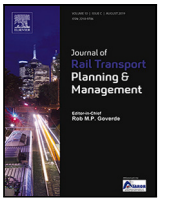
Journal of Rail Transport Planning & Management 36 (2025) 100539

Contents lists available at ScienceDirect

Journal of Rail Transport Planning & Management

journal homepage: www.elsevier.com/locate/jrtpm



A queueing-based approach for timetable-independent railway station performance analysis

Tamme Emunds *, Nils Nießen

Institute of Transport Science, RWTH Aachen, Aachen, Germany

ARTICLE INFO

Keywords:
Railway station capacity
Queueing system
Timetable-independent
Continuous-time Markov chain
Performance analysis

ABSTRACT

Railway stations serve as critical nodes within railway networks, facilitating connections across diverse travel directions. Traditionally, the analytical performance analysis of railway stations has been divided into two distinct components: the examination of stopping tracks and the evaluation of route nodes, the locations within a station where switches determine the direction of travel. This study introduces an innovative Continuous-Time Markov Chain model that represents a comprehensive queueing system encompassing the entire railway station. By deriving timetable-independent performance indicators, this model provides a robust framework for assessing station performance. Consequently, it equips infrastructure operators with a holistic tool for infrastructure planning and evaluation.

1. Introduction

With the rising demand for railway traffic, it is imperative that existing infrastructure undergo continuous evaluation and that large-scale infrastructure projects are planned with a long-term perspective. Railway infrastructure is designed to last for multiple decades, enabling infrastructure managers to make informed decisions based on comparative quality assessments of planned constructions is crucial.

Railway stations serve as critical nodes in the performance evaluation of railway networks. However, timetable-independent performance analyses have typically been conducted only on individual segments of station infrastructure, specifically dividing them into *route nodes* and *track areas*. The former (also known as *switch area*) refers to the section of the station where train routes from entry/exit signals to platform tracks may overlap. In contrast, the latter pertains to the group of platforms where trains stop to perform operational tasks such as (un)loading freight or facilitating passenger exchange.

To measure performance capability of railway infrastructure, three main types of railway capacity (see also Jensen et al., 2020) can be distinguished: *theoretical capacity*, *timetable capacity*, and *operational capacity*. *Theoretical capacity* is defined as the maximum number of trains or train-route inquiries that can be scheduled without conflicts on a given infrastructure, taking into account the driving dynamics and installed railway control systems. *Timetable capacity* (occ. referred to as *maximal capacity*), denotes the highest number of requests that can be scheduled on the infrastructure during the timetabling process while maintaining an acceptable quality level compared to a specified threshold. This includes considerations of driving dynamics, railway control systems, and operating program specifics such as train-mix and arrival processes. *Operational capacity* (occ. referred to as *practical capacity*, Abril et al., 2008) pertains to the maximum number of trains that can traverse the infrastructure with acceptable operational quality relative to a specified threshold. This definition encompasses all factors considered in timetable capacity while additionally accounting for the probability of disturbances and knock-on delays.

* Corresponding author.
E-mail address: emunds@via.rwth-aachen.de (T. Emunds).

https://doi.org/10.1016/j.jrtpm.2025.100539
Received 8 April 2025; Received in revised form 7 July 2025; Accepted 11 August 2025
Available online 21 August 2025

Table 1
Methods for queueing-based railway station performance analysis.

Method	Route node analysis	Track group analysis	Route-based decomposition	Multi-channel model	Solution method
Potthoff (1962)		✓		✓	Iterative formula
Schwanhäußer (1978)	✓		(✓)		Closed-form formula
Hertel (1986)		✓		✓	Closed-form formula
Nießen (2008, 2013)	✓		(✓)	✓	Iterative formula
Schmitz et al. (2017)	✓			✓	Matrix–vector equations
Weik (2020)	✓		(✓)		Matrix–vector equations
Bychkov et al. (2021)	(✓)	(✓)		(✓)	Simulation method
Emunds and Nießen (2024)	✓		✓	✓	Probabilistic model-checking
Introduced here	✓	✓	✓	✓	Probabilistic model-checking

Various methodologies have been employed to analyse the capacity of railway infrastructure. The International Union of Railways (UIC) (2013) introduces the *capacity occupation* metric, describing the share of actually utilized blocking times in a time horizon, by compressing the existing blocking-time staircases of a timetable (Bešinović and Goverde, 2018). Some other methods directly utilize operational data to describe quality during daily operations (Graffagnino, 2012; Armstrong and Preston, 2017; Weik, 2022; Corman and Henken, 2022). Others employ theoretical simulations (Abril et al., 2008; Schmidt and Martin, 2010; Liang et al., 2017; Zieger et al., 2018) or Markov chains (Spanninger et al., 2023) to estimate operational capacity (Abril et al., 2008; Schmidt and Martin, 2010; Liang et al., 2017) or delay propagation effects (Zieger et al., 2018; Spanninger et al., 2023). Another approach involves using Mixed Integer Programming (MIP) to optimize train throughput within the infrastructure while considering different objective functions (Burdett and Kozan, 2006; Harrod, 2009; Mussone and Wolfler Calvo, 2013; Burdett, 2015; Liao et al., 2021). They aim to quantify infrastructure capacity by either utilizing an existing timetable or saturating it with additional train journeys.

Queueing-based methods provide timetable-independent metrics for assessing railway infrastructure performance. They typically describe the performance metric of timetable capacity, as introduced in Wendler (2007). These queueing-based methods can be further categorized based on their analysed infrastructure and solution techniques. Some approaches focus on the performance of railway lines (Schwanhäußer, 1974; Schwanhäußer and Schultze, 1982; Wendler, 2007; Weik and Nießen, 2017), while others examine route nodes (Schwanhäußer, 1978; Nießen, 2008, 2013; Schmitz et al., 2017; Weik, 2020; Emunds and Nießen, 2024) or track groups (Potthoff, 1962; Hertel, 1986). Additionally, some studies analyse queueing networks to determine the capacity of freight railway stations (Bychkov et al., 2021) and railway networks (Huisman et al., 2002; Kazakov et al., 2023).

Table 1 provides a comparative analysis of various queueing-based methodologies for assessing railway station performance. It should be noted that any statement enclosed in parentheses indicates partial support for that feature. For route-based decomposition, this typically implies either the utilization of single-channel systems exclusively (Schwanhäußer, 1978; Weik, 2020) or a decomposition based on serial-route-nodes combined with an analysis of route-based quality metrics (Nießen, 2008, 2013).

The queueing network described by Bychkov et al. (2021) pertains to a freight railway station and, as such, has not been segmented into route nodes and track groups. Additionally, Bychkov et al. (2021) employ simulations to derive results from their formulated queueing network, whereas the other methods listed in Table 1 achieve exact solutions through formulas, matrix–vector equations, or probabilistic model-checking.

The methodologies proposed by Potthoff (1962) and Hertel (1986) are specifically designed to determine the required number of platform tracks at a station. Other approaches focus on analysing the performance of route nodes; some use single-channel systems to approximate complex route nodes (Schwanhäußer, 1978; Weik, 2020), while others introduce multi-channel queueing systems that more directly account for parallel usability (Nießen, 2008, 2013; Schmitz et al., 2017; Emunds and Nießen, 2024).

In this work, we introduce a novel Continuous-Time Markov Chain (CTMC) formulation that describes the system dynamics of a railway station. To the best of our knowledge, this is the first method, describing the timetable-independent performance of an entire railway station in one combined multi-channel queueing system. The main contributions of this study are:

- Presenting a novel exact queueing-based model for an entire railway station,
- Facilitating probabilistic model-checking and numerical root-finding methods to obtain timetable-independent performance metrics,
- Decomposing the queueing process by station routes and service processes to identify bottlenecks, and

- Applying the method to a case study from the perspective of an infrastructure manager, thereby aiding in decision-making between potential operating programs.

This work begins with the introduction of the formulated queueing system model in Section 2. Subsequently, the process for determining timetable capacity is described in Section 3. The case study presented in Section 4 demonstrates the practical applicability of the approach for infrastructure operators. A conclusion and discussion is provided in Section 5.

2. Queueing system model of a railway station

In this section, we introduce the queueing system model of a railway station. First, we discuss queueing systems in general terms (Section 2.1). Next, we present an approximation formula relevant to our model (Section 2.2). Afterwards, we elucidate our approach to partitioning railway stations within this framework (Section 2.3). Finally, we present a novel Continuous Time Markov Chain (CTMC) model for the queueing system of a railway station (Section 2.4).

2.1. General discussion on queueing systems

A queueing system describes the arrival and service process of entities through a system. Since entities may arrive at arbitrary points in time, and their service may take an arbitrary amount of time, both the arrival and service processes are described using probability distributions. Furthermore, a queueing system also includes specifications regarding the number of waiting positions m, where arriving entities are 'parked' until they are serviced, and the number of channels capable of simultaneously servicing different entities (see also Kendall 1953).

The arrival process of such a system can be described with the arrival rate λ and its coefficient of variation v_A, the service process with the service rate μ and its coefficient of variation v_S.

To evaluate the performance of such a system, different indicators can be distinguished. For example, the occupation ratio $\rho = \lambda/\mu$ provides an initial estimation regarding the stability of the system.

As a consequence of the arbitrary arrival and service processes, a queue may form in front of the first server. The average length of this queue, denoted as L, can be utilized to further analyse the system's performance. Furthermore, if the number of waiting positions m is finite, it might occur that the system is incapable of accommodating another arriving request because all waiting positions are already occupied. Consequently, incoming requests are lost, and this can be analysed through the probability of loss, denoted as p_{loss}.

While single-channel systems with exponentially distributed inter-arrival and service times (M/M) can be analysed using a closed-form solution, more advanced models incorporating multiple channels may require a system-specific analysis of the state probabilities within a Continuous-Time Markov Chain (CTMC) formulation. For certain railway infrastructures, such as railway lines (Schwanhäußer, 1974; Fischer and Hertel, 1990; Weik, 2020) and junctions (Schmitz et al., 2017; Emunds and Nießen, 2024), corresponding formulations have already been analysed. However, in this work, we introduce the first CTMC formulation for railway stations (see Section 2.3).

To apply a formulation based on exponential arrival and service processes (M/M) to general independent processes (GI/GI), i.e., with non-unit coefficients of variation $v_A, v_S \neq 1$, an approximation formula can be used.

2.2. Approximation formula

In this work, we use the approximation formula from Hertel (1986), Fischer and Hertel (1990),

$$L(M/M) \cdot \frac{1}{\gamma} \approx L(GI/GI), \tag{1}$$

with the parameters

$$\gamma = \frac{2}{c \cdot v_S^2 + v_A^2} \tag{2}$$

and

$$c = \left(\frac{\rho}{s}\right)^{1-v_A^2} \cdot (1 + v_A^2) - v_A^2, \tag{3}$$

where the scaling factor γ depends on the coefficients of variation, as well as the number of parallel service channels s (fixed at $s = 1$ for our formulation) and the occupation rate ρ.

The next section will introduce our approach for modelling a railway station.

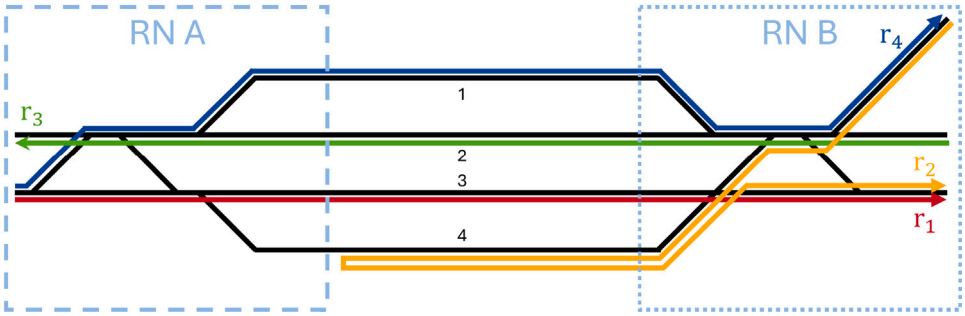


Fig. 1. An example for a railway station infrastructure.

2.3. Approach to railway station partitioning

For the determination of the timetable capacity of a railway station, we decompose a station into its main parts, the switch and track area, and into the requested routes. An example of a railway station can be found in Fig. 1.

It contains four different routes through a railway junction with four tracks. While route r_1 and r_4 enter via route node A and depart via route node B, route r_3 enters the station via route node B and leaves it via route node A. Notably, route r_2 only utilizes route node B since it corresponds to trains that turn on track 4.

To analyse the performance of such a railway station, we introduce multiple queues, one for every route. Therefore, the bottleneck of the station can be identified by comparing the estimated queue-lengths L_r per routes. Since every route corresponds to a queueing system, the parameters of those systems need to be formulated by the route. Assuming the operating programs, i.e. the number of trains per route n_r in a specified time horizon U, are given, the arrival rate of the route is calculated as $\lambda_r = n_r/U$. Similarly, the coefficient of variation $v_{A,r}$ needs to be formulated for the arrival process on this route. For completely random arrival requests, an exponential distribution, with $v_{A,r} = 1.0$, could be assumed (see Wendler 1999, 2007). However, since modern timetables continue to introduce more synchronized route occupation requests, we use a default value of $v_{A,r} = 0.8$ for every route in this work, in correspondence with best-practice recommendations (VIA Consulting & Development GmbH, 2021).

The model of the service process is significantly more complex, since the service of a route in a station infrastructure contains multiple parts. These consist of the *entry blocking*, *stopping* and *departure blocking* processes (this partitioning has already been used in Zwaneveld et al. 2001). For every of those steps in a request's journey through the station infrastructure, different conflicts with other requests may arise. However, those conflicts are usually limited to one route node, or, in the case of entries, to one route node and one platform. Additionally, when modelling *passing*, i.e. not stopping, requests, a conflicts in both route nodes and one platform track might be also feasible. The separation between the two route nodes and the platform tracks is however still useful to get an understanding about the underlying blocking structure of the different service processes.

Fig. 2 illustrates the dependencies between the various arrival and departure processes for the two routes, r_4 and r_3, in the given example. It is important to note that only one queue per route is depicted, positioned just before the entry blocking process, representing the arrival until standstill at the platform track. Additionally, for each route, the different service processes — entry blocking, stopping, and departure blocking — are shown sequentially in this order.

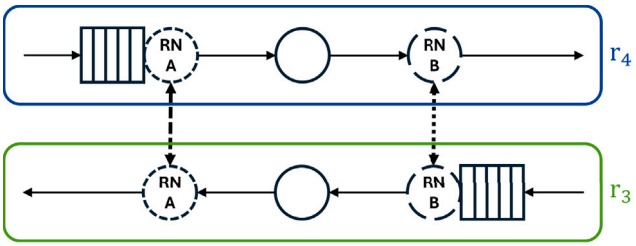


Fig. 2. Dependencies in the queueing systems for two routes in opposite directions.

Since r_4 starts at route node A and continues its journey to route node B, blocking both the entry of r_3 at route node B and its departure from route node A, a dependency between the corresponding service processes is depicted in the diagram.

All three service processes can be modelled with a corresponding service rate for each route: the service rate $\mu_{r,\text{entr}}$ for the entry blocking process, the service rate $\mu_{r,\text{stop}}$ for the stopping process, and the service rate $\mu_{r,\text{dep}}$ for the departure blocking process. Let $b_{r,\text{entr}}, b_{r,\text{dep}}, b_{r,\text{stop}}$ denote the average service times for the entry blocking, departure blocking, and stopping processes for route r, respectively. The corresponding service rates are then calculated as $\mu_{r,\text{entr}} = 1/b_{r,\text{entr}}$ for the entry blocking process and analogously for the departure blocking and stopping processes per route.

The required service times $b_{r,\text{entr}}, b_{r,\text{dep}}$ can be calculated by utilizing any microscopic running time calculation tool, and enhancing it with the corresponding times for route-setting, reaction, approaching, clearing and route-release, forming the blocking-time (see f.e. Hansen and Pachl 2014). To incorporate information about further exclusions in the line sections before or after the analysed stations, infrastructure operators can use the minimum headway times from or to the adjacent passing stations for all sequences of two route, rolling stock combinations sharing more than one signalling block, and averaging over all possible sequences with their relative frequency to obtain the average service time per route.

Since this work utilizes the Hertel formula (Section 2.2), all coefficients of variation for the three service processes, $v_{S,r,\text{entr}}, v_{S,r,\text{stop}}, v_{S,r,\text{dep}}$, need to be combined into a single coefficient of variation $v_{S,r}$ per route. This combined coefficient describes the total variation in all three service processes. We assume statistical independence of the three service processes in order to calculate the combined coefficient of variation. Further details can be found in the Appendix. The variation coefficients of the arrival process v_A and of the combined service process v_S can then be used in the approximation formula to scale the estimated queue-length L_r per route.

Another challenge lies in addressing the mutual exclusions of the service processes for different routes. These can be described using the following formulation as a Continuous-Time Markov Chain.

2.4. Continuous-time Markov chain formulation

A Continuous-Time Markov Chain (CTMC) $MC = (S, T)$ is described by a set of states S and a set of transitions T. By defining an initial state $s_0 \in S$ and transition rates for transitions $t = (s, s') \in T$, a *stationary distribution*, i.e., a probability distribution that remains unchanged as time progresses, can be calculated. This stationary distribution represents the long-term behaviour of the system, where the probabilities of being in each state stabilize over time.

States

A state $s \in S$ needs to contain information regarding the current status $\pi_r \in \Pi_r$ in the service process of all routes and the current length of the queue q_r for all routes. Hence, the state set can be described as

$$S = \left\{ \left(\pi_1, \mathrm{q}_1, \dots, \pi_k, \mathrm{q}_k\right) \mid \left(\pi_r, \mathrm{q}_r\right) \in \Pi_r \times \{0, \dots, m\} \right\}, \tag{4}$$

for a total number of k routes and the set of different service statuses Π_r.

The status $\pi_r \in \Pi_r = \{0, \dots, 4\}$ of a route represents the progress through the station. A status of $\pi_r = 0$ corresponds to a route without any trains on it; after the arrival of a request, the status changes to $\pi_r = 1$. Furthermore, $\pi_r = 2$ indicates that the train is stopping, and $\pi_r = 4$ signifies the departure of the train. The remaining status, $\pi_r = 3$, is achieved only when a train has already completed its required stopping time but needs to wait until the departure route can be occupied.

Note that the status set of a passing route r' can be restricted to $\Pi_{r'} = \{0, 1, 4\}$ since the stopping status $\pi_{r'} = 2$ and consequently the waiting for departure status $\pi_{r'} = 3$ are omitted. This implies that both the arrival and departure routes of a passing train must be unoccupied for service to commence.

In the following, we extend the notation to $\pi_r(s)$, which describes the status of route r in state s, and $\mathrm{q}_r(s)$, which denotes the queue-length of route r in state s.

Furthermore, since the state set S corresponds to a composition of the status and queue-length settings for every route $r \in R$, certain combinations can be ruled out to further restrict the state set. For example, a state s cannot contain $\pi_{r_3} = 1$ and $\pi_{r_4} = 4$ if the entry blocking of route r_3 and the departure blocking of route r_4 are conflicting. In the implementation of this work, this process is automatically handled by the parsing algorithms of the model files describing the CTMC, by analysing the reachability of the different states as defined by the transitions T.

Transitions

A transition $t = (s, s') \in T$ between two states $s, s' \in S$ describes the transfer from state s to state s'. Its duration can be expressed with the transition rate, which depends on the type of the transition. The different types can include an *arrival* to a route, one of the various *service processes* for a route, or additionally, a *choice* of which route to serve next.

When representing the transitions in a figure, not all dimensions of the state space can be represented, in the following we therefore select two dimensions per figure. Furthermore we distinguish between invariant and conditional transitions to enhance the representations. For a fixed route r, *invariant* transitions (s, t) correspond to transitions that do not depend on parameters $\pi_{r'}(s), q_{r'}(s)$ of other routes $r' \neq r$. Conversely, *conditional* transitions do depend on other parameters.

An arrival transition $s \to s'$ utilizes the arrival rate λ_r of route r and defines

$$\pi_r(s') = 1, \tag{5}$$

if no conflicting route is currently in service in the corresponding route node. Otherwise, it defines

$$\mathrm{q}_r(s') = \mathrm{q}_r(s) + 1, \tag{6}$$

as long as $\mathrm{q}_r(s') \leq m_r$ holds and the arrival can therefore still be modelled.

Fig. 3 gives a representation of both possible arrival transitions. Either the entry blocking service can directly start ($\pi_r(s') = 1$, Fig. 3(a)) or it is blocked by service on another route, therefore staying in the same service phase ($\pi_r(s') = 0$) and adding the train

to the queue ($q_r(s') = 1$), see Fig. 3(b). The arrival transition is therefore of the conditional transition type. In order to represent all possibilities of a conditional transition in a single figure, a combined representation, where every conditional transition might be taken from the origin state s, depending on the status $\pi_{r'}(s)$ of the service process in conflicting (and not necessarily visualized) routes r', can be used. An example of the discussed arrival transition is shown in Fig. 3(c).

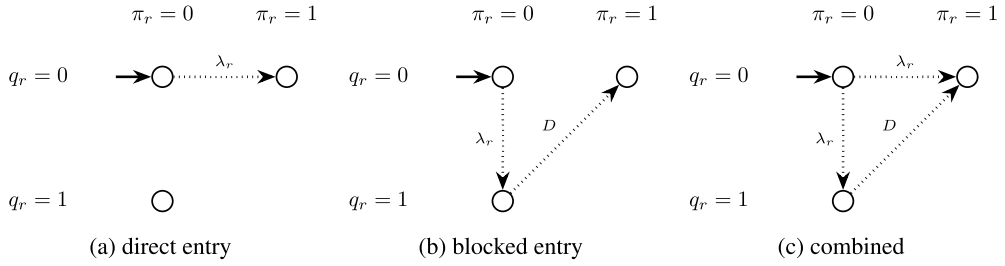


Fig. 3. The two possibilities of the arrival transition to an empty route ($q_r = 0$) and a combined representation.

For the service transitions, the entry blocking process is using the service rate $\mu_{r,\text{entr}}$ for the transition $s \to s'$ with $\pi_r(s) = 1$ and

$$\pi_r(s') = 2. \tag{7}$$

The stopping transition $s \to s'$ with rate $\mu_{r,\text{stop}}$ depends on conflicts between the departure blocking of route r and the arrival or departure blockings of other routes r' in state s, with $\pi_r(s) = 2$. If no conflicts are present, it maps the status of s' to

$$\pi_r(s') = 4. \tag{8}$$

Otherwise, it maps the status to

$$\pi_r(s') = 3. \tag{9}$$

The final service transition $s \to s'$ is of the kind departure blocking. It corresponds to the transfer between s with $\pi_r(s) = 4$, to s', fulfilling

$$\pi_r(s') = 0, \tag{10}$$

with a rate of $\mu_{r,\text{dep}}$.

Notably, states s with $\pi_r(s) = 3$ do not have an outgoing service transition for the route r. This is due to their origin in a conflict with the departure blocking of route r. Hence, their departure service is initiated only after the initial conflicting service has been completed. Since other conflicting routes might also be waiting for their next service step, a choice transition is used to model a random decision on which route to serve next. Therefore, every state s with a route r, for which $\pi_r(s) = 3$ holds, has an outgoing choice transition to s' with

$$\pi_r(s') = 4, \tag{11}$$

using an artificial constant transition rate D. This rate is chosen sufficiently high so that the induced delay $1/D$ to the system is not significant.

Similarly, states s with $\pi_r(s) = 0$ and $\mathrm{q}_r(s) > 0$ also have an outgoing choice transition (with rate D) to a state s' with an activated entry blocking

$$\pi_r(s') = 1, \tag{12}$$

and hence one less item in the queue

$$\mathrm{q}_r(s') = \mathrm{q}_r(s) - 1 \tag{13}$$

for route r.

Fig. 4 highlights π_{r_1} and π_{r_4} in an example describing the entry blocking service of two conflicting routes.

For the initial state s with $\pi_{r_1}(s) = 0$ and $\pi_{r_4}(s) = 0$, either the transition (s, t), with $\pi_{r_1}(t) = 1$ and $\pi_{r_4}(t) = 0$, or the transition (s, t'), with $\pi_{r_1}(t') = 0$ and $\pi_{r_4}(t') = 1$, is selected first.

Since all choice transitions use the same constant transition rate D, every possible choice in such a state is used with equal probability, and hence independent on the sequence of preceding events.

Example

All transitions share the characteristic that only one route is affected by each transition. For the arrival transitions, only one route gains another item in its system. Similarly, for the service transitions, the various service operations are carried out on a single specified route. For the choice transitions, either the commencement of the entry blocking or departure blocking is initiated for a single route.

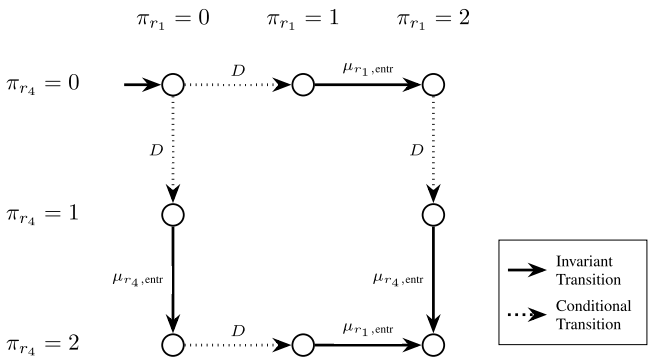


Fig. 4. An example for the choice transitions between the entry processes of two conflicting routes r_1 and r_4, assuming non-empty queues ($q_{r_1}, q_{r_4} > 0$) for both routes.

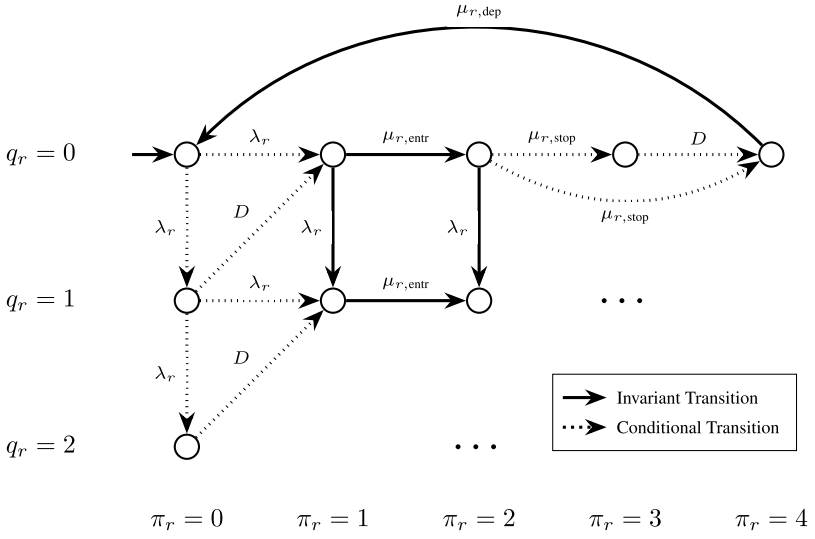


Fig. 5. A CTMC for the arrival, entry, stopping and departure process on one route with $m = 2$.

The different states and transitions can therefore be easily visualized when limiting the visualization to a fixed route r. Fig. 5 thus shows the dependencies between the states of a single route.

For an initial state s with $\pi_r(s) = 0$ the arrival transition (s, t) is an example of a conditional transition, as further explained in Fig. 3. However, for states s' with any other current service state $\pi_r(s') > 0$, the arrival transition (s', t') to route r always adds to the corresponding queue $q_r(t') = q_r(s') + 1$.

Another conditional transition (s, t) is that of the stopping process. Depending on whether or not the departure blocking process is in conflict with any other routes, the transition's destination state t either fulfils $\pi_r(t) = 3$ (conflict) or $\pi_r(t) = 4$ (no conflict).

Furthermore, both the departure blocking process and correspondingly, the entry blocking process are invariant to other states. They simply connect states s and t with

$$\pi_r(t) = \begin{cases} 2 & ;\text{if } \pi_r(s) = 1, \\ 0 & ;\text{if } \pi_r(s) = 4. \end{cases} \tag{14}$$

All choice transitions (s, t) are conditional since they are usually applied when conflicts between routes may exist. They correspond to the choice of the next service transition, an example can be found in Fig. 4. Consequently, any type of service transition on a route r holds only if no other route r' is currently in a conflicting entry blocking, stopping or departure blocking process.

3. Determining the timetable capacity of railway stations

While the previous section introduced the Continuous-Time Markov Chain model, this section focuses on obtaining the timetable capacity of a railway station. After an introduction to timetable capacity (Section 3.1), the handling of the described CTMC model is discussed — from the required data (Section 3.2), to calculating the queue-length estimations (Section 3.3), defining the thresholds (Section 3.4) and outlining the procedural flow (Section 3.5). Furthermore, Section 3.6 gives insights into the scalability of the introduced model.

3.1. The timetable capacity performance metric

Since the timetable of a railway system is more dynamic than its infrastructure, timetable-independent performance metrics are of high significance. In this work, we employ the notation of timetable capacity, as articulated in Wendler (2007). This capacity metric can be understood as the performance potential of railway infrastructure during the timetabling process, wherein train operators request resources for their planned train journeys from the infrastructure manager. It is important to note that multiple journeys may request identical or conflicting infrastructure routes for the same point in time, necessitating the rescheduling of certain journeys to achieve a conflict-free schedule. For each route, rescheduled journeys can be represented as a queue awaiting access to the requested infrastructure route. At any given point in time, the queue-length of a given route corresponds to the number of requests pending access to the infrastructure. Although the proposed method technically pertains to the allocation of occupation requests within a railway schedule, we sometimes refer to these requests as 'trains' for the sake of readability.

In this approach, the arrival times and required occupation durations of requests are subject to uncertainties and can thus be effectively modelled using queueing systems (see Section 2). Consequently, the estimated value of the average queue length per route during the timetabling process provides an estimation of infrastructure quality. To ensure adequate quality, thresholds for this queue length can be established (see Section 3.4). The maximum number of requests, for which the timetabling process meets this quality requirement for a given traffic distribution, is denoted as the timetable capacity of the railway infrastructure.

The following section outlines the data requirements necessary to determine the timetable capacity of a railway station.

3.2. Required data

To determine the timetable capacity of a railway station using the introduced model, some data needs to be provided in order to obtain the results.

First, the time horizon U, which describes the analysed time span in minutes, needs to be specified. Furthermore, the station infrastructure I is required. This consists of a set of routes R, a set of route nodes G, a set of platform tracks P, and a function describing the conflicts between routes and their service processes:

$$C : R \times \Pi \times R \times \Pi \to \{0,1\}, (r,\pi,r',\pi') \mapsto \begin{cases} 1 & \text{, conflict between } r \text{ in } \pi \text{ and } r' \text{ in } \pi' \\ 0 & \text{, otherwise.} \end{cases} \tag{15}$$

Notably, $\Pi = \{1,2,4\}$ describes the three service processes: entry blocking (1), stopping (2), and departure blocking (4) for a route.

Furthermore, additional characteristics for each route need to be provided. Specifically, the arrival rate λ_r and the service rates $\mu_{r,\text{entr}}$, $\mu_{r,\text{stop}}$, and $\mu_{r,\text{dep}}$ are required as input to the CTMC.

However, since we are optimizing for the maximum number of requests $n_{\max}$, the total train count n_{total} changes during the optimization process (see Section 3.5). Hence, the number of trains per route n_r, used for calculating $\lambda_r = n_r/U$, is determined using a base count factor ψ_r with $n_r = \psi_r \cdot n_{\text{total}}$. This factor ψ_r thus describes the distribution of traffic across individual routes and is assumed to be fixed for a given instance. Hence, we denote the collection $\{\psi_r | r \in R\}$ of all base count factors as *traffic distribution*.

In addition, the service times $b_{r,\text{entr}}, b_{r,\text{stop}}, b_{r,\text{dep}}$ for each service step need to be provided in order to calculate the corresponding service rates. An infrastructure operator would typically determine these using blocking time theory on microscopic infrastructure data.

Since different rolling stock configurations may traverse a given route, a set of rolling stock configurations T needs to be specified. This work utilizes the different routes to decompose the queueing system of the considered junction; hence, the required arrival and service processes are formulated per route and thus always describe a weighted average over the rolling stock configurations. To enhance readability, we assume that the aggregation of the different rolling stock configurations per route has been completed prior to the procedure described in this work. Therefore, we differentiate only between routes.

To establish the queue-length threshold per route (see Section 3.4), it is necessary to specify the share of passenger trains on each route $p_{\text{pt},r}$. Moreover, to describe non-exponential arrival or service processes, the corresponding variation coefficients must also be included.

Table 2 lists all the variables and notations, highlighting the required input data as well as the default values used in this work. Notably, the default values correspond to the coefficients of variation for both, the arrival and the service processes. These default values are recommended in general planning rules (see f.e. VIA Consulting & Development GmbH 2021) for the determination of timetable capacity.

3.3. Acquiring queue-length estimations

The formulated Continuous-Time Markov Chain (see Section 2.4) $MC = (S,T)$ can be used to calculate the steady-state probabilities $p(s)$ of every state $s \in S$. In combination with the queue lengths $\mathrm{q}_r(s)$ per route r for each state, the expected queue length per route

$$L_r = \sum_{s \in S} p(s) \cdot \mathrm{q}_r(s) \tag{16}$$

can be obtained.

Table 2
List of variables and notations.

Variable/notation	Description	Input data	Default value
Infrastructure and Time Horizon			
U	Time horizon	✓	–
I	Infrastructure	✓	–
R	Set of routes	✓	–
G	Set of route nodes	✓	–
P	Set of platform tracks	✓	–
C	Conflicts between routes	✓	–
n_{total}	Total number of requests in time horizon	–	–
Queueing and Service Process			
m	Number of waiting positions	–	6 (excl. Section 4.3)
λ	Arrival rate	–	–
μ	Service rate	–	–
ρ	Occupation ratio ($\rho = \lambda/\mu$)	–	–
L	Average length of the queue	–	–
p_{loss}	Probability of loss	–	–
$p(s)$	Steady-state probability distribution over states	–	–
Scaling Variables			
γ	Scaling factor in approximation formula	–	–
c	Intermediate variable for calculating γ	–	–
s	Number of parallel service channels in the approximation formula	–	fixed at 1
v_A	Coefficient of variation of the arrival process	–	–
v_S	Coefficient of variation of the service process	–	–
Route-Specific Parameters			
L_r	Estimated queue-length per route	–	–
ψ_r	Base count factor per route	✓	–
n_r	Number of trains per route in a specified time horizon ($n_r = \psi_r \cdot n_{\text{total}}$)	–	–
$b_{r,\text{entr}}$, $b_{r,\text{stop}}$, $b_{r,\text{dep}}$	Average service times for entry blocking, stopping, and departure blocking processes respectively	✓	–
$v_{A,r}$	Coefficient of variation for the arrival process on a specific route	✓	0.8
$v_{S,r,\text{entr}}$, $v_{S,r,\text{stop}}$, $v_{S,r,\text{dep}}$	Variation coefficient for entry blocking, stopping, and departure blocking processes respectively	✓	0.3
$v_{S,r}$	Combined coefficient of variation for all three service processes on a specific route	–	–
Additional CTMC Parameters			
π_r	Status in the service process for route r (e.g., 0: no service; 1: entry; 2: stopping; 3: waiting to depart; 4: departing)	–	–
$\pi_r(s)$	Status of route r in state s	–	–
$q_r(s)$	Queue length of route r in state s	–	–
m_r	Maximum queue length (number of waiting positions) for route r	–	–
λ_r	Arrival rate for a specific route ($\lambda_r = n_r/U$)	–	–
$\mu_{r,\text{entr}}$, $\mu_{r,\text{stop}}$, $\mu_{r,\text{dep}}$	Service rates for entry blocking, stopping, and departure blocking processes respectively	–	–
D	Artificial constant transition rate used to model random decisions between conflicting routes.	–	600

While some research utilizes the underlying Kolmogorov equations to directly obtain the state probabilities via matrix–vector equations (Weik, 2020) or iterative formulas (Nießen, 2008), recent publications use probabilistic model checking (Emunds and Nießen, 2024) to directly calculate the estimated queue lengths per route. In this work, the latter approach has been adopted by implementing Python scripts that utilize the Python interface `stormpy` (Junges and Volk, 2023) for the probabilistic model checker `STORM` (Hensel et al., 2022).

The expected queue lengths per route L_r obtained in this manner needs to be scaled with the approximation formula (see Section 2.2) and compared to a threshold to ensure sufficient timetabling quality.

3.4. Defining thresholds

In published research, various thresholds have been established to ensure adequate quality during the planning phase. While the International Union of Railways (UIC) (2013) introduces maximum occupation rates, the largest European infrastructure manager, DB InfraGO (2022), employs thresholds based on the expected queue length.

The formula to determine the maximum admissible queue length, $L_{\text{limit, r}}$, was developed in Schwanhäußer and Schultze (1982) by conducting a survey under railway dispatchers to obtain thresholds for the quality of a railway line. As each route could be seen as a single service channel within the railway station, we utilize this threshold independently for each route in this work.

Depending on the proportion of passenger traffic $p_{\text{pt},r}$ on a route r, the limit is defined as

$$L_{\text{limit, r}} = 0.479 \cdot \exp(-1.3 \cdot p_{\text{pt},r}) \tag{17}$$

. Computing this limit for railway stations with passenger traffic only ($p_{\text{pt},r} = 1$), results in a maximum queue-length of $L_{\text{limit, r}} = 0.1305$ per route.

In this study, this limit is compared to the estimated queue length per route L_r using the quality factor

$$\text{qf}_r = \frac{L_r}{L_{\text{limit, r}}}, \tag{18}$$

which quantifies the quality on route r. The quality threshold is thus defined by a quality factor of $\text{qf}_r = 1$.

In the subsequent section, the procedure for determining the timetable capacity of a railway station by comparing the estimated queue lengths per route (see Section 3.3) to the here introduced limits $L_{\text{limit},r}$ is described in detail.

3.5. Procedure

Given an instance, describing a railway station with the Infrastructure $I = (R, G, P, C)$ over the time horizon U, and the desired train mix, in detail the share of passenger traffic $p_{\text{pt},r}$ and the base count factor ψ_r per route r, the timetable capacity can be determined with the following general steps.

1. Determine the required minimum headway times h_{r_i,r_j} for every sequence of routes in their arrival or departure process to/from the corresponding platform track $p_{r_i,r_j} \in P$ in all route nodes G.
2. Determine the required stopping times per route.
3. Determine the maximum number of trains $n_{\max}$ in multiple steps for each considered n_{total}:
    - 3.1. Calculate the number of trains per route $n_r = n_{\text{total}} \cdot \psi_r$.
    - 3.2. Determine the required mean service times $b_{r,\text{entr}}$, $b_{r,\text{stop}}$, and $b_{r,\text{dep}}$ on the routes by averaging the minimum headway and stopping times of the routes.
    - 3.3. Determine variation coefficients $v_{A,r}$ for the arrival processes on the routes.
    - 3.4. Determine variation coefficients $v_{S,r,\text{entr}}, v_{S,r,\text{stop}}, v_{S,r,\text{dep}}$ for the service processes on the routes.
    - 3.5. Formulate the CTMC in a `PRISM` model file.
    - 3.6. Obtain L_r for every route from the CTMC and scale according to Section 2.2.
    - 3.7. Calculate the quality factor qf_r for each route.
    - 3.8. Determine $\text{qf}_{\max} = \max_r(\text{qf}_r)$.
    - 3.9. Select next considered value of n_{total} if precision has not been met.

To calculate the quality factors qf_r in step 3.6, the expected queue lengths L_r need to be obtained from the formulated CTMC. For this purpose, the CTMC is described in a `PRISM` model file (`.pm`) (Kwiatkowska et al., 2011) and evaluated with the probabilistic model-checking tool `stormpy` (Junges and Volk, 2023).

One evaluation step can take several seconds up to minutes, depending on the size of the CTMC. It is therefore crucial to restrict the number of iterations of step 3 by utilizing efficient methods for determining the next considered n_{total} (step 3.9).

The function

$$\phi : \mathbb{R} \to \mathbb{R},\ n_{\text{total}} \mapsto \phi(n_{\text{total}}) = \text{qf}_{\max} - 1 \tag{19}$$

can be formulated to describe the difference between the maximum quality factor $\text{qf}_{\max}$ in step 3.8 and the desired value of 1.

With ϕ, determining timetable capacity can be formulated as a root-finding problem, i.e., finding $n^*_{\text{total}} = n_{\max}$ such that $\phi(n^*_{\text{total}}) = 0$. This problem can be solved numerically, for example with Brent's method (Brent, 1973). In this work, we utilize the `scipy` (Virtanen et al., 2020) implementation of Brent's method with a relative tolerance of rtol = 0.001 and an absolute tolerance of xtol = 0.001 as the precision. Details about the implementation and the precision parameters are available at the `scipy` documentation (SciPy Developers, 2023).

The next section discusses the scalability of this procedure.

3.6. Scalability

As mentioned in the previous section, the computational requirements of the introduced procedure mainly depend on two factors: The number of necessary iterations of step 3 and the size of the CTMC model. The necessary iterations can be minimized with the discussed numerical solution to a root-finding problem, but the required computational power for a single iteration depends on the size of the CTMC model, in particular on the number of elements in the state set.

Recall the definition of the state set

$$S = \left\{ \left(\pi_1, \mathrm{q}_1, \dots, \pi_k, \mathrm{q}_k \right) \mid \left(\pi_r, \mathrm{q}_r \right) \in \Pi_r \times \{0, \dots, m\} \right\}, \tag{20}$$

from Section 2.4. All states $s \in S$ therefore correspond to a selection of $\pi_r \in \Pi_r \subseteq \{0, \dots, 4\}$ and $\mathrm{q}_r \in \{0, \dots, m\}$ for each route r out of k possible routes, i.e. k choices between $\left|\Pi_r\right| \leq 5$ and $m+1$ elements. Therefore, the upper bound

$$|S| \leq \left|\Pi_r\right|^k \cdot (m+1)^k \leq (5\ m+5)^k \tag{21}$$

to the number of states $|S|$ can be formulated.

Hence, the number of states grows exponentially with the number of routes k. Furthermore, for a fixed number of routes k^*, the number of states $|S|$ is bounded by a polynomial in $\mathcal{O}\left((5\ m)^{k^*}\right)$. Note that the factor 5^{k^*} describes the number of π_r-combinations for all k^* routes and it can be further restricted depending on the conflict function C in the route nodes of the analysed railway station. In comparison, the factor $(m+1)^{k^*}$ is more tight since every queue length in $\{0, \dots, m\}$ is obtainable for every route, independent of the service status π_r and queue length of other routes.

Therefore, to enhance the scalability of the introduced model, it is preferable to decrease the number of modelled routes k, for example by enhancing the abstraction level or decomposing one station into multiple track groups. If the number of routes is fixed, another approach to reduce the state set is to minimize the modelled queue length per route, possibly individually for every route and in dependence to other influencing factors.

Given the substantial computational requirements, a small representative railway station has been chosen to illustrate the application of the proposed procedure in the next section.

4. Case study

This case study analyses the performance capabilities of the station depicted in Fig. 1. One aspect of the analysis examines two different operational scenarios, varying the stopping positions of two routes (Section 4.1). In a subsequent sensitivity-analysis (Section 4.2), the influence of the duration of the different service processes on timetable capacity is tested. Furthermore, Section 4.3 investigates models with different values for the maximum number of waiting positions m regarding their performance.

4.1. Operational scenario analysis

In this analysis, two different operational scenarios are compared for the passenger railway station described in Fig. 1. The corresponding infrastructure example can be found in Fig. 6.

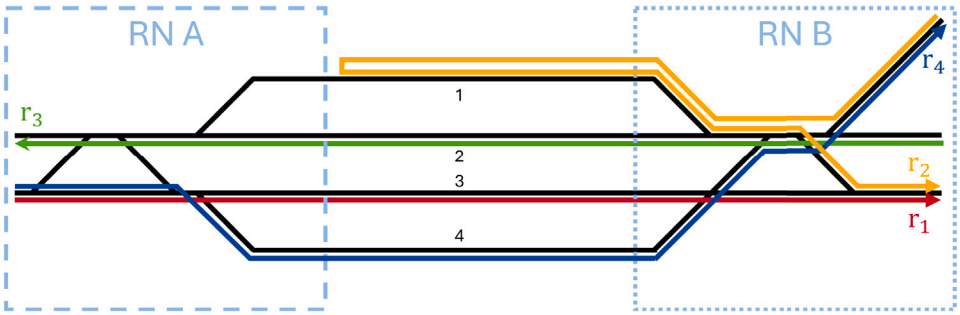


Fig. 6. The railway station infrastructure with swapped stopping platform tracks for routes r_2 and r_4.

In the original Scenario 1, route r_2 requests a stop at platform track 4 and r_4 at platform track 1. In Scenario 2, these two tracks have been swapped in order to reduce the number of conflicts during the entry process of all routes. Further details can be found in Table 3.

Notably, the track swap between the two routes leads to different entry and departure conflicts for all routes. In Scenario 2, for example, r_3 can depart without any conflicting route, whereas it is excluded by r_4 in Scenario 1. However, in Scenario 2, r_3 now conflicts not only with the entry of r_2 but also with its departure.

The service times for the entry, departure, and stopping processes are assumed to be equal for both scenarios. While the service times would typically be calculated using microscopic infrastructure and rolling stock data for a real-world application, we omit

Table 3
Route node and platform usage per scenario.

Route	Scenario 1					Scenario 2				
	Entry route node	Platform track	Departure route node	Entry conflicts	Departure conflicts	Entry route node	Platform track	Departure route node	Entry conflicts	Departure conflicts
r_1	A	3	B	r_4	r_2	A	3	B	r_4	r_2 , r_4
r_2	B	4	B	r_1 , r_3 , r_4	r_1	B	1	B	r_3 , r_4	r_1 , r_3 , r_4
r_3	B	2	A	r_2 , r_4	r_4	B	2	A	r_2 , r_4	–
r_4	A	1	B	r_1 , r_3	r_2 , r_3	A	4	B	r_1	r_1 , r_2 , r_3

Table 4
Blocking and stopping times in minutes.

Route	Entry blocking time	Stopping time	Departure blocking time
r_1	0.5	2.0	0.5
r_2	1.5	4.0	1.0
r_3	0.5	2.0	0.5
r_4	1.0	2.0	1.0

those details in this work and assume values instead. More information on the calculation of entry and departure blocking times can be found in Hansen and Pachl (2014). Table 4 lists the assumed service times in minutes.

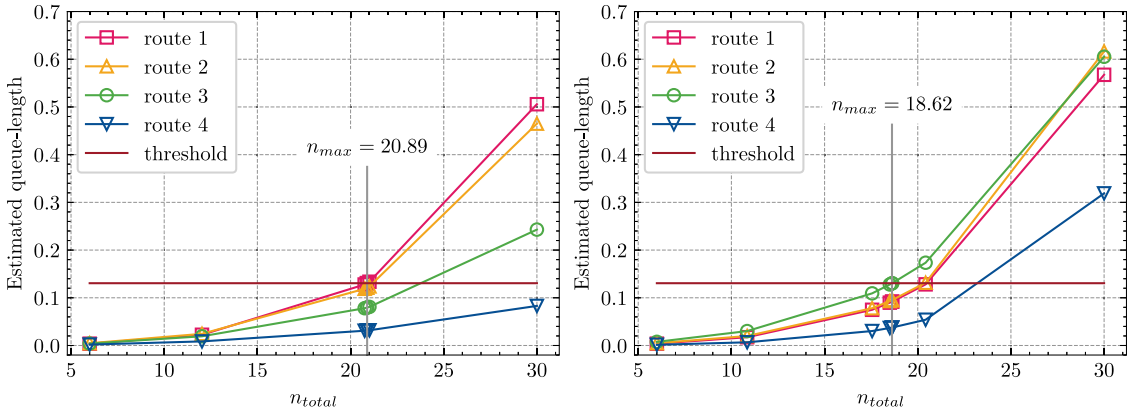
Notably, the arrival and departure blocking times for routes r_1 and r_3 are assumed to be 0.5 min each, as these routes do not need to traverse any switch branches. Furthermore, the stopping time of r_4 is assumed to be 4 minutes due to the necessity of turning on the platform track.

For all routes, only passenger traffic ($p_{\mathrm{pt},r} = 1$) has been assumed. Furthermore, we assume a traffic distribution of

$$\psi_r = \begin{cases} \frac{1}{3} & ;\text{if } r \in \{1,3\}, \\ \frac{1}{6} & ;\text{otherwise}, \end{cases} \tag{22}$$

thereby doubling the amount of traffic on the main line routes r_1 and r_3.

Using the default values for the variation coefficients of the arrival and service processes (see Section 3.2) and $m = 6$ waiting positions per route, the procedure from Section 3.5 has been performed for both scenarios. The resulting queue lengths (scaled according to Section 2.2) at each step of the method are depicted in Fig. 7 and compared with the threshold $L_{\mathrm{limit,\,r}} = 0.1305$ for passenger stations.



(a) Original Route Configuration (b) Swapped Stopping Tracks

Fig. 7. Comparison of the queue-length estimations for all routes in both scenarios.

While Fig. 7(a) depicts the estimated queue lengths for the original Scenario 1, Fig. 7(b) shows the estimated queue lengths for the scenario with swapped platform tracks.

For Scenario 1 (Fig. 7(a)), the longest average queue is estimated for route r_1, closely followed by the average queue length of r_2. Route r_3 falls in the mid-field, while route r_4 appears to be the least occupied. Therefore, the bottleneck can be assumed to be on route r_1, resulting in a maximum capacity of $n_{\max} = 20.89$ trains per hour.

In contrast, for Scenario 2 with swapped platform track stopping positions, r_3 is the most utilized route. While r_1 and r_2 follow closely, r_4 remains the least occupied route. Utilizing the threshold for full-passenger-traffic routes (Section 3.4), the timetable capacity of $n_{\max} = 18.62$ trains per hour can be concluded for this scenario.

The differences in the performance capabilities of the two scenarios can be explained by the high number of conflicts resulting from the re-routing of r_2. In the first scenario, the entry of route r_3 is only influenced during the entry of r_2, and the departure of route r_1 is influenced by both the entry and departure of r_2.

In the second scenario, the entry process of r_3 is affected by both the entry and departure processes of r_2. While the influences on r_1 decrease, the increased conflicts of r_2 with the entry service process of r_3 significantly impact the queue length of r_3, making it the bottleneck for the entire station.

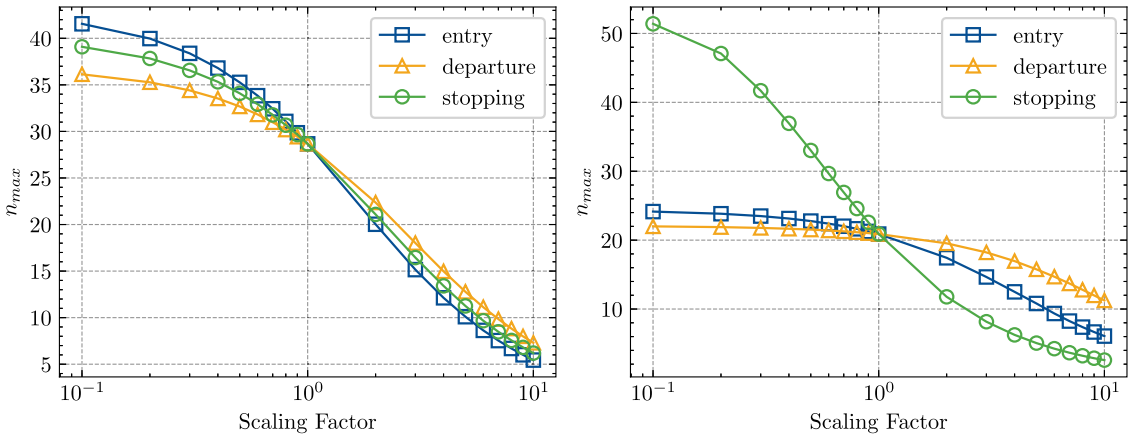
An infrastructure operator should hence opt for the first scenario to fully utilize the stations performance capabilities.

4.2. Influence of service processes

From the preceding results, the hypothesis that the duration of the entry blocking process is more influential to the timetable capacity of the railway junction, than the departure blocking process can be made. In this Section, a sensitivity analysis of the railway junction from scenario 1 to the duration of the service processes is performed.

For this, the service times $b_{r,\text{entr}}$, $b_{r,\text{stop}}$, $b_{r,\text{dep}}$ for the entry blocking, stopping and departure blocking process have each been scaled individually with a factor in the set $\{0.1, 0.2 \dots, 0.9, 1, 2, \dots, 9, 10\}$. The factor of both remaining service times has been set to 1. With this setting, the influence of the nominal value of the entry blocking time, stopping time or departure blocking time can be analysed. Note that the presented method utilizes an average service duration for each service step per route and does not distinguish between different train sequences. To obtain detailed results regarding headway times between specific conflicting train sequences, other methods, such as microscopic simulation tools, are to be preferred.

Fig. 8 shows the resulting timetable capacity for the different scaling factors per service type.



(a) Uniform Service Times (b) Original Service Time Configuration

Fig. 8. Sensitivity of the timetable capacity to service time scaling.

To identify the sensitivity independently of the service time settings of the example railway station, Fig. 8(a) assumes a base service time of 1 minute for every service process and scales it according to the scaling factor.

This sensitivity analysis reveals a significantly higher influence of the entry process duration, followed by the stopping duration, and finally the duration of the departure process.

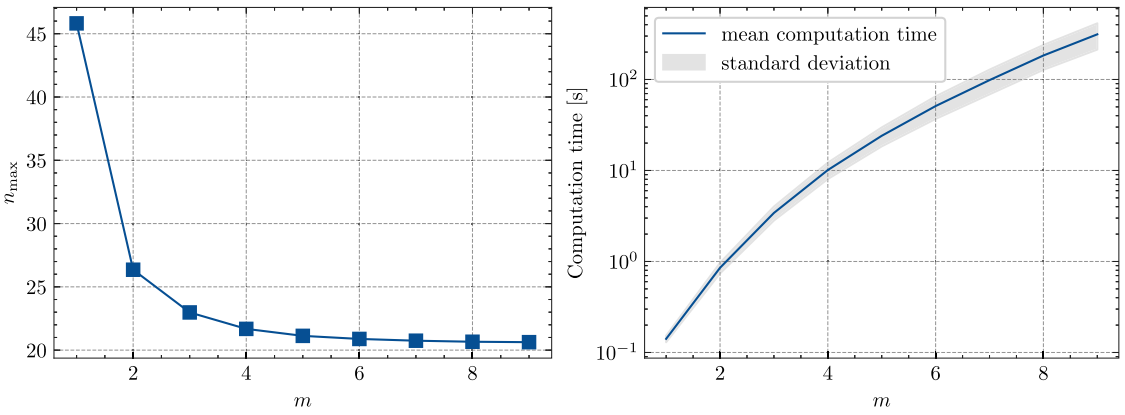
In comparison, the scaling of the stopping time has the highest impact for the original service time configuration (Fig. 8(b)). This can be explained by its significantly higher base setting of 2–4 minutes in the original configuration, compared to 0.5–1.5 minutes for the entry and departure blocking times. However, this assumption seems reasonable to allow a sufficient stopping duration for passenger exchange.

To increase the capacity of the example railway station, infrastructure operators should therefore focus on limiting the stopping time on the platform track and ensure sufficient railway control systems for a quick release of the entry route, in order to shorten the required blocking times in the entry process.

In the final analysis, the influence of the number of waiting positions m to the approximation quality of the introduced approach is investigated.

4.3. Dependence on the number of waiting positions

The performance of the modelled queueing system is not only influenced by the arrival and service process, but additionally by the maximum number of requests that may be queued up for service (see also Section 2.1). While a constant number of $m = 6$ waiting positions per route has been assumed in the previous analysis, this section examines the impact of different numbers of waiting positions.



(a) Timetable Capacity (b) Computation Times

Fig. 9. Comparison of Capacity and Computation Times for Different Number of Waiting Positions m.

For this, the timetable capacity of the original scenario 1 has been calculated for any $m \in \{1, \dots, 9\}$. Fig. 9 shows the resulting capacity and computation times for the different values.

The size of the state set S grows with the number of waiting positions per route, as every possible value combination must be considered. Consequently, the computation time increases substantially with increasing m (Fig. 9(b)).

In contrast, for sufficiently small occupation ratios ρ, the probability of reaching a high number of parallel waiting requests per route might be insignificantly small, resulting in only incremental improvements for high values of m. For the analysed station example, the relative difference in timetable capacity between $m = 1$ and $m = 6$ exceeds 119%, whereas the difference between $m = 6$ and $m = 9$ is less than 2% of the value at $m = 6$.

Nonetheless, a higher number of available waiting slots increases the accuracy of the method. In order to fulfil the threshold to the expected average queue length (Section 3.4), the probability of experiencing such high queue lengths within the timetabling process needs to be negligible. Therefore, the determined capacity value decreases, in order to reduce the occupation rate per route and hence the probability of conflicting requests, directly influencing the amount of waiting trains.

In summary, a small number of waiting positions per route may result in faster solving times but leads to a significant overestimation of timetable capacity with the introduced approach.

5. Conclusion

In this work, a queueing-based method to evaluate the performance of an entire railway station has been introduced. For this purpose, the station infrastructure has been partitioned along its routes, while the service process of each route has been further decomposed into entry, stopping, and departure processes.

This approach has been used to develop a novel Continuous-Time Markov Chain (CTMC) model that describes the timetabling process in railway stations (Section 2). With its multidimensional state-space, conflicts arising in all three service parts of different routes can be modelled. Additionally, the average number of trains per route waiting for their scheduled infrastructure occupation can be analysed.

By applying probabilistic model-checking, this number can be computed based on a fixed total train number, considering fixed traffic distributions and infrastructure properties (Section 3). A comparison with the formulated quality thresholds allows for the assessment of route-based quality factors to identify bottlenecks in the analysed infrastructure. Furthermore, numerical root-finding methods can be utilized to determine the timetable capacity of a given railway station's infrastructure.

The applicability of the introduced approach to practical questions posed by infrastructure operators has been discussed through a case study presented in Section 4. In this case study, the performance limits of two different operating procedures on an example infrastructure were analysed. The influence of different phases in the service life-cycle of a train within the railway station was also investigated. Finally, the introduced approach was applied for different numbers of waiting positions, highlighting the high computational requirements for an increasing number of waiting positions.

This observation supports the theoretical analysis of scalability discussed in Section 3.6, which associates computational complexity with the size of the state space in the CTMC. The state space not only depends on the number of waiting positions but also grows exponentially with an increase in the number of routes within the railway station. While the current implementation is effective for smaller stations, tackling more complex nodes will require more sophisticated abstraction methods.

Other than enhancing abstraction levels or further decomposing the station, employing adaptive techniques to individually select the number of modelled waiting positions for each route might significantly enhance the scalability of the proposed model. This approach would involve estimating the error associated with formulations that use a fixed number of waiting positions per route. Such an estimation would facilitate the establishment of lower bounds on timetable capacity, thereby providing soundness

guarantees. Consequently, this would enable the determination of the minimum number of waiting positions per route necessary to achieve a specified level of accuracy in the method.

Further exacerbating the need for better scalability is the issue of including an occupancy-based track assignment. In the current solution, the track-choice is fixed for every route and therefore for the traffic requests from train operators. Obtaining some flexibility here would certainly extend applicability, but it would correspond to introducing additional states in the CTMC formulation to monitor the occupation status of every track. Utilizing its occupation status, a track could be assigned when dequeueing a request to start its entry process. The choice between multiple available tracks could potentially be sorted by given track priorities per route. This procedure would not only necessitate to include information about track occupation into the state space and to reformulate transitions to reserve/repeal tracks when starting the entry process or terminating the departure process. In addition, conflicts between routes depend on the chosen tracks, hence further research would need to carefully model the dynamic parallel service capabilities of entry/arrival processes with flexible track assignments.

Regarding the accuracy, two main issues could be addressed. Firstly, the description of non exponential service processes is currently modelled by scaling results with the formula in Section 2.2. By utilizing phase-type distributions for these processes, more accurate models could be built. However, this would also require additional states in the CTMC, which again negatively influences the scalability.

The second main accuracy extension is to further develop implemented quality thresholds. In its current formulation, the limit formula (17) has been obtained with railway line performance in mind. While it may be applicable to the individual routes of a railway station or junction, further research is needed to evaluate its precision for those application scenarios. In addition, current developments towards more synchronized timetables might necessitate the formulation of enhanced quality metrics, which could address challenges posed by concentrated requests within short time intervals and ensure that the infrastructure allows for synchronized dwell times.

While the introduced method is designed to cope with uncertainties in the long-term strategic planning phase through its focus on abstract stochastic processes, it is not suitable to determine the performance of specific timetables. For such analyses, a simulation approach would be more appropriate, as it can provide detailed insights into timetable dependencies and evaluate performance metrics under various operational scenarios.

In summary, while future research can significantly enhance this initial method for comprehensive station performance analysis, the introduced model already serves as a valuable decision support tool for infrastructure managers. It extends existing methodologies by enabling a combined analysis of entire railway station infrastructure in the strategical planning phase.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Germany – 2236/2. Computational Experiments were performed with computing resources granted by RWTH Aachen University under projects rwth1413 and rwth1635.

Appendix. Combining coefficients of variation

The coefficient of variation of a probability distribution X is described as

$$\upsilon_X = \frac{\mathrm{std}(X)}{\mathrm{E}(X)}, \tag{23}$$

with the standard deviation $\mathrm{std}(X)$ and expected value $\mathrm{E}(X)$. For multiple stochastic independent distributions $X_1, \dots, X_n$ the expected value of the combination can be given as

$$\mathrm{E}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathrm{E}(X_i) \tag{24}$$

and the Variance as

$$\mathrm{Var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathrm{Var}(X_i). \tag{25}$$

Let the expected value $\mathrm{E}(X_i)$ and the coefficients of variation υ_{X_i} be given, then the standard deviation can be calculated with

$$\mathrm{std}(X_i) = \upsilon_{X_i} \cdot \mathrm{E}(X_i). \tag{26}$$

Using $\mathrm{Var}(X_i) = \mathrm{std}(X_i)^2$, the combined variance can be given as

$$\mathrm{Var}(X) = \sum_{i=1}^{n} (\upsilon_{X_i}^2 \cdot \mathrm{E}(X_i)^2). \tag{27}$$

Hence, the combined coefficient of variation is obtained by

$$v_X = \frac{\mathrm{std}(X)}{\mathrm{E}(X)} = \frac{\sqrt{\sum_{i=1}^{n} (v_{X_i}^2 \cdot \mathrm{E}(X_i)^2)}}{\sum_{i=1}^{n} \mathrm{E}(X_i)}. \tag{28}$$

Data availability

Data will be made available on request.

References

Abril, M., Barber, F., Ingolotti, L., Salido, M., Tormos, P., Lova, A., 2008. An assessment of railway capacity. Transp. Res. Part E: Logist. Transp. Rev. 44 (5), 774–806.

Armstrong, J., Preston, J., 2017. Capacity utilisation and performance at railway stations. J. Rail Transp. Plan. Manag. 7 (3), 187–205.

Bešinović, N., Goverde, R.M.P., 2018. Capacity assessment in railway networks. In: Borndörfer, R., Klug, T., Lamorgese, L., Mannino, C., Reuther, M., Schlechte, T. (Eds.), Handbook of Optimization in the Railway Industry. vol. 268, Springer International Publishing, pp. 25–45.

Brent, R., 1973. Algorithms for Minimization Without Derivatives. Prentice-Hall, Englewood Cliffs NJ.

Burdett, R.L., 2015. Multi-objective models and techniques for analysing the absolute capacity of railway networks. European J. Oper. Res. 245 (2), 489–505.

Burdett, R.L., Kozan, E., 2006. Techniques for absolute capacity determination in railways. Transp. Res. Part B: Methodol. 40 (8), 616–632.

Bychkov, I., Kazakov, A., Lempert, A., Zharkov, M., 2021. Modeling of railway stations based on queuing networks. Appl. Sci. 11 (5), 2425.

Corman, F., Henken, J., 2022. Estimating aggregate railway performance from realized empirical data: Literature review, a test case and a research roadmap. J. Rail Transp. Plan. Manag. 22, 100316.

DB InfraGO, 2022. Richtlinie 405 Fahrwegkapazität. (German only).

Emunds, T., Nießen, N., 2024. Evaluating railway junction infrastructure: A queueing-based, timetable-independent analysis. Transp. Res. Part C: Emerg. Technol. 165, 104704.

Fischer, K., Hertel, G., 1990. Bedienungsprozesse im Transportwesen: Grundlagen und Anwendungen der Bedienungstheorie. Transpress-Verlag (German only).

Graffagnino, T., 2012. Ensuring timetable stability with train traffic data. In: Brebbia, C.A., Tomii, N., Tzieropoulos, P., Mera, J.M. (Eds.), Computers in Railways XIII. In: WIT Transactions on The Built Environment, WIT PressSouthampton, UK, pp. 427–438.

Hansen, I., Pachl, J., 2014. Railway Timetabling and Operations: Analysis, Modelling, Optimisation, Simulation, Performance, Evaluation. Eurail Press.

Harrod, S., 2009. Capacity factors of a mixed speed railway network. Transp. Res. Part E: Logist. Transp. Rev. 45 (5), 830–841.

Hensel, C., Junges, S., Katoen, J.-P., Quatmann, T., Volk, M., 2022. The probabilistic model checker storm. Int. J. Softw. Tools Technol. Transf. 24 (4), 589–610.

Hertel, G., 1986. Analytische Modellierung und formelmässige Behandlung von Standard-Bedienungssystemen mit störanfälligen Kanälen (Ph.D. thesis). Hochschule für Verkehrswesen "Friedrich List", Dresden.

Huisman, T., Boucherie, R.J., Van Dijk, N.M., 2002. A solvable queueing network model for railway networks and its validation and applications for the Netherlands. European J. Oper. Res. 142 (1), 30–51.

International Union of Railways (UIC), 2013. Capacity, second ed. In: UIC Code, (406), UIC, Paris, France.

Jensen, L.W., Schmidt, M., Nielsen, O.A., 2020. Determination of infrastructure capacity in railway networks without the need for a fixed timetable. Transp. Res. Part C: Emerg. Technol. 119, 102751.

Junges, S., Volk, M., 2023. Stormpy - Python bindings for storm: Version 1.7.0. URL https://moves-rwth.github.io/stormpy/index.html.

Kazakov, A., Lempert, A., Zharkov, M., 2023. An approach to railway network sections modeling based on queuing networks. J. Rail Transp. Plan. Manag. 27, 100404.

Kendall, D.G., 1953. Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov chain. Ann. Math. Stat. 338–354.

Kwiatkowska, M., Norman, G., Parker, D., 2011. PRISM 4.0: Verification of probabilistic real-time systems. In: Gopalakrishnan, G., Qadeer, S. (Eds.), Proc. 23rd International Conference on Computer Aided Verification. CAV'11, In: LNCS, vol. 6806, Springer, pp. 585–591.

Liang, J., Martin, U., Cui, Y., 2017. Increasing performance of railway systems by exploitation of the relationship between capacity and operation quality. J. Rail Transp. Plan. Manag. 7 (3), 127–140, URL https://linkinghub.elsevier.com/retrieve/pii/S2210970617300458.

Liao, Z., Li, H., Miao, J., Corman, F., 2021. Railway capacity estimation considering vehicle circulation: Integrated timetable and vehicles scheduling on hybrid time-space networks. Transp. Res. Part C: Emerg. Technol. 124, 102961.

Mussone, L., Wolfler Calvo, R., 2013. An analytical approach to calculate the capacity of a railway system. European J. Oper. Res. 228 (1), 11–23.

Nießen, N., 2008. Leistungskenngrößen Für Gesamtfahrstraßenknoten (Ph.D. thesis). RWTH Aachen University (German only).

Nießen, N., 2013. Waiting and loss probabilities for route nodes. In: International Conference on Railway Operations Modelling and Analysis. RailCopenhagen.

Potthoff, G., 1962. Verkehrsströmungslehre - Die Zugfolge auf Strecken und in Bahnhöfen. Trans Press (German only).

Schmidt, C., Martin, U., 2010. Erhöhung der Effektivität und Transparenz bei Leistungsuntersuchungen mit Simulationsverfahren. Eisenbahntech. Rundsch. (ETR) 07-08/2010, 463–468, (German only).

Schmitz, C., Weik, N., Zieger, S., Nießen, N., Schmeink, A., 2017. Markov models for the performance analysis of railway networks. In: International Conference on Railway Operations Modelling and Analysis. RailLille, Lille, France, p. 23.

Schwanhäußer, W., 1974. Die Bemessung der Pufferzeiten im Fahrplangefüge der Eisenbahn (Ph.D. thesis). Verkehrswissenschaftliches Institut der Rheinisch-Westfälischen Technischen Hochschule Aachen (German only).

Schwanhäußer, W., 1978. Die ermittlung der leistungsfähigkeit von großen fahrstraßenknoten und von teilen des eisenbahnetzes. Arch. Für Eisenbahntechnik 1978 (38), 7–18 (German only).

Schwanhäußer, W., Schultze, K., 1982. Ermittlung von Qualitätsmaßstäben für die Berechnung der Leistungsfähigkeit eines Streckenabschnittes und Entwicklung eines Rechenverfahrens zur Ermittlung von Endverspätungen: Forschungsarbeit für die Deutsche Bundesbahn. (German only).

SciPy Developers, 2023. scipy.optimize.brentq — SciPy v1.11.4 manual. Available at: https://docs.scipy.org/doc/scipy-1.11.4/reference/generated/scipy.optimize.brentq.html. (Accessed 25 June 2025).

Spanninger, T., Büchel, B., Corman, F., 2023. Train delay predictions using Markov chains based on process time deviations and elastic state boundaries. Mathematics 11 (4), 839.

VIA Consulting & Development GmbH, 2021. LUKS Handbuch (Software Handbook, German only).

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., SciPy 1.0 Contributors, 2020. SciPy 1.0: Fundamental algorithms for scientific computing in Python. Nature Methods 17, 261–272.

Weik, N., 2020. Long-Term Capacity Planning of Railway Infrastructure – A Stochastic Approach Capturing Infrastructure Unavailability (Ph.D. thesis). RWTH Aachen University.
Weik, N., 2022. Macroscopic traffic flow in railway systems — A discussion of the applicability of fundamental diagrams. J. Rail Transp. Plan. Manag. 23, 100330.
Weik, N., Nießen, N., 2017. A quasi-birth-and-death process approach for integrated capacity and reliability modeling of railway systems. J. Rail Transp. Plan. Manag. 7 (3), 114–126.
Wendler, E., 1999. Analytische Berechnung Der Planmässigen Wartezeiten Bei Asynchroner Fahrplankonstruktion (Ph.D. thesis). RWTH Aachen University.
Wendler, E., 2007. The scheduled waiting time on railway lines. Transp. Res. Part B: Methodol. 41 (2), 148–158.
Zieger, S., Weik, N., Nießen, N., 2018. The influence of buffer time distributions in delay propagation modelling of railway networks. J. Rail Transp. Plan. Manag. 8 (3–4), 220–232.
Zwaneveld, P.J., Kroon, L.G., Van Hoesel, S.P., 2001. Routing trains through a railway station based on a node packing model. European J. Oper. Res. 128 (1), 14–33.