Journal of Rail Transport Planning & Management 37 (2026) 100558

Contents lists available at ScienceDirect

Journal of Rail Transport Planning & Management

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

Full Length Article

Reliability evaluation of rail freight transport processes — An agent-based approach combining microscopic rail simulation and freight network logistics

Jakob Geischberger [a],*, Norman Weik [b], Jürgen Pannek [c]

[a] German Aerospace Center (DLR), Institute of Transportation Systems, Lilienthalplatz 7, 38108, Braunschweig, Germany
[b] Technical University of Munich, TUM School of Engineering and Design, Department of Mobility Systems Engineering, Professorship Design and Operation of Public Rail Transport Systems (RTS), Parkring 35, 85748, Garching, Germany
[c] Technical University of Braunschweig, Institute for Intermodal Transport and Logistics Systems (ITL), Hermann-Blenk-Straße 42, 38108, Braunschweig, Germany

ARTICLE INFO

Keywords:
Reliability
Railway operation
Transportation chains
Network simulation
Punctuality
Rail freight logistics

ABSTRACT

In densely operated European railway networks the trade-off between network utilization and service quality leads to pondering traffic volume against the reliability of services. The perspective on service reliability, however, tends to vary between stakeholders. In this paper, we investigate and assess reliability in railway networks, particularly focusing on the comparison between railway operators and users in freight transport systems: While operators mainly focus on the punctuality of trains, users typically perceive reliability on the level of transportation chains. To combine both perspectives, a new agent-based simulation approach is developed that combines a microscopic domain-specific traffic control simulation incorporating the essential properties of the signaling system with a logistical perspective featuring the flow of goods within the network. The study outlines the methodical steps of model-setup, as well as the implementation on the backbone of an agent-based, open-source simulation environment. The methodology is applied to the German railway network and effects on the two different reliability perspectives are assessed and compared. A significant deviation of transport reliability from train punctuality is observed, highlighting the benefit of incorporating the end-customer perspective in reliability evaluation. It is also found that higher amounts of segments per transportation chain lead to significantly lower reliability.

1. Introduction

As an efficient and climate-friendly alternative to congested road transportation, railway systems in central Europe have seen a significant increase in demand in recent years. The effect is particularly pronounced in the rail freight segment, where an increase by 18% and 20%, respectively, has been reported for France and Germany (European Union Agency for Railways, 2024), for instance. In view of strategic initiatives to further strengthen rail freight transportation and to increase its modal share from 18 to 30% – laid down, for example, in the European Rail Master Plan (Europe's Rail Joint Undertaking, 2021) or the European Rail Freight Vision (Rail Freight Forward, 2024) – this development is expected to continue.

The increase of traffic volume meets a system, where even today, many railway lines, stations, and core facilities are operated close to or beyond their designed capacity. As railway infrastructure construction or extension projects are time-consuming and

* Corresponding author.
E-mail address: jakob.geischberger@dlr.de (J. Geischberger).

https://doi.org/10.1016/j.jrtpm.2025.100558
Received 10 January 2025; Received in revised form 23 September 2025; Accepted 19 November 2025
Available online 21 November 2025

slowly adapt to changing demand and system requirements, further stresses on rail transportation infrastructure and rolling stock, especially on central corridors and in bottleneck areas, are almost inevitable.

Given the complex interactions between timetabling, railway signaling and train control, the increasing load on the network will almost certainly give rise to increasing delays — especially within mixed-traffic networks, where heterogeneous traffic with different driving characteristics and product types need to be synchronized. As a result, a trade-off between network utilization and service quality needs to be negotiated factoring traffic volume against the reliability and punctuality of services.

The *reliability* of transport systems is a core component of the rail service concept and an important aspect of customer satisfaction with railway operations (Vromans et al., 2006). Analogously to the rise in capacity, the improvement of the reliability of railway transportation is also part of the strategical objectives formulated in the European Rail Master Plan (Europe's Rail Joint Undertaking, 2021) and identified as a core driver of competitiveness of rail freight transportation (Federal Ministry of Transport and Digital Infrastructure, 2017). As many different stakeholders are involved in railway planning and operation, there generally exists no unanimous definition of reliability (Lenze and Nießen, 2021). Different parties have different objectives and possibly diverging perspectives on reliability relating to their business models.

Comparing two particularly important perspectives, those of train operators and users, operators tend to focus on the operational reliability of train runs, which is the basis for keeping vehicle and staff rotations stable. Rail users (for example passengers, or logistics companies), by contrast, tend to evaluate the reliability of operations based on trips, or *transportation chains*.

As a result, the main interest of train operators is the punctual and reliable transportation of goods or passengers, referring to the product offered to the customer, i.e. the rail-operated trip segments. One important reliability evaluation criterion of railway operation companies thus is the punctuality *of their offered train services*, for example the share of trains operated "on time" i.e. within certain timetable windows (compare Section 2). This perspective does not fully capture the perspective of railway transportation users, as trips or transportation chains may involve multiple interconnected services offered by different operators. In this context, reliability refers to the reliability of the transportation *process*, focusing on the reliability at key points such as departure, arrival, or critical transfers, rather than the delay of individual trains within the transportation chain. To harmonize competing interests and to improve the reliability of railway systems from both introduced perspectives, measures of very different character and scope can be considered, including, but not limited to, infrastructural enhancements, technological improvement (for example increased traction of trains or enhanced command and signaling), timetabling (for example harmonization) or optimizing freight production systems. Depending on the perspective, the effectiveness and viability of different measures may vary. At the same time, systematic changes to the rail system often involve cost-intensive construction or difficult implementation phases, thus decreasing the reliability of operation before an enhancement can be observed. As data from observations may not be available immediately, a detailed understanding involving predictive modeling of the impact of measures on railway operations is required.

The effects of system adjustments to improve reliability are typically assessed using simulation tools: By reproducing railway operations in a digital copy of the system, scenarios for system modifications can be evaluated and compared to a reference case (for example the current system). The results then serve as a method of decision support, helping to decide which measure to take in order to improve and maximize the reliability of the system. However, only very few approaches exist that both incorporate a detailed rail-specific representation of train control actions and the means to analyze logistical processes on the level of transport chains. Conventional operational railway simulations have a strong focus on the very detailed railway-specific modeling and assessment of trains and their interactions governed by the signaling system. They typically do not include transportation processes or flows of goods. Thus, they cover the train-centered reliability perspective of operators and consequently the impact of potentially reliability-enhancing measures can only be evaluated from their perspective, such that the impact of system modifications on the market situation of rail transportation, which is the focus of policies for strengthening rail transportation (for example (Federal Ministry of Transport and Digital Infrastructure, 2017)), can only be measured indirectly, in an ex-post evaluation.

To directly assess reliability effects for the users, we extend existing approaches and propose a new integrated logistical railway simulation methodology that evaluates the reliability of railway operation not only on the level of trains, but also on the level of transportation chains. The agent-based approach is capable of both modeling railway specifics on the one hand and transportation chains on the other hand. The resulting simulation methodology can then be used to more precisely assess the impact of transport policies by providing a direct relationship to both the operators' and the customers' perspective on reliability. It hence provides a valuable tool to pinpoint directions for improving the customer experience in using railway transportation. In detail, the contributions made by our approach are:

- Introducing a new agent-based rail logistics simulation procedure, which combines microscopic rail-specific train control simulation with a logistical perspective on the flow of passengers and goods within the network. The approach is based on an open-source and open-access simulation environment, such that the methodological assumptions are transparent and accessible to all stakeholders in the rail industry.
- The approach allows to assess and compare the reliability of railway operation both based on the perspective of railway operators and system users by building on an agent-based representation of both trains and shipment units (goods).
- As a result, complex transport chains in large-scale national or international railway networks can be simulated while keeping the detailed interactions between trains from railway operations simulation. In particular, models with a different model granularity (for example macroscopic origin–destination-based (OD-based) logistics information or planned strategic national timetable concepts and train slots) can be mapped to and fused with microscopic train interaction processes.

In the following, we first introduce related work in Section 2, particularly focusing on existing evaluation methodologies in railway operations management and logistics and their scopes and limitations in modeling reliability in railway systems. Afterwards, the proposed methodology is outlined in Section 3, showing the strategic value and how the identified demand for a new methodology is met by the interaction of different components. Moreover, the build-up of the resulting simulation model on a detailed, technical level is shown. In Section 4.1, the proposed methodology is applied on the use case of the German Railway network. Section 5 summarizes and gives an outlook on further development.

2. Related work

In the following we review related work on reliability assessment in rail transportation, starting with an introduction to metrics and definitions for reliability in the railway context.

Reliability has different definitions and performance indicators. In the context of railway systems, reliability can, for instance, refer to as the share of cancellations of trains, where higher shares mean lower reliability (compare, for example, practices in the UK (Office of Rail and Road, 2024)). Moreover, punctuality and reliability are often used in a closely related way or even synonymously. Typically, train punctuality is defined as the train being within a certain margin of the scheduled arrival or departure time, i.e. trains are considered punctual if they arrive or depart below a maximum delay. Different operators and countries apply different delay thresholds to measure punctuality. In Germany, for example, trains are considered punctual if they have a maximum delay of 5:59 min, while passenger punctuality – including transfers – is defined as an arrival delay below 15 min (Deutsche Bahn, 2024). In the UK, as a further example, three respective thresholds for train punctuality are defined: a delay below one minute is considered “on time”, complemented by the punctuality thresholds of three and 15 min (Office of Rail and Road, 2024). The resulting calculated percentage share of punctual trains obeying these delay thresholds is then sometimes referred to as reliability and is also used for communication purposes. The Swiss operator SBB additionally publishes connection-based punctuality, accounting for the share of realized connections. Furthermore, consignment punctuality is communicated for freight rail services, where punctuality is defined on a wagon-basis according to its respective delivery delay. Currently, wagons with a delay below 20 min are considered punctual (10 min for express services) (SBB, 2024). In contrast, the Central Japan Railway Company communicates an absolute average punctuality number for its Shinkansen service (Central Japan Railway Company, 2024). For a further recent discussion on rail punctuality indicators see Blayac and Stéphan (2021). The U.S. Federal Railroad Administration uses On-Time Performance (OTP) as the performance indicator for reliability (U.S. Department of Transportation Federal Railroad Administration, 2013), which in fact is very close to the above mentioned punctuality thresholds. They explicitly differentiate between ”end-point” and ”all-station” OTP. Rietveld et al. (2001) propose alternative approaches to define reliability in the context of railways, for example the probability of an early departure or the mean difference between scheduled arrival time and arrival time in operations (the latter was also taken up by Vromans (2005)).

Coming from the context of logistics, more specifically the field of intermodal transportation networks, reliability can be defined as the probability that transportation units can successfully be transported from the origin to the destination (compare Niu et al. (2024)). There, the corresponding logistics model is modeled as a cost-related multi-distribution multi-state flow network. Zhang et al. (2018) observe a two-fold-structure of reliability definitions in intermodal transport, pointing out that existing literature either focuses on optimizing container terminals or transport choice modes. They opt for the indicator “on-time delivery” as a widely-used probability-based network reliability indicator, measuring the percentage of shipments arriving within the planned transportation time. Shams et al. (2017) give an extensive overview of studies also using on-time delivery indicators. They compare them to both mean–variance based approaches that measure reliability as the standard deviation of travel time (for example as adopted for the Netherlands in de Jong et al. (2014)) and schedule based approaches (early and late schedule delay, as introduced by Small (1982)).

In the field of railway logistics, Kramarz et al. (2022) measure reliability of intermodal railway networks as a combination of the number of canceled trains and an average delay time of trains. Zhang et al. (2019) define travel time reliability as “the probability of the goods that can be transported in specified OD within expected travel time by railway”. This concept is applied to railway transportation chains and analyzes interdependencies to other indicators (for example buffer times) with a detailed focus on the segmentation of the transportation chains. Travel time reliability is then calculated using an Hasofer Lind-Rackwitz Fiessler (HL-RF) algorithm. Moreover, Meng et al. (2020) propose a more general approach to the definition of reliability, defining it as ability or possibility of systems to perform without failure. On the basis of that, railway systems are divided into stations and sections and reliability is defined by the reachability between two stations. Table 1 gives an overview of the mentioned reliability definitions and indicators from both perspectives.

In the methodical context of simulation, the optimization of transportation chains has been a widely researched field. Specifically, agent-based approaches can be found, for example, in Holmgren et al. (2012). There, a combined physical (infrastructural) simulator and decision making simulator is presented, implementing production chains in detail. The application of the micro-level approach to a Baltic Sea scenario shows its suitability to study effects of different measures to the system (for example political or infrastructural), but does not specifically focus on railway transportation. Similarly, an agent-based simulation approach to optimize parameters of a railway transportation chain is proposed in Matsiuk et al. (2021). The AnyLogic-based approach combines discrete-event and agent-based principles and focuses on estimating the demand for locomotives and cars. AnyLogic is also the backbone of the approach proposed in Gong et al. (2019), a Multi-Agent GIS (Geographic Information System) Simulation for Railway Logistics Optimization. Furthermore, in Mancera et al. (2015) an activity-based MATSim framework to optimize single-wagon load railway operation is proposed. While these approaches feature an in-detail depiction of rail-bound transportation chains on a logistical level, train-specific behavior is not modeled microscopically on the basis of signaling.

Table 1
Reliability indicators.

Concept	Source	Reliability indicator	Context
A	Office of Rail and Road (2024)	Share of cancellations of trains	Railways
B	Office of Rail and Road (2024), Deutsche Bahn (2024), SBB (2024)	Share of trains within defined punctuality threshold	Railways
C	U.S. Department of Transportation Federal Railroad Administration (2013)	On-Time Performance	Railways
D	Rietveld et al. (2001)	Probability of early departure	Railways
E	Rietveld et al. (2001)	Mean difference between scheduled arrival time and arrival time in operations	Railways
F	Niu et al. (2024)	Probability that transportation units can successfully be transported from the origin to the destination	Logistics
G	Zhang et al. (2018), Shams et al. (2017)	On-time delivery	Logistics
H	Shams et al. (2017), de Jong et al. (2014)	Standard deviation of travel time	Logistics
I	Small (1982), Shams et al. (2017)	Early and late schedule delay	Logistics
J	Kramarz et al. (2022)	Combination of the number of canceled trains and an average delay time	Logistics
K	Zhang et al. (2019)	"The probability of the goods that can be transported in specified OD within expected travel time by railway"	Logistics
L	Meng et al. (2020)	Reachability between two stations	Logistics

Focusing specifically on the perspective of users, prior research has been performed on passenger delay. In Nielsen et al. (2009), different methods to calculate passenger delays are presented and discussed, with an empirical evaluation of real-life data of the Copenhagen area as well as a link with a simulation approach to forecast delays. Landex (2012) proposes a framework where a traditional railway simulation is coupled with a passenger delay model in order to evaluate network effects on passengers specifically. Moreover, a further overview of and distinction between different methods of calculating passenger delays is given. A MATSim framework to calculate passenger door-to-door delays based on Automatic Vehicle Location data is presented in Paulsen et al. (2018). It is applied to a use case of the Copenhagen area. A special weight on the coupling between passenger and train flows characterizes the approach presented in Zhang et al. (2023a). Moreover, Jiang et al. (2012) present a simulation model specifically designed to investigate the relationship between passenger and train delays. While these approaches share a focus on passenger journeys, very little research has been observed on freight transportation and the delay of transportation chains as presented in this paper. Moreover, we propose an integrated approach to simulating trains and transportation chains in one environment, rather than combining different approaches.

Next, our approach is related to conventional railway operation simulation tools, focusing on a very detailed, microscopic modeling of railway driving dynamics and operations. These tools are frequently used to asses changes to railway operations, for example, in infrastructure, command and signaling or on a planning horizon. By modeling a digital twin, the predicted effect of a certain measure can be evaluated beforehand and decision support can be provided. Frequently, these tools are also used to assess the effect of both planned and unplanned operational disruptions – for example construction work or weather-related implications – and create alternative timetables. Examples for commercial railway simulation tools are – amongst others – RailSys (Rail Management Consultants International GmbH, 2024), OpenTrack (Huerlimann and Nash, 2004) or Trenissimo (de Fabris et al., 2018). These tools are typically used analyzing smaller lines or stations rather than whole networks and do not feature agent-based modeling of passengers or goods.

3. Methodology

The methodology presented in this paper describes how the microscopic railway simulation tool is set up based on a macroscopic timetable. It features both trains and agent-based transportation chains in order to microscopically simulate and evaluate railway transportation scenarios on a network-wide basis. This section first outlines basic requirements and explains the tool choice against the background of other approaches presented in Section 2. Afterwards, the according steps to create the model are shown.

3.1. Basic requirements of the rail logistics simulation tool

In Section 1, two important perspectives (operators and users) on railway network operations and their respective view on the reliability of operations have been discussed: While operators mainly focus on punctual *trains*, users have a wider perception of reliable transportation, mainly focusing on reliable *transportation chains* – independent from singular trains involved.

At the same time, Section 2 describes how microscopic railway simulation tools fall short on representing both perspectives, as the focus of most tools is on operator-centered evaluation. Still, these tools are able to assess changes in train-centered indicators but typically do not include evaluation possibilities on a journey or transportation chain level.

The main objective of the methodology we propose is to evaluate railway operations reliability from both perspectives. This especially refers to the impact of changes on the system and their evaluation and comparison from both perspectives. For that purpose, at first a reliability concept suiting and covering both perspectives needs to be identified and defined. This concept is chosen based on Table 1: The table shows that all of the concepts are either based on the general realization of a transport (in the sense of taking place or not being canceled: concepts A, F, J, K, L) or on a comparison of a scheduled timestamp with an observed one (B, C, D, E, G, H, I, J). Our approach does explicitly not focus on train cancellations but rather concentrates on effects in *realized* transportation. For that reason, concepts A, F, J, K and L are not taken into consideration further. Amongst the other concepts, some express reliability via absolute delay numbers (concepts E, H and I), while others calculate a share of trains staying within a certain delay threshold (concepts B, C and G). As transportation chains rather focus on the arrival of trains and goods than departures (see below), concept D is also excluded.

For the reliable operation of transportation chains there is no strict necessity for trains to be punctual. However, punctuality is sufficient for the reliability of transportation chains. Given fixed train frequencies and transshipment times in nodes, delays of single trains exceeding a certain buffer threshold are highly probable of transferring not only their own delay to their transported goods, but also adding additional waiting times to the transportation chain. This effect may occur several times within a transportation chain, adding up along its run. It is worth noticing that this effect may also appear in the opposite direction, when delayed trains might replace scheduled services earlier than planned (depending on transshipment times). In any case, these effects accumulate to the arrival time of the transported good at its final destination.

Given these considerations, we suggest that for the transportation of goods the main interest for the user of the system is whether the arrival of the good at its destination is within a timespan defined by that user. This focus on the arrival of goods is supported by the fact that freight operations are indifferent to discomfort in changing routes or trains (in contrast to passengers). Moreover, the resulting predictability assures a satisfactory incorporation of the transportation chain in other processes (for example further processing or consumption) from the perspective of the user. For that reason, we propose to define the reliability of railway transportation chains on whether this defined time-window for the arrival of goods is met or not. This subsequently leads to a share of trains that fulfill an arrival within the window, while others do not (this explicitly refers to both early and late arrivals). Hence, this definition shows strong similarity to logistical on-time delivery indicators (G), which focus on the concept of an arrival time window. The proposed approach also takes up the character of concepts B and C, while referring to transports and not trains. At the same time, this arrival rate can be compared to common thresholds for train punctuality. Still, to assure an even stronger comparability between the perspectives and to gain a deeper understanding of underlying coherences, we propose to additionally evaluate absolute arrival delay numbers. This will also help to identify and pinpoint local effects.

Core of the proposed methodological framework is a microscopic railway simulation model that covers railway specifics on the one hand and the feature of modeling transportation chains on the other hand by merging goods with trains in an agent-based manner. The output needs to contain agent-specific arrival time stamps regarding both the trains and the transportation chains which are then compared to schedule data (again for both levels) and further processed and evaluated (see above). Furthermore, as long-distance trains (especially within the freight segment) involve long train-runs and transportation chains cover even wider distances by combining them, the simulator needs to cover large networks (such as whole countries) and at the same time feature acceptable computational times. Moreover, transportation demand is frequently characterized by a different granularity of data than timetabling data. Additionally, the proposed tool is meant to analyze reliability not only in current operations, but also for future concepts. For these reasons, input of data of different granularity (for example macroscopic timetables) needs to be processable. This corresponds to fully representing the above mentioned perspective of users (and possibly their limited quality or availability of data). Finally, the comparison of different perspectives and stakeholders implies good or even unlimited access to the tooling for involved (or other) parties. For this reason, open access was incorporated as a further requirement of the proposed methodological framework.

3.2. Overview of methodology

Fig. 1 gives an overview of the proposed methodology to set up a simulation model capable of microscopically simulating railway transportation chains. Here, different methodological steps (gray) interact in order to create the model that allows railway network evaluations regarding both trains and transportation chains. It can be seen that network infrastructure and information on the network timetable are needed as elementary input. Those two data sources are merged and processed in a sequential row of steps: First, their granularity is aligned by mapping macroscopic timetable data and the microscopic infrastructure needed. Next, a realistic railway operation flow is assured by smoothing operation within nodes. This is achieved by solving station-wise sub-problems of the train platforming problem by means of mixed integer programming (MIP). Remaining routing conflicts are solved within the next step. A consistency check is performed before the resulting railway simulation model is finally set up. In a parallel set of steps, the respective input demand is processed to agent trips. These are merged with the train runs in the model, leading to an agent-based railway model.

In the implementation, SUMO (Lopez et al., 2018) – Simulation of Urban Mobility – is used as the simulation backbone of the methodology. SUMO is a microscopic, agent-based traffic simulation tool developed to build digital twins of various mobility scenarios. It does include, but is not limited to, railway operation. Frequent application scenarios include urban traffic simulation involving trajectories of several means of transport, as well as pedestrians, bicycles and other active modes. Therefore, it is also designed to model intermodal mobility and transportation. Within railways, it features command and signaling to model block occupation and train headways. As a microscopic tool, it further includes switches and track infrastructure, train stops, driving dynamics and different routing approaches to model railway operations. With respect to different types and technological depth of command and signaling as well as the implementation of railway-specific operational guidelines, SUMO does currently not bring the

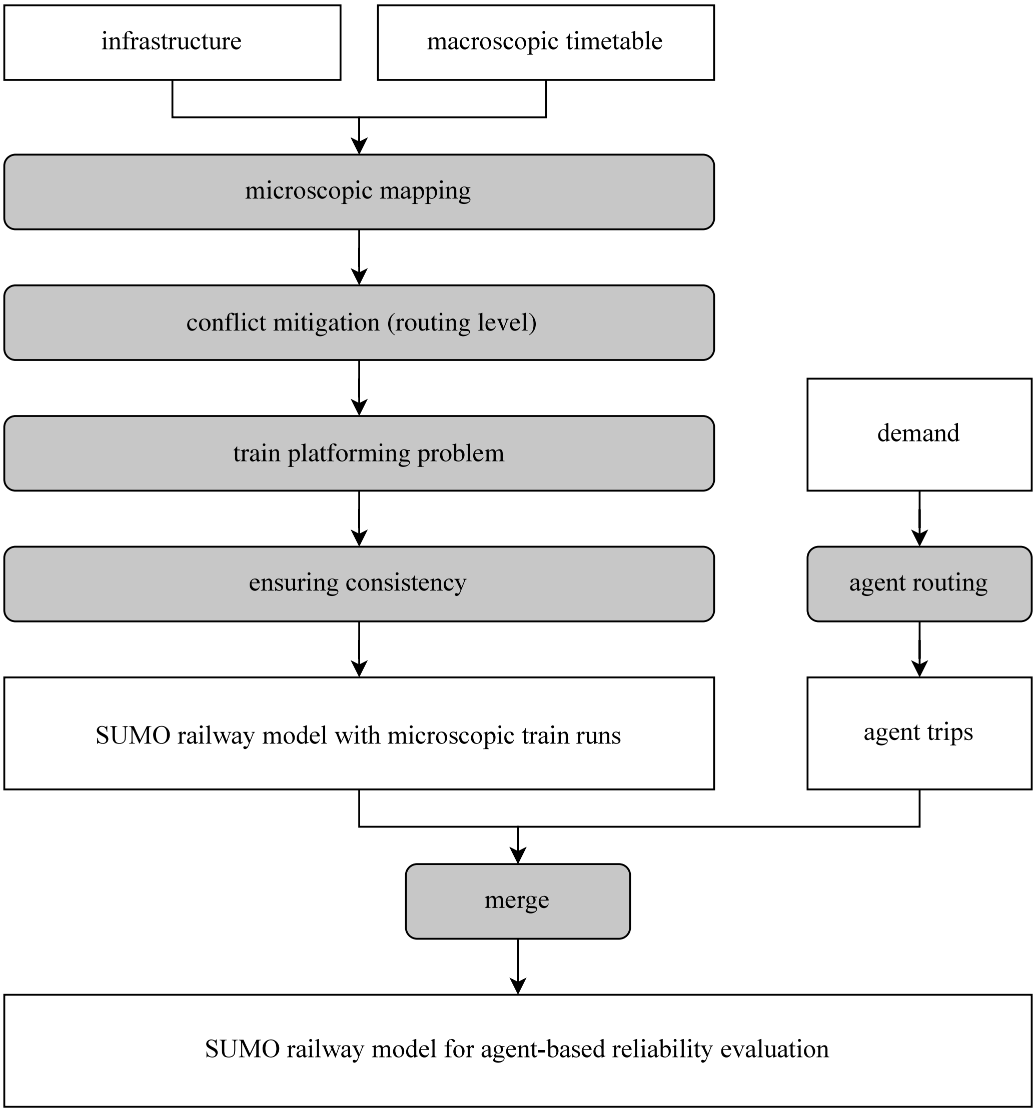

Fig. 1. Overview of methodological approach to create a railway simulation model for agent-based reliability evaluation.

requirements to simulate railway operations on the same professional level as commercial tools developed for the railway industry. For Example, axle counters for track vacancy detection are assumed after every signal and switch. Consequently, the partial release of driveways is assumed to be applied in the most efficient fashion. This might underestimate infrastructure occupation times at locations where local infrastructure characteristics may not correspond to these assumptions. This also applies to dispatching routines, which base on train succession orders but have not implemented passings or platform changes. However, this may be considered tolerable given the simulator features performant large-scale railway operation concepts (also applied in industry projects (Erdmann, 2022)) on the one hand and agent-based intermodal transportation modeling on the other hand. Geischberger and Weik (2022) presents a detailed requirement analysis and explanation how SUMO meets criteria to model simulation networks in the intersection of railway simulation and logistics. Given its design as a digital twin for multimodal urban simulation, it is highly performant in simulating railway systems on the network scale. It can also be used as a backend simulator, for example for reinforcement learning (compare Erdmann (2022)), accepting control inputs via a so-called TraCI (Traffic Control Interface) client (German Aerospace Center (DLR) et al., 2008–2024). This allows to couple the simulation core with customized approaches, for example for mapping macroscopic timetables on microscopic infrastructure or routing trains and agents, as shown in the following sections. It also enables actively manipulating simulation during runtime. Moreover, it enables individual, agent-based output processing to match the above defined reliability indicators. Finally, SUMO is an open-source simulation environment that features a strong link to Open Street Map data (OpenStreetMap, 2024) input and processing, making it ideal as a transparent and highly

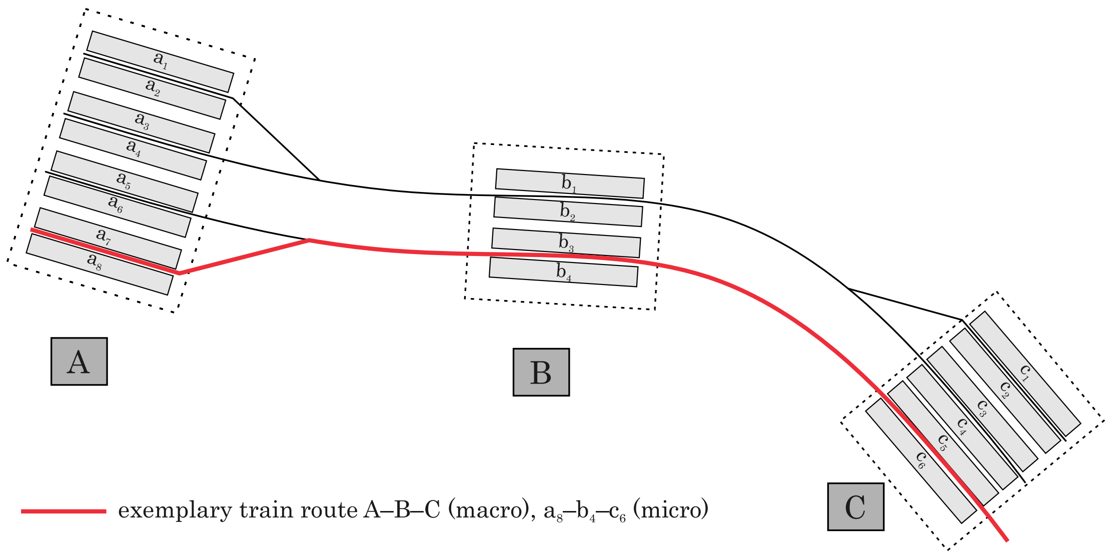

Fig. 2. Schematic view on macro and micro stops on an exemplary infrastructure.

adjustable simulation backbone for our proposed methodology. In the following, the involved steps of the methodology are explained in detail.

3.3. Input data and model setup

Infrastructure is downloaded from OpenStreetMap (OSM) (OpenStreetMap, 2024) via overpass-turbo API (Raifer, 2024), focusing on infrastructure elements with tag “railway”=“rail” (“railway”=“stop” for stations) within the desired bounding-box or region. It is worth noticing that – depending on the region – other keys (for example “public transport”) or values (for example “light rail”) may be imported, too. Next, the downloaded data is transformed by SUMO netconvert functionality (German Aerospace Center (DLR) et al., 2001–2024) into a so-called network file, which serves as infrastructure input for the SUMO backend. This file consists of nodes and edges (representing tracks), as well as railway-specific infrastructure elements, for example signals. Edges can be augmented by characteristics, such as, for example, allowed maximum speed or further OSM data tags considering the usage of infrastructure, for example “service”=“yard”. Railway stops are automatically transformed into model-stops, so called additionals. Every stop is assigned to a specific edge of the network.

Fig. 2 shows the basic concept of infrastructure interpretation, with A, B and C being exemplary macroscopic stations (for example Berlin Main Station), further on referred to as macro stops. Every macro stop has a an assigned number of microscopic stop IDs with an explicit ID, such as a_1 or b_2. These are further on called micro stops.

A configuration was chosen where bidirectional routing is allowed on all edges, given the differing directionality of input data from OSM. Also – despite many countries having a defined main direction of train runs – tracks may be frequently operated in both directions in daily railway operation. The resulting default bidirectionality of edges brings a core challenge in model build-up, as fastest, shortest and wanted path may differ (see Sections 3.4 and 3.5). The SUMO simulator therefore proposes a concept of freely-editable edge-priorities. These edge priorities are assigned to the infrastructure elements in such a way that they manipulate routing as wished by users. Bidirectionality also results in two micro stops per physical platform, one per direction (compare Fig. 2).

Next, timetable data is processed (which, in general, can be acquired from public data sources as well). In case of macroscopic data, a transfer to microscopic train paths is performed (see next Section 3.4). Train runs according to the timetable are then transferred into simulation vehicles, calling at defined micro stops and taking a path (set of edges) through the network. The next sections explain how to create these train paths in such a way that their route is valid, plausible and free of conflicts.

3.4. Microscopic mapping

A core advantage of the proposed methodology is that it can merge microscopic OSM infrastructure-data with macroscopic network timetable data. This is of special importance when dealing with early states of timetabling, infrastructure changes or future concepts in general. More precisely, the sole information on which order of stations needs to be served by a train at what time, is sufficient. Information on specific dwelling platform numbers or station paths are not required.

We apply a routing-algorithm involving TraCI (German Aerospace Center (DLR) et al., 2008–2024) and sumolib routing functionalities (both basing on a Dijkstra-Algorithm) (German Aerospace Center (DLR) et al., 2011–2024) that allows a step-by-step depth-search to identify routable, conflict-free and plausible train paths on a network-level. This is accomplished by a physically motivated metaheuristic as follows:

Definitions (in accordance to Fig. 2):

- M is the set of all macro stops within the network
- $M_x \subset M$ are the macro stops of train x according to the timetable, more precisely the sequence of macro stops the train x calls at: $M_x := (A, B, C, D \dots)$ where $A, B, C, D, \dots \in M_x$ are the respective macro stops
- $A = \{a_1, a_2, \dots\}, B = \{b_1, b_2, \dots\}, \dots$ are sets of micro stops assigned to exactly one macro stop
- The segment-router $sr(A, B)$ is a proposed function to identify valid and plausible paths between two macro stops by trying to find the fastest (or alternatively the shortest) route between two respective micro stops on an n:n-regime, where
  - $A, B \in M_x$
  - P is the set of all fastest paths between two macro stops A, B
  - P_{pr} is the set of fastest paths between two macro stops A, B that are plausible and routable
  - P_{pro} has identical values with P_{pr}, but is ordered by path time
  - $sr(A, B) = P_{pro}(A, B) = \{[a_1, b_1], [a_1, b_2], \dots\}$ with $a_1, a_2, .. \in A$, $b_1, b_2, .. \in B$. This results in every element of P (and P_{pr}, P_{pro} respectively) being a path between to micro stops a_i and b_j (with $i \in \{0, \dots, |B|\}, j \in \{0, \dots, |B|\}$). Paths are calculated based on edge-wise travel times.
  - Routability shall be further referred to as a characteristic of a set of two micro stops a_i and b_j (with $i \in 0, \dots, |A|, j \in 0, \dots, |B|$), where a path between a_i and b_j can be found given the defined routing settings; every path is a connected, ordered sequence of edges and has at least an assigned length and time, where neither may be 0 or inf..
  - Route plausibility shall be further referred to as a characteristic of a set of two micro stops a_i and b_j (with $i \in \{0, \dots, |A|\}, j \in \{0, \dots, |B|\}$), where a path between a_i and b_j can be found given the defined routing settings and the length and/or distance of the route need to be below a certain threshold (which, for example, can be derived from average values or percentual comparisons)
  - Output of the function $P_{pro}(A, B)[0]$ is the first Element of P_{pro} and can therefore be considered the fastest valid and plausible route between two macro stops
  - If - instead of a macro stop - a micro stop is given as one of the input variables, a search only from/to this specific stop (and not the according macro stop) is performed
- S_x is the sequence (ordered set) of micro stops that train x calls at in the simulation environment: $S_x = \{a_x, b_x, c_x, d_x \dots\}$ with $a_x \in A, b_x \in B, c_x \in C, d_x \in D, \dots$.

Output.

The method shall calculate S_x, as shown for an exemplary path in Fig. 2. This is done by a means of a depth-search-procedure as follows:

Procedure (trainwise).

The merge of M_x with S_x by iterative calls of $sr()$ is described in pseudo code in Fig. 3. As shown in the code, the macro-to-micro mapping consists of three main steps:

In step 1, a startpath is identified (see lines 1 and 2 of Fig. 3). This is done separately from the main body, as the algorithm is conceptualized as a depth-first approach – benefiting from a good start to the search – and this is the only step of the procedure where an n:n routing needs to be performed. In addition, the first segment of train runs may likely start in highly complex stations with many different routing options. Thus, other plausibility criteria may be applied for the first segment of train runs, such as predefined tracks or station areas or stricter route length deviations from average (or minimal) values, to reduce the amount of paths found. Experience shows that finding a good initial path can take as long as finding the whole rest of the route. Moreover, in case the first one does not exhibit a useful result, the proposed methodology considers different initial paths.

In step 2, an iterative procedure is initialized where the segment router tries to find microscopic paths from macro stop to macro stop by assigning micro stops to the route. Starting with the second micro stop (defined from the start path, see Fig. 3, line 3), it tries to find paths from the last fixed micro stop to the micro stop assigned to the next macro stop. This can be seen in line 5, where the $i - 1$ references the step before.

In step 3, if a plausible route from the last fixed stop to a next stop (also called "candidate stop") is found (Fig. 3, line 6), an additional route search is initiated, to ensure that a way can be found from the candidate stop to the macro stop thereafter as well (lines 7 and 8). Only then the candidate stop is fixed (line 9), otherwise the candidate stop is discarded and the respective stop deleted from further routing options (lines 11–13). If no route is found, the algorithm will "go back" in the path of micro stops and try an alternative stop (line 17). This behavior is similar to and inspired by depth-first tree searches, where branches are taken or discarded according to results in iteration.

The result of the algorithmic procedure is an ordered set of micro stops S_x for train x, that guarantees a connected path from the first macro stop to the last macro stop, calling at respective intermediate stops. Given the infrastructure characteristics of the model (such as edge speeds or priorities, see above), the path will be most likely the fastest plausible path. However, it does not yet assure that two further requirements for smooth and realistic train operations simulation are met: first, paths need to be non-conflicting on a routing level. Second, within stations an optimal (or at least feasible) track occupation needs to be generated. This is tackled in the next two sections, respectively.

```
Startpath = sr(Mx1,Mx2)
Sx=[Startpath]
i=2
while i <len(Mx) and i >0:
        segment_route=sr(Sx[i-1],Mx[i]) #sr chooses microstop in Mxi
        if segment_route:
                segment_route_ahead=sr(Mx[i],Mx[i+1])
                if segment_route_ahead:
                        Sx.append(segment_route)
                        i=i+1
                else:
                        i=i-1
                        delete(segment_route)
                        #found route is discarded
                        #opted microstop is removed from Mx[i] options
        else:
                i=i-1
```

Fig. 3. Pseudo code for mapping macroscopic timetables to microscopic train paths.

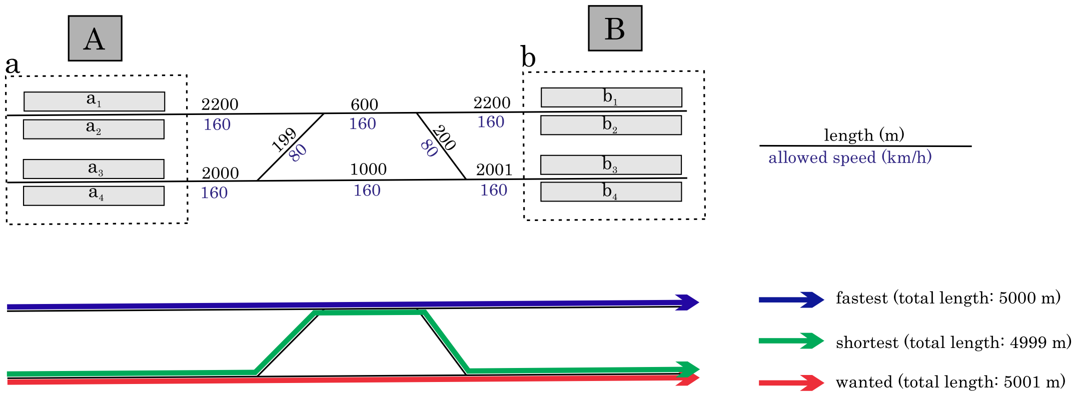

Fig. 4. Shortest, fastest and wanted for route between two exemplary stations.

3.5. Conflict mitigation (routing level)

Routing trains within a large railway network simulation is a procedure of high complexity. As stated before, it is far from trivial to assure trains will not have conflicting microscopic pathways and will find their way as smooth and realistic as possible, without stopping or – in the worst case – ending up in a deadlock situation. It is worth noticing, that "conflicting paths" is further on interpreted in an infrastructural and topological sense, but explicitly not in the dimension of time. This means that topological paths of trains are considered "conflicting" if they partially require the same infrastructure resources — regardless of their scheduled time. The authors are aware that "classical" conflict definition, identification and resolution refers to both time and path. However, in many occasions it may be assumed that the given mapped input timetable was constructed with non-conflicting time slots (for example by specialized timetable software). It is further stated that non-conflicting paths of trains result in a high probability of non-conflicting time slots as well.

Path conflicts mainly arise due to the fact that both fastest and shortest path routing operations may not provide plausible results. Most countries have clear regulations on either the left or right track as main direction (given standard two-track operation) which is, unfortunately, not always compatible with the concepts of shortest or fastest path as shown in Fig. 4.

It can be clearly seen for this example that shortest, fastest and wanted route may at times differ. Due to edge lengths the shortest path consists of edges in opposite main direction and will thus lead to conflicts with the main opposite direction. This can be observed especially within sharp turns. The fastest path may not include such a track change (due to edge speeds) but still may not be in accordance with the desired main direction (here right). As following the main direction is highly probable once the start path is directed accordingly, this particular effect may be lowered within the algorithmic path finding (see above). Still, many cases can be observed, where fastest paths lead to conflicts as well. Also, a core problem is the high sensitivity of paths being found and that found and wanted paths may only vary by a few seconds or meters, respectively. More generally, the main reason for these outlined issues lies in connectivity of infrastructure elements and data quality. Data of low quality and consistency may worsen these issues significantly (still, many cases can be observed independent from data quality).

Consequently, we propose an alternative way of finding plausible routes: based on the results from the previous methodical step, a both valid and plausible route found by the introduced segment-router needs to be tested for routing conflicts. This is accomplished by comparing it with the route of the respective opposite train, fictitiously assuming that they are "passing each other". Only if the edges of the pathways of the respective trains do not match, the train paths are fixed. If edges match, however, two options are left that need to be tackled: first, infrastructure may be single-tracked. This case can be excluded as part of the edge priority calculation (see above), where information on every timetable-segment and whether it is single-tracked or not is obtained. Second, conflicts may be the reason for matching edges within the paths of two trains going in opposite direction. In this case, conflicting edges are identified and prepared for another routing step: by making them momentarily unavailable to the router (for example, by setting their speed to 0), routes without conflicting edges may be found. This procedure is similar to the "Penalty Method" used to calculate alternative paths in street networks (Bader et al., 2011). It can be tried in both directions and significantly lowers the amount of routing conflicts within the network. As a result, as many non-conflicting alternative train paths as possible can be acquired per train (and its respective counterpart in opposite direction).

It is worth noticing that non-conflicting paths may also be acquired using MILP-based routing schemes (Borndörfer et al., 2016). However, given the vast amount of edges and trains within large networks and the complex implementation of main directions on bi-directional paths, this was not further considered. Instead, the proposed train-by-train path-finding scheme with subsequent individual path conflict mitigation exhibits benefits due to comparably low computation times given by its local focus.

3.6. Train platforming problem

Resulting from the previous step, unnecessary path conflicts have been minimized. Still, realistic large-scale railway-operation is not fully assured. Due to the nature of the described path finding algorithm, trains with the same or similar series of macro stops may be assigned to the same micro stops, leading to unrealistic over-occupation of few stops within stations. As a result, train stops within stations should be varied with the target of smoothing junction-internal operations, while at the same time fixing the previously identified paths on lines between stations. This problem is also often referred to as the train platforming problem (TPP) (Zhang et al., 2023b). Within our methodology, the allocation of tracks to trains is accomplished by means of a MIP (mixed-integer programming) approach inspired from literature Caprara et al. (2011) and outlined below. It is worth noticing that the proposed MIP is not operated for strict enhancement purposes but rather for its capability of quickly gaining both feasible and operatively suitable solutions. In the context of conventional optimization, smooth operation within stations is commonly reached by maximizing buffers between train stops on the same platforms, thus rendering robustness as high as possible (see Widmann and Weik (2024) for a deeper insight on robust station routing).

$$\max \sum_{p\in P}\sum_{x\in X} w_{x,p}\, b_{x,p} \tag{1}$$

$$\text{s.t.} \sum_{p\in P} b_{x,p} = 1 \qquad \forall\, x \in X \tag{2}$$

$$\sum_{\tilde{x}\in F_{\tilde{x},p}} b_{\tilde{x},p} \le 1 \qquad \forall\, p \in P,\ x \in X \tag{3}$$

$$b_{x,p} \le r_{x,p} \qquad \forall\, x \in X,\ p \in P \tag{4}$$

$$b_{x,p} \in \{0,1\} \qquad \forall\, x \in X,\ p \in P \tag{5}$$

The model refers to a railway network with a set of trains X, where all trains have a unique train id x. For a railway station, the set of all platforms is defined as P, where individual platforms are assigned a platform id p. The parameter $r_{x,p}$ refers to the topological reachability of a platform p by train x, where 1 means the platform is reachable and 0 means the platform is not reachable for train x. $F_{x,p}$ is further defined as the set of all trains that have a time-overlap in platform p (also derived from input). Table 2 gives an overview of sets, parameters and decision variables used in the model.

The objective function maximizes the train occupation per station by nesting the sum of calling trains at a micro stop over all micro stops, multiplied by a weighting factor. The binary occupation variable $b_{x,p}$ is introduced that can be either 1 (meaning train x calls at platform p) or 0 (meaning that train x does not call at platform p; see Eq. (5)). The weighting factor $w_{x,p}$ is designed in a way close to operational practicability. Therefore, the sequence of reachable platforms p per train x where $r_{x,p} = 1$ (compare Eq. (4)) is sorted as follows: the first item of the sequence is the platform assigned by the algorithmic routing (see Section 3.4). Subsequently, alternative tracks to the geometrical "right" of the assigned platform are added, taking the orientation of the train into account. They are added in a way such that their "rightness" increases. Only then, alternatives to the left of the assigned track are added by decreasing rightness. This results in a sequence that is sorted in accordance to highest aptness in terms of practical application, where the alternative to the right is considered the best alternative, followed by platforms even further to the right. Platforms to the left are considered the least apt alternatives, as they have a higher probability of encountering trains running in

Table 2
Sets, parameters and decision variables of the MIP model.

sets	description
X	set of all trains in the network
P	set of all platforms within a station
$F_{x,p}$	set of all trains that have a time-overlap in platform p (input)
parameters	**description**
x	train id
p	platform id
$w_{x,p}$	weighting factor
$r_{x,p}$	reachability of platform p by train x (1 = reachable, 0 = not reachable; from input)
decision variables	**description**
$b_{x,p}$	occupation of platform p by train x (1 = occupied, 0 = not occupied)

the opposite direction (assuming a right-drive regime). The weighting factor is calculated as 1 divided by the index of an alternative in the sequence of alternatives and affects the objective in such a way as to choose the most apt alternatives where possible.

The set of constraints (2) to (4) is formulated as follows: Constraint (2) assures that every train calls at exactly one micro stop. The information on conflicting trains at a certain station is processed as an input for the MIP. Constraint (3) implements this into the optimization problem: $F_{x,p}$ is a set of trains that have an overlap by time according to the stop occupation data which resulted from the previous methodological steps. The sum of overlapping trains in a station $F_{x,p}$ is designed as less or equal to 1 (inequality formulation). This means that overlaps are prohibited for the occupation planning.

A core challenge within the train platforming problem is that alternative stops are not only bound to infrastructural restrictions but also depend on the route of a train. This means that – given the intertwingularity of railway infrastructure – not all available alternative platforms within a station will be valid alternatives, given the topological macroscopic route and timetable of a specific train (this especially applies to stations where lines merge). Here, a synergy between routing and optimization can be exploited: train-specific, station-wise platform alternatives are identical to the above defined micro stops. For this reason, alternative micro stops per macro stops and train can be identified as a second output of the route-finding algorithm proposed in Section 3.4. Within the mathematical model, this is transferred into the variable $r_{x,p}$, as introduced in inequality constraint (4). According to this constraint, a platform p being not reachable by a train x (and thus $r_{x,p} = 0$) implies that x will not be allowed to call at p (and thus $b_{x,p} = 0$).

As a result of the methodological steps outlined so far, all train runs according to the macroscopic input timetable are transferred to connected and plausible microscopic paths. These paths are then minimized for path-conflicts on an edge basis. Last, the called micro stops per macro stop are varied, making sure that no overlapping platform occupations occur. At the same time, train paths between stations remain fixed, assuring operation as smooth as possible. As a result, the train model creation is accomplished.

3.7. Consistency

Consistency is an important requirement to a subsequent validation of the model and its respective behavior. Consistency is referred to both the used and pre-processed input data as well as the model itself.

With respect to input data, consistency is ensured by using data sets (compare Section 3.3) that are produced under the same circumstances and conditions. Furthermore, heterogeneous data sources are transformed to a common reference station nomenclature. Analogously, the definition, derivation and use of the introduced concept of macro and micro stops was consistently applied throughout different data sets, even though some sources may have different terms for it (for example "platform", "track", "station"). Moreover, driving dynamics profiles of SUMO vehicles are calibrated and aligned to those used in timetabling.

With respect to the modeling procedure, compatibility and consistency is ensured by the fact that the model is based on a fundamental infrastructure graph representation ("ground truth"), which is used by all subroutines (or sub processes) of the simulation, including routing functionalities. This also includes the above introduced train platforming problem (compare Section 3.6), where train paths from the infrastructure graph were transformed into variables. Moreover, analyzed train simulation data represent a result of the calibrated SUMO model.

3.8. Agent routing

In reference to the toolchain described in Fig. 1 and Section 3.1, the last sections described the detailed process on how to create the train level of the simulation model. Next, this level is merged with the level of logistic flows to implement the concept of transportation chains. This is accomplished as follows:

We rely on demand structure in the form of origin–destination relations (also often referred to as OD relations). These datasets contain information on the amount of load units being transported from a certain origin location (further on called source) to a destination location (further on referred to as sink) within a defined time span. However, this information does not yet contain information on which route the respective transport takes through the network to fulfill the demand. Moreover, demand data may potentially be of different granularity, typically more rough than the information level of service: for example, demand sources might be cities or regions (sometimes also referred to as a demand cell) rather than specific platforms or tracks within specific stations.

The resulting agent routing scheme applied as part of the proposed methodology is characterized as follows: First, demand of a certain cell is allocated to the respective stations within the cell. Stations are considered eligible for a demand source or sink if they are scheduled within freight train timetables. If no explicit freight train timetables exist, eligible stations may be chosen by economic characteristics, such as freight hubs, production plants or other factors that suggest suitability for freight transport supply. In case of more than one eligible station per cell, demand is equally distributed between the respective stations. As a result, OD demand is transformed from a cell-to-cell level to a station-to-station-level, or – in accordance to Section 3.4 – to a macro stop-to-macro stop level.

It is assumed that both demand sources and sinks are indifferent between micro stops within macro stops. This means that transportation chains may start/end in any train calling at any platform within the macro stop defined as source or sink. Therefore, all micro stops within a macro stop are linked with a so-called "demand-edge" that represents the demand function of the macro stop as a sink or source. Next, a fastest path search on a graph representation of the macro timetable is performed, where agents from one demand edge within the infrastructure are routed to another demand edge. The resulting transportation chain involves different train services and is calculated by fastest total travel time. In order to do so, an underlying production scheme needs to be specified: transshipment stations are – analogously to sink and source stations – defined according to freight train timetables, more precisely stations where both freight train arrivals and departures exist. This is again accomplished by setting demand edges within transshipment stations (see above). Moreover, fixed or variable transshipment times and time spans may be defined by the characteristics of the demand edges within transshipment stations. In simulation, the transshipment process is performed on a first-in first-out (FIFO) basis, meaning that all agents are assigned the same priority and trains are loaded based on when the agents arrived at the transshipment stations.

As a result, every single element of the OD matrix can be assigned to a specific set of trains with in-between-transshipments. The resulting transportation chain (sometimes also referred to as *trip* in the context of intermodal mobility simulation) is characterized by a sequence of train rides and transshipments in order for the good to reach its sink. Thus, agents and trains are merged. Every agent trip has a predefined route that is statically computed in pre-processing, where both origin and destination as well as transshipment stations are explicitly set on a micro stop basis. Moreover, specific intended trains with departure and arrival information are assigned to the trip. Once a trip cannot be upheld during simulation (for example due to high delays), agents will take the next available train run to their next transshipment station or final destination.

4. Results and discussion

The methodical steps described in the sections before lead to a simulation model that has both a train and a transportation chain level and enables reliability analyses on both levels. The methodology is applied to the use case of Germany: infrastructure is imported and processed according to the described process steps. Next, publicly available macroscopic timetable data (Bundesministerium für Digitales und Verkehr, 2024) is mapped on the infrastructure, train path conflicts are resolved and station track occupations optimized. For a realistic network occupation, both freight and passenger trains were imported and processed. Load units are considered wagons or containers, depending on the production system. All freight train types can transport all agent types and are assigned a specific capacity on how many agents can be loaded. The train operation model is populated with agents, where sources and sinks were derived from freight timetable data and demand quantities on these relations were assigned randomly. The resulting transportation chains were limited to a maximum of three transhipment processes. Transhipment times were defined to 60 min, with a minimum of 30 min in case of delay. The resulting two-leveled model is simulated for a 24 h time span - a full operational day. Computation was carried out on a compute server with an Intel(R) Xeon(R) Platinum 8358 CPU and 32 GB of RAM. Using SUMO version 1.21, an average computation time of 98 min per simulation cycle was observed. One computation cycle contains 24 h of microscopic railway simulation of the German national railway network as well as agent-specific modeling of transported goods.

To evaluate operational reliability, train entry delays were injected according to DB guidelines (DB NETZ AG, 2008). Currently, the model does not explicitly include dwell-time excess or specific runtime variations. However, given the size of the network with its vast intrinsic complexity of train interactions and the nature of microscopic time-step-based simulation, these effects can be observed indirectly via secondary delays. For example, trains may depart or arrive late as a reaction to a red signal due to blocked driveways. Regarding minimum dwell times, the following behavior has been implemented: All passenger trains have a train-type specific minimum dwell time (2 min for long distance trains, 0.7 min for regional trains, 0.25 min for commuter trains) that may be smaller or equal to the scheduled dwell time. This leads to highly delayed trains stopping for their minimum dwell time, whereas punctual trains may stop longer than that depending on the scheduled dwell time.

4.1. Results

Fig. 5 shows the resulting reliability, as defined in Section 3.1. Here, the arrival time window for transportation chains is defined with a lower bound of −20 min, meaning that goods may be a maximum of 20 min early to meet the arrival time window. The upper bound of the arrival window is varied, as an acceptable maximum delay may vary from user to user. This means that along the x-axis the time span of the window is successively expanded. This leads to more and more train and transportation chain arrivals meeting the window.

Fig. 5 shows four plots with different delay injection levels. Here, the delay scenarios were defined based on the above mentioned DB guideline but in a discretized derivation as in Haehn (2022). Based on that, scenario 1 is considered the reference scenario,

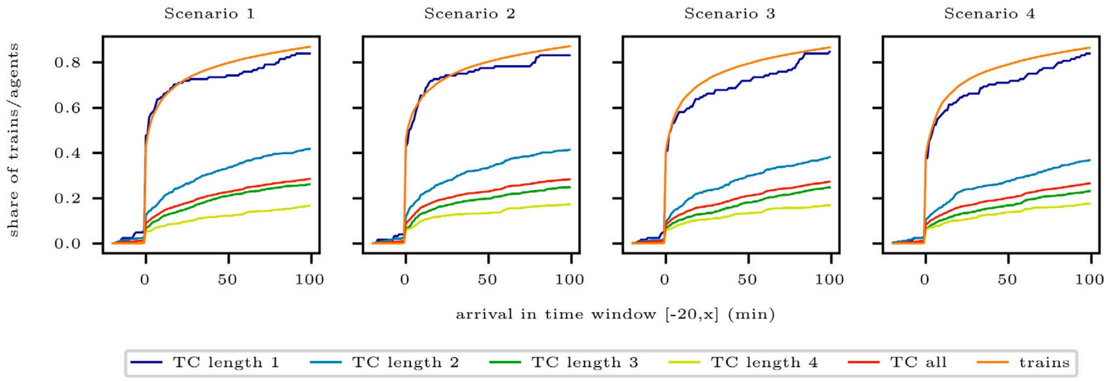

Fig. 5. Train and transportation chain (TC) reliability by arrival within defined time window.

Table 3
Delay scenarios adapted from DB NETZ AG (2008), Haehn (2022).

Category	Entry delays (min)				
	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Probabilities
Long dist.	(0, 1, 4, 10)	(0, 0.5, 2, 5)	(0, 1.5, 6, 15)	(0, 2, 8, 20)	(0.5, 0.17, 0.18, 0.15)
Regional	(0, 1, 4)	(0, 0.5, 2)	(0, 1.5, 6)	(0, 2, 8)	(0.5, 0.32, 0.18)
Commuter	(0, 1, 4)	(0, 0.5, 2)	(0, 1.5, 6)	(0, 2, 8)	(0.8, 0.16, 0.04)
Freight regional	(0, 10, 30)	(0, 5, 15)	(0, 15, 45)	(0, 20, 60)	(0.6, 0.2, 0.2)
Freight long dist.	(0, 10, 30)	(0, 5, 15)	(0, 15, 45)	(0, 20, 60)	(0.5, 0.25, 0.25)

with scenario 2 having injected delays reduced by 50% while scenarios 3 and 4 have a rise in injected delays of 50% and 100%, respectively. Percentage steps were chosen in accordance with Sipilä (2015), but adapted to fit the used approach. Table 3 shows an overview of the scenarios.

For all scenarios, the curves can be seen as accumulative curves and therefore rise continuously. The punctuality threshold for trains is aligned with the arrival time window for goods. This means that, for example, the values at x=20 base on an arrival window for both levels going from −20 to 20. Referring to this example, the orange curves show the share of trains that arrive within that window, which might also be referred to as the "20-minute-punctuality" (compare Sections 2 and 3.1). Analogously, the value of the other curves at for example x=20 show the share of transported goods that arrive within the defined window of [−20,20]. This corresponds to a railway system user that defines that shipped/ordered goods may be early by a maximum of 20 min and late by a maximum of 20 min as well.

For all scenarios, it can be seen that the orange curve sharply rises around the value of 0 which means that a good share of trains have no or little delay, while only very few arrived early. The curve then changes to a less steep ascend, meaning that expanding the window does not add as large train numbers to the window anymore as to the beginning.

The curve colored red shows the average reliability of transportation chains, while other colors distinguish curves by the amount of segments (for example a transportation chain with length 2 (light blue) means that one transshipment takes place). While they all ascend steadily, the values for length 2, 3 and 4 are significantly far below those of the train punctuality, which is due to the fact that these transportation chains involve more than one train run and, hence, the delay of trains is transferred to the transportation chains multiple times. This is supported by the dark blue curve that shows the behavior of direct transportation (length 1). Interestingly, it aligns to the punctuality of trains quite narrowly, pinpointing the effect of direct transportation in comparison to transshipment. Moreover, reliability decreases with increasing length (and thus complexity) of transportation chains.

Between the scenarios, comparably little effect of delay injection variation can be observed for the punctuality of trains (orange). It is worth noticing that the punctuality depicted here is accumulated over all train categories. As the used delay injection guideline uses severely higher delays for freight trains and these were increased over the scenarios by percentage, still the absolute surplus of injected delays are higher for freight trains in scenarios 3 and 4 compared to passenger trains. Again, the unreliability of transportation chains only slightly increases within the defined window, as can be observed from the red curve for all transportation chains. This, however, does not necessarily exclude a stronger negative impact of higher delay levels outside the window. Moreover, direct transports are slightly more sensible to injected delays (dark blue curve), with a visible descent in scenarios 3 and 4. It can also be seen from the curves that the remaining share of transported goods and trains, respectively, do not arrive within the defined time window. This means they might be even earlier or have a delay larger than 100 min.

To assess this in more detail, the density of absolute arrivals of trains and goods is compared by the example of the base scenario 1 in Fig. 6. The density plot at the top shows the punctuality of trains analogously to Fig. 5. It can be seen that the vast majority of trains arrive within the defined maximum time window [−20,100] of Fig. 5, which is taken up by the vertical dashed black lines. More specifically, a clear peak in density reflecting no or little delay can be observed.

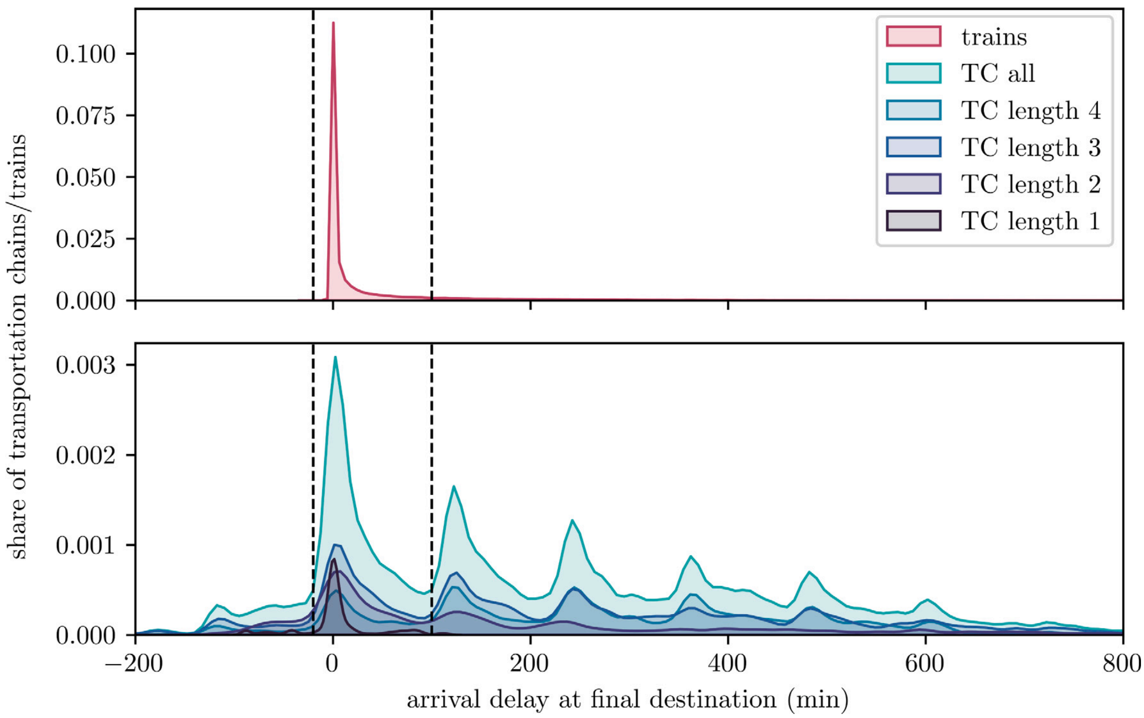

Fig. 6. Arrival delay densities for trains and transportation chains for scenario 1.

The lower part of the figure shows arrival time and hence reliability distributions by their density for transportation chains. Here, the green plot shows the distribution for all agent arrivals, while the remaining colors differ by the amount of transshipments, for example the black curve shows the behavior of direct transportation. The normation was chosen in an accumulative way, meaning that the densities of all other curves add up to the curve of all transportation chains. For that reason another scale to the y-axis was chosen compared to the top figure to better visualize effects on transportation chains.

It can be observed clearly that the arrival time of transportation chains is spread far wider than the punctuality of trains, with a considerable amount of transportation chains not meeting the requested window. This means that the used delay injection guidelines in combination with the chosen transhipment parameters lead to a high share of unreliable transportation chains, often delivering goods significantly late, despite showing solid punctuality rates for trains. The low freight transportation reliability corresponds with statistics from 2021 regarding unpunctual freight trains in Germany, according to which 42% of German freight trains (not goods) had a delay of 1 h or higher (Bundestag, 2022). Moreover, reliability is highly impacted by the length of the transportation chain, as can be seen by the layered characteristic of the blue and green curves: the more segments a transportation chain has, the higher the share of unreliable transports is. The peaks which roughly occur every two hours can be explained by the basic frequency of the underlying timetable, where the next alternative transportation in case of missed transshipment may be 120 min later. Moreover, a fair share of goods arrives early. This might be rather surprising at first sight, but can be explained by the mentioned buffer of 30 min in transshipment times and a certain share of highly delayed trains that replace scheduled transportation chain plans.

Coming to local effects, Fig. 7 compares two map representations of the reliability of transportation chains: The left-hand figure shows all final destinations of goods (sinks) - derived from timetable data and the random demand - and the respective arrival delay of transportation chains ending at these locations. This represents the perspective of the receiver of transported goods, hence the arrival delay of a transported good is plotted to the sink of the transportation chain. The right-hand figure depicts a comparison to the perspective of the sender, the initiator of the transportation process. Hence, the right side of the double figure assigns arrival delays of goods at the final destination to the source of the transportation chain. The figure adds understanding of local effects to the systemic reliability analysis shown in Fig. 5 by comparing both the initiation and reception of goods.

Both figures show the arrival delay of goods per sink (and source, respectively) averaged over all arriving goods. For the sinks, it can be seen that a low share has especially high mean arrival delays of transportation chains (red), while the majority of locations have medium arrival delays. Another considerable amount of stations is assigned no arrival delay or even early arrival (blue and green dots). This refers to the transportation chains where goods arrived according to their plans, while the other chains (red) were not performing as scheduled initially. In comparison, the right plot largely shows similar structure.

Four combinations between the two plots exist: First, stations show low arrival delays from both the perspective of senders and receivers. Second, stations show high arrival delays in both figures, accounting for high unreliability of transportation in both ways. Third, stations may show low receiving, but high sending reliability and – fourth – vice versa. The figure shows all four

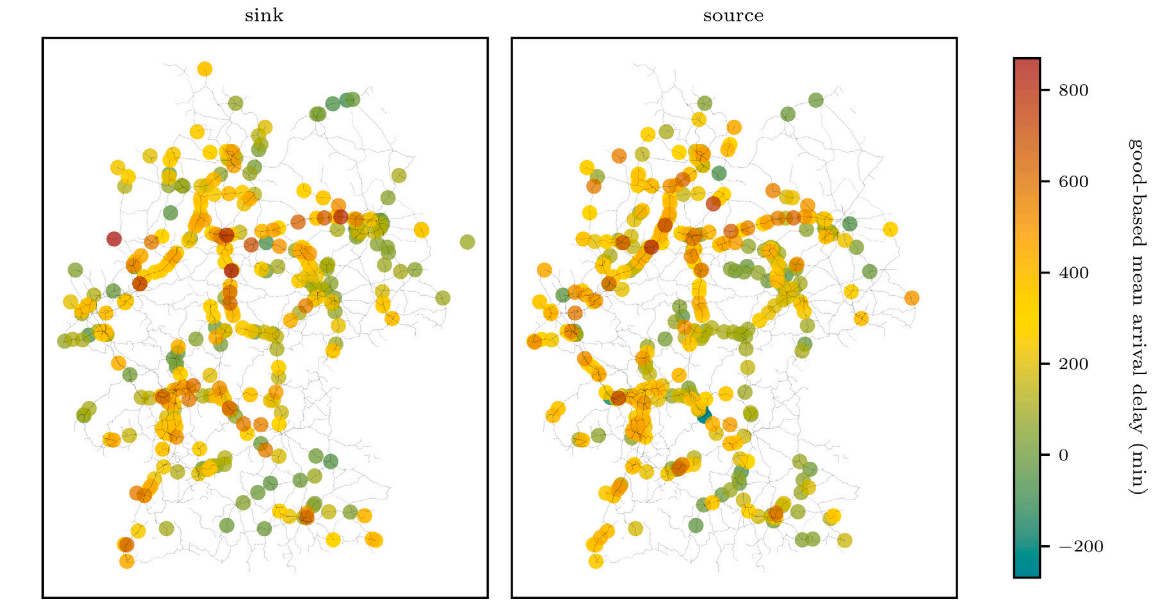

Fig. 7. Agent-based mean arrival delay of transportation chains aggregated at demand sink (left) and source (right)

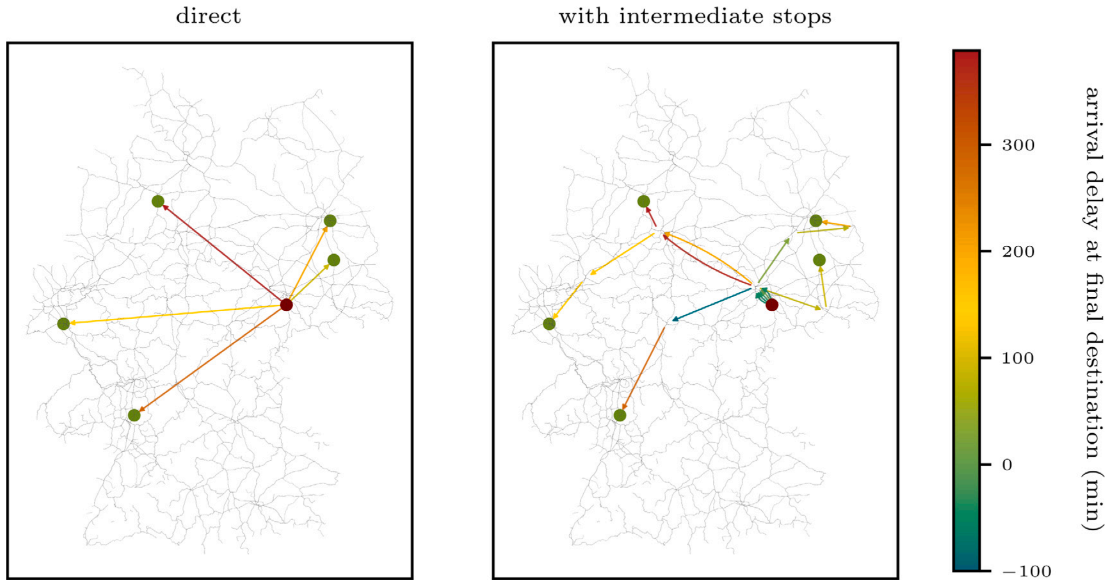

Fig. 8. Five exemplary transportation chains, all starting from one source (red) and ending at sinks (green). Reliability is compared between direct (left) and segment-wise (right) assignment of delays.

cases and helps understanding where exactly unreliable transportation chains affect the users of the railway system the most. It is worth noticing that the created random demand was deliberately not forced to be symmetric (similar to real-life demand patterns, for example in production supply chains). Moreover, the figure does explicitly not show where transportation chains "broke", so a high arrival delay does not necessarily mean that goods had to wait there for a long time but rather that users (both senders and receivers) of the system may be particularly unsatisfied at that location.

The exact course of transportation chains is exemplary shown in Fig. 8. To the left, a source (red) and five selected sinks (green) of a transportation chain were directly linked with a pseudo-link that represents direction and arrival delay of the whole transportation process, independent from eventual transshipment processes. This again takes up the perspective of the users of the railway system where only the arrival delay at the final destination is of relevance. To the right, the same transportation process is broken down to its segments, meaning that links stand for actual train runs and their color for the punctuality of trains, respectively.

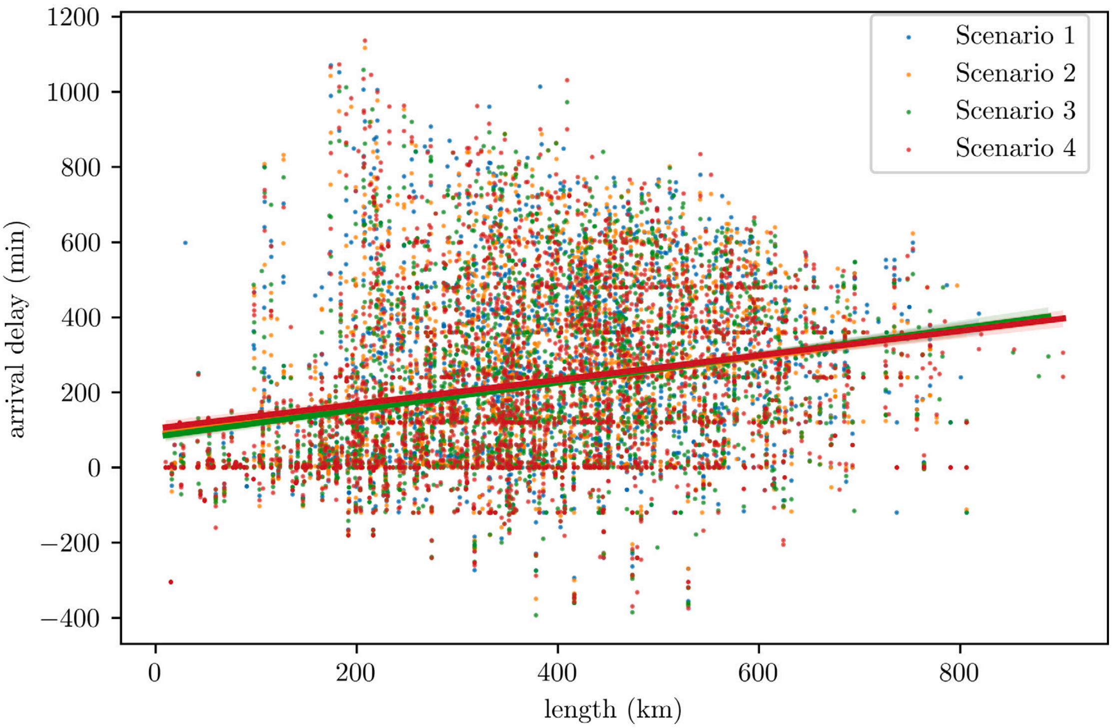

Fig. 9. Correlation of reliability and length of transportation chains over all delay scenarios.

Here, the core effect of the correlation between trains and transportation chains can be seen: in case only one involved train is delayed, the respective delay is transferred to the transportation chain as a whole. Punctual trains may have no effect, as indicated by the green and blue links in the right figure.

Last, the correlation between the length and reliability of transportation chains is shown in Fig. 9. Here, lengths correspond to physical train paths, not air distances. Again, the figure includes all four delay scenarios. A linear regression curve is estimated, showing very similar behavior between the scenarios: All four trend lines come with a rise of arrival delay (and hence lower reliability) of transportation chains when transportation chains are longer. This corresponds to what might intuitively be the case for rail freight transportation, as longer runs come with a higher probability of interacting with other trains and gaining secondary delays. However, the delay data structure shows a high variance and therefore does by far not show a correlation as clear as between arrival delays and the amount of segments of the transportation chain (compare Fig. 6).

4.2. Discussion

Summarizing, the findings underline the importance for joint modeling of (microscopic) rail traffic control systematics and macroscopic network structures for macroscopic supply-chain reliability in rail freight transportation. Extending existing open-source tools and libraries, a comprehensive approach for agent-based modeling of rail freight logistics has been demonstrated. Allowing for a mapping of macroscopic network timetables to microscopic simulation infrastructure, the approach is particularly suitable for early planning stages in timetable and freight logistics planning, where the proposed toolchain can reveal reliability implications and shortcomings of envisioned traffic concepts. Second, the study can be seen as an impulse to more strongly integrate the perspective of end-customers in rail freight transportation planning and operations. As explained in Section 1 – and shown in Fig. 5 – two perspectives on the same transportation process may involve very different evaluation metrics and parameters. Given the complexity of every-day railway operations, the study hence deepens the practical understanding of involved stakeholders, their preferences and respective views on the transportation process (as shown in Fig. 8). Third, in view of the strong effects of transport segmentation (see Fig. 6) on transport reliability, trans-shipments have been identified as key to transport reliability in the system — which, in combination with strong perturbations currently observed in Europe's heavily congested rail transport system, will introduce high delay risks. This highlights the need to develop solutions to target supply-chain reliability at this level, where the current approach is believed to yield important insights for scenario-based analysis and optimization of macro-economic transport planning decision making.

5. Overview and outlook

In this study we propose a railway simulation methodology that is able to assess the reliability of railway operations not only on the level of trains, but also on the level of transportation chains. This is accomplished by merging trains with units of goods in an agent-based approach. The resulting simulation model is capable of modeling railway specifics (for example signaling) and at the same time logistic flows and how they interact. This extends previous approaches and enables an assessment of railway operation reliability from the perspective of the users of the railway system. A further feature of the proposed methodology is the implementation of future timetable concepts by mapping macroscopic timetable data to microscopic infrastructure and omitting path conflicts. The resulting model is able to simulate large-scale networks and is based on an open source simulation environment.

Reliability of transportation chains is defined by arrival within a time window defined by the user of the system. This time window was aligned with punctuality thresholds used in the railway industry. Results show that transportation chain reliability is far below the respective train punctuality window, given an initial delay setting according to German DB guidelines. This can be explained by the structure of transportation chains, as they typically involve not only one train run but rather a sequence of trains. Hence, complexity is larger within the level of transportation chains. This is confirmed by simulation results, showing a far more dispersed distribution of arrival delays with transportation chains compared to trains. Rising delay levels in the simulation lead to slight rises in overall train unpunctualities, also affecting the unreliability of transportation processes. Results also highlight local effects, mapping the reliability of transportation chains to both their sinks and sources, pinpointing potentially low user satisfaction rates from both the perspective of senders and receivers. Hence, these spots should be initial points for further system improvements. Moreover, results show that short transportation chains are slightly more likely to be reliable. However, a far stronger correlation can be observed between the reliability of transportation chains and the amount of involved transshipment processes. It is also confirmed that high delays in only one segment of a transportation chain affects the transportation process as a whole.

Further research includes the implementation of extended train disposition algorithms to model more sophisticated train interaction (for example train passings). Moreover, transportation demand patterns can be included using real-live data and coupled with a more detailed train capacity modeling, for example by linking it with a CVRP (capacitated vehicle routing problem) algorithm. Furthermore, the transshipment processes within the freight hub station can be included into the simulation routines, also considering different priorities. These enhancements will help to further refine the proposed methodology and strengthen reliability evaluation methodologies in railway systems towards their users.

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

The data used is open source.

References

Bader, Roland, Dees, Jonathan, Geisberger, Robert, Sanders, Peter, 2011. Alternative route graphs in road networks. In: Marchetti-Spaccamela, Alberto, Segal, Michael (Eds.), Theory and Practice of Algorithms in (Computer) Systems. In: Lecture Notes in Computer Science, vol. 6595, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 21–32. http://dx.doi.org/10.1007/978-3-642-19754-3_5.

Blayac, Thierry, Stéphan, Maïté, 2021. Are retrospective rail punctuality indicators useful? Evidence from users perceptions. Transp. Res. Part A: Policy Pract. 146, 193–213. http://dx.doi.org/10.1016/j.tra.2021.01.013.

Borndörfer, Ralf, Klug, Torsten, Schlechte, Thomas, Fügenschuh, Armin, Schang, Thilo, Schülldorf, Hanno, 2016. The freight train routing problem for congested railway networks with mixed traffic. Transp. Sci. 50 (2), 408–423. http://dx.doi.org/10.1287/trsc.2015.0656.

Bundesministerium für Digitales und Verkehr, 2024. Daten des Zielfahrplans Deutschlandtakt im Format railML TT. URL: https://www.govdata.de/web/guest/suchen/-/details/daten-des-zielfahrplans-deutschlandtakt-im-format-railml-ttebb5b.

Bundestag, Deutscher, 2022. Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Wolfgang Wiehle, Dr. Dirk Spaniel, René Bochmann, weiterer Abgeordneter und der Fraktion der AfD – Drucksache 20/382 –: Aktuelle Probleme des Schienengüterverkehrs. 03.02.2022. URL: https://dserver.bundestag.de/btd/20/006/2000603.pdf.

Caprara, Alberto, Galli, Laura, Toth, Paolo, 2011. Solution of the train platforming problem. Transp. Sci. 45 (2), 246–257. http://dx.doi.org/10.1287/trsc.1100.0366.

Central Japan Railway Company, 2024. About the Shinkansen: Punctuality. URL: https://global.jr-central.co.jp/en/company/about_shinkansen/.

DB NETZ AG, 2008. Richtlinie 405 Fahrwegkapazität.

de Fabris, Stefano, Medeossi, Giorgio, Montanaro, Giuliano, 2018. TRENISSIMO : Improving the microscopic simulation of railway networks. In: Passerini, G., Mera, J.M., Tomii, N., Tzieropoulos, P. (Eds.), Computers in Railways XVI : Railway Engineering Design and Operation. In: WIT Transactions on The Built Environment, WIT PressSouthampton UK, pp. 199–211. http://dx.doi.org/10.2495/CR180181.

de Jong, Gerard, Kouwenhoven, Marco, Bates, John, Koster, Paul, Verhoef, Erik, Tavasszy, Lori, Warffemius, Pim, 2014. New SP-values of time and reliability for freight transport in the Netherlands. Transp. Res. Part E: Logist. Transp. Rev. 64, 71–87. http://dx.doi.org/10.1016/j.tre.2014.01.008.

Deutsche Bahn, 2024. Pünktlichkeitswerte. 21.06.2024. URL: https://www.deutschebahn.com/de/konzern/konzernprofil/zahlen_fakten/puenktlichkeitswerte-6878476.

Erdmann, Jakob, 2022. Simulation als Baustein für KI Anwendung in der Disposition von digitalen Eisenbahnknoten. In: SYMPOSIUM DER ZUG ZUR DIGITALISIERUNG, 30.11.2022. URL: https://elib.dlr.de/191378/1/ZzD_Simulation_KI_Dispo.pdf.

European Union Agency for Railways, 2024. ERA Railway Factsheet: Germany and France. URL: https://www.era.europa.eu/assets/factsheets.

Europe's Rail Joint Undertaking, 2021. Master plan. October 2021. URL: https://rail-research.europa.eu/wp-content/uploads/2022/03/EURAIL_Master-Plan.pdf.

Federal Ministry of Transport and Digital Infrastructure, 2017. Rail freight masterplan. June 2017. URL: https://bmdv.bund.de/SharedDocs/EN/publications/rail-freight-masterplan.pdf?__blob=publicationFile.

Geischberger, Jakob, Weik, Norman, 2022. Combining operative train simulation with logistics simulation in SUMO. SUMO Conf. Proc. 3, 145–157. http://dx.doi.org/10.52825/scp.v3i.112.

German Aerospace Center (DLR), et al., 2001–2024. SUMO documentation: netconvert. URL: https://sumo.dlr.de/docs/netconvert.html.
German Aerospace Center (DLR), et al., 2008–2024. SUMO traci documentation. URL: https://sumo.dlr.de/pydoc/traci.html.
German Aerospace Center (DLR), et al., 2011–2024. SUMO sumolib documentation. URL: https://sumo.dlr.de/pydoc/sumolib.html.
Gong, Weiwei, Zhou, Lingyun, Ye, Fei, 2019. Multi-agent GIS simulation for railway logistics optimization. In: The 4th International Conference on Intelligent Transportation Engineering. IEEE, Piscataway, NJ, pp. 64–68. http://dx.doi.org/10.1109/ICITE.2019.8880169.
Haehn, Rebecca, 2022. Optimisation and Analysis of Railway Timetables Under Consideration of Uncertainties (Dissertation). http://dx.doi.org/10.18154/RWTH-2022-10063.
Holmgren, Johan, Davidsson, Paul, Persson, Jan A., Ramstedt, Linda, 2012. TAPAS: A multi-agent-based model for simulation of transport chains. Simul. Model. Pract. Theory 23, 1–18. http://dx.doi.org/10.1016/j.simpat.2011.12.011.
Huerlimann, D., Nash, A., 2004. Railroad simulation using OpenTrack. WIT Trans. Built Environ. 74, 10.
Jiang, Zhi-bin, Li, Feng, Xu, Rui-hua, Gao, Peng, 2012. A simulation model for estimating train and passenger delays in large-scale rail transit networks. J. Central South Univ. 19 (12), 3603–3613. http://dx.doi.org/10.1007/s11771-012-1448-9.
Kramarz, Marzena, Przybylska, Edyta, Wolny, Maciej, 2022. Reliability of the intermodal transport network under disrupted conditions in the rail freight transport. Res. Transp. Bus. Manag. 44, 100686. http://dx.doi.org/10.1016/j.rtbm.2021.100686.
Landex, A., 2012. Network effects in railways. 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. 391–401. http://dx.doi.org/10.2495/CR120331.
Lenze, Wiebke, Nießen, Nils, 2021. Quality in rail timetabling - A detailed review of stakeholders' interests compared with current practices in Germany. J. Rail Transp. Plan. Manag. 20, 100282. http://dx.doi.org/10.1016/j.jrtpm.2021.100282.
Lopez, Pablo Alvarez, Wiessner, Evamarie, Behrisch, Michael, Bieker-Walz, Laura, Erdmann, Jakob, Flotterod, Yun-Pang, Hilbrich, Robert, Lucken, Leonhard, Rummel, Johannes, Wagner, Peter, 2018. Microscopic traffic simulation using SUMO. In: 2018 IEEE Intelligent Transportation Systems Conference. IEEE, Piscataway, NJ, pp. 2575–2582. http://dx.doi.org/10.1109/ITSC.2018.8569938.
Mancera, Albert, Bruckman, Dirk, Weidmann, Ulrich, 2015. Single wagonload production schemes improvements using GüterSim (Agent-based Simulation Tool). Transp. Res. Procedia 10, 615–624. http://dx.doi.org/10.1016/j.trpro.2015.09.015.
Matsiuk, Viacheslav, Galan, Olga, Prokhorchenko, Andrii, Tverdomed, Volodymyr, 2021. An agent-based simulation for optimizing the parameters of a railway transport system. In: ICT in Education, Research and Industrial Applications ; Volume 1: Main conference, Ph.D. Symposium, and Posters. URL: http://ceur-ws.org/Vol-3013/#20210121.
Meng, Xuelei, Wang, Yahui, Jia, Limin, Li, Lei, 2020. Reliability optimization of a railway network. Sustainability 12 (23), 9805. http://dx.doi.org/10.3390/su12239805.
Nielsen, Otto A., Landex, Otto, Frederiksen, Rasmus D., 2009. Passenger delay models for rail networks. In: Wilson, Nigel H.M. (Ed.), Schedule-Based Modeling of Transportation Networks. In: Operations Research, Computer Science Interfaces Series, vol. 46, Springer, New York, NY, pp. 1–23. http://dx.doi.org/10.1007/978-0-387-84812-9_2.
Niu, Yi-Feng, Xiang, Hai-Yan, Xu, Xiu-Zhen, 2024. Expected performance evaluation and optimization of a multi-distribution multi-state logistics network based on network reliability. Reliab. Eng. Syst. Saf. 251, 110321. http://dx.doi.org/10.1016/j.ress.2024.110321.
Office of Rail and Road, 2024. Passenger rail performance: 1 January to 31 March 2024. 30 May 2024. URL: https://dataportal.orr.gov.uk/media/jwfpdpty/performance-stats-release-jan-mar-2024.pdf.
OpenStreetMap, 2024. OpenStreetMap: OpenStreetMap is open data, licensed under the Open Data Commons Open Database License (ODbL) by the OpenStreetMap Foundation (OSMF). URL: https://www.openstreetmap.org.
Paulsen, Mads, Rasmussen, Thomas Kjær, Nielsen, Otto Anker, 2018. Modelling railway-induced passenger delays in multi-modal public transport networks: An agent-based Copenhagen case study using empirical train delay data. In: 14th Conference on Advanced Systems in Public Transport. CASPT 2018.
Raifer, Martin, 2024. overpass turbo. URL: https://overpass-turbo.eu/.
Rail Freight Forward, 2024. 30 by 2030: Rail Freight strategy to boost modal shift: European Rail Freight Vision 2030. URL: https://www.railfreightforward.eu/sites/default/files/usercontent/white_paper-30by2030-150dpi6.pdf.
Rail Management Consultants International GmbH, 2024. RailSys. 21.06.2024. URL: https://www.rmcon-int.de/railsys/.
Rietveld, P., Bruinsma, F.R, van Vuuren, D.J, 2001. Coping with unreliability in public transport chains: A case study for Netherlands. Transp. Res. Part A: Policy Pract. 35 (6), 539–559. http://dx.doi.org/10.1016/S0965-8564(00)00006-9.
SBB, 2024. Punctuality. URL: reporting.sbb.ch.
Shams, Kollol, Asgari, Hamidreza, Jin, Xia, 2017. Valuation of travel time reliability in freight transportation: A review and meta-analysis of stated preference studies. Transp. Res. Part A: Policy Pract. 102, 228–243. http://dx.doi.org/10.1016/j.tra.2016.08.001.
Sipilä, Hans, 2015. Simulation of rail traffic: Methods for timetable construction, delay modeling and infrastructure evaluation: Diss. (sammanfattning)–Stockholm: Kungliga Tekniska högskolan, 2015. TRITA-TSC-PHD, vol. 15:001, Architecture and the Built Environmetn KTH Royal Institute of Technology, Stockholm, URL: http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-168032.
Small, Kenneth A., 1982. The scheduling of consumer activities: Work trips. Am. Econ. Rev.
U.S. Department of Transportation Federal Railroad Administration, 2013. Reliability and on-time performance. URL: https://railroads.fra.dot.gov/sites/fra.dot.gov/files/fra_net/3260/FRA%20Reliability%20and%20On-Time%20Performance_061713.pdf.
Vromans, Michiel Johannes Cornelius Maria, 2005. Reliability of railway systems:=Betrouwbaarheid van spoorwegsystemen: Zugl.: Rotterdam, Erasmus Univ., Diss., 2005. In: ERIM Ph.D. Series Research in Management, vol. 62, Erasmus Research Inst. of Management (ERIM), Rotterdam.
Vromans, Michiel J.C.M., Dekker, Rommert, Kroon, Leo G., 2006. Reliability and heterogeneity of railway services. European J. Oper. Res. 172 (2), 647–665. http://dx.doi.org/10.1016/j.ejor.2004.10.010.
Widmann, Philipp, Weik, Norman, 2024. Robust Routing in Railway Stations – Generalizing and Assessing Robustness Concepts Using Absorption Graphs. http://dx.doi.org/10.2139/ssrn.5000863.
Zhang, Rong, Jian, Wenliang, Tavasszy, Lóránt, 2018. Estimation of network level benefits of reliability improvements in intermodal freight transport. Res. Transp. Econ. 70, 1–8. http://dx.doi.org/10.1016/j.retrec.2018.09.002.
Zhang, Rong, Li, Lu, Jian, Wenliang, 2019. Reliability analysis on railway transport chain. Int. J. Transp. Sci. Technol. 8 (2), 192–201. http://dx.doi.org/10.1016/j.ijtst.2018.11.004.
Zhang, Ping, Yang, Xin, Wu, Jianjun, Sun, Huijun, Wei, Yun, Gao, Ziyou, 2023a. Coupling analysis of passenger and train flows for a large-scale urban rail transit system. Front. Eng. Manag. 10 (2), 250–261. http://dx.doi.org/10.1007/s42524-021-0180-2.
Zhang, Bojian, Zhang, Yongxiang, D'Ariano, Andrea, Bosi, Tommaso, Lu, Gongyuan, Peng, Qiyuan, 2023b. Optimal platforming, routing, and scheduling of trains and locomotives in a rail passenger station yard. Transp. Res. Part C: Emerg. Technol. 152, 104160. http://dx.doi.org/10.1016/j.trc.2023.104160.