applied sciences

MDPI

Article

Investigation of Railway Network Capacity by Means of Dynamic Flows

Dominik Nikolayzik, Maren Maus * and Nils Nießen

Institute of Transport Science, RWTH Aachen University, Mies-van-der-Rohe-Straße 1, 52074 Aachen, Germany; dominik.nikolayzik@rwth-aachen.de (D.N.)
* Correspondence: maus@via.rwth-aachen.de

Abstract: Capacity calculations are essential for the long-term planning of railway infrastructure. Many of the methods currently used in practice calculate characteristic capacity values separately for the single elements (mainly lines and nodes) of a railway network. Approaches that consider the entire network to account for interactions between the elements often rely on many assumptions, which renders a direct practical application difficult. This paper therefore introduces a linear optimization model that comprises railway-specific constraints such as minimum headway times, line capacities, and route conflicts. Under the consideration of these constraints, the developed model permits the calculation of a network-wide capacity. The model is validated on a sample network that is based on a real network of the German railway infrastructure.

Keywords: capacity; railway networks; network effects; optimization; dynamic multicommodity flows; long-term planning

check for updates

Citation: Nikolayzik, D.; Maus, M.; Nießen, N. Investigation of Railway Network Capacity by Means of Dynamic Flows. *Appl. Sci.* **2024**, *14*, 8233. https://doi.org/10.3390/app14188233

Academic Editors: Valerio De Martinis and Raimond Matthias Wüst

Received: 30 July 2024
Revised: 5 September 2024
Accepted: 9 September 2024
Published: 12 September 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Ensuring climate-friendly and sustainable mobility requires strong traffic growth on the railways. However, in many places, numerous line sections and nodes are already operating at high capacity today. Therefore, the railway infrastructure must be optimally used, upgraded, and expanded. Future bottlenecks in the network must be detected at an early stage to derive expansion and new construction measures. Railway capacity studies are a means of identifying existing bottlenecks on the one hand and making a statement about feasible train numbers before the realization of a measure on the other hand. Capacity analyses are thus a key component for the planning and dimensioning of railway infrastructure.

Depending on the investigated problem, different capacity calculation methods are applied. The internationally widely applied UIC Code 406 [1] introduced by the International Union of Railways (UIC) determines the capacity usage of a specific timetable and is mostly limited to lines [2]. Simulation methods are also based on a specific timetable while being mainly used for the identification of bottlenecks, the assessment of timetable stability and thus for the comparison of different timetable or infrastructure options [3]. They model interactions between different infrastructure elements and are hence also suitable to analyze railway networks. Since they rely on microscopic infrastructure data and are therefore very complex and computationally time-consuming, the network size that can be examined is limited.

Analytical methods offer a possibility for capacity determination that is independent of a concrete timetable. These methods are mainly used for the assessment of different infrastructure and operating program variants and have short computing times due to being based on rough information on the operating program instead of concrete schedules [4]. Analytical methods are hence especially suitable for long-term infrastructure planning, which is the application area this paper focuses on.

Recognized analytical capacity calculation methods, however, determine capacities for individual building blocks such as line sections, track groups, and route nodes separately, as well as according to different quality standards. As a result, the respective capacities are not comparable with one another and do not sufficiently consider interactions between the single elements. However, since the capacity of an individual element changes depending on the surrounding elements, it is of particular interest for long-term infrastructure planning to make statements about the interaction of elements in larger contiguous areas. Simulations as well as queueing theory, which is used as the main tool in analytical capacity assessments, quickly reach their computational limits when considering networks. Therefore, most alternative approaches to capacity determination in networks make use of other techniques such as very macroscopic mathematical model descriptions or the formulation of optimization models that are solved numerically. Even though these models are more suitable for network descriptions than models based on queueing theory, their complexity also increases rapidly. For this reason, current optimization approaches are largely based on strong simplifications or consider only the details that are relevant regarding the specifically investigated problem.

In this paper, an optimization model that aims at determining the usable line capacities in a railway network under consideration of influences by surrounding infrastructure elements and of railway-specific details is presented. These details include, for example, the modeling of minimum headway times on lines and route conflicts in nodes. Furthermore, the model ensures that the load on each line can maintain given quality values through the application of the STRELE-formula [5,6]. This practice guarantees that the developed model is in line with state-of-the-art methods in Germany and therefore prepares for a potential practical implementation.

To introduce the model for capacity determination and discuss its contribution, this paper is structured as follows: First, an overview of the currently existing approaches for capacity calculation is given in Section 2, before the model description and the details of the mathematical formulation are presented in Section 3. In Section 4, the application of the model to a sample network is described, and the results obtained are discussed. Finally, a summary as well as suggestions for future research are provided in Section 5.

2. Literature Review

This section provides an overview of different capacity calculation methods, where first a summary of the state-of-the-art methods for determining the capacity of individual elements is given. Further, it is described which approaches have been pursued to calculate a network-wide capacity. The focus is on analytical methods and optimization approaches relying on methods from flow theory, while a complete overview of methods for capacity analysis in railways can be found in [7].

One widely used method to determine capacity consumption on lines is the UIC Code 406 compression method [1], which was developed with the aim of standardizing capacity assessment across different countries. It is applied to a specific timetable and based on the occupancy rate, which is obtained from the total occupancy derived via the compression of the single train rides' blocking time stairways in relation to the investigation period. To ensure a certain quality, the UIC defined permissible standard values for the occupancy rate depending on the type and usage of the line. For the assessment of node capacity, however, the method is not unrestrictedly suitable [2].

Analytical methods are often applied to determine capacities independently of an actual timetable. These methods mostly rely on the division of the infrastructure into nodes and lines and are based on queueing theory. In the case of nodes, the corresponding queueing system is solved for the scheduled waiting times occurring in the investigated infrastructure. For lines, in Germany, unscheduled waiting times are determined using the STRELE-formula [5,6]. In both cases, the resulting waiting times serve as quality criteria for which a defined quality standard can be applied to relate to the optimal number of trains, i.e., the infrastructure element's capacity.

Since the respective capacities for nodes and lines are not comparable and do not sufficiently consider interactions between the single elements, the question arises whether capacities can also be determined network-wide to be used as a basis for long-term infrastructure planning. There are already a few approaches for deriving such a capacity, selected ones of which are presented below.

In [8], queueing theory is applied to calculate unscheduled waiting times in a network, for example. However, the calculations assume simplified station layouts and there is no possibility for trains to pass each other. In [9], the methods for lines and nodes described above are modified such that the resulting capacities are approximately comparable. Subsequently, the train rides in the network are iteratively optimized based on the calculated capacities for the single elements. Another approach that explicitly includes interactions between the individual infrastructure elements on a corridor is presented in [10]. Here, individual infrastructure sections are modeled by queueing systems, the throughput for a corridor is calculated, and the influence of travel time variability on it is examined. Due to the quickly increasing complexity of the problem, the approach is only suited for corridors, though, and cannot be applied to branched networks. Also, in [11], interactions are considered by investigating the influence of junctions on the capacity of lines by means of Markov models. Within the scope of the analysis, different operating strategies are mapped and compared with each other. This approach is extended in [12] to include the state of the infrastructure. In [13], the queueing approach is also extended to calculate capacity indicators for networks. Also, for the model developed in this approach, drawbacks consist in limited network sizes and a substantial generalization of the underlying infrastructure models, for example.

In addition to analytical calculation methods, there are approaches that rely on the formulation of an optimization model to calculate a network-wide capacity. Here, the idea is that calculating a network capacity can be traced back to calculating a maximum flow in a restricted flow network. Network flow problems in general are a well-known class of combinatorial optimization problems. Within such a problem, a so-called flow network is used to map the flow of objects such as vehicles or goods in a network consisting of vertices and directed arcs. There are numerous types and applications of network flow problems. In the most general maximum flow problem, the goal is to transport as many flow units as possible from one certain vertex to another while capacity restrictions on the arcs are respected and flow conservation in the network holds. Applications of maximum flow problems range from capacity assessments (e.g., of an oil pipeline network or an urban street network) to more abstract issues like scheduling or matching problems. There are excellent surveys in the literature of the various formulations and applications of network flows. Two examples are given by [14] and [15]. The first reference especially examines the vast area of static network flows, while the latter focuses on the topic of dynamic network flows.

To calculate the network capacity using network flows in the railway context, the underlying flow network is defined in such a way that vertices correspond to the stations or junctions, and arcs represent the line sections in between.

An important and practice-oriented part of the flow network in the railway context is the existence of travel times assigned to the arcs. Since the travel time of an arc depends on which train is running on it, another important aspect of the study is the distinction between different trains. To map these properties, for the application in the railway context, the standard maximum flow problem is extended to a dynamic multicommodity maximum flow problem.

Among the previous applications of dynamic multicommodity flow problems or of optimization methods in general, there are examples where capacity studies in railway networks have been conducted. The most relevant approaches for this study are described in more detail below.

In [16], a study on the topological features of transportation networks is conducted. Since insights into network structure provide clear indications of flow properties, this approach establishes a solid graph theoretical foundation for flow analysis.

A direct application of flow analysis to a railway network is presented in [17]. It estimates the network capacity by finding the maximum flow through the network applying the Ford–Fulkerson [18] and Gomory–Hu methods. However, a notable limitation is the lack of accounting for time dependency such as travel times on the edges.

In [19], a model is developed in the form of an integer linear program whose binary decision variables are defined based on the arcs of a time-expanded network. If the value of a variable for a specific line at a specific time or for a certain station at a certain time equals one, then a train is running on this line or stopping at this station at that time. Maximizing the sum over these variables calculates the maximum number of trains that can run on a sequence of line sections with stations in between. In this way, the capacity of an entire railway line is examined. Various limiting aspects are considered by the restrictions defined in the linear program. These include, for example, that a certain train mix must be maintained on the line, and that the number of station tracks that can be used in parallel by stopping trains is never exceeded.

The methodical approach of [20] is similar. Again, the capacity is calculated for a single railway line, which consists of a sequence of line sections with stations in between. However, compared to [19], further railway-specific restrictions are considered. Minimum headway times applicable to the line sections, and the minimum and maximum dwell times of the trains in the stations are modeled as additional influencing factors on the capacity. The level of the calculated capacity is also affected by the specification of temporary disruptions on individual line sections for specific time intervals. In addition, each train is assigned an earliest departure time and a latest arrival time in a discrete time horizon. Further binary variables are used to depict whether a train reaches its destination within the specified time horizon or not. As part of the model's constraints, it is required that a certain minimum percentage of trains arrive at their destination on time. Finally, it is compulsory that freight trains are only scheduled up to a certain maximum percentage of the train mix. Similar to the previously described approach, one main drawback of the method is that it is only applicable to railway lines and does not extend to more complex and branched networks.

Since requests for freight train paths are often received at very short notice, a model with which as many train paths for freight traffic as possible can be planned ad hoc into an existing timetable is developed in [21]. A larger network area is investigated, which is modeled using a time-expanded network. The resulting optimized timetable can then be used to determine a maximum number of trains. The results are strongly dependent on the existing timetable. An investigation based on the same methodology is presented in [22]. Here, too, the usable residual capacities in an existing timetable are determined for a network. First, the number of trains is maximized, and, in a further step, the travel times of the trains are minimized. The capacity of the network can also be inferred from the optimized train journeys. It should be noted, though, that the results are only valid for the respective input timetable, meaning that the method is not suitable for planning processes.

In [23], the authors concentrate on the differences between freight and passenger traffic from another perspective. They especially investigate the effect of traffic separation on capacity. Therefore, they optimize train flows, i.e., routes are not fixed and minimize different objective functions such as travel time or energy intensity.

Another capacity study that takes infrastructure as well as vehicle circulation into account using a hybrid time-space network is given in [24]. Capacity is defined as maximal transportation, which does include the number of passengers instead of absolute train numbers with the result of a saturated timetable.

The approach described in [25] focuses on the differences in train properties such as different speed or stopping patterns. It uses multiobjective optimization to create a saturated timetable whereas the application of the solutions in practice requires further research according to the authors.

Also [26] includes different properties of trains. For fixed sets of trains, the time that is needed for all of them to pass the network is calculated. Hereby, different sequences within the set are considered and capacity of a set is defined by the time in which a certain percentage of all sequences can pass the network. This hence makes the approach timetable independent with the limitation that the sets are fixed beforehand.

In summary, the methods currently used in practice to determine capacity are very much geared to the subdivision of the railway network into its single elements. The capacity parameters used to derive the number of trains that can be operated differ for the individual elements. To include interactions between elements in capacity assessments, several studies, a selection of which is described above, present approaches that consider longer line sections or the network as a whole. However, these methods often rely on many assumptions and were mostly used to investigate specific problems. For example, some of the methods are limited to sequences of line sections and do not extend to branched networks. Others are based and very much dependent on a specific input timetable or relate to a capacity definition that is based on passenger numbers. They are therefore only suitable to a limited extent as a method for calculating a generally defined capacity in the context of long-term infrastructure planning. This paper, in contrast, introduces an optimization model to calculate a network-wide capacity that is independent of an actual timetable and includes a realistic representation of railway-specific constraints. These constraints concentrate on capacity determining properties. The objective of the optimization is to find the maximum number of trains that can run through the network. The capacity term thereby does not refer to this single number for the network. Instead, the resulting capacities for the single line sections, i.e., the extent to which they can use their capacities under the consideration of the surrounding network is analyzed. A further aim is the facilitation of an application in practice that fits in with the state-of-the-art methods, which is why the model is tested on a sample network based on a real network in Germany.

3. Optimization Model for Determining a Network-Wide Capacity

3.1. Modeling of the Railway Network

In order to investigate the capacity of a railway network using network flows, we define a directed graph $G = (V, A)$ with vertices V and arcs A. Vertices may represent either railway stations or junctions, whereas arcs depict the line sections between the stations. Each arc that connects two vertices $i, j \in V$ is given as tuple $(i, j) \in A$. Because every line can be operated in either direction, we always define both a forward arc $(i, j) \in A$ and a backward arc $(j, i) \in A$ for each connection of two vertices $i, j \in V : i \neq j$. This applies to both single-track and double-track line sections. The capacity of single-track sections is represented in the model under the assumption that only one train can run on the section at any time. For this assumption to have a realistic effect on the investigation, the modeled network should therefore only contain shorter single-track line sections. Lines that have only minor connection points to the rest of the network can be neglected.

The network G is divided into two subsets: the core network G_c on the one hand and the so-called linking lines G_l on the other hand. The network capacity defined in this paper is calculated for the core network $G_c = (V_c, A_c)$. The linking lines $G_l = (V_l, A_l)$ comprise all line sections $A_l \subset A$ entering or leaving the core network as well as virtual vertices $V_l \subset V$ at the end of these line sections. They are defined to connect the evaluated core network to the surrounding railway network to depict interactions and dependencies to the entire system.

3.2. Input Sets and Parameters of the Model

3.2.1. Different Trains and Train-Related Parameters

In the railway context of this study, we consider different trains denoted by $k \in K$. K is the total set of trains, which is split into the disjoint subsets K_{sch} and K_{pot}. K_{sch} is defined as the set of trains that are scheduled by the model in any case, while K_{pot} describes the set of trains that can potentially be routed in addition. This division is made to have a fixed

number of scheduled trains on the one hand and to allow for the possibility of saturation by additional trains for capacity calculation on the other hand. To determine the capacity of the core network, the model developed in this paper therefore aims to schedule as many additional trains from the set K_{pot} as possible. Note that the presented optimization model only leads to a feasible solution if K_{sch} does not yet contain too many trains, i.e., if residual capacities are available.

Each train $k \in K$ is assigned a route number $n(k)$, an origin vertex $o(k) \in V$, a destination vertex $d(k) \in V$, and its route through the network given as set of arcs $A^k \subset A$ and set of vertices $V^k \subset V$. Each train is assigned a route through the network that it is permitted to travel, for example with regard to the permissible clearance gauge as well as length and weight restrictions. We also introduce the set N of occurring route numbers ($N := \{n(k) \mid k \in K\}$), and, for each route number $n \in N$, the set K_n, which comprises all trains that share route number n ($K_n := \{k \in K \mid n(k) = n\}$). Trains that share the identical route number are trains of the same type that are assigned the same origin and destination vertices, and the same route through the network—they only differ in their given index k (which can also be interpreted as train identification number). To allow the result of the optimization model to contain timetable-like structures, each route number $n \in N$ (for which trains are defined in the set K_{sch}) is assigned a number f_n that specifies how often trains of this route number should run per hour. Note that the total number of trains of route number n defined in the set K_{sch} must equal the number f_n multiplied by the number of hours covered by the time horizon of the investigation.

Furthermore, for each train $k \in K$, minimum and maximum dwell times ($w^k_{min,\,i}$, $w^k_{max,\,i}$) are specified at each station $i \in V^k$ of its route. By setting the maximum dwell time to zero, a train can be forced not to stop at a station. In contrast, a train can be forced to stop at a station by setting the minimum dwell time to at least one.

In addition, each train $k \in K$ is assigned an individual travel time $\tau^k_{ij} \in \mathbb{N}$ on each arc $(i, j) \in A^k$ of its route. Since each train can have its own travel time on each arc of the network, the model can consider trains with different characteristics. Travel times on linking lines $(i, j) \in A_l$ are zero for all trains $k \in K$.

3.2.2. Network-Related Parameters

Since there is no established method for determining node capacities (cf. Section 1), the railway-specific dependencies in nodes are approximated using two different types of constraints in the model. Each vertex of the core network $i \in V_c$ is assigned a capacity $\kappa_i \in \mathbb{N}_0$ and a set of route conflicts C_i, which can also be empty. The capacity describes the maximum number of available tracks in a station that can be occupied by stopping trains at the same time. When determining this number, it must be ensured that at least two tracks (one in each direction) are available for trains that pass the station without stopping (e.g. for the purpose of an overtaking), if applicable. The set of route conflicts contains individual route conflicts, each of which applies to all trains that belong to the respective route number. One conflict $c \in C_i$ is structured as $c = ((j,\, n,\, \mu),\, (j',\, n',\, \mu'))$, where j and j' are vertices adjacent to i, n and n' are two route numbers and μ and μ' denote station arrivals or departures ($\mu,\, \mu' \in \{'a',\, 'd'\}$). For example, a conflict that is given as $((p,\, P4,\, 'd'),\, (q,\, F2,\, 'a'))$ indicates that all trains belonging to route number $P4$ ($k \in K_{P4}$) departing from station i towards station p are in conflict with all trains that belong to route number $F2$ ($k' \in K_{F2}$) arriving in station i from station q. In a feasible solution of the model (feasible schedule), the routes of two trains k and k' of each possible train pair $(k,\, k') \in K_{P4} \times K_{F2}$ must be set at a certain minimum time interval in station i.

For the arcs of the core network $(i, j) \in A_c$ the parameter $s_{ij} \in \{0, 1\}$ indicates whether the associated line section is a single-track ($s_{ij} = 1$) or double-track line section ($s_{ij} = 0$). For double-track line sections, an average minimum headway time $\overline{h}_{ij}$ plus a minimum buffer time b_{ij}, and an hourly capacity u_{ij} are defined. For single-track line sections, an hourly capacity u_{ij}, which describes the maximum number of trains that can enter the line section from both directions within an hour is specified. Furthermore, each arc

$(i,j) \in A_c$ is assigned a set of trains K_{ij} whose routes contain the arc (i,j), i.e., all trains that will use this arc within the time horizon in case they are scheduled by the model ($K_{ij} := \left\{k \in K \;\middle|\; (i,j) \in A^k\right\}$).

3.2.3. Time-Expanded Network

Before the model can be formulated in detail, the time component is first added by initializing the time-expanded network $G^T = (V^T, A^T)$, which represents the investigated network. The time horizon T, which is needed to construct the time-expanded network, is split into discrete time steps $t \in \{0,1,\ldots,T\}$. Some of the constraints of the model presented in Section 3.3.3 require that minute steps are used for the discrete time steps and that the selected time horizon covers a full hour period.

For each time step there is a copy of the original set of vertices V such that a vertex in the time-expanded network is given as tuple $(i,t) \in V^T$, which represents station (or junction) $i \in V$ at time $t \in \{0,1,\ldots,T\}$. Furthermore, for every vertex $i \in V$ there is a source (i,α), in which a train starting at node i begins its journey through the network, and a sink (i,ω), in which a train ending in node i completes its journey. These vertices allow for individual start and end times of the train paths.

The node set of the network is thus given by:

$$\begin{aligned} V^T := \; & \{(i,t) \mid i \in V,\, t \in \{0,1,\ldots,T\}\} \\ & \cup \{(i,\,\alpha) \mid i \in V\} \\ & \cup \{(i,\omega) \mid i \in V\} \end{aligned} \tag{1}$$

The arc set is defined as follows:

$$\begin{aligned} A^T := & \left\{\left((i,t),\left(j,t+\tau_{ij}^k\right)\right) \;\middle|\; (i,j) \in A, t \in \left\{0,1,\ldots,T-\tau_{ij}^k\right\}\right\} \\ & \cup\{((i,t),(i,t+1))|i \in V_c, t \in \{0,1,\ldots,T-1\}\} \\ & \cup\{((i,\alpha),(i,t))|i \in \mathrm{V}, t \in \{0,1,\ldots,T\}\} \\ & \cup\{((i,t),(i,\omega))|i \in \mathrm{V}, t \in \{0,1,\ldots,T\}\} \end{aligned} \tag{2}$$

There are three different kinds of arcs defined in the time-expanded network G^T. The first subset describes travel arcs, which enable the actual movement of a train k between two vertices i and j at some point in time t. The second subset contains holdover arcs which are defined for every vertex of the core network ($i \in V_c$) to depict train stops. Lastly, to connect the sources and sinks to their corresponding vertices, auxiliary arcs are defined for each vertex $i \in V$ and time step $t \in \{0,1,\ldots,T\}$.

Figure 1 shows an exemplary railway network that forms the basis for the time-expanded network in Figure 2. This example is to illustrate the structure of the time-expanded network in the model.

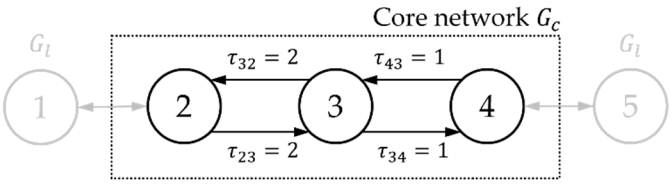


Figure 1. Exemplary railway network consisting of two line sections. The core network G_c is highlighted by the dotted box. The linking lines G_l (in gray color) contain the arcs $(1,2)$, $(2,1)$, $(4,5)$ and $(5,4)$ as well as the virtual vertices 1 and 5. Uniform travel times τ_{ij} are denoted on the arcs.

The colored parts in Figure 2 highlight different possible train journeys, with one color corresponding to one train path. The orange train no. 1 (k = 1), for instance, enters the core network at station 2 at time step 1. It stops there for one time step until it leaves for station 3 where it arrives at time step 4. After another dwell time of two time steps, it is

parked on a siding in station 3. The green train no. 4 (k = 4) enters the core network at station 4 at time step 0 and continues its journey towards station 3 right away. In station 3, it has a dwell time of one time step before it departs for station 2 where it directly leaves the core network towards 1.

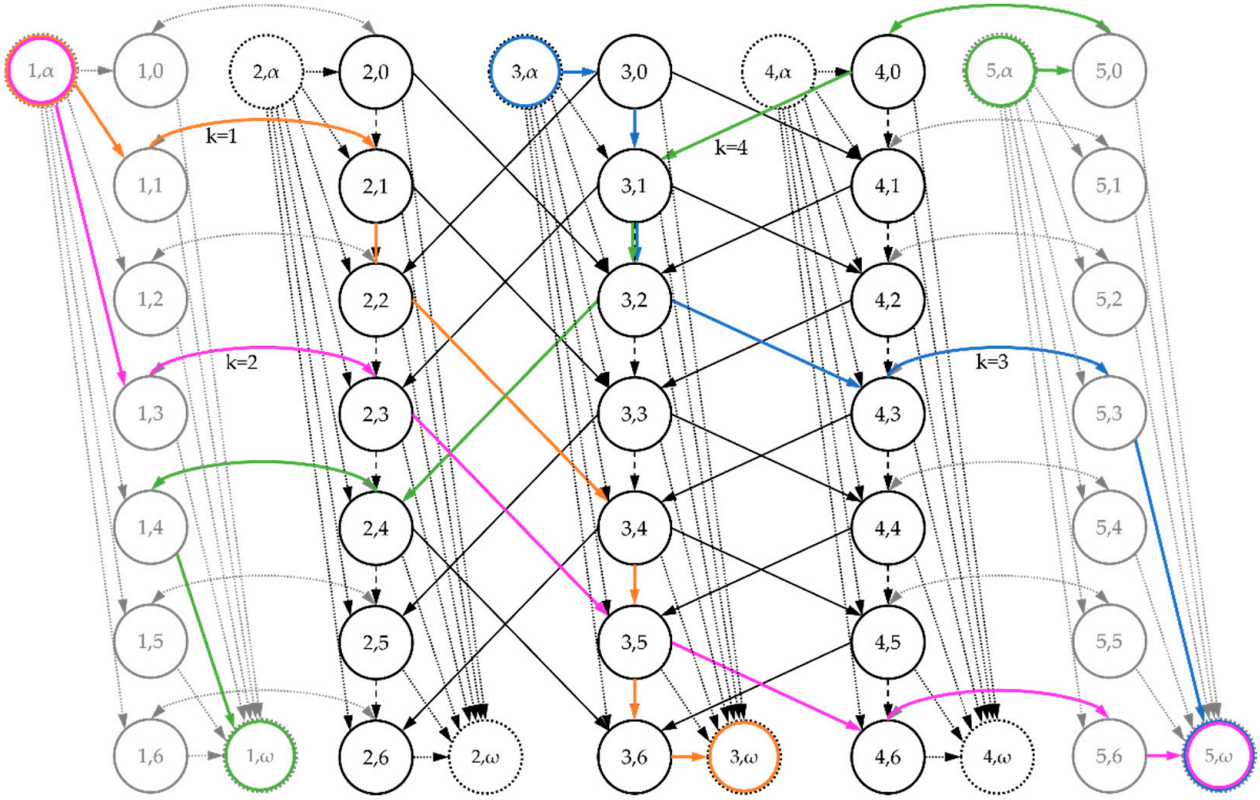


Figure 2. The time-expanded network corresponding to the railway network in Figure 1 with time horizon $T = 6$. Travel arcs are defined according to their assigned travel times and depicted with solid lines. Holdover arcs plotted with dashed lines connect the time copies of the core network's vertices. Auxiliary arcs are illustrated with dotted lines.

Note that in this graphical example, route conflicts and minimum headway times are not considered. For example, if arc $(2,3)$ is assigned an average minimum headway time of $\overline{h}_{23} = 3$ and a minimum required buffer time of $b_{23} = 1$, the model uses appropriate constraints to ensure that in a feasible solution there is a minimum time distance of four time steps between two train journeys on the arc $(2,3)$. So, in this graphical example, it would not be possible for train no. 2 (k = 2) to follow train no. 1 (k = 1) on the line section between stations 2 and 3 directly in the next time step.

Table 1 summarizes all parameters that are used in the developed model.

Table 1. Overview of the sets and parameters used in the model.

Symbol	Meaning
G	Railway network ($G = (V, A) = G_c \cup G_l$)
G_c	Investigated core network ($G_c = (V_c, A_c)$)
G_l	Linking lines ($G_l = (V_l, A_l)$)
G^T	Time-expanded network ($G^T = (V^T, A^T)$)
V , V_c , V_l , V^T	Set of vertices in G, G_c, G_l, G^T
A, A_c, A_l, A^T	Set of arcs in G, G_c, G_l, G^T

Table 1. *Cont.*

Symbol	Meaning
T	Time horizon of the examination
$\delta^+((i,t))$, $\delta^-((i,t))$	Set of outgoing/ incoming arcs of vertex (i,t) in G^T
(i,α)	Vertex in G^T called source of vertex $i \in V$ ($\delta^-((i,\alpha)) = \varnothing$)
(i,ω)	Vertex in G^T called sink of vertex $i \in V$ ($\delta^+((i,\omega)) = \varnothing$)
κ_i	Capacity of vertex $i \in V_c$ (number of available tracks for stopping trains)
C_i	Set of (potentially occurring) route conflicts between trains arriving and/or departing in vertex $i \in V_c$; One conflict $c \in C_i$ is given as tuple $((j,\ n,\ \mu),\ (j',\ n',\ \mu'))$.
$\mu \in \{'d','a'\}$	Parameter to indicate station departures ($'d'$) and arrivals ($'a'$)
K	Set of trains ($K = K_{sch} \cup K_{pot}$)
K_{sch}	Set of trains that must be scheduled by the model in any case
K_{pot}	Set of trains to be potentially planned in addition
$n(k)$	Route number of train k
N	Set of all occurring route numbers ($N := \{n(k) \mid k \in K\}$)
K_n	Set of all trains that have route number n ($K_n := \{k \in K \mid n(k) = n\}$)
$o(k)$, $d(k)$	Origin/ destination vertex of train k ($o(k),\ d(k) \in V$)
f_n	Departure frequency (trains per hour) of route number n
V^k, A^k	Set of vertices/ arcs that train k must pass throughout its journey iff scheduled ($V^k \subset V$, $A^k \subset A$)
$A^{T,k}$	Set of relevant arcs for train k in G^T ($A^{T,k} \subset A^T$), i.e., arcs corresponding to its assigned route (derived from V^k and A^k)
$w^k_{min,i}$, $w^k_{max,i}$	Minimum and maximum dwell times of train k in station $i \in V^k$
τ^k_{ij}	Travel time of train k on arc $(i,j) \in A^k$
K_{ij}	Set of trains whose routes contain arc (i,j) ($K_{ij} := \left\{k \in K \,\middle|\, (i,j) \in A^k\right\}$)
$s_{ij} \in \{0,1\}$	Information if arc $(i,j) \in A_c$ corresponds to a single-track ($s_{ij} = 1$) or double-track ($s_{ij} = 0$) line section
$\overline{h}_{ij}$	Average minimum headway time on arc $(i,j) \in A_c$ (only in case $s_{ij} = 0$)
b_{ij}	Minimum required buffer time on arc $(i,j) \in A_c$ (only in case $s_{ij} = 0$)
u_{ij}	Hourly capacity of arc $(i,j) \in A_c$ (in case $s_{ij} = 1$: sum of both directions)

3.3. Mathematical Formulation of the Model

3.3.1. Decision Variables

For every train $k \in K$ and each of its relevant arcs $a \in A^{T,k}$ in the time-expanded network, we define a binary routing variable $x^k_a \in \{0,1\}$ that indicates whether train k uses arc a ($x^k_a = 1$) or not ($x^k_a = 0$). The decision $x^k_a = 1$ for a travel arc of the form $a = \left((i,t), \left(j, t + \tau^k_{ij}\right)\right)$ means that train k leaves station (or junction) i at time t and arrives in j exactly τ^k_{ij} time steps later.

Furthermore, for every train $k \in K$, we define a binary variable $y^k \in \{0,1\}$ that states whether the network's capacity suffices to schedule train k under the valid restrictions within the given time horizon ($y^k = 1$) or not ($y^k = 0$).

3.3.2. Objective Function

This paper's approach to calculate the capacity of the investigated core network is to maximize the sum over all y—variables.

$$\max \quad z = \sum_{k \in K} y^k \tag{3}$$

The optimization model therefore attempts to schedule as many of the defined trains as possible, more precisely all trains in K_{sch} (ensured by appropriate constraints) and as many trains as possible in K_{pot}.

3.3.3. Constraints

Scheduled trains. All trains in K_{sch} are scheduled.

$$\begin{gathered} y^k = 1 \\ \forall\, k \in K_{sch} \end{gathered} \tag{4}$$

Departure frequencies of scheduled trains. For each route number n, for which trains are defined in the set K_{sch}, it is ensured that within each hour of the time horizon, a certain number of trains f_n start their journey at their (common) origin vertex. The formulation of these constraints assume that one time step corresponds to one minute.

$$\begin{gathered} \sum_{k \in K_n \cap K_{sch}} \sum_{t=z \cdot 60}^{(z+1) \cdot 60 - 1} x^k_{((o(k),\alpha),(o(k),t))} = f_n \\ \forall\, n \in N, \forall\, z \in \left\{0, 1, \ldots, \tfrac{T}{60} - 1\right\} \end{gathered} \tag{5}$$

Start and end of train journeys. Each train (scheduled by the model, i.e. $y^k = 1$) starts its journey in the source of its origin vertex and ends it in the sink of its destination vertex.

$$\begin{gathered} \sum_{t=0}^{T} x^k_{((o(k),\alpha),(o(k),t))} = \sum_{t=0}^{T} x^k_{((d(k),t),(d(k),\omega))} = y^k \\ \forall\, k \in K \end{gathered} \tag{6}$$

Flow conservation. Flow conservation holds for every vertex (except for the sources and sinks) and every train in the time-expanded network.

$$\begin{gathered} \sum_{a \in \delta^+((i,t)) \mid a \in A^{T,k}} x^k_a - \sum_{a \in \delta^-((i,t)) \mid a \in A^{T,k}} x^k_a = 0 \\ \forall\, (i,t) \in V^T :\ t \in \{0, 1, \ldots, T\},\ \forall\, k \in K \end{gathered} \tag{7}$$

Train routes through the network. Each (scheduled) train follows its assigned route through the network.

$$\begin{gathered} \sum_{t=0}^{T - \tau^k_{ij}} x^k_{((i,t),(j,t+\tau^k_{ij}))} = y^k \\ \forall\, k \in K,\ \forall\, (i,j) \in A^k \end{gathered} \tag{8}$$

Available station tracks. Station capacities are respected at any time.

$$\begin{gathered} \sum_{k \in K \mid i \in V^k} x^k_{((i,t),(i,t+1))} \leq \kappa_i \\ \forall\, i \in V_c,\ \forall\, t \in \{0, 1, \ldots, T-1\} \end{gathered} \tag{9}$$

Dwell times in stations. Minimum and maximum dwell times of (scheduled) trains in stations are complied with.

$$\begin{gathered} w^k_{min,\,i} \cdot y^k \leq \sum_{t=0}^{T-1} x^k_{((i,t),(i,t+1))} \leq w^k_{max,\,i} \\ \forall\, k \in K,\ \forall\, i \in V^k \end{gathered} \tag{10}$$

Route conflicts in stations. As described in Section 3.2.2, the parameter μ can take the values $'a'$ or $'d'$ depending on whether it indicates a station arrival or departure. Because each route conflict is defined for exactly two route numbers, each of which can be represented in the conflict by an arrival or departure, four cases need to be distinguished in the formulation of the constraints below.

Two station arrivals (route number n from direction j and route number n' from direction j') must be at least Δt time steps apart.

$$\begin{gathered}\sum_{k\in K_n}\sum_{t'=t\mid \tau^k_{ji}\le t'\le T}^{t+\Delta t-1} x^k_{((j,t'-\tau^k_{ji}),(i,t'))}\\ +\sum_{k\in K_{n'}}\sum_{t'=t\mid \tau^k_{j'i}\le t'\le T}^{t+\Delta t-1} x^k_{((j',t'-\tau^k_{j'i}),(i,t'))}\le 1\\ \forall i\in V_c,\ \forall((j,n,\mu),(j',n',\mu'))\in C_i:\ \mu=\mu'=' a',\ \forall t\in\{0,1,\ldots,T\}\end{gathered}\tag{11}$$

One station arrival (route number n from direction j) and one departure (route number n' in direction j') must be at least Δt time steps apart.

$$\begin{gathered}\sum_{k\in K_n}\sum_{t'=t\mid \tau^k_{ji}\le t'\le T}^{t+\Delta t-1} x^k_{((j,t'-\tau^k_{ji}),(i,t'))}\\ +\sum_{k\in K_{n'}}\sum_{t'=t\mid 0\le t'\le T-\tau^k_{ij'}}^{t+\Delta t-1} x^k_{((i,t'),(j',t'+\tau^k_{ij'}))}\le 1\\ \forall i\in V_c,\ \forall((j,n,\mu),(j',n',\mu'))\in C_i:\ \mu='a'\ \wedge\ \mu'='d',\ \forall t\in\{0,1,\ldots,T\}\end{gathered}\tag{12}$$

One station departure (route number n in direction j) and one arrival (route number n' from direction j') must be at least Δt time steps apart.

$$\begin{gathered}\sum_{k\in K_n}\sum_{t'=t\mid 0\le t'\le T-\tau^k_{ij}}^{t+\Delta t-1} x^k_{((i,t'),(j,t'+\tau^k_{ij}))}\\ +\sum_{k\in K_{n'}}\sum_{t'=t\mid \tau^k_{j'i}\le t'\le T}^{t+\Delta t-1} x^k_{((j',t'-\tau^k_{j'i}),(i,t'))}\le 1\\ \forall i\in V_c,\ \forall((j,n,\mu),(j',n',\mu'))\in C_i:\ \mu='d'\ \wedge\ \mu'='a',\ \forall t\in\{0,1,\ldots,T\}\end{gathered}\tag{13}$$

Two station departures (route number n in direction j and route number n' in direction j') must be at least Δt time steps apart.

$$\begin{gathered}\sum_{k\in K_n}\sum_{t'=t\mid 0\le t'\le T-\tau^k_{ij}}^{t+\Delta t-1} x^k_{((i,t'),(j,t'+\tau^k_{ij}))}\\ +\sum_{k\in K_{n'}}\sum_{t'=t\mid 0\le t'\le T-\tau^k_{ij'}}^{t+\Delta t-1} x^k_{((i,t'),(j',t'+\tau^k_{ij'}))}\le 1\\ \forall i\in V_c,\ \forall((j,n,\mu),(j',n',\mu'))\in C_i:\ \mu=\mu'='d',\ \forall t\in\{0,1,\ldots,T\}\end{gathered}\tag{14}$$

Capacity of single-track line sections. Hourly capacity constraints for single-track line sections are respected. The formulation of these constraints assume that one time step corresponds to one minute.

$$\begin{gathered}\sum_{k\in K_{ij}}\sum_{t'=t\mid t'+\tau^k_{ij}\le T}^{t+60-1} x^k_{((i,t'),(j,t'+\tau^k_{ij}))}\\ +\sum_{k\in K_{ji}}\sum_{t'=t\mid t'+\tau^k_{ji}\le T}^{t+60-1} x^k_{((j,t'),(i,t'+\tau^k_{ji}))}\le u_{ij}\\ \forall(i,j)\in A_c:\ s_{ij}=1,\ \forall t\in\{0,1,\ldots,T\}\end{gathered}\tag{15}$$

At most one train travels on a single-track line section (in either direction) at a time.

$$\begin{gathered}\sum_{k\in K_{ij}}\sum_{t'=t-\tau^k_{ij}+1\mid 0\le t'\le T-\tau^k_{ij}}^{t} x^k_{((i,t'),(j,t'+\tau^k_{ij}))}\\ +\sum_{k\in K_{ji}}\sum_{t'=t-\tau^k_{ji}+1\mid 0\le t'\le T-\tau^k_{ji}}^{t} x^k_{((j,t'),(i,t'+\tau^k_{ji}))}\le 1\\ \forall(i,j)\in A_c:\ s_{ij}=1,\ \forall t\in\{0,1,\ldots,T\}\end{gathered}\tag{16}$$

Capacity of double-track line sections. Hourly capacity constraints for double-track line sections are satisfied. The formulation of these constraints assume that one time step corresponds to one minute.

$$\begin{aligned} &\sum_{k \in K_{ij}} \sum_{t'=t \mid t'+\tau_{ij}^k \leq T}^{t+60-1} x^k_{((i,t'),(j,t'+\tau_{ij}^k))} \leq u_{ij} \\ &\forall (i,j) \in A_c : \; s_{ij} = 0, \; \forall t \in \{0,1,\ldots,T\} \end{aligned} \tag{17}$$

At most one train enters a double-track line section in one direction within the interval of the average minimum headway time plus the minimum necessary buffer time.

$$\begin{aligned} &\sum_{k \in K_{ij}} \sum_{t'=t \mid t'+\tau_{ij}^k \leq T}^{t+\bar{h}_{ij}+b_{ij}-1} x^k_{((i,t'),(j,t'+\tau_{ij}^k))} \leq 1 \\ &\forall (i,j) \in A_c : \; s_{ij} = 0, \; \forall t \in \{0,1,\ldots,T\} \end{aligned} \tag{18}$$

4. Exemplary Application to a Sample Network

4.1. Definition of the Sample Network and the Train Routes

To validate the developed optimization model, it is tested using a sample network that is based on a real network in North Rhine-Westphalia, Germany. The core network consists of ten vertices connected by eleven line sections, each of which is depicted by two arcs, as described in Section 3.1. The linking lines contain eight virtual vertices connected to the core network by ten line sections (see Figure 3).

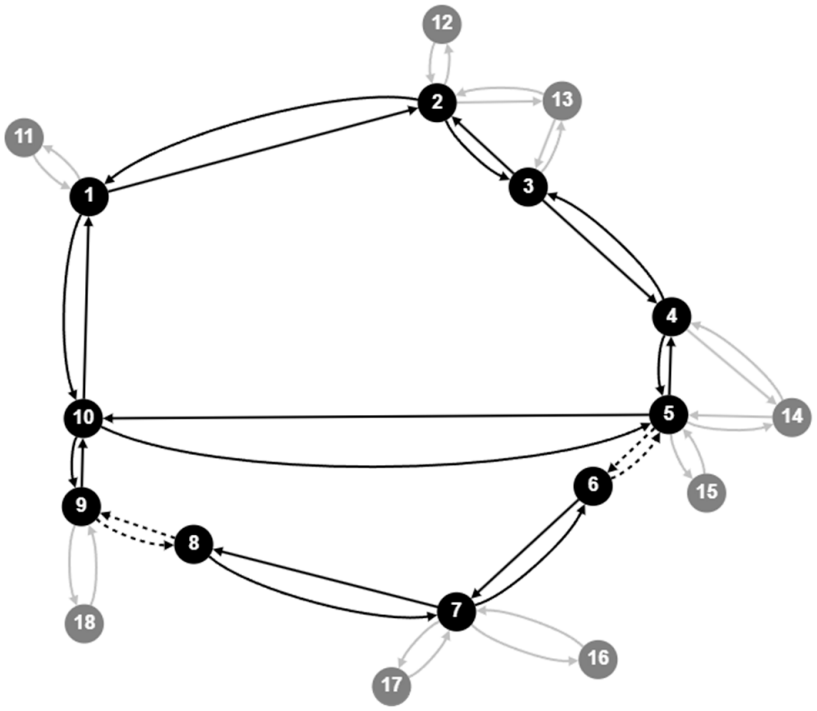


Figure 3. Investigated sample network comprising 18 vertices and 42 arcs in total. The evaluated core network is shown in black, with solid arcs referring to double-track and dashed arcs indicating single-track line sections. The linking lines (also containing the virtual nodes) are depicted in gray color [27].

Apart from two operationally relevant junctions (vertices 3 and 4 in Figure 3), all vertices of the core network correspond to stations, whereas smaller junctions are neglected in favor of computability. Most of the line sections are double-track line sections, but the network also comprises two short single-track line sections, so that all aspects covered by

the model can be tested in the validation process. The distances covered by the different line sections lie between three and fifteen kilometers.

To apply the model to the sample network, it is first necessary to distinguish the different route numbers and, based thereon, to determine the sets of scheduled trains K_{sch} and of potentially to be planned trains K_{pot}.

In the presented case, we consider 13 local passenger and three freight train route numbers in each of the two directions. All the considered route numbers run on at least one arc of the core network. The set of scheduled trains is defined over a specified number of trains per route number and hour. Here, the freight train route numbers are included in K_{sch} once per hour and direction and the passenger route numbers are included according to their frequency in the current timetable. Note that the exact schedule itself is not fixed and only frequencies are extracted from the timetable. The defined trains of a route number can be distributed arbitrarily by the model within the hourly slices of the time horizon according to their frequency.

In the set of potentially to be planned trains K_{pot}, only trains of the freight train route numbers are considered, as this is a realistic scenario in the actual practical application. Usually, timetables for passenger trains are fixed and additional (short-term) slots saturating the timetable are then mostly allocated to freight trains.

For all passenger and freight train route numbers, the departure frequencies and the assigned routes through the sample network of Figure 3 are listed in Table 2.

Table 2. Route numbers of passenger (P) and freight (F) trains with their departure frequencies and assigned routes through the sample network. The route given refers to the route number first in each line. The route is mirrored for the return direction. The letters N, E, S, and W appended to the route numbers are intended to indicate the cardinal points of the associated trains through the sample network.

Route Number	Frequency (per Hour and Direction)	Route
P1E, P1W	Once	18–9–10–5–14
P2N, P2S,	Once	15–5–4–3–2
P3N, P3S, P8N, P8S	Once	16–7–8–9–10
P4N, P4S	Twice	14–4-3–2–12
P5E, P5W	Once	11–1–10–5–14
P6E, P6W, P11E, P11W	Once	10–1–2–13
P7E, P7W	Three times	10–5–14
P9E, P9W	Once	18–9–10–1–2–13
P10E, P10W	Once	18–9–10
P12E, P12W	Once	17–7–6–5–14
P13E, P13W	Once	7–6–5
F1N, F1S	Once (+ potentially to be planned trains)	16–7–8–9–10–1–11
F2N, F2S	Once (+ potentially to be planned trains)	15–5–4–3–13
F3E, F3W	Once (+ potentially to be planned trains)	18–9–10–1–2–13

With the route numbers defined, it is possible to determine the basic load on the individual arcs of the sample network shown in Figure 3. The resulting numbers of trains are listed in Table 3.

Table 3. Initial train numbers on the arcs of the network shown in Figure 3. These numbers define the train mix used to calculate the input parameters for the optimization model.

Arc $(i,j) \in A_c$	Total Number of Trains per Hour	Number of Local Passenger Trains	Number of Freight Trains
(1,2), (2,1)	4	3	1
(2,3), (3,2)	3	3	0
(3,4), (4,3)	4	3	1
(4,5), (5,4)	2	1	1
(5,6), (6,5)	2	2	0
(6,7), (7,6)	2	2	0
(7,8), (8,7)	3	2	1
(8,9), (9,8)	3	2	1
(9,10), (10,9)	7	5	2
(10,5), (5,10)	5	5	0
(10,1), (1,10)	6	4	2

4.2. Calculation of the Input Parameters

The necessary input parameters for the optimization model are determined with LUKS® [28], which is a software tool for microscopic railway operation studies, that, in addition to the analytical capacity investigation of nodes and lines, also enables the simulation of timetable creation and operational processes. The input parameters include the travel times of the different route numbers listed in Table 2, the average minimum headway times between two trains on each line section and the hourly capacity limits for each line section. The minimum headway times as well as the capacity limits of the line sections thereby depend on the train mix. Since the scenario considered in the present study assumes a fixed set of trains that is filled up with freight trains only (cf. Section 4.1), the train mix in the optimization process changes in each iteration (cf. Section 4.3 for a detailed description of the optimization process). An important note in this context is therefore that the calculation of the input parameters is based on the train mix that results from the initially specified operating program through the set K_{sch} (cf. Table 3). Furthermore, since the time resolution here is set to minute steps, all parameters must be rounded to full minutes.

Travel times are calculated for every arc of the core network and each route number running on it. All dynamic vehicle properties are considered. If the resulting travel times exceed full minutes by more than six seconds (0.1 min), they are rounded up to the next full minute. Otherwise, they are rounded down. This approach ensures that the network-wide capacity is rather underestimated than overestimated in the end. Stops for passenger trains are defined in each station of the core network so that the calculated travel times correspond to realistic values for the modeling. No stops are planned for freight trains, but the optimization model allows them to stop in stations if they should be overtaken by passenger trains (for capacity reasons). To consider the time that is necessary for the start-up process and brake application in case of such stops, the travel times of freight trains on each arc of the core network are blanketly increased by one minute.

The use of headway-specific minimum headway times would quickly increase the complexity of the model and thus the runtime. Therefore, average minimum headway times based on the initial train mix (cf. Table 3) are calculated for each arc of the core network. The average minimum headway times are rounded in the same way as the travel times. The minimum buffer times are set to one minute for all arcs of the core network in accordance with DB regulations [29].

The line capacities are determined using the STRELE-formula, which is integrated in the LUKS® software tool, assuming an optimal operational quality according to the standards applied by DB in Germany. The calculation requires, among others, the definition of delay characteristics for the different route numbers that are assumed to take the standard values defined by DB for highly loaded lines. The resulting capacities are rounded down to the next integer. Instead of the STRELE-formula, other methods can also be used to specify an upper limit for the line capacity. For example, a limit for utilization according to UIC Code 406 could be provided. Through iterative addition of train services (cf. procedure described in Section 4.3), where it is checked in each iteration whether this limit can be adhered to or if it is already exceeded by the scheduled trains, the optimal utilization can be determined.

For nodes, a capacity equivalent to that for lines does not exist. Instead, a realistic limit on the number of trains that can pass a node is approximated by including the number of tracks that can be occupied by stopping trains at the same time. On top of that, route conflicts between the different trains are considered. The minimum time interval between two conflicting trains is set to two minutes (see Section 3.2.2 for details).

Moreover, the following assumptions are made for dwell times. Passenger trains have a minimum dwell time of five minutes at their start and destination stations as well as stations where they change the direction of travel. At other planned stops on their routes, they have a minimum dwell time of one minute and a maximum dwell time of two minutes. Additionally, passenger trains are not allowed to stop where they are not planned to stop. For freight trains, there are no restrictions regarding dwell times.

The following Section 4.3 elaborates on how the optimal number of trains is determined, before the results for the network described in Section 4.1 are given in Section 4.4.

4.3. Implementation of an Iterative Optimization Process

One problem in running the optimization model presented in Section 3.3 only once is given in the definition of an appropriate number of trains in the set K_{pot}. If, on the one hand, the number of trains that are defined to run in addition to the trains scheduled in any case is too small, all trains will be routed, but it remains unclear whether the network's capacity is already exhausted or whether additional trains could be scheduled. If, on the other hand, there are too many trains, the optimization process may not terminate within a reasonable time, since the time-expanded network grows rapidly, and with it the number of variables and constraints in the model.

For that reason, an iterative optimization process is developed in this paper. At the beginning, only the trains scheduled in any case (K_{sch}) are defined, and the optimization model is used to check whether a valid schedule exists. In case all trains have been routed, one train is added to the (initially empty) set K_{pot} for each freight train route number. Subsequently, the model is solved again. That means that this time it is checked whether a valid schedule containing both the trains from K_{sch} and the additionally defined trains from K_{pot} can be generated. For each of the following iterations, it is evaluated for which of the freight train route numbers an additional train could be planned in the last iteration. The route numbers for which this is the case are again included in the set K_{pot} by one additional train in the following iteration. Trains that could additionally be planned in one iteration are also scheduled again in the next iterations. On the contrary, the route numbers for which no additional train could be planned in one iteration are excluded for the following iterations. As soon as there are no more trains that can be scheduled for any of the specified additional route numbers, the program terminates. The total number of trains that could be scheduled determines the capacity of the core network.

The iterative optimization process described is depicted in the flow chart of Figure 4.

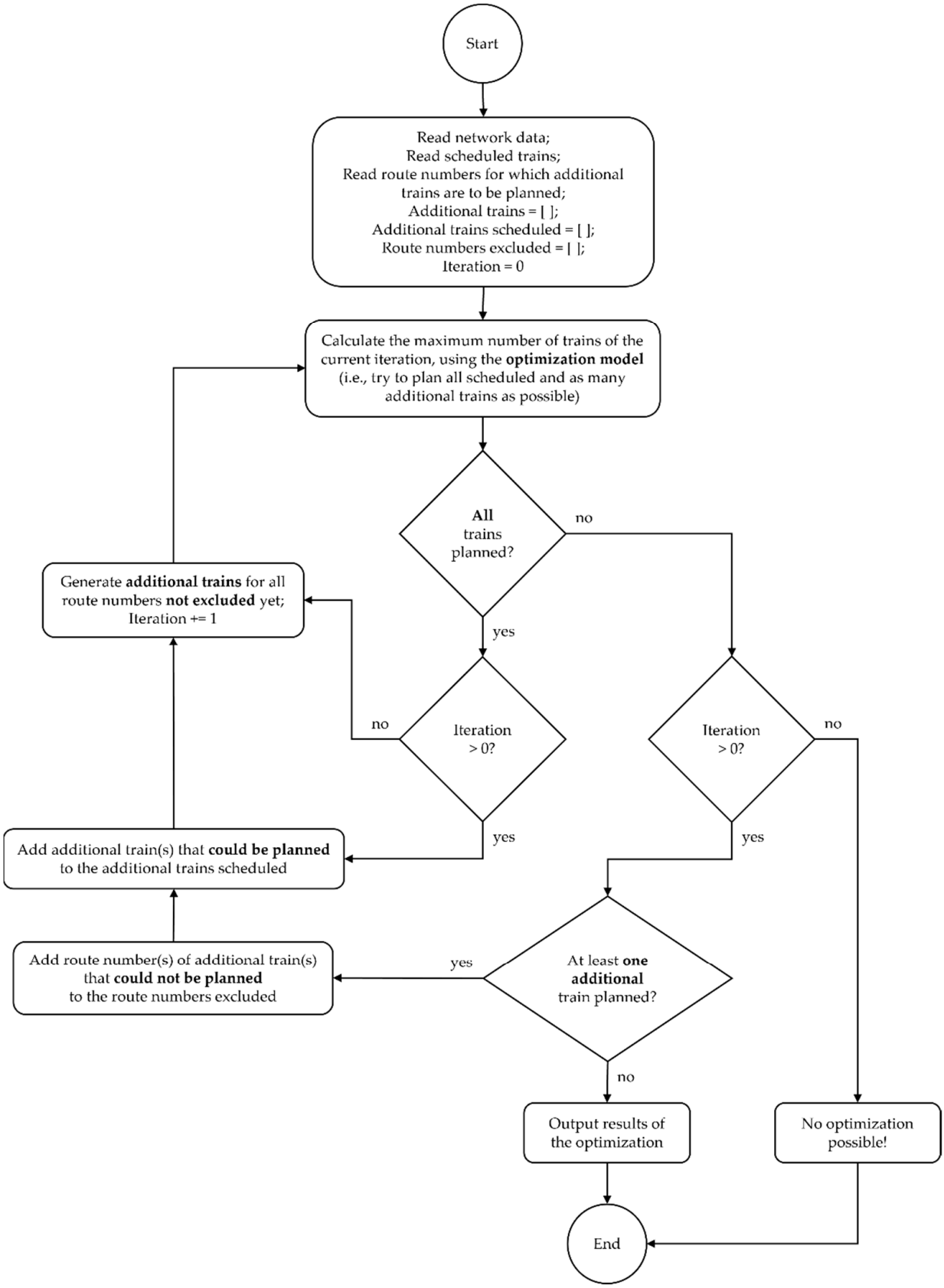


Figure 4. Flow chart of the iterative optimization process applied.

4.4. Results

The optimization model described in the previous sections is implemented in Python 3.9 and solved for the sample network specified in Section 4.1 on the compute cluster of the RWTH Aachen University using Gurobi 9.5. For technical details of the compute cluster see [30]. For the application, a time period of three hours ($T = 180$) is considered. In the following, the results are first analyzed for the entire network as depicted in Figure 3.

For the entire network (see Figure 3), 38 trains of 32 different route numbers are initially scheduled per hour, 6 of which are freight trains. This means that, within the whole investigation period, 114 trains are initially scheduled (cf. Table 3). All freight train route numbers are then considered for the saturation of the network, and in each iteration, it is checked whether another train of each route number can be incorporated. This procedure leads to a total number of 132 trains in the network. Hence, eighteen additional trains that were added in seven iterations could be scheduled.

The runtime of the entire optimization process was approximately 477 min. To achieve this, further accelerating constraints were implemented in the optimization model. Most of these constraints helped the solver make irrelevant but necessary decisions that do not change the objective value of the best possible solution. An example of these constraints is the specification that no trains use holdover arcs of the time-expanded network outside of the core network. One example of constraints that can have a minor impact on the results but contribute significantly to the acceleration of the optimization process, is the specification for trains of passenger route numbers in the set K_{sch} to maintain a minimum time interval of half their frequency at their origin vertices.

Apart from the evaluation regarding the number of totally planned trains, the train numbers can also be evaluated with respect to the single line sections of the considered network. The resulting train numbers can then be compared to the line capacities that were used as input parameters for the model as an upper bound. Figure 5 shows this comparison for the entire network, where the first number indicates the actual number of trains on each line section and the second number indicates the maximum number of trains that could theoretically run on that section according to the calculated line section capacity.

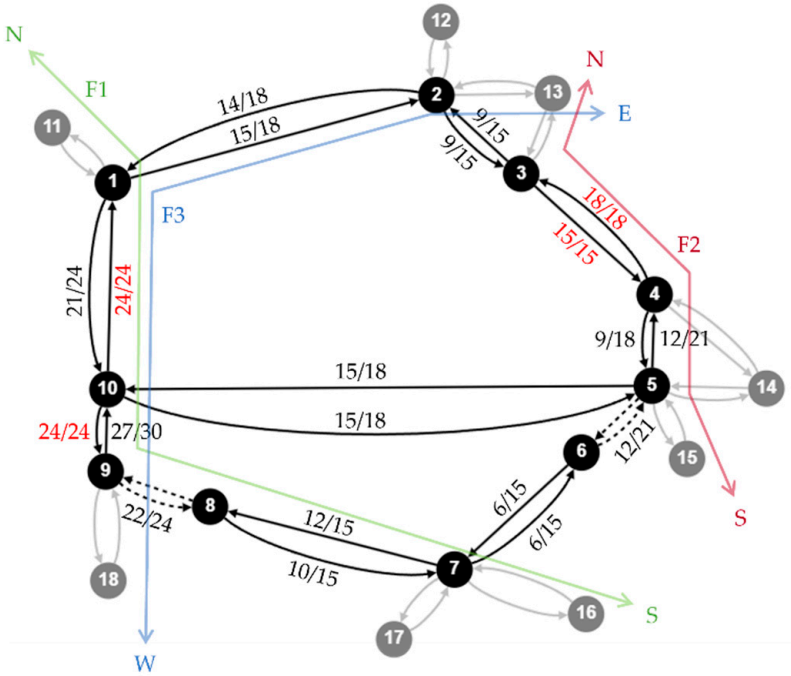


Figure 5. Entire sample network after saturation by the additional freight train route numbers F1, F2 and F3 (in green, dark red and blue color, respectively). Red numbers on the arcs highlight saturated line sections after optimization [27].

To allow for the derivation of network effects, the network is divided into two subnetworks, each of which the analysis is repeated for. The division into the partial networks is illustrated in Figure 6.

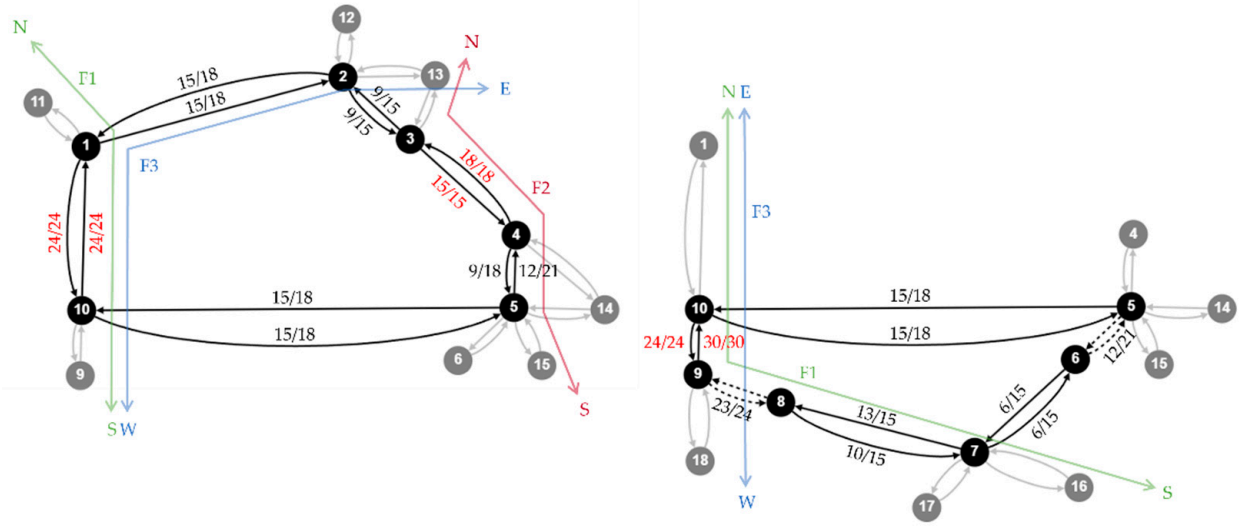


Figure 6. Upper subnetwork (**left** side) and lower subnetwork (**right** side) after saturation in separate analyses. The additional freight train route numbers F1, F2 and F3 are depicted in green, dark red and blue color, respectively. Red numbers on the arcs highlight saturated line sections after optimization [27].

In total, 84 trains were initially planned for the "upper" part of the network and 21 trains were added during the following seven iterations. The runtime of the complete optimization process was approximately 17 min. The subnetwork, consisting of the "lower" part of the entire network, is already saturated after six iterations, and 12 additional trains could be added to the originally planned 78 trains. The runtime of the entire optimization process was approximately 91 min. The load of the respective subnetworks after saturation can also be seen in Figure 6.

In comparison to the analysis of the entire network (cf. Figure 5), it can be seen that in these partial analyses at least as many (and sometimes even more) trains can be scheduled on the individual line sections.

The optimization thus allows the capacities that would be available when looking at a single element to be compared with the actually usable capacity when considering the interaction with surrounding elements. As expected, the usable capacity tends to decrease the larger the network under consideration becomes, as the influences between the elements increase.

5. Conclusions and Discussion

In this paper, an approach to calculate a network-wide capacity was developed. The approach is based on existing optimization methods and considers characteristic capacity values for lines that are determined using established procedures from railway operations science. This ensures that the developed approach is as consistent as possible with existing procedures and quality standards. The derived model maximizes the number of trains in the considered network while ensuring that railway-specific constraints are met. These constraints include, for example, minimum headway times and hourly capacity limits on lines, route conflicts between arriving and departing trains, the number of available tracks, as well as permissible dwell time intervals for trains in stations. As the input of the optimization model, an individual travel time can be specified for each train on each arc of the network; thereby, different train characteristics as well as the maximum permissible speeds of the line sections are reflected in the model.

In addition to the total number of trains determined for the entire network under consideration, the number of trains that can be operated can also be evaluated in relation to individual lines of the network. In this way, the train runs determined for the individual routes can be compared with the capacities that were entered into the model as input parameters.

The application of the model to a sample network illustrated how the model can be used to examine the extent to which the capacities theoretically available on the individual line sections can actually be used, taking network effects into account. The amount by which the usable capacity is smaller also depends on the network size and the utilization of the network at its edges. Therefore, the network boundaries for such a network analysis should be carefully chosen to minimize the interactions that are neglected by the network selection. The exemplary application of the model is intended to show in a comprehensible way that certain sections of a railway network may already be at their capacity limit, while there is still residual capacity on other sections. Such an investigation can therefore be used, especially with regard to economic aspects, to identify crucial bottlenecks in a network and to derive targeted investment decisions that can increase the capacity of the overall system.

The presented study should be considered as a preliminary investigation. Further research is required regarding, for instance, the following aspects: The input parameters of the model are based on the train mix that is given by the initially scheduled trains. However, the train mix actually changes as soon as additional freight trains are added to the network. It should therefore be investigated whether it is possible either to adapt the input parameters in each iteration, or to not exclusively add freight trains but to increase the number of trains in such a way that the initially defined train mix is maintained. Another possibility to extend the model is the inclusion of stochastic travel times to consider deviations from ideal conditions such as adverse weather conditions or time restrictions due to track construction and maintenance. Furthermore, travel times of freight trains should depend on whether a freight train stops or not. One approach would be to define a variable for each freight train that ensures that in the event of an unscheduled stop at a node, the required travel times on the adjacent arcs are increased for this train to adequately model braking and acceleration processes. One more drawback of the presented approach is the rather rough modeling of nodes. It should hence be checked whether the processes in nodes can be modeled in more detail or by node capacities that are derived with a more accurate approach.

A key factor in the capacitive analysis of networks remains that it would be useful to increase the network size compared to the scenario investigated in this study. However, the computing time increases significantly with an increasing network size. Therefore, an important direction for future research is, on the one hand, the development of powerful heuristics. In the best-case scenario, those heuristics can be used to achieve a possibly suboptimal, but nevertheless particularly good result in a significantly reduced computing time. On the other hand, it should be analyzed how the examined network and its associated input parameters can be aggregated into a more macroscopic view, so that fewer variables must be calculated, which also leads to shorter computing times and therefore allows for the examination of larger networks.

Author Contributions: Conceptualization: D.N. and M.M.; methodology: D.N.; software: D.N.; writing—original draft preparation: D.N. and M.M.; writing—review and editing: N.N.; visualization: D.N.; supervision: N.N. All authors have read and agreed to the published version of the manuscript.

Funding: This work is partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—2236/2.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original data presented in the study are openly available in zenodo at https://doi.org/10.5281/zenodo.13304401, accessed on 8 September 2024.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. International Union of Railways. UIC Leaflet 406—Capacity 2013. Available online: https://tamannaei.iut.ac.ir/sites/tamannaei.iut.ac.ir/files/files_course/uic406_2013.pdf (accessed on 8 September 2024).
2. Lindner, T. Applicability of the analytical UIC Code 406 compression method for evaluating line and station capacity. *J. Rail Transp. Plan. Manag.* **2011**, *1*, 49–57. [CrossRef]
3. Watson, R.; Medeossi, G. *Railway Timetabling & Operations. Analysis, Modelling, Optimisation, Performance Evaluation*; Eurail Press: Hamburg, Germany, 2008; Chapter 10: Simulation; pp. 191–213.
4. Nießen, N. *Railway Timetabling & Operations. Analysis, Modelling, Optimisation, Performance Evaluation*; Eurail Press: Hamburg, Germany, 2008; Chapter 6: Queueing; pp. 117–130.
5. Schwanhäußer, W. Die Bemessung der Pufferzeiten im Fahrplangefüge der Eisenbahn. Ph.D. Thesis, RWTH Aachen University, Aachen, Germany, 1974.
6. Weik, N.; Niebel, N.; Nießen, N. Capacity analysis of railway lines in Germany—A rigorous discussion of the queueing based approach. *J. Rail Transp. Plan. Manag.* **2016**, *6*, 99–115. [CrossRef]
7. Sameni, M.K.; Moradi, A. Railway capacity: A review of analysis methods. *J. Rail Transp. Plan. Manag.* **2022**, *24*, 49–57. [CrossRef]
8. Huisman, T.; Boucherie, R.J.; van Dijk, N.M. A solvable queueing network model for railway networks and its validation and applications for the Netherlands. *Eur. J. Oper. Res.* **2002**, *142*, 30–51. [CrossRef]
9. Meirich, C.; Nießen, N. Calculating the maximal number of additional freight trains in a railway network. *J. Rail Transp. Plan. Manag.* **2016**, *6*, 200–217. [CrossRef]
10. Weik, N.; Nießen, N. Quantifying the effects of running time variability on the capacity of rail corridors. *J. Rail Transp. Plan. Manag.* **2020**, *15*, 100203. [CrossRef]
11. Schmitz, C.; Weik, N.; Zieger, S.; Nießen, N.; Schmeink, A. Markov Models for the Performance Analysis of Railway Networks. In Proceedings of the 7th International Conference on Railway Operations Modelling and Analysis, Lille, France, 4–7 April 2017.
12. Weik, N.; Nießen, N. A Quasi-Birth-and-Death Process Approach for Integrated Capacity and Reliability Modeling of Railway Systems. *J. Rail Transp. Plan. Manag.* **2017**, *7*, 114–126. [CrossRef]
13. Kazakov, A.; Lempert, A.; Zharkov, M. An approach to railway network sections modeling based on queuing networks. *J. Rail Transp. Plan. Manag.* **2023**, *27*, 100404. [CrossRef]
14. Ahuja, R.K.; Magnanti, T.L.; Orlin, J.B. *Netw. Flows: Theory, Algorithms, and Applications*, 1st ed.; Springer: Singapore, 1988.
15. Aronson, J.E. A survey of dynamic network flows. *Ann. Oper. Res.* **1989**, *20*, 1–66. [CrossRef]
16. Tarapata, Z. Modelling and analysis of transportation networks using complex networks: Poland case study. *Arch. Transp.* **2015**, *36*, 55–65. [CrossRef]
17. Branishtov, S.A.; Vershinin, Y.A.; Tumchenok, D.A.; Shirvanyan, A.M. Graph methods for estimation of railway capacity. In Proceedings of the 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), Qingdao, China, 8–11 October 2014.
18. Ford, L.R.; Fulkerson, D.R. Maximal Flow Through A Network. *Can. J. Math.* **1956**, *8*, 399–404. [CrossRef]
19. Yaghini, M.; Nikoo, N.; Ahadi, H.R. An Integer Programming Model for Analysing Impacts of Different Train Types on Railway Line Capacity. *Transport* **2014**, *29*, 28–35. [CrossRef]
20. Azadi Moghaddam Arani, A.; Jolai, F.; Nasiri, M.M. A multi-commodity network flow model for railway capacity optimization in case of line blockage. *Int. J. Rail Transp.* **2019**, *7*, 297–320. [CrossRef]
21. Cacchiani, V.; Caprara, A.; Toth, P. Scheduling extra freight trains on railway networks. *Transp. Res. Part B Methodol.* **2010**, *44*, 215–231. [CrossRef]
22. Weik, N.; Hemminki, E.; Nießen, N. The Effective Residual Capacity in Railway Networks with Predefined Train Services. In *Operations Research Proceedings 2019: Selected Papers of the Annual International Conference of the German Operations Research Society (GOR), Dresden, Germany, 4–6 September 2019*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 725–731.
23. Kurhan, M.; Kurhan, D.; Husak, M.; Hmelevska, N. Increasing the Efficiency of the Railway Operation in the Specialization of Directions for Freight and Passenger Transportation. *Acta Polytech. Hung.* **2022**, *19*, 231–244. [CrossRef]
24. Liao, Z.; Li, H.; Miao, J.; Corman, F. Railway capacity estimation considering vehicle circulation: Integrated timetable and vehicles scheduling on hybrid time-space networks. *Transp. Res. Part C Emerg. Technol.* **2021**, *124*, 102961. [CrossRef]
25. Liao, Z.; Li, H.; Miao, J.; Meng, L. Estimating the Railway Network Capacity Utilization with Mixed Train Routes and Stopping Patterns: A Multiobjective Optimization Approach. *J. Adv. Transp.* **2024**, *2024*, 5467767. [CrossRef]
26. Jensen, L.W.; Schmidt, M.; Nielsen, O.A. Determination of Infrastructure Capacity in Railway Networks without the Need for A Fixed Timetable. *Transp. Res. Part C Emerg. Technol.* **2020**, *119*, 102751. [CrossRef]
27. Graph Online. Graph Online. *Create Graph and Find the Shortest Path.* Available online: https://graphonline.ru/en/ (accessed on 15 March 2022).
28. Janecek, D.; Weymann, F. LUKS—Analysis of lines and junctions. In Proceedings of the 12th World Conference on Transport Research, Lisbon, Portugal, 11–15 July 2010.

29. Deutsche Bahn AG. Richtlinie Trassenmanagement (Ril 402) 2009.
30. RWTH Aachen University. RWTH High Performance Computing (Linux). Available online: https://help.itc.rwth-aachen.de/service/rhr4fjjutttf/ (accessed on 9 September 2024).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.