentiles at signalized intersections. In: The 3rd International Symposium on Highway Capacity, Copenhagen, 1998.

Zarrillo, M.L., Radwan, A.E., Al-Deek, H.M., 1997. Modeling traffic operations at electronic toll collection and traffic management systems. Computers & Industrial Engineering 33 (3–4), 857–860.



Andrea Pompigna received the PhD degree in civil, chemical, environmental and materials engineering with specialization in transportation engineering from University of Bologna, Italy. He is a chartered engineer and expert consultant in transportation engineering. He is research collaborator at the Department of Civil, Environmental and Mechanical Engineering, University of Trento, Italy and teaching assistant at the Department of Civil, Chemical, Environmental, and Materials Engineering, University of Bologna, Italy. His research interests are in traffic flow analysis, transportation modelling and simulation, traffic engineering operations and highway capacity, queue models in traffic engineering, time series and econometrics in transportation engineering.

Raffaele Mauro is a full professor of highway and traffic engineering and member of the teaching body of the PhD program at DICAM, University of Trento, Italy. He was the president of the SIIV-Italian Society of Road Infrastructure and founding member of the AISI-Italian Association of History of Engineering. Reviewer and member of the editorial board of several international scientific journals, he is author of 6 books and more than 130 papers published in national and international scientific journals or conference proceedings. He was member of the board of directors of the University of Trento and of the "Autostrada del Brennero SpA".