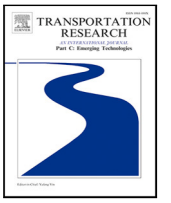
Transportation Research Part C 165 (2024) 104704



Contents lists available at ScienceDirect

Transportation Research Part C

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





Evaluating railway junction infrastructure: A queueing-based, timetable-independent analysis

Tamme Emunds *, Nils Nießen

Institute of Transport Science, RWTH Aachen, Mies-van-der-Rohe Straße 1, Aachen, 52074, Germany

ARTICLE INFO

Keywords:
Junction capacity
Markov Chain
Queueing
Probabilistic model-checking
Railway infrastructure

ABSTRACT

Many infrastructure managers have the goal to increase the capacity of their railway infrastructure due to an increasing demand. While methods for performance calculations of railway line infrastructure are already well established, the determination of railway junction capacity remains a challenge. This work utilizes the concept of queueing theory to develop a method for the capacity calculation of railway junctions, solely depending on their infrastructure layout along with arrival and service rates. The implementation of the introduced approach is based on probabilistic model-checking. It can be used to decide which infrastructure layout to build, i.e. whether an overpass for the analysed railway junction is needed. The developed method addresses the need for fast and reliable timetable-independent railway junction capacity evaluation, catering specifically to the long-term strategic planning of junction infrastructure.

1. Introduction

The amount of goods and passengers transported by rail is predicted to increase significantly, hence infrastructure managers not only need to build new railway infrastructure but also explore enhancements for traffic volume on existing lines and nodes. Both of these tasks require sufficient methods to determine the capacity of all sub-units, i.e. lines, junctions and stations, in a railway network.

While Capacity Determination techniques for railway lines, junctions and stations have already been developed and continue to contribute to more efficient rail transportation today, only some methods describe timetable-independent approaches, allowing for sophisticated infrastructure assessment in the early stages of the planning process, which is traditionally done in multiple steps (UIC, 2013). Strategic design decisions on a network level may take place decades before operation and are an ongoing process for most infrastructure managers. The technical infrastructure planning does take some time and is done several years in advance, while the definitive timetable construction may be done several months or years prior to operation. Therefore, timetable-independent methods determining the capacity of railway infrastructure are of particular importance for the early process stages including ‘strategical network design’ and ‘technical infrastructure planning’.

The main contribution of this work is the introduction of a novel Continuous-Time Markov Chain (*CTMC*) formulation, that

- extends queueing-based capacity research to parallel utilizable infrastructure,
- enables timetable-independent capacity determination by solely depending on route conflicts, traffic shares and route service times, and
- is capable of highlighting the influence of different routes on the performance of analysed infrastructure.

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

https://doi.org/10.1016/j.trc.2024.104704
Received 21 September 2023; Received in revised form 4 June 2024; Accepted 5 June 2024
Available online 22 June 2024
0968-090X/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

It estimates the number of trains waiting for the assignment of their requested infrastructure in the timetabling process, by calculating state probabilities in a CTMC modelling the queuing system of a multi-route railway junction. This work utilizes state-of-the-art probabilistic model-checking software to perform the computations. The presented approach does not require timetable data to determine the capacity of a railway junction, enabling early-stage infrastructure planning. Consequently, technical planning decisions, i.e. whether a railway overpass is to be built, can be assisted by capacity calculations with the described approach. In comparison with other approaches to determine timetable-independent capacity, the presented method does not rely on simulations or sampling of possible sequences for timetable-compression, but rather on queueing systems, formulated for multiple service channels. With its versatile range of parameters, multiple applications of capacity research, i.e. infrastructure planning for a fixed traffic demand or capacity determination of a fixed junction infrastructure, may be realized.

While Section 2 includes a literature review regarding other approaches to the determination of railway capacity, Section 3 introduces formal problem formulations and proposes the new capacity calculation method. In Section 4, the proposed method is compared to a simulation approach concerning computation times and accuracy of the obtained solutions. Two Case Studies, highlighting the applicability of the introduced approach, are performed in Section 5. This work is concluded with a summary and discussion in Section 6.

2. Related work

This section provides an overview of the state-of-the-art for railway capacity analysis. After the definition of some key types of performance analysis and frequently used terms, selected examples of literature are matched regarding their associated methodology and capacity definition.

2.1. Terminology of railway performance analysis

With their various requirements to the grade of detail, different stages of the planning process can require an analysis of diverse definitions of railway capacity. While a first definition of railway capacity as the *maximal number of trains traversing a given infrastructure in a given time period under some fixed assumptions* summarizes the concept in a straightforward manner, the interpretation of the included assumptions determines different levels of capacity. In detail, three types of railway capacity may be distinguished (see also Jensen et al., 2020):

- *Theoretical capacity*: The maximum number of requests (i.e. trains, train-route enquiries), that can be scheduled without conflicts on the given infrastructure under consideration of driving dynamics and installed railway control systems.
- *Timetable capacity* (sometimes referred to as *maximal capacity*): The maximum number of requests, that can traverse the given infrastructure in acceptable quality when compared to a specified threshold, not only considering driving dynamics, railway control systems, but also operating program specific settings, such as train-mix and arrival processes.
- *Operational capacity* (sometimes referred to as *practical capacity*): The maximum number of trains, that can traverse the given infrastructure in acceptable operational quality when compared to a specified threshold, considering driving dynamics, railway control systems, operating programs, and additionally respecting disturbances and (knock-on) delays.

Additionally, research has not only been focused on determining the capacity (of any kind), but also on the calculation of the *capacity utilization*, i.e, the amount of available capacity, consumed in a given timetable.

While methods determining capacity utilization are mostly dependent on a timetable (*timetable dependent*) or at least a fixed order of a given train set, some methods are capable of determining performance indicators without the need for a previously set timetable (*timetable independent*), making them valuable for infrastructure assessment in early stages of infrastructure planning processes. Some approaches build on a given timetable and find maximal subsets of a set of trains that can additionally be scheduled (*timetable saturation*), building a category in between both dependency expressions. In contrast to a timetable, an *operating program* specifies the demand of train types on given lines or routes, hence being indispensable for most performance analyses.

Depending on the stage of railway infrastructure planning, a timetable may already be available, such that methodologies like *timetable compression*, i.e. routing trains through the infrastructure with a separation of only the minimum headway possible, are useable.

The various railway infrastructure planning stages make use of varying capacity definitions and hence distinct methodologies for the analysis of railway infrastructure. They additionally differ in terms of timetable dependency or granularity of the described infrastructure. During the following subsection, the methodologies *max-plus-algebra*, *optimization*, *operational data analysis*, *simulation* and *analytical* are differentiated.

Some additional distinctions of relevant railway performance analysis are made regarding their analysed infrastructure – lines, junctions, stations or networks –, their infrastructure decomposition and their utilized solution methods, f.e. *mixed integer programming* (MIP) or matrix calculations.

2.2. Literature review

The described terminology is used in Table 1 to partition some relevant related research regarding their associated category. Additionally, selected utilization, optimization and delay-propagation methods are briefly introduced in the following, while our

Table 1
Literature considering railway performance estimations.

Methodology	Capacity type	Timetable independent	Junction evaluations	Stochastic arrival process	Stochastic service process	Application specific infrastructure decomposition	Solution technique
Timetable compression	util	no	(yes)	no	no	no	compression
Randomized timetable compression	util	yes	(yes)	no	no	no	combinatorial optimization
Max-plus-algebra	util	(yes)	yes	no	no	no	matrix calculations
Timetable saturation	theo	(yes)	(yes)	no	no	no	mostly MIP
Capacity optimization	theo, util	(yes)	(yes)	no	no	no	mostly MIP
Operational data analysis	op	no	no	(yes)	(yes)	no	stat. analysis and machine learning
Delay propagation analysis	op	(yes)	no	(yes)	yes	no	matrix calculations and iterative formula
Simulation	op, tt	(yes)	yes	(yes)	yes	no	simulation computations
Analytical line capacity	tt, op	yes	no	yes	yes	no	closed formula
(prior) analytical node capacity	tt, op	yes	yes	yes	yes	yes	matrix calculations and iterative formula
Introduced here	tt	yes	yes	yes	yes	no	probabilistic model-checking

Remarks: Please note that a statement in bracelets means that this feature is partially supported, i.e. for some methods within the group or in an incomplete interpretation only. Furthermore, capacity types utilization (util), theoretical (theo), operational (op) and timetable (tt) are abbreviated.

contribution is classified along the strongly related analytical approaches. An additional table, categorizing the considered literature can be found in the Appendix in Table A.13.

Describing the determination of **capacity utilization** and introducing threshold values for sufficient operational quality, the UIC Code 406 (UIC, 2004, 2013) is widely used internationally for capacity assessments on railway lines (Abril et al., 2008; Landex, 2009; Goverde et al., 2013) and stations (Landex and Jensen, 2013; Weik et al., 2020). Utilizing the theory of capacity occupation, Max-Plus Automata (Goverde, 2007; Bešinović and Goverde, 2018) give measures for the assessment of railway infrastructure and timetable structure. While traditional compression methods (UIC, 2004; Goverde, 2007; Abril et al., 2008; Landex, 2009; Goverde et al., 2013) require a timetable to calculate capacity consumption, other approaches have been proposed utilizing randomly generated sequences of train types to overcome timetable dependencies (Jensen et al., 2017, 2020; Weik et al., 2020), focusing on the determination of timetable capacity. An overview and comparison with other capacity consumption methods can be found in Zhong et al. (2023).

Furthermore, **optimization methods** for the estimation of theoretical capacity have been developed. They mostly formulate (linear) mixed integer programming problems, including approaches for railway lines (Harrod, 2009; Yaghini et al., 2014), stations (Zwaneveld et al., 1996, 2001; Delorme et al., 2001; Lusby et al., 2011) and networks (Burdett and Kozan, 2006; Burdett, 2015, 2016). Some approaches rely on solutions to the railway timetabling problem (Cacchiani and Toth, 2012; Leutwiler and Corman, 2022), building timetables utilizing a given infrastructure in a ‘best’, as defined by objective functions, manner. Methods may also estimate the capacity occupation of its solution while creating a timetable (Zhang and Nie, 2016). Heuristic solution techniques for job-shop formulations of the railway timetabling problem have been used for railway junctions and crossovers (Burdett and Kozan, 2009).

Other approaches are based on the saturation of given timetables, which may also include empty schedules, and optimize the amount of additional traffic (Burdett and Kozan, 2006; Harrod, 2009; Lusby et al., 2011; Liao et al., 2021). Expanding their work (Burdett and Kozan, 2006), Burdett (2016) introduces MILP approaches to calculate optimal track sub-divisions on a railway line or determine which connections within a railway network can benefit from track duplications, considering the theoretical capacity of railway infrastructure.

Other optimization methods may incorporate rolling stock information (Liao et al., 2021) to the construction of a saturated timetable or estimate the effects of emerging first-order delays to following trains (Mussone and Wolfler Calvo, 2013), handling the

Table 2
Analytical capacity determination methodology.

Literature	Junction evaluations	Timetable capacity	Route-based decomposition	Route-based quality assessment	Solution method
Potthoff (1970)		✓	✓		Iterative formula
Schwanhäußer (1974)			(✓)*	(✓)*	closed-form formula
Schwanhäußer (1978)		✓	✓		Closed-form formula
Wakob (1984)		✓	(✓)*	(✓)*	Closed-form formula
Wendler (2007)		✓	(✓)*	(✓)*	Matrix–vector equations
Nießen (2008, 2013)	✓	✓		(✓)	Iterative formula
Schmitz et al. (2017)	✓	✓			Matrix–vector equations
Weik et al. (2020)	(✓)	✓	✓		Matrix–vector equations
Introduced here	✓	✓	✓	✓	Probabilistic model-checking

Remarks: (✓)* indicates that the method is used for line capacity and does hence support route-based decomposition and quality assessment natively.

propagation of so called *knock on* delays. For this, Mussone and Wolfler Calvo (2013) extend an approach by De Kort et al. (2003), utilizing max-plus algebras to calculate capacity on lines with some single-track sections, i.e. tunnels.

More detailed insights into operational parameters of specific timetables and detailed infrastructure, rolling stock and delay distribution data can be obtained by utilizing **simulations**. D'Acierno et al. (2019) provide a comprehensive literature review. While being subject to large computational times, simulations are versatile in their use case. They can simulate train operations based on a given timetable or even include arrival and service processes as random variables. The influence on model performance of different parameters can be analysed, i.e. different buffer time distributions to operational capacity estimations (Zieger et al., 2018), analysing the propagation of delays on railway infrastructure.

An investigation of **delay propagation** properties has also been part of further analytical (Goverde, 2010; Büker, 2011; Burdett and Kozan, 2014b,a) research and subject of machine-learning models (Şahin, 2017; Corman and Kecman, 2018), trained on historical operational data. Şahin (2017) calculates probabilities for different delay states with the help of a Markov Chain, while Corman and Kecman (2018) utilize Bayesian networks. Interested readers are referred to Spanninger et al. (2023) for a detailed review.

Recent approaches make additional use of **operational data** to identify capacity bottlenecks. While some research introduces train traffic data mining approaches to analyse actual operation on railway lines (Graffagnino, 2012) or stations (Armstrong and Preston, 2017), others (Weik, 2022; Corman and Henken, 2022) discuss the applicability of macroscopic fundamental diagrams. For this, Weik (2022) introduces a simulation approach to highlight their benefit for further macroscopic research and Corman and Henken (2022) provide an overview of open research applications.

Taking randomly distributed inter-arrival and service times into consideration, **analytical methods**, based on queueing theory, have been developed for an efficient timetable or operational capacity analysis in the early planning stages. Table 2 contains a comparison and summary of their features related to this work for a selection of analytical methods.

In Potthoff (1970) railway lines and stations are dimensioned, analysing loss probabilities depending on filling and service functions utilizing probability distributions for the arrival process of trains to the analysed station. While Schwanhäußer (1974), Wakob (1984), Wendler (2007) introduce measures for an analytical determination of the capacity of lines, Schwanhäußer (1978), Nießen (2008, 2013), Schmitz et al. (2017) analyse junctions, incorporating randomly distributed inter-arrival and service durations.

Schwanhäußer (1974) formulates the STRELE method to calculate expected values for the waiting times of trains in a line infrastructure, which is extended to junction infrastructure in Nießen (2008). They hence give measures to calculate the **operational capacity** of lines and junctions by comparing the estimated waiting times with threshold values (Schwanhäußer and Schultze, 1982).

Potthoff (1970), Schwanhäußer (1978), Wendler (2007), Wakob (1984), Nießen (2008, 2013), Schmitz et al. (2017), Weik (2020), however, implement results for the **timetable capacity** of the analysed infrastructure by matching corresponding limits with estimated waiting times for trains without taking delays into account.

Calculating the expected waiting times via probabilities for the loss of an arriving train, Potthoff (1970), Nießen (2008, 2013) formulate methods for station (Potthoff, 1970) and junction infrastructure. For junction capacity estimations, Nießen (2008, 2013) tackles queueing systems with multiple channels, while Schwanhäußer (1978) approximates the waiting times in a junction via a single-channel system. This single-channel system approximation is based on route-exclusion probabilities and is adapted in Weik (2020), joining it with line-capacity advancements in Wendler (2007), therefore utilizing multi-state service processes and hence more flexible probability distributions when modelling the service process.

Parameter estimations have been done (Wakob, 1984) to obtain a closed formula for approximating the actual waiting times on a railway line with general independent service and inter-arrival time distributions.

Also directly modelling general independent service and arrival processes, Schmitz et al. (2017) formulate multi-dimensional Markov Chains and include phase-type distributions in their approximation. However, their approach is limited to already partitioned junction infrastructure, hence dependent on additional input and deep system knowledge of the algorithm operator. Additionally, they distinguish between train and request types in a queue, resulting in issues with computational memory and scaling, when analysing a more complex infrastructure and/or operating program.

2.3. Motivation and contextualization

Overall, the methodology landscape is still lacking research combining major advantages of analytical and other timetable-independent junction capacity methods: Calculating timetable capacity for multi-channel railway junctions while being easily applicable to a broad range of problem formulations.

This work introduces a novel methodology that calculates timetable capacity with the help of problem statements that solely depend on the routes through the considered infrastructures, the parallel serviceability of those routes and the possible operating programs. A more formal introduction to the considered problem statements can be found in Section 3.2. The introduced approach can therefore be used at early stages in the planning process and does not require any additional system-specific knowledge to further partition the junction infrastructure.

Following other analytical models, timetable capacity is analysed with the help of Continuous-Time Markov Chains. However, the introduced multi-dimensional Markov Chain formulation is obtained by utilizing a novel route-based decomposition. It is therefore easily applicable to more complex junctions by further extending the multi-dimensional state-space with additional dimensions for the modelling of supplementary routes.

While other methods implementing multi-dimensional Markov Chains (Schmitz et al., 2017; Weik, 2020) require a set-up of Matrix-Vector equations at the modelling step and use more general frameworks to solve those equation systems on the equation level, the introduced method applies probabilistic model-checking methodology directly to the obtained Continuous-Time Markov Chain. Hence, the full broadband of graph-theoretical and numerical optimized algorithms for probabilistic model-checking may be utilized to acquire the solution. Furthermore, by explicitly formulating the Continuous-Time Markov Chain, multiple properties, describing different characteristics of the analysed infrastructure, may be checked with the same Markov Chain model. This can be utilized for more efficient parameter iteration, i.e. for different arrival- and service rates in the presented approach, or in further research even for reachability or deadlock analysis of the formulated model, which might be suitable as additional model validation steps.

In this work, infrastructure performance is measured by a timetable-independent approach analysing the timetable capacity, modelled with a multi-dimensional Continuous-Time Markov Chain and calculated with a probabilistic model-checking framework. The later discussed model features a route-based infrastructure decomposition with fixed route allocations, hence being easily applicable to more complex junctions, while respecting arising resource conflicts without the need for additional infrastructure partitioning.

3. A continuous-time Markov chain model for the queueing system in a railway junction

In this section, a CTMC formulation, modelling the timetable process of a railway junction, will be introduced. The first section (Section 3.1) will illustrate an exemplary junction layout and the second section (Section 3.2) will introduce a detailed problem definition. Afterwards, queuing systems in railway line performance modelling will be described (Section 3.3) and the CTMC model for a railway junction will be introduced in Section 3.4. Following Section 3.5, discussing the scalability of the introduced approach, a method to perform the estimations with the modelled CTMC will be proposed (Section 3.6). Furthermore, threshold values for the determination of junction capacity will be discussed in Section 3.7. This Chapter will conclude with sections discussing the methodical procedure (Section 3.8) and contributing factors (Section 3.9) of the introduced approach.

3.1. Junction layout

To evaluate the capacity of a railway system, all sub-units of the network need to be assessed. A common differentiation is made between line and node capacity, while methods describing node capacity can additionally be partitioned into methods describing junction and methods describing station capacity.

A *railway line* is a connection between two origins, usually equipped with single- or double-track infrastructure. If some traffic shares only part of the route and connects to a third origin, a *railway junction* will be installed to divide the traffic regarding its destination. Unlike at junctions, trains may start and end their journey at *railway stations*. They can be used for passenger- and good-exchange or fulfil primarily operational needs, such as ensuring a possibility for over-takings.

In Fig. 1 an infrastructure design for a railway junction of two double-track lines is given. The *main line* spans between the two origins A and B, while the *branching line* connects C with A.

As in most European countries, we consider operation to be defaulting to right-hand traffic, which suggests the usage of route r_1 for the direction A to B and routing traffic from B to A via route r_3 for the main line. Consequently, routes r_2 and r_4 are used accordingly for traffic between A and C.

When dividing traffic on the junction into routes, those four different trails may be considered (Table 3). The infrastructure can be abstracted along those routes to model its operation.

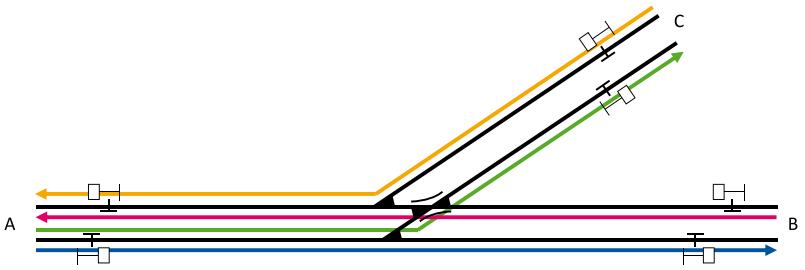


Fig. 1. Track-layout for a double-track junction.

Table 3
Routes through the junction.

Name	Origin	Destination	Conflicts
r_1	A	B	r_2
r_2	A	C	r_1, r_3
r_3	B	A	r_2, r_4
r_4	C	A	r_3

3.2. Problem formulation

A railway junction infrastructure $J = (R, C)$ consists of a set R of k routes $r \in R$ and a *conflict-matrix* $C \in \{0,1\}^{k\times k}$. Two routes $r_i, r_j \in R$ are described as *conflicting*, i.e. they cannot be used at the same time, by denoting $C_{i,j} = 1$, and as not conflicting by $C_{i,j} = 0$. Per default, the same route cannot be used twice at a time, hence $C_{i,i} = 1$ for all $i \in \{1, \dots, k\}$.

An *operating program* on the railway junction J is a set A of *demands* $a = (r, n)$ corresponding to a route $r \in R$ and a total number of trains $n \in \mathbb{R}$ of that demand in the time horizon U. To service a train of demand $a = (r, n) \in A$, the service time $t_{\text{service}}(a)$ may be calculated using microscopic railway software. The service time of demand a is dependent on the infrastructure and safety technology in J (and adjacent railway lines and junctions), as well as some train-specific characteristics, such as braking percentages, acceleration capabilities, or total mass.

The problem of planning of railway junctions can be formulated by two inverse statements:

(I) Given a railway junction J and a set of feasible operating programs $\mathcal{A} = \{A_1, \dots, A_l\}$. Determine $A^* \in \mathcal{A}$, so that the equation

$$\sum_{(r,n)\in A^*} (n) = \max_{A\in\mathcal{A}} \sum_{(r,n)\in A} (n) \tag{1}$$

holds.

(II) Given a set of feasible infrastructure layouts $\mathcal{J} = \{J_1, \dots J_l\}$ with the same set of route R, an operating program A on them and a cost function $\phi : \mathcal{J} \to \mathbb{R}$. Determine $J^* \in \mathcal{J}$, so that the equation

$$\phi(J^*) = \min_{J\in\mathcal{J}} \phi(J) \tag{2}$$

holds.

Note that an objective function can be formulated for both problem statements. Specifically, statement (I) corresponds to finding the operating program with the highest total number of trains, while statement (II) might model the determination of the junction infrastructure layout with the least associated costs (f.e. for infrastructure construction).

While problem statement (I) can be used to assess the theoretical maximal capacity of a railway junction, railway infrastructure operators could be mostly interested in solutions to the statement (II): The operational program can often be assessed beforehand, i.e. when fixed by external factors, such as governments, promoting a modal shift to railways (Europe's Rail, 2015), or increasing demands when establishing new industry locations.

For both problem statements, the selection from the set of feasible solutions regarding their respective objective function is straightforward. However, assessing whether the combination of an operating program A and a junction layout J is feasible requires the determination of performance indicators, which can be compared to thresholds, to verify feasibility.

In this contribution, the timetable capacity of a railway junction is modelled (see Section 2.1 for a definition). It can be interpreted as the capacity of railway infrastructure in the timetabling process, i.e. the process in which the infrastructure operator finds conflict-free allocations for the service of different requests, submitted by transport operators. Since the requests are randomly distributed, all possible timetables that could emerge from the so-distributed request collection are covered. Conflicts may arise with this procedure, as the requests are made randomly, so it may be necessary to set waiting times for some requests. Additionally, some requests might have to be declined, i.e. if an allocation without a conflict is impossible. The estimated number of requests that *wait* for their allocated

Table 4
List of notations.

R	Set of routes
r_i	Route i
$C \in \{0,1\}^{k \times k}$	Conflict matrix
$J = (R, C)$	Junction infrastructure
$\mathcal{J}$	Set of junction infrastructures
$A = \{a \mid a \text{ is demand}\}$	Operating program
$\mathcal{A}$	Set of operating programs
$a = (r, n)$	Demand
U	Time horizon
$t_{\text{service}}(a)$	Service time of a demand a
E_{LW}	Expected queue length
m	number of waiting slots
λ	Arrival rate
μ	Service rate
ET_A	Expected inter-arrival time
ET_S	Expected service time
ρ	Occupancy rate
v_A	Variation coefficient arrival process
v_S	Variation coefficient service process
σ	Standard deviation
p_{loss}	Probability of loss
S	Set of states
$s = (q_1, b_1, \ldots, q_k, b_k)$	State
q_i	Number of requests in queue of route r_i
b_i	Status of Service of route r_i
M	Maximum rate
$MC = (S, T)$	Continuous-Time Markov Chain
T	Set of transitions
$t = (u, v)$	Transition
$L_{W,limit}$	Threshold value for sufficient quality
$h_{i,j}$	headway time of the sequence train j after train i
$\bar{b}$	Mean service time
b_{max}	Maximum mean service time
n_r	number of trains on route r
ψ	Concatenation value
ρ_{conn}	Connected occupation rate

occupancy of a route, can serve as a performance indicator, i.e. the *estimated queue length* E_{LW}, for the infrastructure. While the proposed method technically represents the allocation of occupation requests in a railway schedule, we sometimes refer to those requests as *trains* to facilitate readability.

The method introduced below models infrastructure and operating program in a queuing system and analyzes its characteristics to evaluate the feasibility of such a combination. For further reference, a list of notations has been given in Table 4.

3.3. Queueing system

Queueing theory (Bolch et al., 2006; Zukerman, 2013) has been extensively used to describe the processes in a transportation system. An analysed entity is usually divided into a *service* and a *waiting* part (*queue*). Incoming duties may either be assigned to a processing *channel* or, if none is available, to the next slot in the queue. The Kendall notation (Kendall, 1953) has been developed to abbreviate definitions of Queueing Systems with

$$A/B/n/m.$$

Arrival (A) and service processes (B) can be modelled with arbitrary probability distributions. The described Queuing System can contain one or more service channels (n) and any natural number of waiting slots (m) - including none or infinitely many. In this work, Exponential (M) and General independent (GI) distributions are considered to describe arrival and service processes, but generally, other probability distributions can be utilized as well.

Modelled Queueing Systems are analysed by a set of relevant parameters. For *Markovian* (or *Exponential distributed*) models (*Markov Chains*), the state at any point in time is the only dependency to the future evolution of the process (see Zukerman, 2013, Ch. 2.4). Hence, transition rates between the states in a Markov Chain can be given. We denote *arrival rates* by λ and *service rates* by μ. They can be calculated with the use of expected values for the inter-arrival time ET_A

$$\lambda = \frac{1}{ET_A}, \tag{3}$$

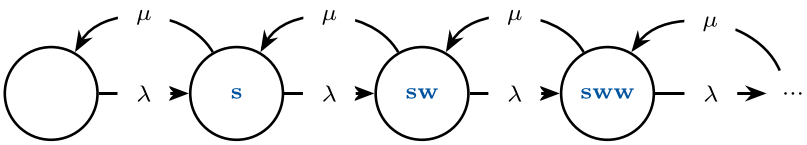


Fig. 2. Markov Chain for a $M/M/1/\infty$ Queueing System.

and service times ET_S

$$\mu = \frac{1}{ET_S}. \tag{4}$$

Further, we describe the *occupancy rate* of a queuing system by

$$\rho = \frac{\lambda}{\mu}. \tag{5}$$

Furthermore, variation coefficients of the arrival process v_A and of the service process v_B may be used for estimating supplementary characteristics. Additional parameters include the *estimated queue length* E_{LW} and the *probability of loss* p_{loss}, describing the probability that incoming requests cannot be considered in the system, as all service and waiting slots may be occupied.

An example of the use of a Queueing System is the determination of the capacity of a railway line (Wakob, 1984; Wendler, 2007), which is usually described with a $GI/GI/1/\infty$ system. As an analytical solution to those systems has not yet been found, one can model it as an $M/M/1/\infty$ system and use an approximation formula (Gudehus, 1976) for the calculation of the expected queue length

$$E_{LW}(GI/GI/1/\infty) \approx E_{LW}(M/M/1/\infty) \cdot c = \frac{\rho^2}{1-\rho} \cdot \frac{v_A^2 + v_B^2}{2} \tag{6}$$

with a factor $c = \frac{v_A^2+v_B^2}{2}$.

Further approximations for the lengths of the queue in general independent arrival and/or service processes have been developed in Wakob (1984), Fischer and Hertel (1990), Wendler (2007), Weik (2020).

Fig. 2 introduces a graphical representation of a Continuous-Time Markov Chain (see Ross, 2014 for a definition) modelling the $M/M/1/\infty$ system. It consists of a root state (without a label) as well as states with one currently serviced unit (denoted by 's') and some units in the queue (each denoted by 'w'). Ordering the states by the number of units in the system, as of being serviced and currently waiting for units, transitions between subsequent states correspond to the arrival of a unit into the Queueing System (denoted by the arrival rate 'λ') or the termination of the service of a unit (denoted by the service rate 'μ').

The following Section generalizes this concept to railway junctions, which allow more complex resource allocations than on a railway line without multiple possible routes.

3.4. Modelling railway junctions

In this work, railway junctions are considered, which differ from railway lines in one central aspect: It may be feasible for multiple trains to use the infrastructure at the same time — depending on the routes, they are scheduled to take.

3.4.1. States

Modelling those infrastructures as a Continuous-Time Markov Chain $MC = (S, T)$, used states $s \in S$ need to be distinguishable regarding the currently serviced and waited-for route(s).

The general set of possible states

$$\hat{S} = \left\{ \left(q_1, b_1, \dots, q_k, b_k\right) | q_i \in \{0, \dots, m\}, b_i \in \{0, 1\} \right\} \tag{7}$$

can be obtained by utilizing the information in the set of routes R of a junction $J = (R, C)$. A State $s = \left(q_1, b_1, \dots, q_k, b_k\right) \in \hat{S}$ contains information regarding the number of trains $q_i \in \{0, \dots, m\}$ waiting in the queue for route r_i and whether a route r_i is currently in service $b_i \in \{0, 1\}$.

The state-space $\hat{S}$ can be further restricted to

$$S = \left\{ \left(q_1, b_1, \dots, q_k, b_k\right) \in \hat{S} | \sum_{i=1}^{k} \sum_{j=1}^{k} \left(C_{i,j} b_i b_j\right) = 0 \right\} \tag{8}$$

by applying the conflicts described in the conflict matrix C. Furtheron, entries q_i or b_i in a State $s = \left(q_1, b_1, \dots, q_k, b_k\right) \in S$ are also referenced by $q_{s,i}$ or $b_{s,i}$.

3.4.2. Transitions

Transitions $(u, v) = t \in T$ between two states $u, v \in S$ can correspond to either the *arrival* of a train for route r_i, the *completion of service* on route r_i or the *choice* which holding train to service next.

The used transition rates of arrival-transitions are given by the arrival rate λ_i, of service-transitions by the service rate μ_i, and of choice-transitions by the maximum rate M. Those transition rates can be obtained by the described operational program A and the time horizon U.

An arrival transition between

$$u = \left(q_1, b_1, \dots, q_i, b_i, \dots, q_k, b_k\right) \tag{9}$$

and

$$v = \left(q_1, b_1, \dots, q_i + 1, b_i, \dots, q_k, b_k\right) \tag{10}$$

utilizes the arrival rate

$$\lambda_i = \frac{n_{r_i}}{U}, \tag{11}$$

where the total number of trains using route r_i, $n_{r_i} = \sum_{(r_i,n)\in A}(n)$, is divided by the time horizon U.

Transitions for a service process can be modelled between

$$u = \left(q_1, b_1, \dots, q_i, 1, \dots, q_k, b_k\right) \tag{12}$$

and

$$v = \left(q_1, b_1, \dots, q_i, 0, \dots, q_k, b_k\right), \tag{13}$$

utilizing the service rate

$$\mu_i = \frac{1}{\frac{\sum_{(r_i,n)\in A}(t_{\text{service}}((r_i,n))\cdot n)}{n_r}}, \tag{14}$$

which corresponds to the reciprocal of the average service times of requests on the route r_i, weighted by the number of trains in this demand.

Choice-transitions $t = (u, v)$ exclusively start at states $u \in S$ with

$$\sum_{i=1}^{k} b_{u,i} = 0 \tag{15}$$

and at least two conflicting sets of routes $R_i, R_j \subset R$, with

$$\begin{aligned} q_{u,r_o} &> 0, \ \forall r_o \in R_i \\ q_{u,r_p} &> 0, \ \forall r_p \in R_j, \end{aligned} \tag{16}$$

which are not operable simultaneously, and end at states $v \in S$, with at least one serviced route.

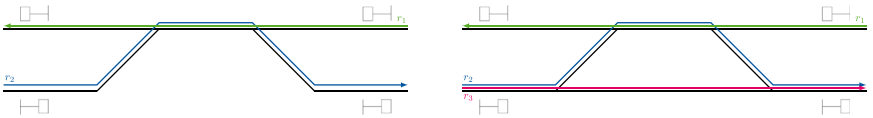
Hence, if the system is in a state s_0 where multiple routes are serviceable next, there is one choice-transition $(s_0, s_{\text{service},r_i})$ for every possible choice of the next serviced route r_i starting at this state s_0 to a corresponding state s_{service,r_i}. They correspond to the choice of which route to operate next when multiple options are possible. While they are modelled by ordinary transitions in the Continuous-Time Markov Chain, they connect to all possible route options utilizing identical transition rates, giving all options the same probability.

Since the choice transitions should not induce additional time in the system, the maximum rate M should be large enough such that the additional expected time in the system per choice transition $1/M$ is sufficiently small. In this work, choice transitions with identical transition rates are included to model the different decisions between the route(s) to be serviced next. Hence, the obtained results can be assumed to be independent of disposition strategies. The maximum rate is further chosen as $M = 600$, corresponding to a rate of 10 per second, which is deemed a competitive approximation.

3.4.3. Example

The introduced Continuous-Time Markov Chain $MC = (S, T)$ can be used to model Queueing Systems with more complex service regulations. In Fig. 3, two different examples of railway track layouts are presented. While layout 3(a) corresponds to a short single-track segment with two possible routes, layout 3(b) illustrates the infrastructure of a crossover segment with three routes, two of which (r_1 and r_3) being operable simultaneously.

Figs. 4 and 5 give graphical representations of Continuous-Time Markov Chains modelling the single track segment (Fig. 3(a)) and the crossover segment (Fig. 3(b)). Starting at the initial state, arrivals to their routes r_i are possible with their respective arrival rate λ_i. Since the Queueing System is empty at the initial state, the first arriving train can start its service time immediately after arriving. While every other train, arriving in the single track system (Fig. 3(a)) during the service time of the first train, will be allocated to a waiting slot (indicated by '**w**'), the assignment of a second train in the Queueing System for the crossover segment does depend on the used routes. If the first train uses r_1 or r_3 and the second train uses the remaining route, the second train can



(a) Two routes on a short single track segment (b) Three routes on a crossover segment

Fig. 3. Conflicting routes in railway infrastructure.

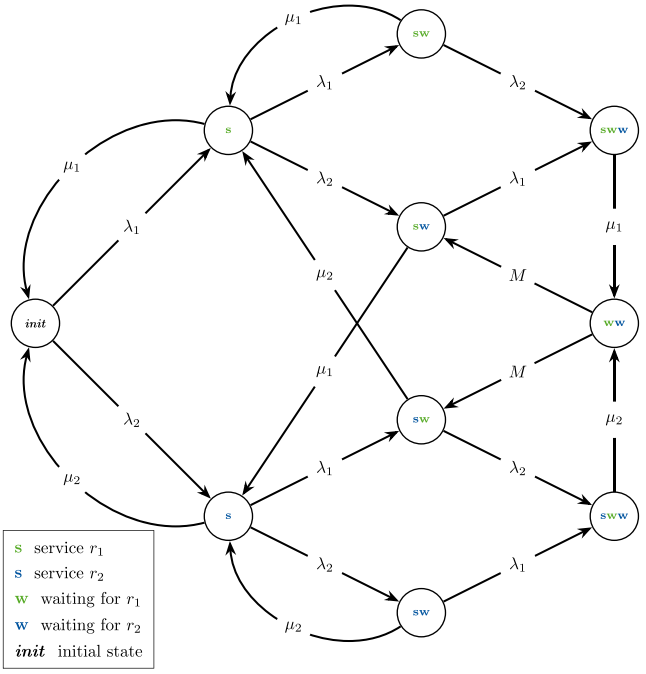


Fig. 4. Markov Chain modelling the Queueing System of the single track segment in Fig. 3(a).

Table 5
Number of states by the number of waiting positions in Queuing Systems for the examples in Fig. 3.

Number of waiting positions m	1	2	4	8	16
Single track segment	10	23	67	227	835
Crossover segment	28	89	397	2261	15 013

traverse the block-section immediately (states indicated with '**ss**'). If the two trains share the same route or one of them is using route r_3, the second train has to be assigned to a waiting slot.

Those implications of route exclusions continue in both graphs and all combinations of arriving and serviced trains are modelled with one limitation: While in theory, unlimited state spaces and therefore the modelling of $m = \infty$ waiting positions are possible, an analysis of those models is very complex and in practical operation, the number of waiting trains will always be limited due to space restrictions. Hence, the examples in Figs. 4 and 5 contain only one waiting slot per route. Trains arriving to a route while all waiting slots are occupied are therefore neglected.

Table 5 lists the number of states for the two models, for a selection of values of m. Since the probability of loss $p_{loss} = p_{loss}(m)$ is decreasing with m, a sufficient number of waiting positions may be calculated by identifying a limit value p^*_{loss} and determining the lowest $m \in \mathbb{N}$ satisfying

$$p_{loss}(m) \leq p^*_{loss} \tag{17}$$

iteratively or by testing relevant expressions of m.

A Model for the junction in Fig. 1 does incorporate more complex route conflicts and additional parameters, leading to substantially higher state numbers, i.e. $|S| = 128$ for an instance with only one waiting slot $m = 1$ per route. Hence, we refrain from including a graphical representation of the induced Continuous-Time Markov Chain. In the online repository (Emunds and Nießen, 2023) however, a full description of the used model, obtained with the definition in Section 3.4, is given.

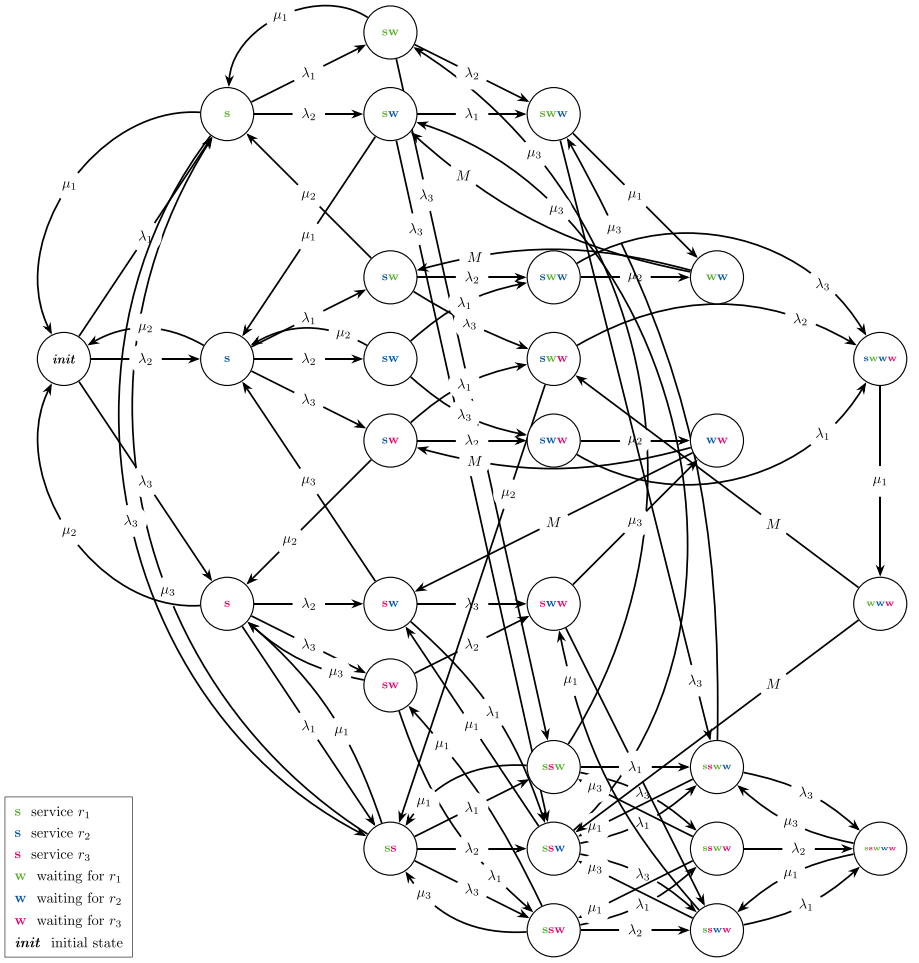


Fig. 5. Markov Chain modelling the Queueing System of the crossover segment in Fig. 3(b).

3.5. Scalability

To further analyse the scalability of the introduced model, an upper bound for the number of states can be given.

According to the definition in Section 3.4.1, a state consists of $k = |R|$ variables $q_1, \dots, q_k$, with $q_i \in \{0, \dots, m\}$ modelling the amount of requests waiting for route r_i, and $k = |R|$ variables $b_1, \dots, b_k$, with $b_i \in \{0, 1\}$ modelling whether route r_i is currently utilized.

Hence, the amount of states $|S|$ is limited by

$$|S| \leq (m+1)^k \cdot 2^k \ . \tag{18}$$

However, not every combination of values is possible, due to the limitations induced by the conflict matrix (see Section 3.4.1). Let L indicate the maximum number of parallel useable routes, then the number of potential combinations of parallel used routes can be bounded by k^L. Therefore, the total number of states is bounded by

$$|S| \leq (m+1)^k \cdot k^L \ . \tag{19}$$

Since the required time and memory to solve a Continuous-Time Markov Chain scales with the number of states in a Markov-Chain, this number can be interpreted to estimate the scalability of the introduced model. According to the Benchmark in Budde et al. (2021), the largest solvable CTMCs possess state spaces in the order of magnitude of 10^7 up to 10^8 states. Under the assumption, that a maximum of $L = 3$ routes are useable in parallel, a model with $k = 8$ routes and a maximum of $m = 5$ waiting positions per route would require up to $8.6 \cdot 10^8$ states according to the bound in Eq. (19) and it therefore might be barely solvable.

Section 5 introduces two different case studies, the first of which analyses a complex railway junction, including 8 routes, with $L = 4$ being parallel useable. To efficiently analyse this junction, the number of waiting positions per route is decreased to $m = 4$

in the model, making it more likely for a request to be declined. To ensure the soundness of the approximation, the probability of loss p_{loss} is checked on the calculated solutions.

For some even more complex junctions or when analysing entire railway stations, other measures might be needed to restrict the size of the corresponding model. For instance, it might be feasible to divide junctions with more than 8 routes into multiple smaller junctions that can be analysed independently, depending on the conflict matrix. This and the development of other approximation techniques are subject to further research.

However, the introduced method is capable of modelling railway junctions with up to 8 routes without any adjustments. The applicability can be further described by studying the number of routes in railway junctions. In the case of two double-track railway lines merging in a railway junction, the number of routes will probably be between 4 and 6, one for every combination of directions. Similarly, an intersection between two double-track railway lines might consist of 4 to 12 routes. Depending on the number of parallel useable routes, those infrastructures can be analysable, for example by reducing the amount of modelled waiting positions per route.

In any case, the modelled CTMC is utilized to determine the estimated queue length of every route.

3.6. Determination of the estimated queue length

While closed-form analytical solutions exist for the estimated queue length E_{LW} for some Queueing Systems (i.e. Eq. (6), more examples in Fischer and Hertel (1990)), the analysis of more complex systems remains challenging.

3.6.1. Utilizing steady-state distribution

For multidimensional Markov Chains, recent research (Schmitz et al., 2017; Weik, 2020) utilizes Matrix-Vector equations to determine the steady-state distribution, i.e. probabilities $p : S \to [0, 1]$ that the system is in state $s \in S$ in the long run.

Given those probabilities , the expected queue length E_{LW,r_i} of a route r_i can be calculated by summing the probabilities of all states containing elements in the queue

$$E_{LW,r_i} = \sum_{\Psi_i(s)>0} p(s) \cdot \Psi_i(s), \tag{20}$$

utilizing the function $\Psi_i : S \to \mathbb{N}^0$, giving the number of elements in the queue of the route r_i for a state $s \in S$.

This approach has been used extensively for the determination of the timetable capacity of railway lines (Weik and Nießen, 2017) and junctions (Schmitz et al., 2017) and even been applied to model infrastructure unavailability with an additional stochastic process (Weik, 2020).

However, in this work, the techniques of probabilistic model-checking are explored regarding their potential for the determination of the estimated queue length.

3.6.2. Probabilistic model-checking

Much scientific research is based on a representation of natural or technical processes in a model. To analyse the properties of such systems, theoretical computer scientists developed formal methods to analyse their behaviour. By specifying the desired requirements in a formal temporal logic, the satisfaction of those requirements to a formalized model can be verified with the help of model-checking techniques.

Since in reality, many processes contain some kind of probabilistic behaviour, an extension, *probabilistic model-checking* has been introduced. This technique can analyse different types of Markov models, which are a way to include probabilistic behaviour into the representation of analysed processes. Since a detailed introduction and explanation of the techniques of probabilistic model-checking would extend the scope of this work, interested readers are referred to the articles of Kwiatkowska et al. (2010) and Katoen (2016).

This work utilizes probabilistic model-checking because of its capability and extensive development for the analysis of Continuous-Time Markov Chains, a modelling approach that has been thoroughly utilized for the analytical determination of railway capacity. The model formulation as a Continuous-Time Markov Chains has been described in Section 3.4 and can be represented in a model file, f.e. in the PRISM modelling language (Parker et al., 2000; Kwiatkowska et al., 2011). An example of a model file of the junction considered in this work has been provided in the online repository (Emunds and Nießen, 2023).

Additionally to the model file, a formal description of the property has to be established. The probabilistic model-checking tools utilized for this work are capable of parsing properties in the PRISM modelling language. This language can be used to declare *path formulae* and *state formulae*, which allow for a formalization of the required logical expressions. Let X_{MC} be the random variable describing the state of the Continuous-Time Markov-Chain MC. The probability that the Markov-Chain corresponds to a state $s \in S$ in the long run is then given by

$$Pr(X_{MC} = s) = p(s) \tag{21}$$

By formulating a state formula Φ, i.e. a condition that can be satisfied on a state, the probability that this condition is satisfied in the model in the long run can be given by

$$Pr(X_{MC} = s, \Phi(s)) = \sum_{s \in S, \Phi(s)} p(s). \tag{22}$$

As an example, the condition whether a state representing a system case, in which at least one train is waiting for the service at route r_1, can be given as $\Phi_{waiting,1}(s) = q_{s,r_1} > 0$ for the formulated model.

The probability $Pr(X_{MC} = s, \Phi_{waiting,1}(s))$ that the system is in a state that satisfies this condition can then be obtained by checking the corresponding property with a probabilistic model-checking tool.

An important extension for properties is information about *rewards*. A reward can be formulated as a function $r : S \to \mathbb{R}, s \mapsto r(s)$ mapping the reward $r(s)$ to every state. Probabilistic model-checking tools can be utilized to determine the expected value of a reward under a certain condition. Regarding the application discussed here, the number of trains on a waiting slot for route r_i in a state $s \in S$ of the Continuous-Time Markov Chain (S, T) from Section 3.4 can be given as $r_{qL,r_i} : S \to \mathbb{R}, s \mapsto r_{qL,r_i}(s)$ with

$$r_{qL,r_i}(s) = \begin{cases} q_{s,r_i} & q_{s,r_i} > 0 \\ 0 & otherwise. \end{cases} \tag{23}$$

and checked with a corresponding property on the Continuous-Time Markov Chain from Section 3.4, to determine the estimated queue length E_{LW,r_i} for route $r_i \in R$.

More information regarding property definition can be found in Parker et al. (2000) and the properties utilized in this paper are accessible via the online repository Emunds and Nießen (2023).

3.6.3. Approximating non-exponential processes

Notice that the calculation of the expected queue length E_{LW,r_i} on an arbitrary route r_i in Sections 3.6.1 and 3.6.2 relies on the use of Markov Chains, solely capable of modelling queueing system of type $M/M/s/\infty$.

To analyse systems with general independent (GI) arrival or service processes, factors, utilizing the variation coefficient of the described process, can be introduced. According to Fischer and Hertel (1990), the expected length of a queue in a Queueing system of type $GI/GI/s/\infty$ can be analysed by using a $M/M/s/\infty$ system and modifying the results

$$E_{LW,r_i}(M/M/s/\infty) \cdot \frac{1}{\gamma} \approx E_{LW,r_i}(GI/GI/s/\infty) \tag{24}$$

accordingly, using

$$\gamma = \frac{2}{c \cdot v_B^2 + v_A^2} \tag{25}$$

and

$$c = \left(\frac{\rho}{s}\right)^{1-v_A^2} \cdot (1 + v_A^2) - v_A^2. \tag{26}$$

Here, v_A and v_B correspond to the coefficients of variation for the arrival and the service process respectively.

Following the Eqs. (24)–(26), a greater coefficient of variation in either the arrival or the service process yields a larger expected queue length. In addition to the introduced coefficients of variation, factor γ depends on the occupancy rate ρ (see identity (5)) and the number of parallel service channels s. Here, the queue length E_{LW,r_i} corresponds to arrivals on one route r_i only, $s = 1$ can hence be fixed.

This modification can be implemented in the design of threshold values that have been developed for the capacity analysis of railway infrastructure.

Regarding the modelled railway infrastructure (see Section 3.4), E_{LW,r_i} has to be calculated for every route r_i. Using the threshold values introduced in the next section, every E_{LW,r_i} can be verified regarding its sufficiency for acceptable operating quality on the infrastructure.

3.7. Threshold values

Different threshold values for theoretical and operational capacity have been discussed in the literature. In the UIC Code 406 (UIC, 2013) maximum occupancy rates have been introduced to limit capacity utilization. Potthoff (1970) introduces limits for the loss probabilities in railway stations, which are still in practical use for the determination of track numbers in a railway station, i.e. at the German infrastructure manager (DB Netz AG, 2022).

In this work, the threshold value $L^*_{W,limit}$, introduced in Schwanhäußer and Schultze (1982) and likewise still in practical use (DB Netz AG, 2022), is utilized. It corresponds to the maximum number of trains that are to wait in the analysed line section at any given time for a sufficient performance of the infrastructure. With the ratio of passenger trains in all considered trains p_{pt}, the threshold value

$$L^*_{W,limit} = 0.479 \cdot \exp(-1.3 \cdot p_{pt}) \tag{27}$$

can be specified. A threshold value of the approximating queuing system

$$L_{W,limit} \approx \gamma \cdot L^*_{W,limit} = \gamma \cdot 0.479 \cdot \exp(-1.3 \cdot p_{pt}) \tag{28}$$

can be obtained by utilizing (24) and (27).

Hence, the performance of railway infrastructure is judged based on the arrival and service processes, including their rates and variation, as well as on the operating program during the analysed time horizon.

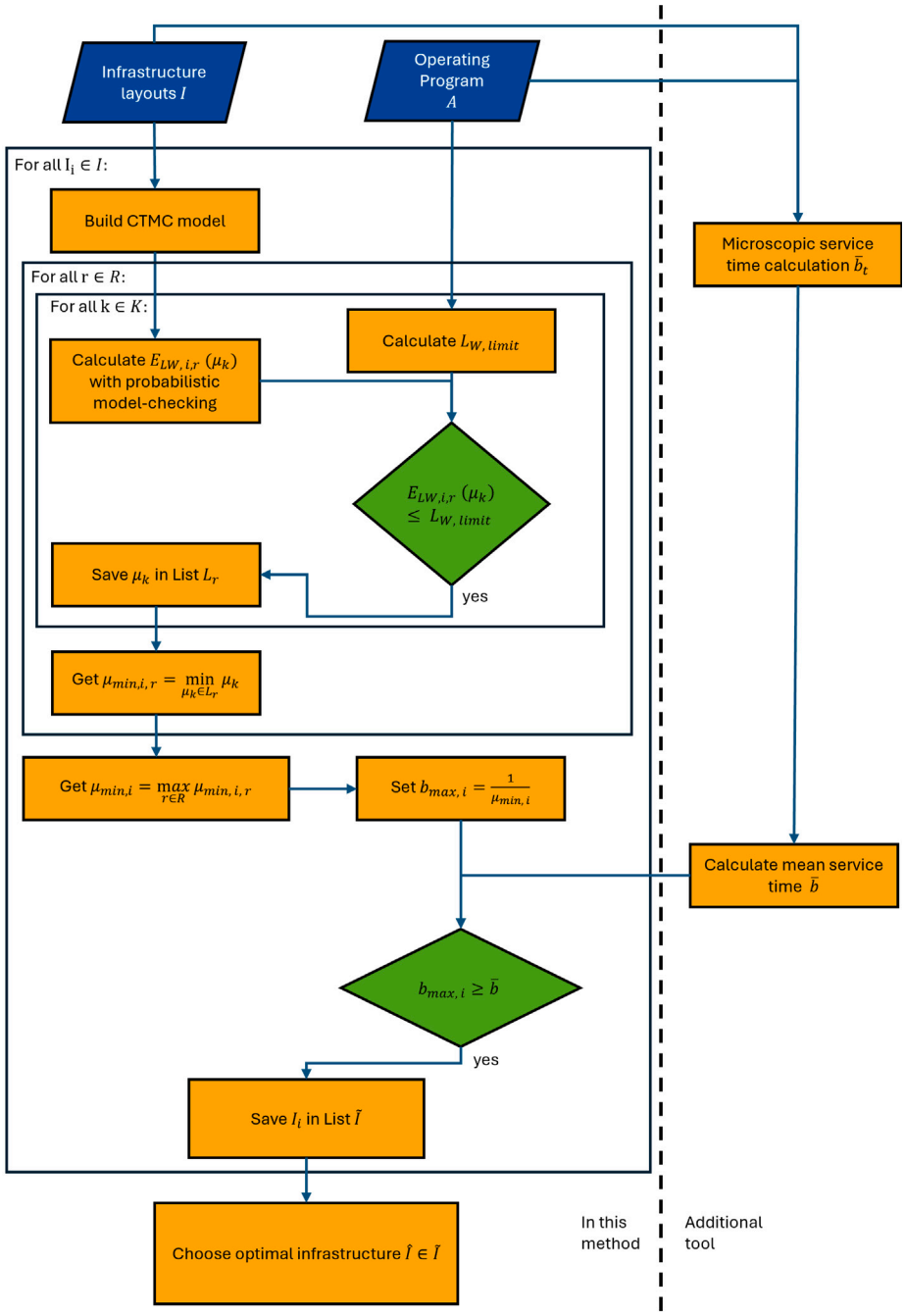


Fig. 6. Application of the introduced method to choose an optimal infrastructure layout.

3.8. Methodical procedure

Two different problem formulations that can be solved with the introduced method have been defined in Section 3.2. The Problem of planning an infrastructure out of a set of possible layouts (II) can be solved by iterating through all possible service times while calculating the corresponding estimated queue lengths at a given route $E_{LW,r}$ by utilizing the introduced Continuous-Time Markov Chain model. A detailed flow chart is given in Fig. 6.

The calculation utilizes an input consisting of a set of infrastructure layouts I and an operating program A, corresponding to a set of demand tuples (r, n). For every infrastructure layout $I_i \in I$, a Continuous-Time Markov Chain model can be formulated according to the definitions in Section 3.4. By iterating over all routes $r \in R$ in this infrastructure layout and all mean service rates

Table 6
Requirements for the described method.

Parameter	Required for capacity determination (I)	Required for infrastructure planning (II)
Routes	✓	✓
Route exclusions	✓	✓
Share of passenger trains	✓	✓
Operating program(s)	(+)	✓
Estimated service times	✓	(+)
Microscopic infrastructure data	(+)	(+)
Train-specific parameters	(+)	(+)
Train control system	(+)	(+)

Remarks: Please note that a strong requirement is indicated by ✓, while an enhancement to the qualitative decision making, f.e. in later planning stages, is indicated by (+).

μ_k for $k \in K$ in an index-set K, the estimated queue length $E_{LW,i,r}(\mu_k)$ can be obtained from probabilistic model-checking, yielding the expected value of the in Section 3.6.2 formulated reward $r_{qL,r}$.

Utilizing the operating program A, a threshold to the queue length $L_{W,limit}$ (see Section 3.7) can be obtained. If the calculated $E_{LW,i,r}(\mu_k)$ is below the computed threshold of $L_{W,limit}$, the assumed service rate μ_k can be saved in a List L, from which, after the value iteration over all μ_k has been completed, the minimum service rate $\mu_{min,i,r}$ for the route r can be obtained.

Since this minimum service rate $\mu_{min,i,r}$ corresponds to the minimum rate with which service of the requests on this route r has to be completed, it gives a lower bound to the total minimum service rate $\mu_{min,i}$ in the infrastructure layout I_i. By building the maximum $\mu_{min,i} = \max_{r \in R} \mu_{min,i,r}$ of all route-based lower bounds, the total minimum service rate can be obtained.

Therefore, by calculating $b_{max,i} = 1/\mu_{min,i}$, an upper bound to the mean service time over all routes in the infrastructure, that needs to be fulfilled for a sufficient quality, can be obtained. Please note that $b_{max,i}$ does not necessarily correspond to practical realizable minimum headway or service times, but rather to a limit to the average service time value. Hence, after verifying that the mean service time in the junction $\bar{b}$ is smaller than the obtained maximum mean service time $b_{max,i}$ for every infrastructure layout $I_i \in \tilde{I} \subset I$, the better layout $\hat{I}$ can be selected.

This architecture translates directly to the determination of the timetable capacity of a railway junction (I), by iterating over all possible arrival rates λ_k and utilizing a given share of passenger trains directly to specify the threshold $L_{W,limit}$. In that case, the estimated service times are needed to determine the fixed service rates μ_r for every route $r \in R$ in the analysed junction infrastructure.

In both cases, the quality of a junction is determined by the calculation of an estimated queue length $E_{LW,r}$ for a given route $r \in R$ and comparing it to a threshold value. An example can be found in the online repository Emunds and Nießen (2023).

While some parts are essential for the construction of the introduced model, others, such as additional microscopic service time calculations, that can be used to calculate the precise mean service time of the different infrastructure layouts, can assist in the determination of the optimal infrastructure layout.

Different parameters may contribute to the formulation and calculation of those processes and therefore influence the timetable capacity of a railway station.

3.9. Contributing factors

The introduced method needs to take different parameters into account when assisting in the planning process of a railway junction. For example, not only the number of trains on every route in the analysed time horizon but also train-specific characteristics may influence the achieved quality, and not all parameters might be known at the time of early planning stages.

In Table 6, different parameters and their requirement for the two different planning problems – capacity determination (I) and infrastructure planning (II) – are described.

While the knowledge of routes, their exclusions and the share of passenger trains are needed for both problem formulations, estimated service times are strongly required for the determination of the timetable capacity (I) for a considered infrastructure only. If an infrastructure layout has to be selected (II), an estimation of service times on every route/train combination is not necessary but can be used to evaluate other planning decisions. The knowledge of the detailed operating program is required for the infrastructure planning (II) formulation only. For the determination of the timetable capacity (I) only the share of passenger trains is required to be given, as it is needed in the definition of a threshold for sufficient quality (see Section 3.7). However, it might be advisable to define a set of possible operation programs, to allow small deviations in the share of passenger trains, which would otherwise be fixed to the given value and hence substantially restrict the solution space.

Additional parameters, such as microscopic infrastructure data, train-specific parameters and the considered train control system may be utilized in combination with a microscopic running and blocking time calculation. Blocking times can be utilized to determine the minimum headway time between the service of two consecutive route/train combinations. Further information regarding their definition and calculation may be found in related literature (Happel, 1959; Hansen, 2010). By building the weighted average $\bar{b}_a$ of all minimum headway times h_{t_1,t_2} between a route/train combination t_1 and any successor t_2, the service time $t_{\text{service}}(t_1) = \bar{b}_{t_1}$ can

Table 7
Model assumptions.

Parameter	Assumed value	Reasoning	Further information
Amount of maximal parallel occupied waiting positions per route m	5 (or 4 in Section 5.1)	Limiting the size of the Markov Chain	Section 3.4.3
Maximum Rate M	600	Corresponding to 10 ticks per second	Section 3.4.2
Variation coefficient arrival process v_A	0.8	Standard value in literature	Wendler (1999)
Variation coefficient service process v_B	0.3	Standard value in literature	Schwanhäußer (1974)
Threshold for sufficient quality $L^*_{W,limit}$	$0.479 \cdot \exp(-1.3 \cdot p_{pt})$	Standard value in literature and practice	Section 3.7

replace the service time estimations for every route/train combination in the railway junction or yield key points for the evaluation of the method's results.

This enables decision-makers to compare the impact of different parameters, such as infrastructure layout, speed restrictions or train control systems and choose the combination that meets all their requirements. As such, the capacity of a railway junction can be increased in manifold ways, f.e. by decreasing the utilized service time, either through relaxed speed restrictions or by implementing alternative train control systems, and by choosing a different infrastructure layout that reduces the number of route exclusions.

With the approach described in this work, all considered parameters may be analysed regarding their influence on the formulated problems by using the described input data in the modelling of a queueing system and analysing its characteristics.

Additionally to the parameters in Table 6, which may be extracted from input data, some choices regarding the model itself have to be made. In particular, Table 7 gives an overview of the modelling constants and their assumed value in the remainder of this work.

Here, the maximal amount of parallel occupied waiting positions per route is set to $m = 5$ to restrict the number of states in the underlying Continuous-Time Markov Chain and hence speed up the computation. To model near instantaneous transitions over the choice transitions (see Section 3.4.2), the choice rate per minute has been set to $M = 600$.

Other parameters have been fixed according to standard values in literature, such as $v_A = 0.8, v_B = 0.3$ for the coefficients of variation of the arrival and service process (see Wendler, 1999; Schwanhäußer, 1974). This assumption is due to the reasoning that the arrival process is not as random as the use of an exponential distribution presumes since traffic demands usually underlie some kind of periodicity. The service times are even more stable as they are majorly determined by a fixed set of physical properties, such as the train control system and train-specific parameters, which are usually similar due to the use of homologous rolling stock for one railway line. The threshold for sufficient quality is assumed to be given by $L^*_{W,limit} = 0.479 \cdot \exp(-1.3 \cdot p_{pt})$ following DB Netz AG (2022) (see also Section 3.7).

In addition to those fixed value presumptions, further modelling assumptions have been made in the remainder of this work. Thus, the arrival and service processes have been supposed to be independent and exponentially distributed for the initial calculation with the Continuous-Time Markov Chain. A more general distribution has been assumed to be approximable by utilizing the factors in Section 3.6. Furthermore, the threshold values (see Section 3.7) have only been formulated for railway line capacity; they have been assumed to also give a reasonable quality limit for the estimated queuing length of a route in a railway junction.

To validate the introduced Continuous-Time Markov Chain model for the queueing system of a railway junction, Section 4 compares simulation results with analytical results obtained by applying the described method.

4. Model performance

Using the formulation of Section 3.4, a Continuous-Time Markov Chain for a railway junction can be obtained. In this work, a maximum of $m = 5$ waiting slots per route has been utilized, as a trade-off between tractability (dependent on the model size) and accuracy of the model (see Section 3.4.3). By calculating the estimated length of a queue and comparing it to the obtained threshold values (see Section 3.7), the capacity of modelled railway junctions may be assessed. To ensure the quality of the obtained model, the validity of the generated queue length estimations has to be verified.

Aiming to survey solely the performance of the solution process using the introduced Continuous-Time Markov Chain, $M/M/s/\infty$ systems have been considered. For this, simulations of a railway junction with multiple incoming railway lines have been built and run on sample data.

4.1. Simulation architecture

The Simulations have been implemented in Python 3.10.9 (Van Rossum and Drake, 1995; Python Software Foundation, 2022), utilizing SimPy (Team SimPy Revision, 2023) and SciPy (Virtanen et al., 2020). A model of the junction in Fig. 1 has been built, including four different routes with route-specific arrival and service processes.

To estimate inter-arrival times for every route, a pseudo-random number generator yields the next inter-arrival time within a specified exponential distribution with an expected value $ET_{A,r}$ equal to the reciprocal value of the mean arrival rate $\lambda_r = \frac{1}{ET_{A,r}}$ of the modelled route. Trains are stored in a first-in-first-out queue for the route and serviced according to the service time acquired by a second pseudo-random number generator, utilizing another exponential distribution with an expected value ET_S equal to the mean service rate $\mu = \frac{1}{ET_S}$ of the system. During the service of a route r, it is ensured, that no conflicting route $r' \in R$ can start service by utilizing shared resources for every pair of routes $(r, r') \in R \times R$.

An implementation can be found in the online repository (Emunds and Nießen, 2023). By performing Kolmogorov–Smirnov tests on the sampled inter-arrival and service times, the correspondence of the generated data with the implied distributions has been verified. The obtained p-values, describing the goodness of the fit for every single simulation run, are summarized in the online repository. Furthermore, the validity of the simulated schedule has been proven for every simulation iteration by computing the number of conflicting trains with a suitable algorithm.

To assess the performance of the simulated process, a snapshot of every route's queue length is taken in every simulated minute. Hence, the mean length of a queue can be obtained easily for every simulation run.

4.2. Validation

Both, implementations of the simulation and the queueing-length estimations with the formulated Continuous-Time Markov Chain, have been run on a single core of an Intel Core i7-10700 Processor (2.90 GHz), utilizing a maximum of 32 GB working memory.

Two different simulation setups have been considered:

1. A simulation with no limit to the number of trains being able to wait in the queues of their requested routes
2. A simulation with a maximum of 5 trains per route being able to wait at the same time for the service on their requested routes

In the second simulation and in the analytical setting, an arriving train is rejected, i.e. not inserted into the respective queue, if its arrival time lies within a time frame where 5 trains are already waiting for the release of the requested route. Here, all routes are capable of having 5 trains waiting for service, hence a total maximum of 20 trains may wait at the same time.

All three solution approaches, both simulation setups and the analytical method, have been set to compute the estimated queue length $E_{LW,r}$ for the route r_3 (see Table 3), conflicting with routes in direction of A and C (see Fig. 1). Route r_3 has been selected as it is one of the routes with the most conflicts (with r_2 and r_4) and it would directly benefit from an overpass construction.

Since those computations have been done to compare the results and running times of the approaches, only a small set of 10 different service rates, between 0.1 and 1.0 trains per minute, has been considered. An arrival rate of 0.1 trains per minute has been set for every route in all computed instances.

Each simulation has been run 100 times for every considered service time and for every run 22 h have been simulated. From those 22 h, the first and last have not been considered, resulting in an evaluated time of 20 h per simulation and 2000 h in total for every service time investigated.

For both simulation setups, the mean computing time per hour and the mean total run time per 22-hour simulation have been evaluated. Those can be found in Fig. 7, which additionally includes the running time of the analytical solution for reference.

The computing time results in Fig. 7 indicate that the analytical approach is faster than the tested simulation implementation, even if compared to the run of a single simulation hour. Noting the logarithmic scale, for small service rates the analytical approach is faster by a factor of 5 to 10, depending on the considered service rate. Since simulations have to be conducted multiple times to receive sufficient results, a comparison with the total simulation times might be more realistic, yielding factors of 10^3 up to 10^4.

In general, the computing times of the simulation approach without a limit are significantly increasing with a decreasing service rate. This is probably due to a higher amount of trains in the network at any given time, as the mean service time increases with a decreasing service rate, leading to more trains waiting for the release of their requested route. Similarly, the simulation runs with an unlimited m required up to 50% more computing time compared to the simulation scenarios with a limit of $m = 5$.

Additional observations can be made by evaluating the accuracy of the introduced analytical approach. This can be done by comparing the results obtained from the simulation methods with the results of the queueing-based analytical approach. In Fig. 8 the results obtained from the computations E_{LW,r_3} are depicted by including the exact analytical results and the standard deviation area of the simulation results, i.e. the area $\left[\overline{E}_{LW} - \sigma, \overline{E}_{LW,r_3} + \sigma\right]$, surrounding the mean $\overline{E}_{LW,r_3}$ by the standard deviation σ. Supplementary boxplots, including a study varying the number of trains per route for a fixed service rate, are available in Appendix B.

Noting the logarithmic scale in Fig. 8, the results show significant differences between the simulation setups. For the setup with a limit of $m = 5$ on the number of trains in a queue (Fig. 8(b)), matching the setting for the queueing-based analytical approach, the results of the simulation and the analytical solution are similar. Contrarily, when comparing the analytical solution to the simulation setup without a limit on the number of trains in a queue (Fig. 8(a)), serious discrepancies between the simulation and the analytical results can be found for service rates of less than 0.3.

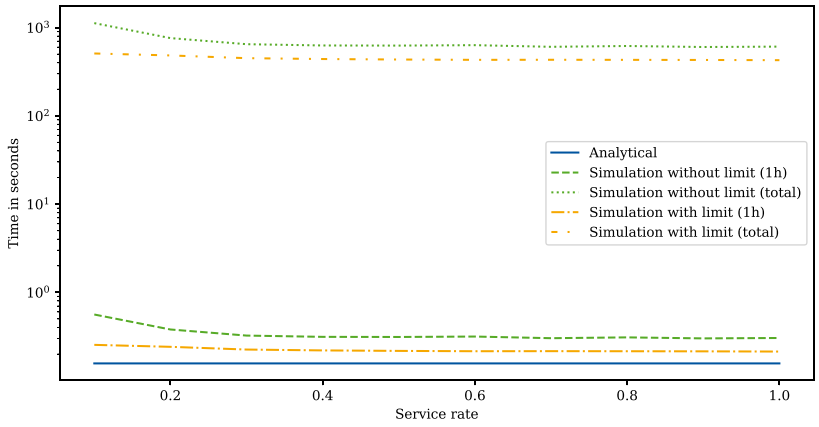
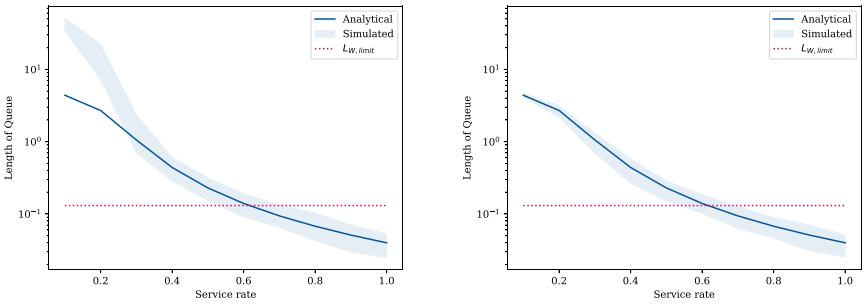


Fig. 7. Comparison of the required computation times.



(a) Simulation with unlimited m (b) Simulation with a limit of $m = 5$

Fig. 8. Accuracy of the analytical approach when compared to the confidence interval of the performed simulations.

Hence, an assumption, that a high probability of the queue length exceeding the limit of 5 requests exists, can be made for the considered railway junction, utilizing the fixed predefinitions and an averaged service rate of less than 0.3. Consequently, arriving trains are likely to have to wait, resulting in very poor service quality.

Taking the limit for $M/M/s/\infty$ queueing systems (see Section 3.7), in Fig. 8 with '$L_{W,limit}$' annotated, into account, the accuracy of the introduced queueing-based analytical approach is sufficient for the use in practical applications.

While the introduced analytical method is sufficiently accurate with only one computation, the considered simulation methodology would require multiple computations to describe the queuing process satisfactorily, since it can only compute samples for process iterations and does not consider the stochastic process in its entirety. Furthermore, the simulations would need to be rerun for every change in parameters, and since both described problem statements need to evaluate the queue lengths at multiple parameter configurations, such as various service or arrival rates, this would substantially increase the total computation time. Therefore, the case studies in the following section have been conducted with the introduced analytical approach.

5. Computational study

The introduced method for the calculation of queue lengths can be used to guide infrastructure managers when planning railway junctions. In this section, two different railway junctions have been considered to highlight the application of the proposed method for both introduced problem statements (see Section 3.2).

While Section 5.1 determines the timetable capacity of a given railway junction, Section 5.2 focuses on the selection of an infrastructure layout for fixed operating program demands.

5.1. Timetable capacity analysis

To describe the performance capabilities of a complex railway junction in the timetabling process, its timetable capacity can be calculated with the introduced method. In this section, a junction located in the east of Paris, referred to as the *triangle of Gagny*, is analysed.

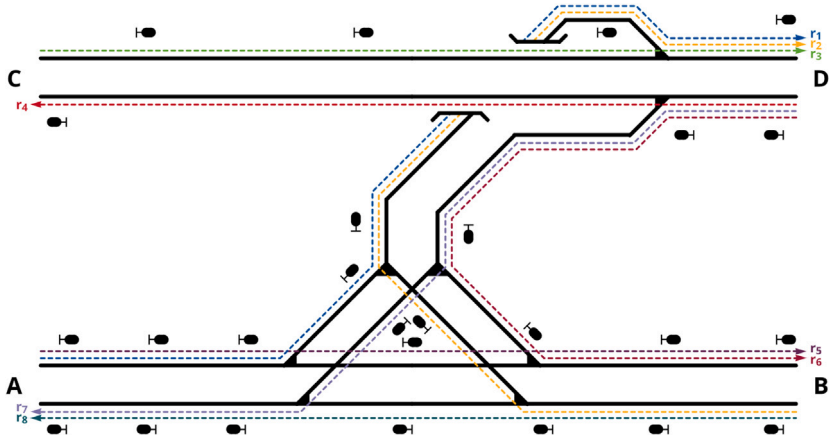


Fig. 9. A track-layout plan for the triangle of Gagny. Own illustration.
Source: Adapted from Pellegrini et al. (2014).

Table 8
Traffic share $p_{i,\theta}$ of routes r_i in every scenario $\theta \in \Theta$.

	1	2	3	4	5	6	7	8
a	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8
b	1/12	1/12	1/6	1/6	1/6	1/12	1/12	1/6
c	1/6	1/6	1/12	1/12	1/12	1/6	1/6	1/12

Since the junction has already been subject to railway-specific research, i.e. the design of real-time traffic management algorithms in Pellegrini et al. (2014), a track diagram (see Fig. 9) could be obtained from the existing literature. The junction infrastructure

$$J_{\text{Gagny}} = \left(R = \{r_1, \dots, r_8\}, C\right) \tag{29}$$

can be described with the conflict matrix

$$C_{i,j} = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}. \tag{30}$$

We assume that an infrastructure manager plans to utilize all 8 routes with passenger traffic only and discuss three different distributions of traffic shares on requested routes.

The set of possible scenarios $\Theta = \{a, b, c\}$ includes the following scenarios:

(a) a homogeneous distribution of requests to all 8 routes,
(b) a heterogeneous distribution, assuming two times more traffic on the direct routes (r_3, r_4, r_5, r_8),
(c) a heterogeneous distribution, assuming two times more traffic on the branching routes (r_1, r_2, r_6, r_7).

The three distributions are listed in Table 8 by their traffic share $p_{i,\theta}$ in the scenario θ.

These scenarios are further investigated by estimating the queue length of every route for a set $\mathcal{A}_\theta$ of possible operating programs for every scenario $\theta \in \Theta$. A possible operating program $A \in \mathcal{A}_\theta$ consists of demands $(r_i, n_i) \in A$, so that the numbers of trains per routes n_i correspond to the traffic share distribution of the scenario θ.

For this example, we consider the set $\mathcal{A}_\theta = \left\{A_\theta(c) | c \in \{1, \dots, 100\}\right\}$ of operating programs

$$A_\theta(c) = \left\{\left(r_i, c \cdot p_{i,\theta}\right) | r_i \in R\right\} \tag{31}$$

for every scenario θ, corresponding to a total number of 1 up to 100 trains per hour in the system.

While the arrival rates λ_r can be obtained from the operating program A by applying Eq. (11), the service rates are calculated from the service times of every route in the infrastructure. Usually, this would be done through a microscopic minimum headway time calculation. Unfortunately, no microscopic data has been made available for this example, therefore the minimum headway times have been estimated in Table 9, assuming only one type of rolling stock per route.

Table 9
Assumed minimum headway times $h_{i,j}$ in seconds for the junction of Gagny.

i \ j	1	2	3	4	5	6	7	8
1	180	180	140		130			
2	180	180	140		90		90	130
3	90	90	90					
4				90		90	90	
5	110	90			110	110	90	
6				160	160	160	160	
7		90		180	90	160	180	180
8		110					110	110

Table 10
Service time $b_{i,\theta}$ of route r_i in every scenario $\theta \in \Theta$ in minutes.

	1	2	3	4	5	6	7	8
a	2.63	2.25	1.5	1.5	1.7	2.67	2.44	1.83
b	2.5	2.17	1.5	1.5	1.72	2.67	2.46	1.83
c	2.75	2.33	1.5	1.5	1.69	2.67	2.43	1.83

The service time $b_{i,\theta}$ of a route r_i can be obtained by building the weighted average

$$b_{i,\theta} = \sum_{j=1}^{8} h_{i,j} \cdot w_{i,j,\theta}, \tag{32}$$

with the weights

$$w_{i,j,\theta} = \begin{cases} p_{i,j,\theta} & C_{i,j} = 1 \\ 0 & otherwise, \end{cases} \tag{33}$$

corresponding to the probability

$$p_{i,j,\theta} = \frac{p_{j,\theta}}{\sum_{j,\ C_{i,j}=1}(p_{j,\theta})} \tag{34}$$

that route r_j follows route r_i in scenario θ if r_j has a conflict with r_i. The calculated service times $b_{i,\theta}$ are listed in Table 10.

Utilizing the service time of a selected scenario θ, the service rate μ_i can be calculated with $\mu_i = 1/b_{i,\theta}$.

Together with the arrival rate λ_i, a CTMC, modelling the timetabling process on the triangle of Gagny, can be described according to the model in Section 3.4. Since the evaluated junction infrastructure utilizes a total of $|R| = 8$ routes, only $m = 4$ waiting positions have been assumed for every route in the queueing system. Here, the maximum rate is modelled with $M = 600$, similar to the models utilized for validation in Section 4. In total, the analysed $MC_{\text{Gagny}} = (S, T)$ consists of $|S| = 1.25 \cdot 10^7$ states and $|T| \approx 1.35 \cdot 10^8$ transitions.

The solving process has been automated using Python 3.10.9 (Van Rossum and Drake, 1995; Python Software Foundation, 2022) and probabilistic model-checking Software. For this work, the model-checker *Storm* (Hensel et al., 2022) and its Python-Interface *Stormpy* (Junges and Volk, 2023) have been utilized. Used Models have been formulated according to Section 3.4, expressed in the PRISM modelling language (Parker et al., 2000; Kwiatkowska et al., 2011). Detailed information regarding the solving process and the used model files can be found in the online repository (Emunds and Nießen, 2023). The solution time for a model is between 15 and 20 min on six cores of an Intel Xeon Platinum 8160 Processor (2.1 GHz), utilizing 23 400 MB working memory.

The results of the probabilistic model-checking algorithms are the expected queue lengths $E_{LW,r_i}(M/M/s/m) = E^*_{LW,i}$ on the routes r_i when describing exponentially distributed inter-arrival and service times. By assuming v_A and v_S according to Table 7, the expected queue lengths $E_{LW,r_i}(GI/GI/s/m) = E_{LW,i}$ can be estimated with Eq. (24).

In Fig. 10, the estimated queue lengths for all 8 routes are depicted for the three scenarios. Furthermore, the threshold

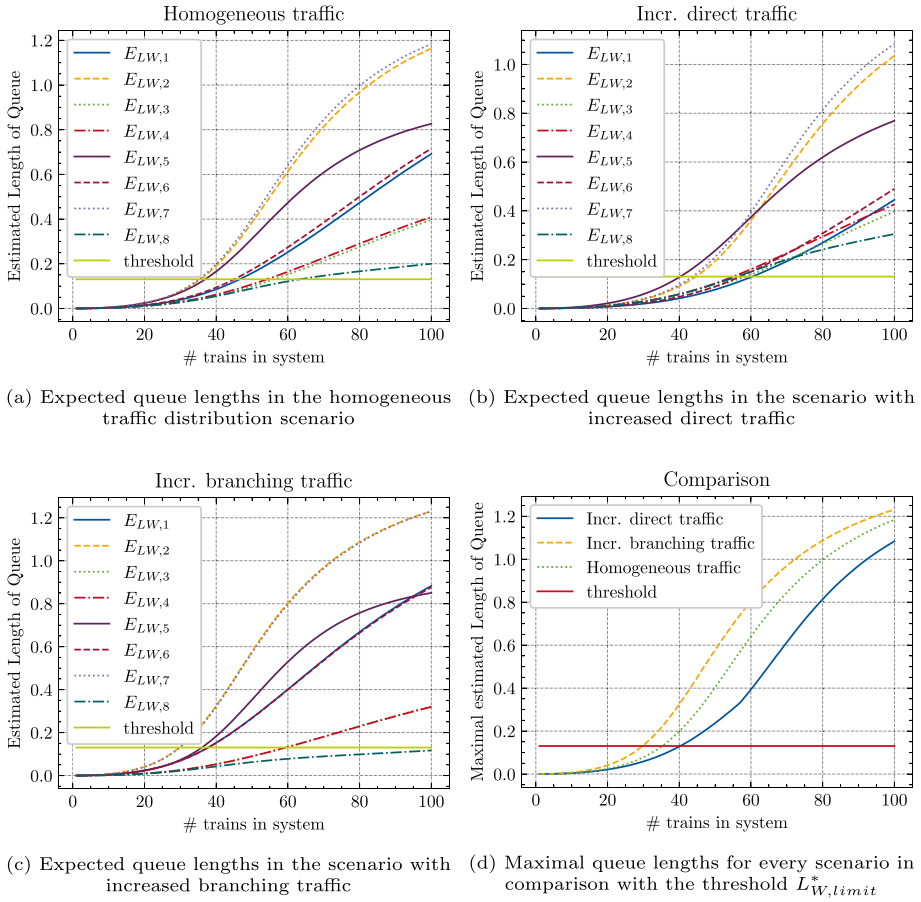
$$L^*_{W,limit} = 0.479 \cdot \exp(-1.3 \cdot 1) \approx 0.131 \tag{35}$$

for the queue length in a system with only passenger trains (see Section 3.7) is shown.

Since the traffic distribution varies between the calculated scenarios, curve characteristics differ for the plotted queue length estimations.

In the homogeneous traffic scenario, the longest queue is estimated for route r_7 (crossing multiple other routes on its way from D to A). Another crossing route, r_2, has a similar progression. As the straight route r_5 has multiple conflict points with crossing routes, its estimated queue length is higher than the one of the other branching (r_1 and r_6) and direct (r_3, r_4 and r_8) routes.

In the scenario with increased direct traffic, route r_5 is the route with the highest estimated queue length for a total number of trains of less than 60. If more than 60 requests are to be scheduled in an hour, the routes r_7 and r_2 are estimated with a higher queue length than r_5.



(a) Expected queue lengths in the homogeneous traffic distribution scenario

(b) Expected queue lengths in the scenario with increased direct traffic

(c) Expected queue lengths in the scenario with increased branching traffic

(d) Maximal queue lengths for every scenario in comparison with the threshold $L^*_{W,limit}$

Fig. 10. Obtained estimated queue lengths for every scenario (a–c) and comparison of the maximal queue length of any route per scenario (d).

Table 11
Timetable capacity for all considered scenarios and selected parameters of the significant operating program A^*_θ.

	n_{total}	n_1	n_2	n_3	n_4	n_5	n_6	n_7	n_8	Significant route	$p_{\text{loss,max}}$
a	35	4.38	4.38	4.38	4.38	4.38	4.38	4.38	4.38	r_7	0.0103
b	40	3.333	3.33	6.67	6.67	6.67	3.333	3.333	6.67	r_5	0.0118
c	29	4.83	4.83	2.42	2.42	2.42	4.83	4.83	2.42	r_2	0.0098

For the scenario with more traffic on the branching routes, $E_{LW,2}$ outgrows the queue length of every other route, while $E_{LW,7}$ is only marginally smaller.

When comparing the maximal queue length $\max_i E_{LW,i}$ of all routes for a given scenario with the threshold $L^*_{W,limit}$ for sufficient quality (Fig. 10(d)), the maximal number of requests per hour can be assessed by selecting the corresponding operating program A^*_θ. This timetable capacity differs for the different scenarios, significant values can be read in Table 11.

Summarizing the results, timetable capacity differs depending on the selected traffic share. While only $m = 4$ waiting positions per route have been included in the created CTMC due to memory restrictions, the maximum probability of a loss $p_{\text{loss,max}}$ on all routes has been calculated as ≈ 0.01 for every considered scenario. This can serve as evidence for the soundness of the presented model, even though the model has been restricted to this amount of waiting positions.

In conclusion, the introduced method enables infrastructure managers to estimate the performance of junction infrastructure even before planning a detailed timetable, by restricting the required input to route-exclusions and traffic share.

5.2. Infrastructure layout planning

The introduced timetable capacity analysis allows the infrastructure operator to quantify the number of plannable trains in the timetable process considering a fixed infrastructure layout. It therefore utilizes pre-specified service times for each different route through the infrastructure.



Fig. 11. Track-layout for a double-track junction with an overpass.

Table 12
Number of trains by traffic type for the considered operating programs.

Operating program	# regional trains	# high speed trains	# freight trains	Share of passenger trains p_{pt}
Low intensity mixed	2	0	1	0.67
Long distance	0	4	0	1
Local train	4	1	0	1
Freight train	0	0	5	0
High intensity mixed	4	2	2	0.75
Urban railway	10	0	0	1

However, infrastructure operators might be interested in not only evaluating a single infrastructure layout regarding its capacity but also in choosing one layout out of a collection. With the following application of the introduced method, different infrastructure layouts can be compared with the performance indicator *maximum mean service time*. That indicator describes an upper bound to the average service time, that a request in a given operating program may utilize.

This section therefore particularly focuses on choosing the right infrastructure layout for a given operating program, i.e. problem statement (II) in Section 3.2.

5.2.1. Setup

In detail, a case study, deciding whether or not an overpass should be built for a junction with a given operating program has been conducted. An overpass is a way to reduce the number of route conflicts in a railway junction, an example for the track layout of the junction in Fig. 1 is given in Fig. 11.

To show the applicability of the described analytical approach, 23 different operating programs have been considered. Note the definition of an operating program in Section 3.2 as the train count on a route per U. In total, the operating program of the railway junction is assumed to have the same number of trains on the routes of the main line and the same number of trains on the routes of the branching line. Thus, all combinations of 23 different operating programs for both to the railway junction adjacent railway lines have been built and analysed for their peak traffic hour, resulting in a total number of $23^2 = 529$ different examples. A detailed description of all considered operating program combinations can be found in Appendix D, we restrict ourselves to six exemplary railway lines here.

The operating programs have been selected according to different types of adjacent railway lines, we distinguish between *mixed traffic lines*, *local train lines*, *long-distance train lines*, *freight train lines*, and *urban railway lines*. Additionally, two different loads have been considered for the mixed traffic line. Table 12 lists the selected operating programs and their corresponding share of passenger traffic p_{pt}. Since we refrain from distinguishing different service rates per train type, the influence of the train mix in the operating program is mainly the definition of route-specific arrival rates and the determination of the share of passenger trains for the calculation of the threshold $L_{W,limit}$ (see Section 3.7).

In Fig. 12 the considered combinations of those operating programs have been listed. In every entry, the top bar corresponds to the operating program on the *main line*, i.e. the routes from A to B (r_1) and from B to A (r_3), while the bottom bar corresponds to the operating program on the *branching line*, i.e. the routes from (r_4) and to (r_2) C.

The total number of trains on a route n_r is utilized to determine the arrival rate $\lambda_r = \frac{1}{n_r}$ for this route. Next, Continuous-Time Markov Chains (see Section 3.4) can be formulated for any possible service rate $\mu \in (0, 1]$, using the defined arrival rates λ_r and the maximum number of waiting positions per queue $m = 5$, while also considering conflicting routes. To ensure sufficient precision while also maintaining efficient computability, all service rates $\mu \in [0.01, 1]$ with a step size of 0.01 have been taken into consideration in this work.

Hence, 100 Models have been analysed for every route in both junction infrastructure layouts with/without an overpass for every considered combination of operating programs. The solving process has been automated utilizing the same software as in the case study in Section 5.1, similarly, the model files are available in the online repository (Emunds and Nießen, 2023).

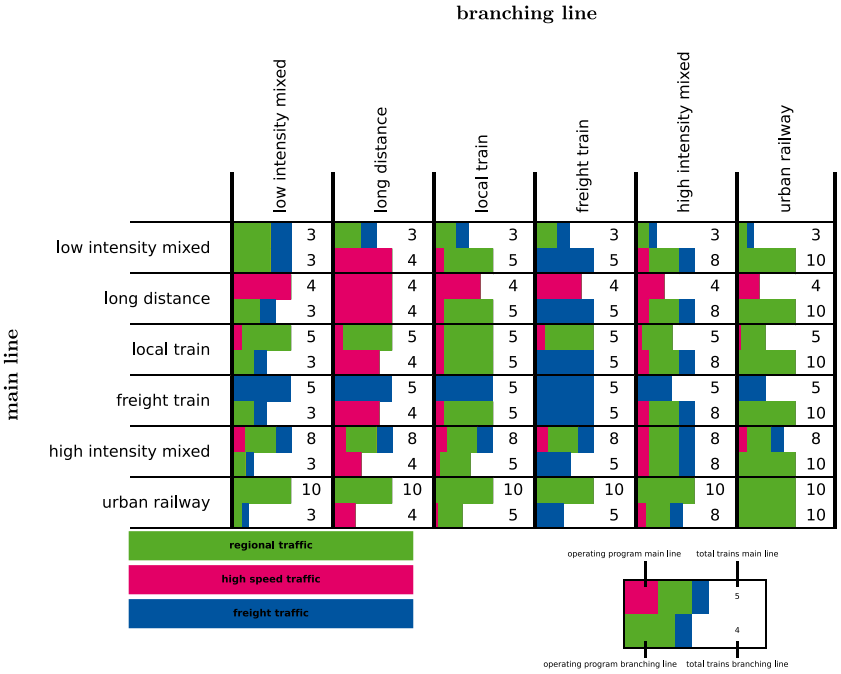


Fig. 12. Overview of the considered operating program combinations.

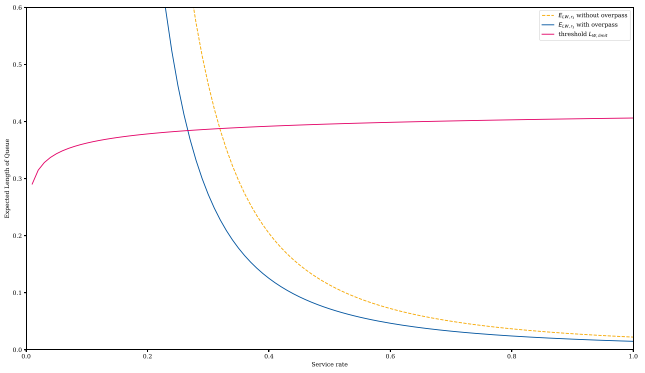


Fig. 13. Estimated queue length on route r_3 for the combination of a local train main line and a long-distance train branching line.

Aiming to gain insights into the performance benefit of the infrastructure layout including the overpass (see Fig. 11), the expected queue lengths at every route have been analysed. To facilitate a sufficient infrastructure quality, the most limiting service rate, hence the highest service rate of all routes in the junction, has been chosen.

Resulting from the solving process is a grid of expected lengths of the queue for both infrastructure layouts for every considered service rate $\mu \in [0.01, 1]$. In Fig. 13 the expected queue length at route r_3 from B to A for both railway junction layouts of a local train main line and a long-distance train branching line is recorded. Selected values are listed in Appendix C.

Both operating programs consider passenger transport only, hence a ratio of passenger trains in all considered trains of $p_{pt} = 1$ can be used for the calculation of the threshold value $L_{W,limit}$.

The combination is assumed to have to deal with an hourly load of 5 trains on the main line, from A to B and back, as well as 4 additional trains on the branching line, from A to C and back. This resumes in an arrival rate of $\lambda_r = 0.083$ for the routes $r \in \{r_1, r_3\}$ and of $\lambda_{r'} = 0.067$ for the routes $r' \in \{r_2, r_4\}$.

When observing the calculated estimations for the queue length E_{LW,r_3} at route r_3, the difference between the results for the infrastructure layouts with or without an overpass can be analysed. With a rising mean service rate μ in the infrastructure, the difference between the estimated queue length between the two infrastructure layouts drops significantly. This can be explained by the fact, that resource conflicts become less likely when less service time is consumed by the requests, indicated by a higher service rate. Hence the omission of the additional route conflict in the scenario with an overpass is much less relevant for high service rates.

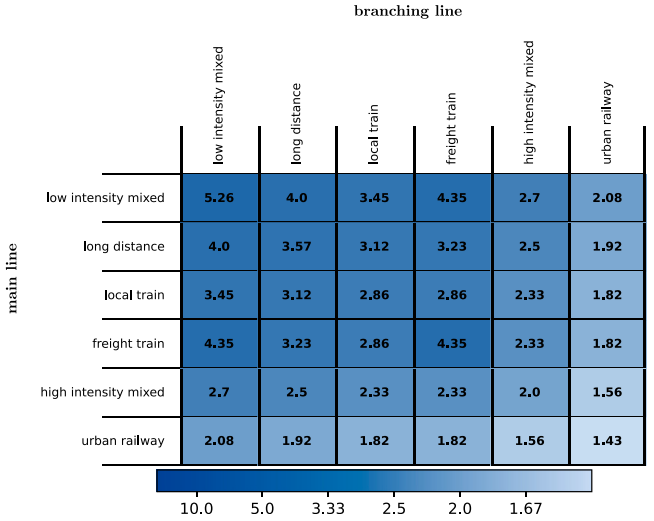


Fig. 14. Calculated maximum mean service times in minutes for the scenario without an overpass.

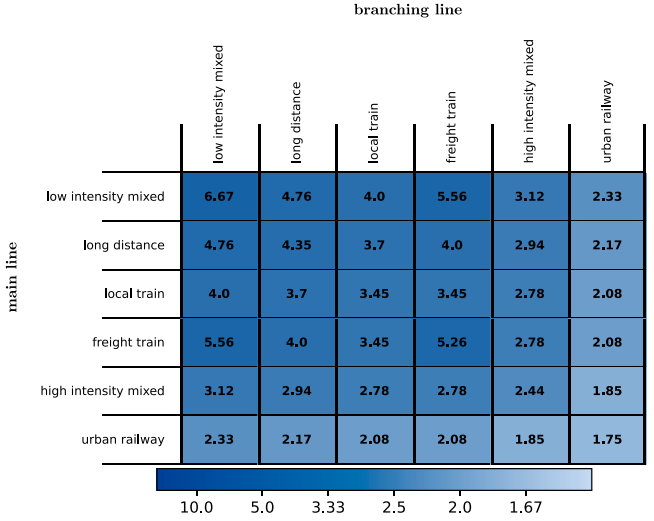


Fig. 15. Calculated maximum mean service times in minutes for the scenario with an overpass.

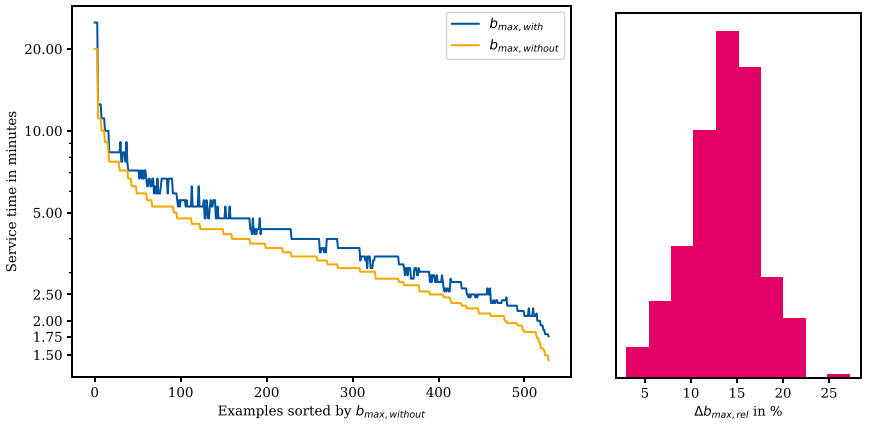
Additionally, the threshold $L_{W,limit}$ is depicted, modified according to Section 3.7, depending on the occupancy rate of $\rho = \frac{\lambda_{r_3}}{\mu}$ for route r_3. The assumed parameters are listed in Table 7.

Utilizing the grid of resulting queue length estimations for every route and the calculated threshold, a minimum mean service rate $\mu_{\min}$, needed for sufficient infrastructure quality, may be obtained. Hence, the maximum mean service time $b_{\max} = \frac{1}{\mu_{\min}}$ for the given operation program on the main and branching lines can be derived. This maximum mean service time can be used to investigate the needed infrastructure by comparing it to the actual achieved service times on the analysed junction, which are subject to train and control system specific parameters.

5.2.2. Results

The introduced derivation of a maximum mean service time has been applied to all 529 considered operating program combinations for the infrastructure settings with and without an overpass.

In Figs. 14 and 15 the results for the selected 36 combinations are shown in heat-maps for both considered railway junction layouts. While both infrastructure settings share the same global distributions of service time requirements, results show that under



(a) Distribution of maximum mean service times on all examples

(b) Relative difference in maximum mean service times

Fig. 16. Comparison between the obtained maximum mean service times of the two infrastructure scenarios.

the same load, as fixed by the operational program of the main and branching line, a railway junction including an overpass structure allows for a higher mean service time, while achieving the same operational quality as a railway junction without the overpass structure.

Considering the example of a railway junction (see Fig. 13) with a local train main line (5 trains per direction) and a long-distance train branching line (4 trains per direction), the following results are obtained. To implement an infrastructure layout without an overpass a mean service time of at most 3.12 min needs to be fulfilled. Furthermore, the infrastructure layout with an overpass would allow for a mean service time of at most 3.7 min. By constructing an overpass, trains would be allowed to utilize 0.58 min (18.59%) more service time.

With the calculated maximum mean service times, the infrastructure operator can specify precise planning guidelines for infrastructure planners when to plan for an overpass structure, or when to decrease average mean service times by implementing more efficient railway control systems. After developing the microscopic railway junction design scenario, infrastructure planners can decide whether the specified upper bound is adhered to with a microscopic minimum headway time calculation.

Further investigating the difference between the two infrastructure settings, Fig. 16 includes the maximum mean service times (Fig. 16(a)) for all 529 computed combinations of main and branching lines (see also Appendix D) and a histogram, representing the distribution of relative differences in the calculated maximum mean service times (Fig. 16(b)).

Taking the logarithmic scale in Fig. 16(a) into account, a wide range of calculated maximum mean service times required for sufficient operational quality can be recognized. While some operating programs, i.e. examples with a low number of total trains in the considered peak hours, are granting sufficient operational quality even for mean service times as high as 10 to 20 min, the majority of considered examples require a mean service time of under 5 min, with very densely operated main and branching lines demanding maximum mean service times of even under 2 min.

By including an overpass in the layout of a planned railway junction, the required maximum mean service times can be decreased by a significant margin. For the majority of considered examples, the relative difference in the maximum mean service time lies between 10% and 20%, but the achieved reduction of maximum mean service time is diverse. Some examples show only a marginal difference between the considered infrastructure layouts, others report maximum mean service time decreases by up to 25%.

Crucially, service times are dependent on various factors, some of which are not under the influence of infrastructure managers. Hence, substantial limitations to the lower bound of service times are relevant, affected by physical properties such as length or acceleration and braking performance of the rolling stock.

In conclusion, differences between the considered infrastructure layouts have been analysed with the introduced method. For this, the calculation of many different operating programs and service rate configurations has been conducted, resulting in substantiated estimations of infrastructure quality requirements. The results clearly show the applicability of the introduced approach to quantify the effect of overpass structures in railway junctions.

5.2.3. Comparison with capacity utilization

The introduced method calculates the estimated queue length on a route through the junction infrastructure for different service times. It compares the results to the threshold of the maximum queue length (see Section 3.7) to obtain the maximum possible mean service time on all routes b_{max} for sufficient quality. While this threshold is in practical use for the timetable capacity calculation in

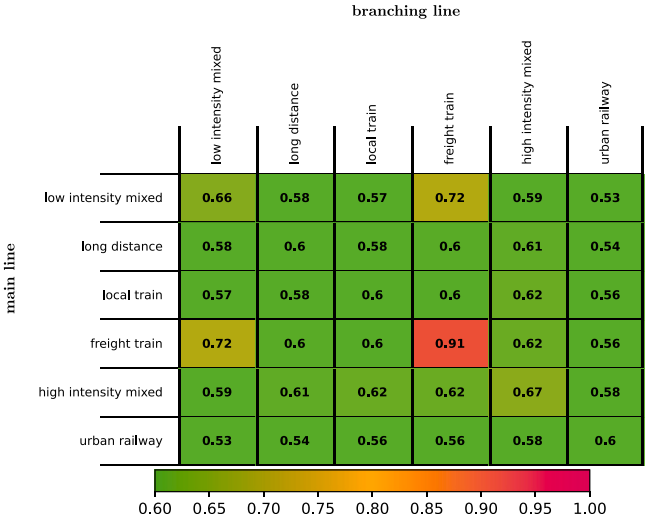


Fig. 17. Occupancy rates (following Eq. (38)) for the scenario without an overpass, assuming mean service times from Fig. 14.

Germany (DB Netz AG, 2022), internationally the UIC 406 Method (UIC, 2013) is applied more often, planning railway infrastructure according to the capacity utilization of a fixed timetable.

For this, the occupation rate ρ_{UIC} for the infrastructure is calculated by compressing the service times of all trains $b^* = \sum_i b_i$ and dividing it by the time period t_U

$$\rho_{UIC} = \frac{\sum_i b_i}{t_U}. \tag{36}$$

For this method, an occupation rate between 60% and 80%, $\rho_{UIC} \in [0.6, 0.8]$, is defined as a sufficient quality in railway node switch areas.

While the introduced methodology is timetable-independent, the UIC

Method is based on a fixed timetable and can therefore utilize the order of train runs and routes to calculate the minimum headway times between subsequent train runs on a railway junction. However, in previous research (Schwanhäußer, 1978; Hansen and Pachl, 2014) a timetable-independent method to calculate the *connected occupation rate* ρ_{conn} of a railway junction by utilizing the *concatenation value* ψ has been proposed. The concatenation value of a railway junction can be obtained with the operating program $A = \{(r_i, n_i) | i \in R\}$ by calculating the frequency of a sequence i-j, $p_{i,j} = (n_i * n_j)/(n^2_{total})$, with the total number of trains in the operating program n_{total} and

$$\psi = \sum_{i,j} p_{i,j} \cdot C_{i,j}, \tag{37}$$

applying the conflicts described in the conflict-matrix C (Potthoff, 1970). The connected occupancy rate can further be obtained with the knowledge of a mean service time b with

$$\rho_{conn} = \frac{b \cdot n_{total} \cdot \psi}{t_U}. \tag{38}$$

Fig. 17 depicts the connected occupancy rates of a railway junction, considering the maximum mean service rates computed in Section 5.2.2 for the infrastructure setup without an overpass. It indicates the similarity between the utilized threshold, describing the limit on the queue length on the route r_3, and the capacity utilization threshold formulated in UIC (2013).

For the considered examples, most calculated maximum mean service times correspond to capacity occupations in the desired range of 60% to 80%. However, for operating programs with a lot of freight trains, the occupation factors are substantially higher, reaching up to 91%. This can be explained by the dependence of the utilized threshold $L^*_{W,limit}$ from the ratio of passenger trains p_{pt} in the operating program.

Some maximum mean service time results yield capacity occupation factors of less than 0.6, with the lowest value at 0.53. This can be observed for some operating programs with a very small calculated maximum mean service time, such as the instances for urban railways. In those cases, the amount of traffic on the main line is greater than the amount of traffic on the branching line. Hence, the measured quality on the main line is worse than in the combined junction. Therefore, a small maximum mean service time is needed to adhere to the quality threshold on the main line.

Similar observations can be made for the results with an overpass and when considering all operating program combinations (see Appendix E).

Noting the difference in the objectives, timetable-independent vs. timetable-dependent analysis, and the additional train mix sensitivity of the utilized threshold, the introduced method yields plausible results for the considered operating program combinations. Since the introduced method provides route exact results, it is sensitive to differences in the occupation of different routes through the infrastructure. Therefore, the calculated values of all routes should be considered for infrastructure planning, in particular for instances with diverging main and branching line operating programs.

6. Discussion and outlook

This work introduced a novel method for analysing the timetable capacity of railway junctions based on queueing theory. It is applicable for solving both formulated problem statements (Section 3.2), timetable capacity determination of a given railway junction infrastructure (I) and planning of junction infrastructure for a fixed operating program (II). By modelling railway junction routes as parts of a queueing system, while respecting their parallel service possibilities and considering resource conflicts, timetable independent analyses of examined infrastructure are enabled.

Extending classical queueing theory concepts, well established for railway line performance analysis (see Section 3.3), timetable capacity is determined by comparing queue length estimations through a novel Continuous-Time Markov Chain formulation representing the considered railway junction (see Section 3.4) with threshold values (see Section 3.7) depending on parameters of considered service and arrival process distributions along with operating program specifics. In this work, estimations of queue lengths have been carried out by probabilistic model-checking software, enabling a fast and reliable computation for complex infrastructure dependencies.

With the introduced CTMC, railway junctions of up to 8 routes can be described without further amendments, however, railway line intersections of more than 2 lines and railway stations may consist of a higher number of routes. While the capabilities of the presented modelling approach will advance with increasing memory availability in future computer architecture generations, it is unlikely that those performance developments can fully account for the route-dependent exponential scaling of the introduced state set. It is therefore crucial to enforce the development of approximation techniques and decomposition approaches in further research.

The performance of the introduced approach has been studied by exemplary comparing the computation times and results of the analytical solution with simulations (see Section 4). When compared to simulations with an unlimited number of waiting positions, differences between the calculated queue lengths can be identified. Considering that this difference decreases with a reduction of the occupation ratio, the accuracy of the introduced method is accurate enough for a sufficient analysis within the neighbourhood of and below the threshold values, which correspond to the quality, where practical operation of railway traffic usually takes place. The conducted simulations indicate the high accuracy of the introduced method when limiting the number of waiting positions. Furthermore, the described calculation process determines the analytically correct solution to the presented model, while also being significantly faster than the applied simulations, which can only combine individual representations of the randomized timetabling process.

In an implementation of the novel analytical model for two separate case studies (see Section 5), the two problem formulations, determining the timetable capacity of a railway junction and selecting an infrastructure layout for a given operating program, have been tested. As an initial example, the timetable capacity of a complex railway junction in the west of Paris has been estimated for different traffic share scenarios. Additionally, a case study describing the selection of sufficient junction infrastructure, assuming 529 different operating program combinations, has been conducted, leading to substantive estimations regarding the effect of overpass structure on the timetable capacity of railway junctions. In a comparison with an approach calculating the capacity occupation of the junction infrastructure, the route-sensitive accuracy of the obtained results has been highlighted. Hence, the introduced method has been proven to be able to create accurate solutions for manifold application scenarios, ranging from an abstract analysis of the general effect of overpass infrastructures to a practical implementation in the planning process of railway infrastructure operators when evaluating the performance of infrastructure layouts.

Even though the concept is already applicable, additional research could still yield substantial benefits. As such, the modelling of general independent stochastic processes with Markov processes and approximation factors might be improvable, i.e. by including phase-type distributions. Similarly, utilized thresholds for the expected length of a queue have been introduced for use in the context of railway lines, updating these to innovative measures could improve the real-world applicability. Additionally, the introduced concept could be extended to implementations for railway stations and eventually railway networks. By including delay distributions and propagation in the model, similar measures for operational capacity could be enabled.

However, by utilizing the developed timetable independent method, infrastructure managers can identify bottlenecks in the early stages of the planning process. With the achieved computing time benefits when compared to simulation approaches, they are set up to give junction designers substantiated indicators for the required infrastructure. Hence, the introduced analytical timetable independent approach has been proven to form a valuable addition to the capacity method landscape.

CRediT authorship contribution statement

Tamme Emunds: Writing – original draft, Software, Methodology, Formal analysis, Conceptualization. **Nils Nießen:** Writing – review & editing, Supervision, Conceptualization.

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.

Data availability

Research Data has been made available at an online repository.

Acknowledgements

The authors thank Mr. Alexander Bork for his invaluable insights regarding probabilistic model-checking techniques in general and regarding the software Storm in specific. Furthermore, the authors thank Mr. Tobias Müller and Dr. Andreas Pfeifer for their supervision and guidance regarding practical implementations as well as for providing application-relevant case examples. Additionally, the authors thank DB Netz AG for the opportunity to apply the described theory in a practical project.

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

Appendix A. Literature summary

See Table A.13.

Table A.13
Literature considering railway performance estimations.

Reference	Capacity type	Timetable dependence	Methodology	Infrastructure type	Infrastructure decomposition	Solution technique
UIC (2004)	utilization	dependent	compression	line	block-section	closed-form formula
UIC (2013)	utilization	dependent	compression	line and junction	block-section and route-based	closed-form formula
Abril et al. (2008)	utilization	dependent	compression	line	block-section	closed-form formula
Landex (2009)	utilization	dependent	compression	line	block-section	closed-form formula
Landex and Jensen (2013)	utilization	dependent	compression	junction	block-section	closed-form formula
Goverde et al. (2013)	utilization	dependent	compression	line	train-path	closed-form formula
Jensen et al. (2017)	utilization	independent	compression	single-line network	train-path	combinatorial algorithms
Jensen et al. (2020)	utilization	independent	compression	single-line network	train-path	combinatorial algorithms
Weik et al. (2020)	utilization	independent	compression	station	train-path	closed-form formula
Goverde (2007)	utilization	dependent	max-plus-algebra	network	timed-events	matrix-calculations
Bešinović and Goverde (2018)	utilization	dependent	max-plus-algebra	network	timed-events	matrix-calculations
De Kort et al. (2003)	operational	independent	max-plus-algebra	single-line network	timed-events	matrix-calculations
Harrod (2009)	theoretical	timetable-saturation	optimization	line	train-path	MIP
Yaghini et al. (2014)	theoretical	independent	optimization	line	timed-events	MIP
Zwaneveld et al. (1996)	theoretical	independent	optimization	station	route-based	MIP
Zwaneveld et al. (2001)	theoretical	independent	optimization	station	route-based	MIP
Delorme et al. (2001)	theoretical	independent	optimization	junction	block-section	combinatorial algorithms
Lusby et al. (2011)	(theoretical)	timetable-saturation	optimization	junction	block-section	MIP
Burdett and Kozan (2006)	theoretical	timetable-saturation	optimization	network	train-path	MIP

(*continued on next page*)

Table A.13 (*continued*).

Burdett (2015)	theoretical	independent	optimization	network	train-path	combinatorial algorithms
(Cacchiani and Toth, 2012)	(theoretical)	independent	optimization	–	–	mostly MIP
Leutwiler and Corman (2022)	(theoretical)	independent	optimization	network	timed-events	disjunctive programming
Zhang and Nie (2016)	utilization	independent	optimization	line	timed-events	combinatorial algorithms
Liao et al. (2021)	theoretical	timetable-saturation	optimization	line	timed-events	MIP
Mussone and Wolfler Calvo (2013)	utilization	independent	optimization	network	train-path	MIP
Graffagnino (2012)	operational	dependent	operational-data-analysis	line	train-path	statistical analysis
Armstrong and Preston (2017)	utilization	dependent	operational-data-analysis	station	timed-events	statistical-analysis
Weik (2022)	operational	dependent	operational-data-analysis	line	block-section	statistical-analysis
(Corman and Henken, 2022)	operational	dependent	operational-data-analysis	line	train-path	statistical-analysis
Goverde (2010)	(operational)	dependent	max-plus-algebra	network	timed-events	matrix-calculations
Büker (2011)	(operational)	independent	analytical	network	timed-events	iterative formula
Şahin (2017)	(operational)	dependent	analytical	line	timed-events	matrix-calculations
Zieger et al. (2018)	(operational)	independent	simulation	line	train-path	simulation
Corman and Kecman (2018)	(operational)	dependent	operational-data-analysis	line	timed-events	bayesian-network
(Spanninger et al., 2023)	(operational)	–	operational-data-analysis	–	–	–
Potthoff (1970)	timetable	independent	analytical	station	route-based	iterative formula
Schwanhäußer (1974)	operational	independent	analytical	line	train-path	closed-form formula
Schwanhäußer (1978)	timetable	independent	analytical	junction	route-based	closed-form formula
Wakob (1984)	timetable	independent	analytical	line	train-path	closed-form formula
Wendler (2007)	timetable	independent	analytical	line	train-path	matrix-calculations
Nießen (2008)	operational	independent	analytical	junction	individual	iterative formula
Nießen (2013)	timetable	independent	analytical	junction	individual	iterative formula
Schmitz et al. (2017)	timetable	independent	analytical	junction	individual	matrix-calculations
Weik (2020)	timetable	independent	analytical	line and junction	route-based	matrix-calculations
Introduced here	timetable	independent	analytical	junction	route-based	prob. model-checking

Remarks: Please note that literature referring to the *(operational)* capacity type is introducing methods for computing delay propagation, which is the basis for calculating operational capacity. Furthermore, literature referring to the *(theoretical)* capacity type is introducing methods for railway timetabling, which can be used to compute theoretical capacity. Literature written within bracelets *(reference)* does not introduce new methods but rather gives a comprehensible review of relevant methodology.

Appendix B. Supplementary boxplots

See Figs. B.18–B.21.

Appendix C. Table with exemplary results

In Section 5.2, Fig. 13 depicted the expected queue length at route r_3 (see Fig. 1) from B to A for both railway junction layouts of a local train main line and a long-distance train branching line. Supplementary, selected results are listed in Table C.14.

Appendix D. Full computational study

The computational study explained in Section 5.2 has been conducted for 529 operating program combinations in total. Since we refrained from showing their results in full detail in the text, some additional figures are given on the next pages. Please note that the results can additionally be found in the online repository (Emunds and Nießen, 2023).

Table D.15 lists all selected operating programs in Section 5.2 and their respective names in the following figures.

In Fig. D.22 all used operating program combinations are depicted. Considered operating programs have been created as samples for demands to railway infrastructure and named in accordance with European Comission (2019), additionally distinguished between

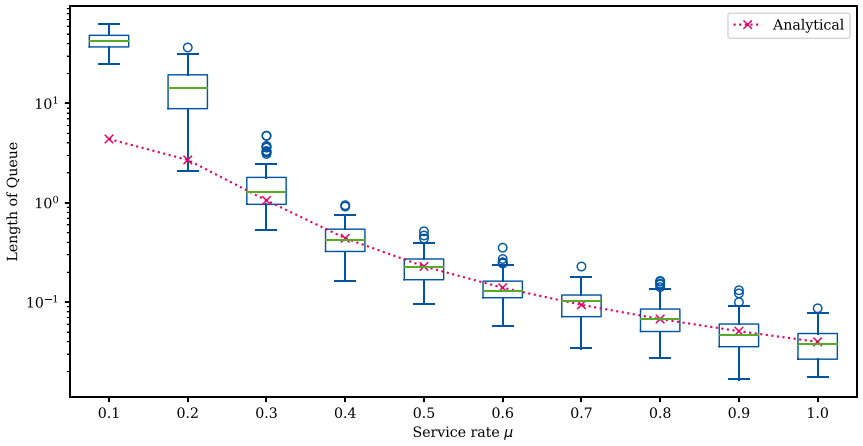


Fig. B.18. Comparison between the analytical solution and the simulations with an unlimited m.

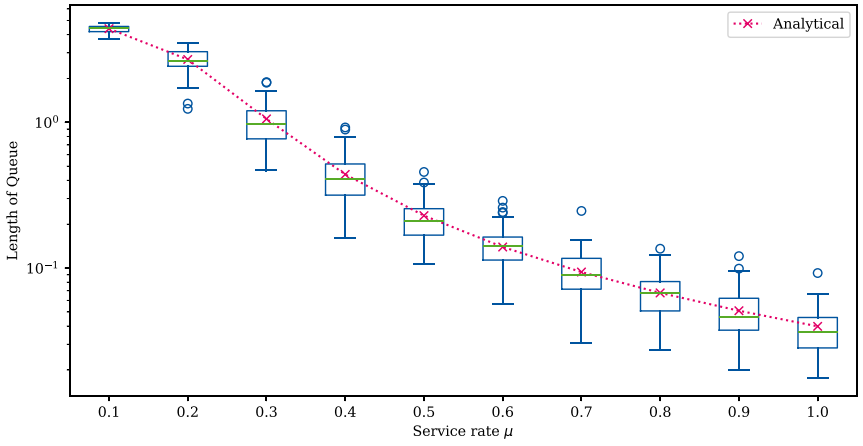


Fig. B.19. Comparison between the analytical solution and the simulations with a limit of $m = 5$.

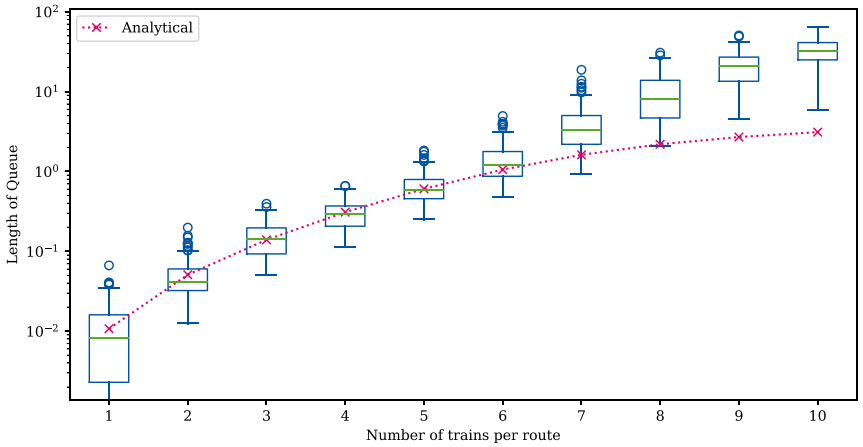


Fig. B.20. Comparison between the analytical solution and the simulations with an unlimited m, assuming different numbers of trains per route and a fixed service rate $\mu_i = 0.3$ for all $r_i \in R$.

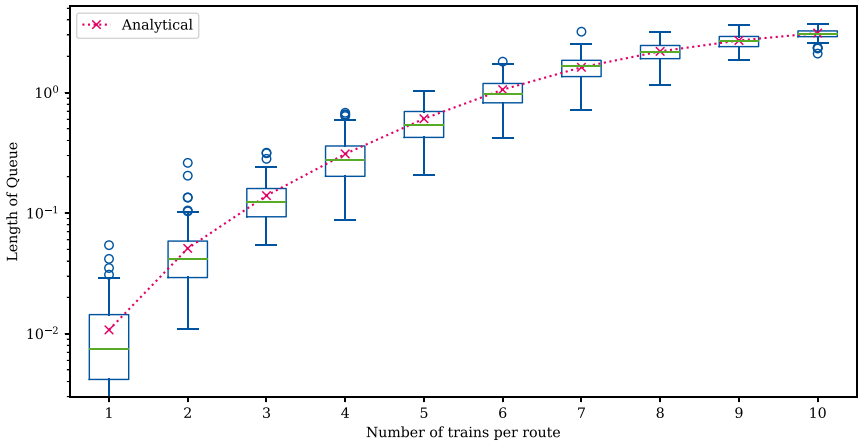


Fig. B.21. Comparison between the analytical solution and the simulations with a limit of $m = 5$, assuming different numbers of trains per route and a fixed service rate $\mu_i = 0.3$ for all $r_i \in R$.

Table C.14
Estimated queue length on route r_3 for the combination of a local train main line and a long-distance train branching line.

μ	$E_{LW,3}$ with overpass	$E_{LW,3}$ without overpass
0.01	4.9362	4.9575
0.1	3.6028	3.9855
0.2	0.9082	1.5143
0.3	0.2723	0.4654
0.4	0.1252	0.2042
0.5	0.0717	0.1135
0.6	0.0465	0.0722
0.7	0.0326	0.0500
0.8	0.0241	0.0367
0.9	0.0186	0.0281
1	0.0147	0.0222

Table D.15
Selected operating programs.

Name of railway line	Name in appendix
Long-distance	P1, (0)
Local train	P3, (0)
Urban railway	P3-P5, (+)
Freight train	F1, (0)
Low intensity mixed	P3_F1, (0)
High intensity mixed	P3_F1, (+)

Table D.16
Selected traffic code parameters from European Comission (2019).

Traffic code	Traffic type	Line speed [km/h]	Useable length of platform [m]	Train length [m]
P1	Passenger traffic	250–350	400	n.a.
P2	Passenger traffic	200–250	200–400	n.a.
P3	Passenger traffic	120–200	400–400	n.a.
P5	Passenger traffic	80–120	50–200	n.a.
F1	Freight traffic	100–120	n.a.	740–1050
F2	Freight traffic	100–120	n.a.	600–1050
F3	Freight traffic	60–100	n.a.	500–1050

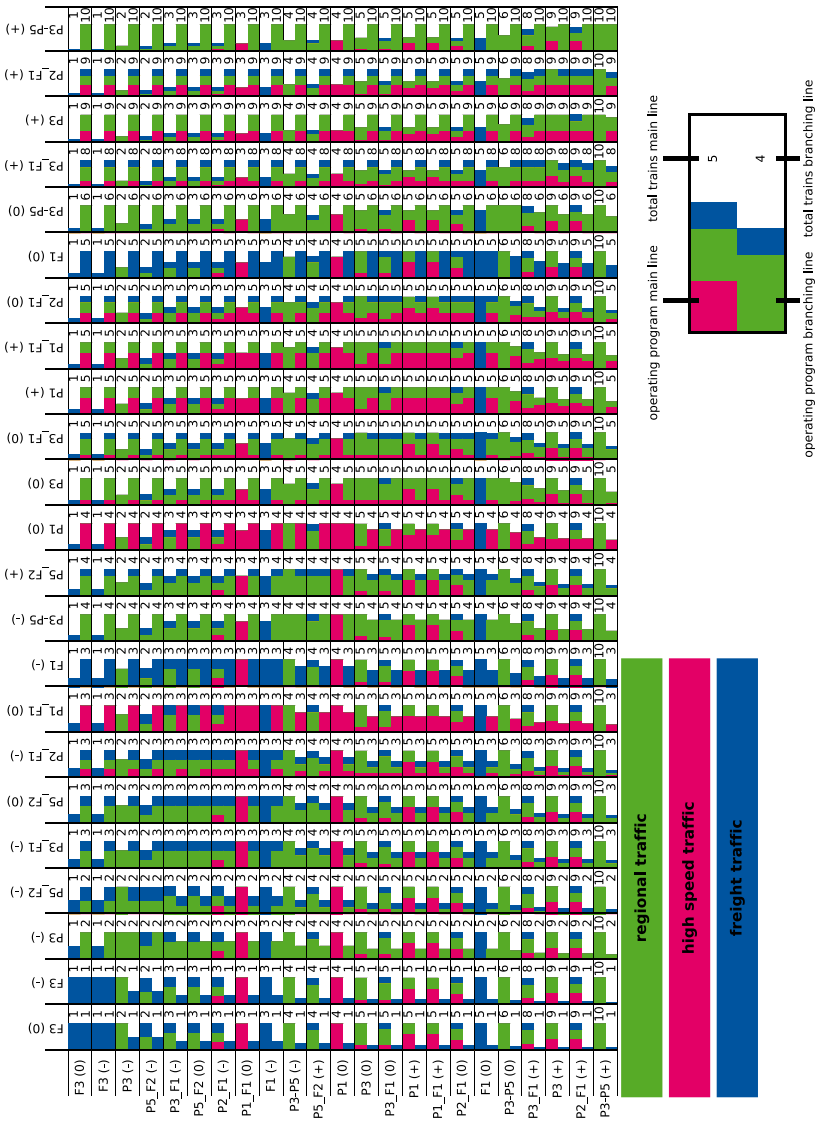


Fig. D.22. Overview of all operating program combinations.

the three categories of exposure *base load* (−), *medium load* (0) and *high load* (+). Table D.16 lists a selection of traffic code parameters as specified in European Comission (2019).

Additionally, the corresponding results are depicted in Figs. D.23 and D.24 for the calculations of maximum mean service times for the two infrastructure settings without and with an overpass.

Appendix E. Full comparison with capacity utilization

See Figs. E.25–E.27.

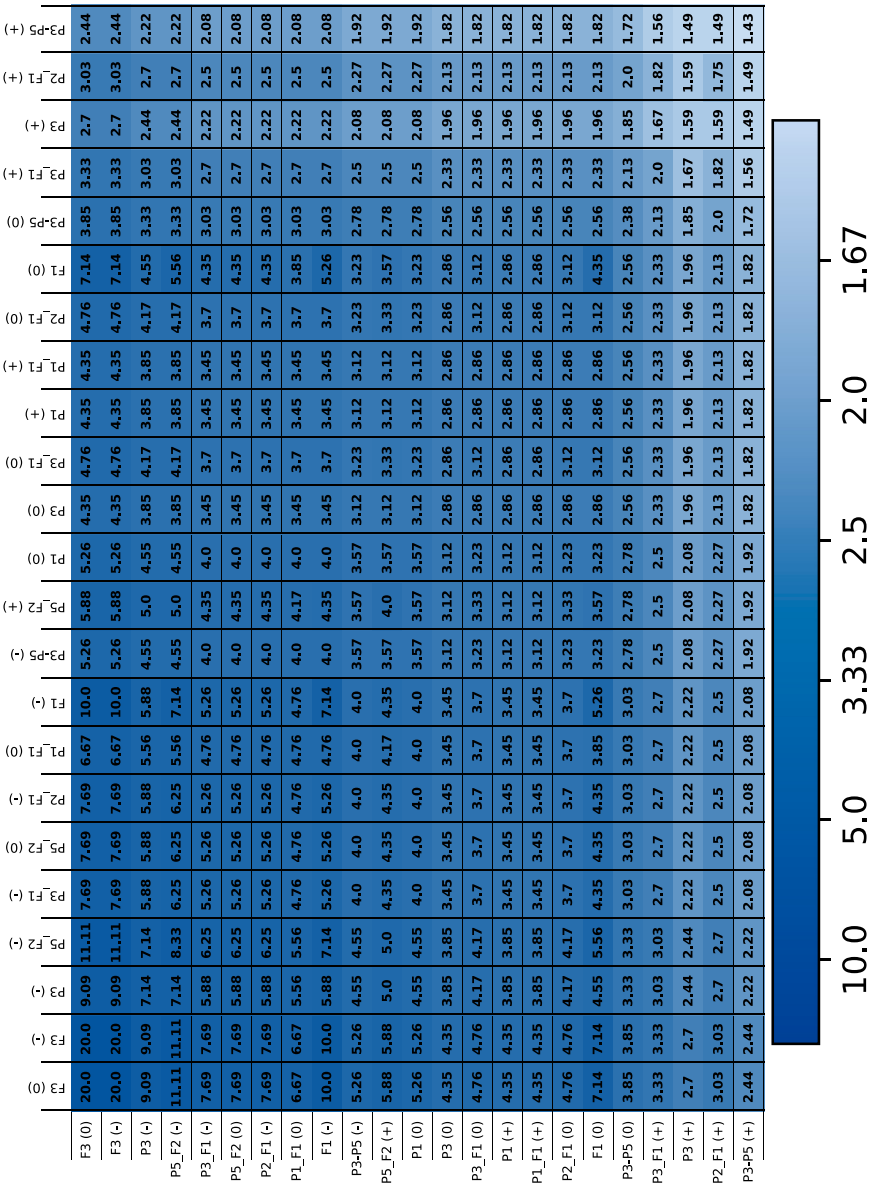


Fig. D.23. Calculated maximum mean service times in minutes for the scenario without an overpass.

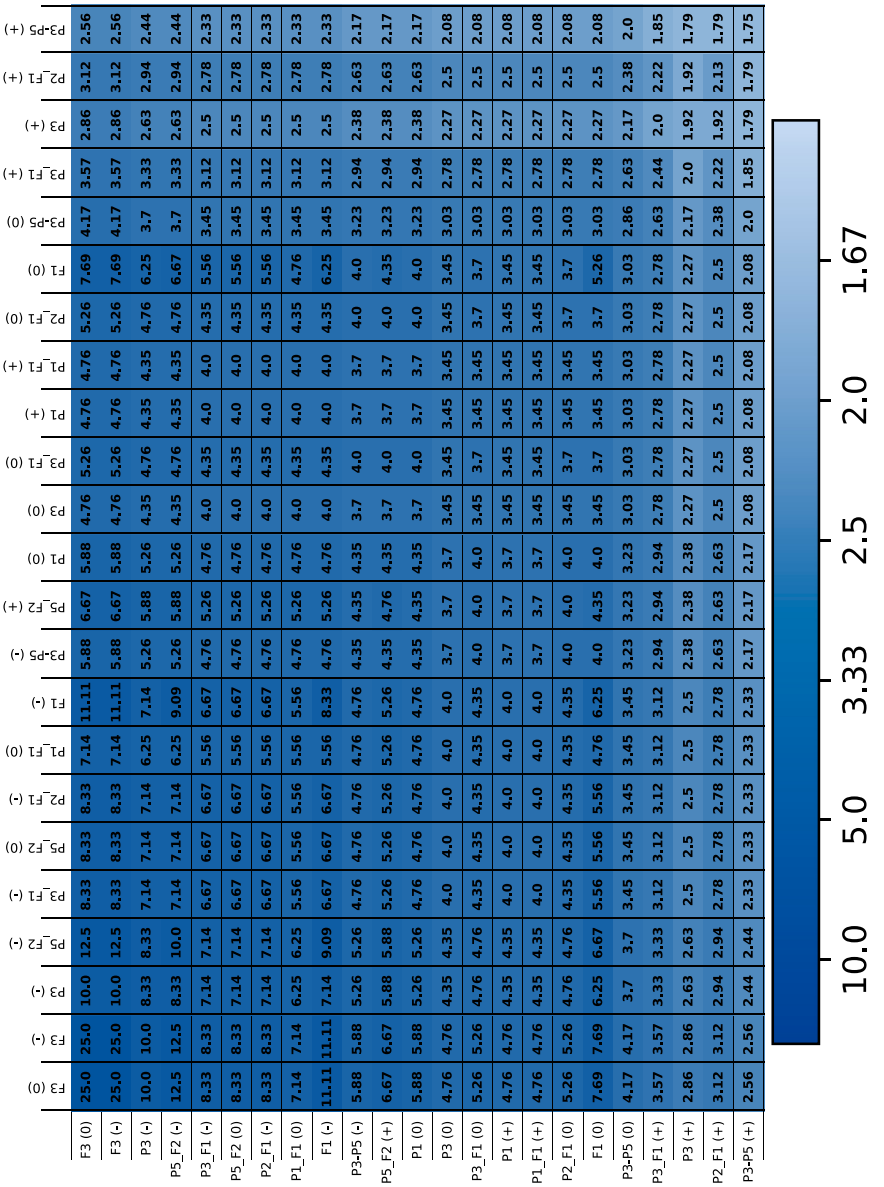


Fig. D.24. Calculated maximum mean service times in minutes for the scenario with an overpass.

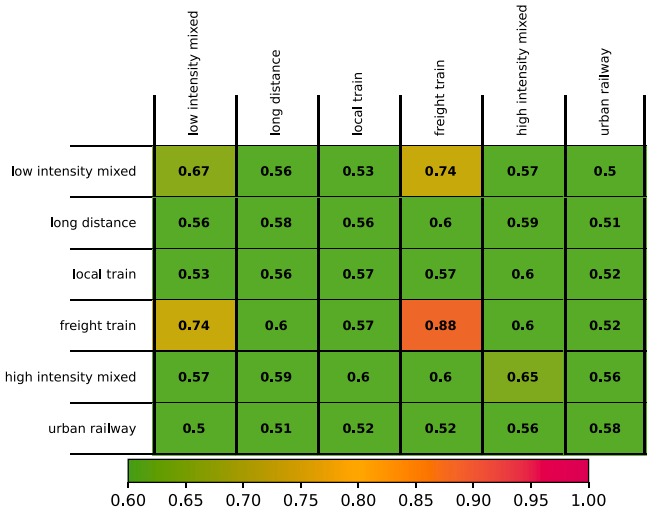


Fig. E.25. Occupancy rates (following Eq. (38)) for the scenario with an overpass, assuming mean service times from Fig. 15.



Fig. E.26. Occupancy rates (following Eq. (38)) for the scenario without an overpass, assuming mean service times from Fig. D.23.

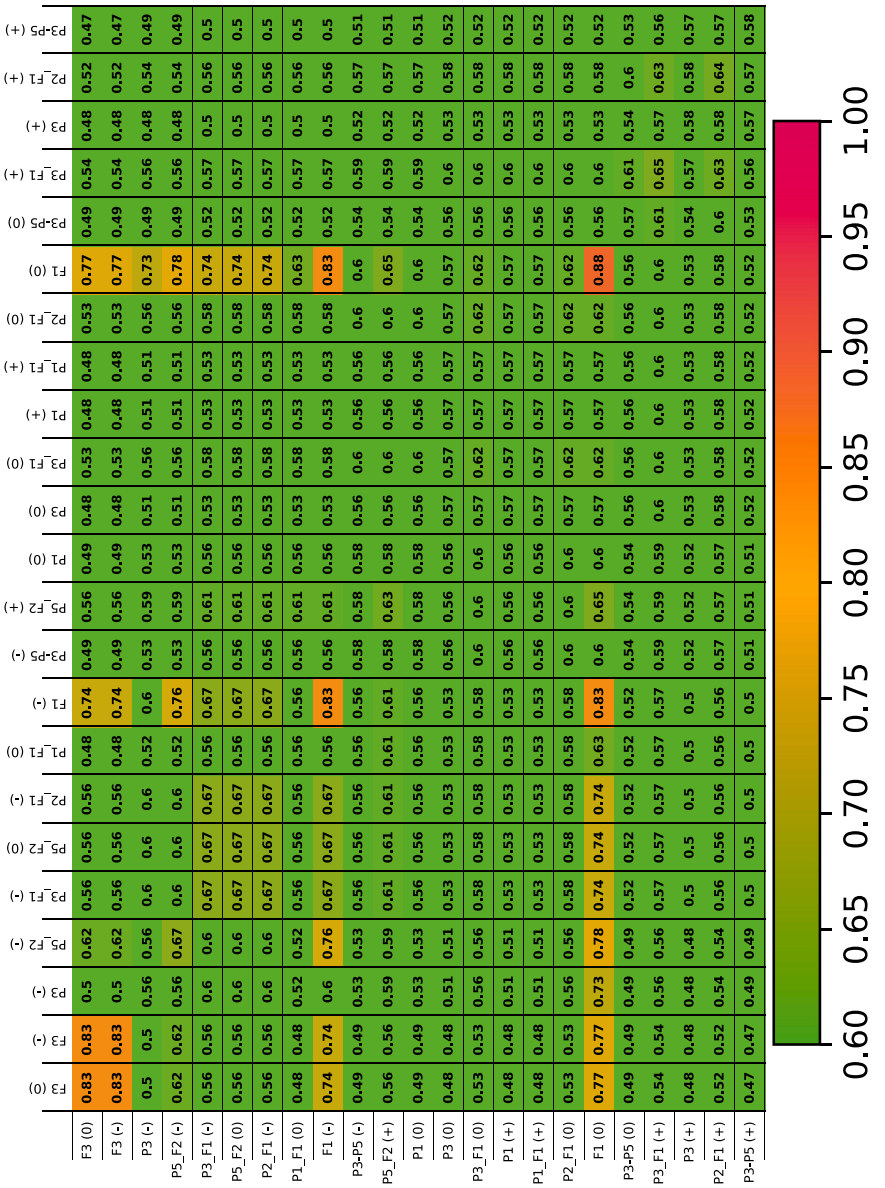


Fig. E.27. Occupancy rates (following Eq. (38)) for the scenario with an overpass, assuming mean service times from Fig. D.24.

References

Abril, M., Barber, F., Ingolotti, L., Salido, M.A., Tormos, P., Lova, A., 2008. An assessment of railway capacity. Transp. Res. E Logist. Transp. Rev. 44 (5), 774–806. http://dx.doi.org/10.1016/j.tre.2007.04.001.

Armstrong, J., Preston, J., 2017. Capacity utilisation and performance at railway stations. J. Rail Transp. Plan. Manage. 7 (3), 187–205. http://dx.doi.org/10.1016/j.jrtpm.2017.08.003.

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. In: International Series in Operations Research & Management Science, vol. 268, Springer International Publishing, Cham, pp. 25–45. http://dx.doi.org/10.1007/978-3-319-72153-8_2.

Bolch, G., Greiner, S., de Meer, H., Trivedi, K.S., 2006. Queueing Networks and Markov Chains: Modeling and Performance Evaluation with Computer Science Applications. Wiley, URL: https://books.google.de/books?id=8Mei8w6YUHYC.

Budde, C.E., Hartmanns, A., Klauck, M., Křetínský, J., Parker, D., Quatmann, T., Turrini, A., Zhang, Z., 2021. On correctness, precision, and performance in quantitative verification. In: Margaria, T., Steffen, B. (Eds.), Leveraging Applications of Formal Methods, Verification and Validation: Tools and Trends. In: Lecture Notes in Computer Science, vol. 12479, Springer International Publishing, Cham, pp. 216–241. http://dx.doi.org/10.1007/978-3-030-83723-5_15.

Büker, T., 2011. Ausgewählte Aspekte der Verspätungsfortpflanzung in Netzen (Ph.D. thesis). Rheinisch-Westfälischen Technische Hochschule Aachen (german only).

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. http://dx.doi.org/10.1016/j.ejor.2015.03.020.
Burdett, R.L., 2016. Optimisation models for expanding a railway's theoretical capacity. European J. Oper. Res. 251 (3), 783–797. http://dx.doi.org/10.1016/j.ejor.2015.12.033.
Burdett, R.L., Kozan, E., 2006. Techniques for absolute capacity determination in railways. Transp. Res. B 40 (8), 616–632. http://dx.doi.org/10.1016/j.trb.2005.09.004.
Burdett, R.L., Kozan, E., 2009. Scheduling trains on parallel lines with crossover points. J. Intell. Transp. Syst. 13 (4), 171–187. http://dx.doi.org/10.1080/15472450903297608.
Burdett, R., Kozan, E., 2014a. Determining operations affected by delay in predictive train timetables. Comput. Oper. Res. 41, 150–166. http://dx.doi.org/10.1016/j.cor.2013.08.011.
Burdett, R.L., Kozan, E., 2014b. Performance profiling for predictive train schedules. J. Rail Transp. Plan. Manage. 4 (4), 98–114. http://dx.doi.org/10.1016/j.jrtpm.2014.11.002.
Cacchiani, V., Toth, P., 2012. Nominal and robust train timetabling problems. European J. Oper. Res. 219 (3), 727–737. http://dx.doi.org/10.1016/j.ejor.2011.11.003.
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. Manage. 22, 100316. http://dx.doi.org/10.1016/j.jrtpm.2022.100316.
Corman, F., Kecman, P., 2018. Stochastic prediction of train delays in real-time using Bayesian networks. Transp. Res. C 95, 599–615. http://dx.doi.org/10.1016/j.trc.2018.08.003.
D'Acierno, L., Botte, M., Pignatiello, G., 2019. A simulation-based approach for estimating railway capacity. Int. J. Transp. Dev. Integr. 3 (3), 232–244. http://dx.doi.org/10.2495/TDI-V3-N3-232-244.
DB Netz AG, 2022. Richtlinie Fahrwegkapazität.
De Kort, A.F., Heidergott, B., Ayhan, H., 2003. A probabilistic (max, +) approach for determining railway infrastructure capacity. European J. Oper. Res. 148 (3), 644–661. http://dx.doi.org/10.1016/S0377-2217(02)00467-8.
Delorme, X., Rodriguez, J., Gandibleux, X., 2001. Heuristics for railway infrastructure saturation. Electron. Notes Theor. Comput. Sci. 50 (1), 39–53. http://dx.doi.org/10.1016/S1571-0661(04)00164-1.
Emunds, T., Nießen, N., 2023. Research data and code: To build or not to build - A queueing-based approach to timetable independent railway junction infrastructure dimensioning. http://dx.doi.org/10.5281/zenodo.8363461.
European Comission, 2019. Technical specification for interoperability relating to the 'infrastructure' subsystem of the rail system within the European union: Comission regulation (EU) 2019/776. Official J. Eur. Union.
Europe's Rail, 2015. Shift 2 rail. URL: https://projects.shift2rail.org/.
Fischer, K., Hertel, G., 1990. Bedienungsprozesse im Transportwesen - Grundlagen und Anwendungen der Bedienungstheorie. Transpress.
Goverde, R.M., 2007. Railway timetable stability analysis using max-plus system theory. Transp. Res. B 41 (2), 179–201. http://dx.doi.org/10.1016/j.trb.2006.02.003.
Goverde, R.M., 2010. A delay propagation algorithm for large-scale railway traffic networks. Transp. Res. C 18 (3), 269–287. http://dx.doi.org/10.1016/j.trc.2010.01.002.
Goverde, R.M., Corman, F., D'Ariano, A., 2013. Railway line capacity consumption of different railway signalling systems under scheduled and disturbed conditions. J. Rail Transp. Plan. Manage. 3 (3), 78–94. http://dx.doi.org/10.1016/j.jrtpm.2013.12.001.
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. http://dx.doi.org/10.2495/CR120361.
Gudehus, T., 1976. Staueffekte vor Transportknoten. Z. Oper. Res. 20 (6), B207–B252. http://dx.doi.org/10.1007/BF01918395.
Hansen, I.A., 2010. Timetable Planning and Information Quality. WIT Press.
Hansen, I., Pachl, J., 2014. Railway Timetabling and Operations: Analysis, Modelling, Optimisation, Simulation, Performance, Evaluation. Eurail Press.
Happel, 1959. Sperrzeiten als Grundlage für die Fahrplankonstruktion. Eisenbahntech. Rundsch. (ETR) 2, 79.
Harrod, S., 2009. Capacity factors of a mixed speed railway network. Transp. Res. E Logist. Transp. Rev. 45 (5), 830–841. http://dx.doi.org/10.1016/j.tre.2009.03.004.
Hensel, C., Junges, S., Katoen, J.P., Quatmann, T., Volk, M., 2022. The probabilistic model checker storm. Int. J. Softw. Tools Technol. Transfer 24 (4), 589–610. http://dx.doi.org/10.1007/s10009-021-00633-z.
Jensen, L.W., Landex, A., Nielsen, O.A., Kroon, L.G., Schmidt, M., 2017. Strategic assessment of capacity consumption in railway networks: Framework and model. Transp. Res. C 74, 126–149. http://dx.doi.org/10.1016/j.trc.2016.10.013.
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. C 119, 102751. http://dx.doi.org/10.1016/j.trc.2020.102751.
Junges, S., Volk, M., 2023. Stormpy - Python bindings for storm: Version 1.7.0. URL: https://moves-rwth.github.io/stormpy/index.html.
Katoen, J.P., 2016. The probabilistic model checking landscape. In: Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science. LICS '16, Association for Computing Machinery, New York, NY, USA, pp. 31–45. http://dx.doi.org/10.1145/2933575.2934574.
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., 2010. Advances and challenges of probabilistic model checking. In: 2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton). pp. 1691–1698. http://dx.doi.org/10.1109/ALLERTON.2010.5707120.
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.
Landex, A., 2009. Evaluation of railway networks with single track operation using the UIC 406 capacity method. Netw. Spat. Econ. 9 (1), 7–23. http://dx.doi.org/10.1007/s11067-008-9090-7.
Landex, A., Jensen, L.W., 2013. Measures for track complexity and robustness of operation at stations. J. Rail Transp. Plan. Manage. 3 (1–2), 22–35. http://dx.doi.org/10.1016/j.jrtpm.2013.10.003.
Leutwiler, F., Corman, F., 2022. A logic-based benders decomposition for microscopic railway timetable planning. European J. Oper. Res. 303 (2), 525–540. http://dx.doi.org/10.1016/j.ejor.2022.02.043.
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. C 124, 102961. http://dx.doi.org/10.1016/j.trc.2020.102961.
Lusby, R., Larsen, J., Ryan, D., Ehrgott, M., 2011. Routing trains through railway junctions: A new set-packing approach. Transp. Sci. 45 (2), 228–245. http://dx.doi.org/10.1287/trsc.1100.0362.
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. http://dx.doi.org/10.1016/j.ejor.2012.12.027.
Nießen, N., 2008. Leistungskenngrößen für Gesamtfahrstraßenknoten (Ph.D. thesis). Verkehrswissenschaftliches Institut der RWTH Aachen.
Nießen, N., 2013. Waiting and loss probabilities for route nodes. In: International Conference on Railway Operations Modelling and Analysis. RailCopenhagen.
Parker, D., Norman, G., Kwiatkowska, M., 2000. PRISM. URL: https://www.prismmodelchecker.org/.

Pellegrini, P., Marlière, G., Rodriguez, J., 2014. Optimal train routing and scheduling for managing traffic perturbations in complex junctions. Transp. Res. B 59, 58–80. http://dx.doi.org/10.1016/j.trb.2013.10.013.
Potthoff, G., 1970. Verkehrsströmungslehre - Die Zugfolge auf Strecken und in Bahnhöfen. Transpress.
Python Software Foundation, 2022. Python 3.10.9. URL: https://docs.python.org/release/3.10.9/index.html.
Ross, S.M.A., 2014. Introduction to Probability Models. Elsevier.
Şahin, İ., 2017. Markov chain model for delay distribution in train schedules: Assessing the effectiveness of time allowances. J. Rail Transp. Plan. Manage. 7 (3), 101–113. http://dx.doi.org/10.1016/j.jrtpm.2017.08.006.
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.
Schwanhäußer, W., 1978. Die Ermittlung der Leistungsfähigkeit von großen Fahrstraßenknoten und von Teilen des Eisenbahnetzes. Archiv Eisenbahntechnik 1978 (38), 7–18.
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. work not formally published.
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. http://dx.doi.org/10.3390/math11040839.
Team SimPy Revision, 2023. SimPy 4.0.1. URL: https://simpy.readthedocs.io/en/4.0.1/.
UIC, 2004. Code 406 - Capacity. International Union of Railways, Paris, France.
UIC, 2013. Code 406 - Capacity: 2nd Edition. Paris, France.
Van Rossum, G., Drake, Jr., F.L., 1995. Python Tutorial. Centrum voor Wiskunde en Informatica Amsterdam, The Netherlands.
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. http://dx.doi.org/10.1038/s41592-019-0686-2.
Wakob, H., 1984. Ableitung eines generellen Wartemodells zur Ermittlung der planmäßigen Wartezeiten im Eisenbahnbetrieb unter besonderer Berücksichtigung der Aspekte Leistungsfähigkeit und Anlagenbelastung (Ph.D. thesis). Aachen, URL: https://publications.rwth-aachen.de/record/74574.
Weik, N., 2020. Long-Term Capacity Planning of Railway Infrastructure – A Stochastic Approach Capturing Infrastructure Unavailability (Ph.D. thesis). RWTH Aachen University, http://dx.doi.org/10.18154/RWTH-2020-06771.
Weik, N., 2022. Macroscopic traffic flow in railway systems — A discussion of the applicability of fundamental diagrams. J. Rail Transp. Plan. Manage. 23, 100330. http://dx.doi.org/10.1016/j.jrtpm.2022.100330.
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. Manage. 7 (3), 114–126. http://dx.doi.org/10.1016/j.jrtpm.2017.06.001.
Weik, N., Warg, J., Johansson, I., Bohlin, M., Nießen, N., 2020. Extending UIC 406-based capacity analysis – New approaches for railway nodes and network effects. J. Rail Transp. Plan. Manage. 15, 100199. http://dx.doi.org/10.1016/j.jrtpm.2020.100199.
Wendler, E., 1999. Analytische Berechnung der planmässigen Wartezeiten bei asynchroner Fahrplankonstruktion: Dissertation. In: Veröffentlichungen des Verkehrswissenschaftlichen Institutes der Rheinisch-Westfälischen Technischen Hochschule Aachen, vol. 55, Verkehrswissenschaftliches Institut der Rheinisch-Westfälischen Technischen Hochschule Aachen, Aachen.
Wendler, E., 2007. The scheduled waiting time on railway lines. Transp. Res. B 41 (2), 148–158. http://dx.doi.org/10.1016/j.trb.2006.02.009.
Yaghini, M., Nikoo, N., Ahadi, H.R., 2014. An integer programming model for analysing impacts of different train types on railway line capacity. Transport 29 (1), 28–35. http://dx.doi.org/10.3846/16484142.2014.894938.
Zhang, X., Nie, L., 2016. Integrating capacity analysis with high-speed railway timetabling: A minimum cycle time calculation model with flexible overtaking constraints and intelligent enumeration. Transp. Res. C 68, 509–531. http://dx.doi.org/10.1016/j.trc.2016.05.005.
Zhong, Q., Xu, C., Yang, R., Zhong, Q., 2023. Equivalences between analytical railway capacity methods. J. Rail Transp. Plan. Manage. 25, 100367. http://dx.doi.org/10.1016/j.jrtpm.2022.100367.
Zieger, S., Weik, N., Nießen, N., 2018. The influence of buffer time distributions in delay propagation modelling of railway networks. J. Rail Transp. Plan. Manage. 8 (3–4), 220–232. http://dx.doi.org/10.1016/j.jrtpm.2018.09.001.
Zukerman, M., 2013. Introduction to queueing theory and stochastic teletraffic models. http://dx.doi.org/10.48550/arXiv.1307.2968.
Zwaneveld, P.J., Kroon, L.G., Romeijn, H.E., Salomon, M., Dauzère-Pérès, S., van Hoesel, S.P.M., Ambergen, H.W., 1996. Routing trains through railway stations: Model formulation and algorithms. Transp. Sci. 30 (3), 181–194. http://dx.doi.org/10.1287/trsc.30.3.181.
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. http://dx.doi.org/10.1016/S0377-2217(00)00087-4.