Deutsche Bahn Schedules Train Rotations Using Hypergraph Optimization

Ralf Borndörfer, Thomas Eßer, Patrick Frankenberger, Andreas Huck, Christoph Jobmann, Boris Krostitz, Karsten Kuchenbecker, Kai Mohrhagen, Philipp Nagl, Michael Peterson, Markus Reuther, Thilo Schang, Michael Schoch, Hanno Schülldorf, Peter Schütz, Tobias Therolf, Kerstin Waas, Steffen Weider

Please scroll down for article—it is on subsequent pages

With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.
For more information on INFORMS, its publications, membership, or meetings visit http://www.informs.org

THE FRANZ EDELMAN AWARD
Achievement in Operations Research

Deutsche Bahn Schedules Train Rotations Using Hypergraph Optimization

Ralf Borndörfer,[a] Thomas Eßer,[b] Patrick Frankenberger,[c] Andreas Huck,[c] Christoph Jobmann,[c] Boris Krostitz,[c] Karsten Kuchenbecker,[c] Kai Mohrhagen,[c] Philipp Nagl,[d] Michael Peterson,[d] Markus Reuther,[e] Thilo Schang,[f] Michael Schoch,[c] Hanno Schülldorf,[c] Peter Schütz,[g] Tobias Therolf,[h] Kerstin Waas,[i] Steffen Weider[e]

[a] Mathematics Department, Freie Universität Berlin, and Network Optimization Department, Zuse Institute Berlin, 14195 Berlin, Germany;
[b] Deutsche Bahn Cargo AG, 55116 Mainz, Germany; [c] Deutsche Bahn AG, 60329 Frankfurt am Main, Germany; [d] Deutsche Bahn Fernverkehr AG, 60326 Frankfurt am Main, Germany; [e] LBW Optimization GmbH, 14195 Berlin, Germany; [f] Deutsche Bahn Netz AG, 60486 Frankfurt am Main, Germany; [g] Deutsche Bahn AG, 60528 Frankfurt am Main, Germany; [h] Deutsche Bahn Regio AG, 68163 Mannheim, Germany; [i] Deutsche Bahn Netz AG, 60326 Frankfurt am Main, Germany

Contact: borndoerfer@zib.de, https://orcid.org/0000-0001-7223-9174 (RB); thomas.esser@deutschebahn.com (TE); patrick.frankenberger@deutschebahn.com (PF); andreas.huck@deutschebahn.com (AH); christoph.jobmann@deutschebahn.com (CJ); boris.krostitz@deutschebahn.com (BK); karsten.kuchenbecker@deutschebahn.com (KK); kai.mohrhagen@deutschebahn.com (KM); philipp.nagl@deutschebahn.com (PN); michael.peterson@deutschebahn.com (MP); reuther@lbw-optimization.de (MR); thilo.schang@deutschebahn.com (TS); michael.schoch@deutschebahn.com (MS); hanno.schuelldorf@deutschebahn.com (HS); peter.schuetz@deutschebahn.com (PS); tobias.therolf@deutschebahn.com (TT); kerstin.waas@deutschebahn.com (KW); weider@lbw-optimization.de (SW)

Abstract. Deutsche Bahn (DB) operates a large fleet of rolling stock (locomotives, wagons, and train sets) that must be combined into trains to perform rolling stock rotations. This train composition is a special characteristic of railway operations that distinguishes rolling stock rotation planning from the vehicle scheduling problems prevalent in other industries. DB models train compositions using hyperarcs. The resulting hypergraph models are addressed using a novel coarse-to-fine method that implements a hierarchical column generation over three levels of detail. This algorithm is the mathematical core of DB's fleet employment optimization (FEO) system for rolling stock rotation planning. FEO's impact within DB's planning departments has been revolutionary. DB has used it to support the company's procurements of its newest high-speed passenger train fleet and its intermodal cargo locomotive fleet for crossborder operations. FEO is the key to successful tendering in regional transport and to construction site management in daily operations. DB's planning departments appreciate FEO's high-quality results, ability to reoptimize (quickly), and ease of use. Both employees and customers benefit from the increased regularity of operations. DB attributes annual savings of 74 million euro, an annual reduction of 34,000 tons of CO_2 emissions, and the elimination of 600 coupling operations in crossborder operations to the implementation of FEO.

Introduction

The Western European railway network is among the densest in the world and is of major importance for passenger and freight transport in the European Union (Figure 1). It supports a broad mix of operations, including InterCity Express (ICE) high-speed connections between major cities in central Europe at up to 300 kilometers per hour (km/h), regional passenger transport such as commuter services, and cargo transport with heavy freight trains.

Deutsche Bahn (DB) is among the leading railway companies in the world, with revenue exceeding that

Figure 1. The Map Shows the Western European Railway Network with Germany (Shaded in Light Gray) at the Center
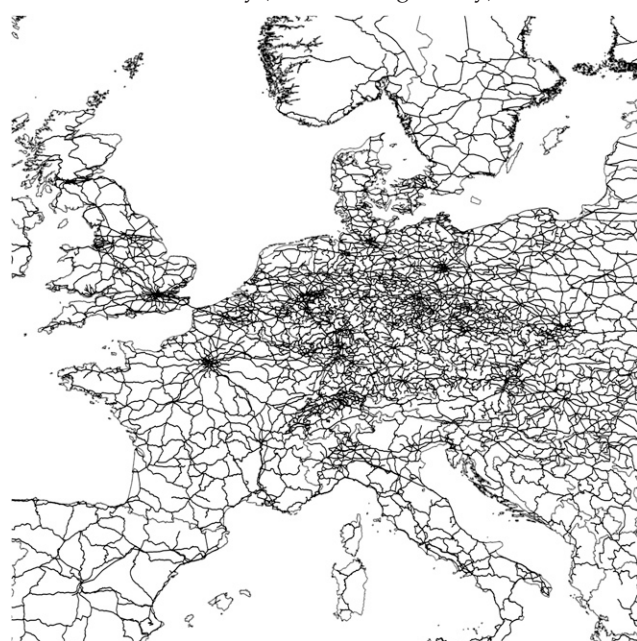
of BNSF and Union Pacific combined. Each day, the company transports about 5.7 million people using approximately 24,000 timetabled trips (DB 2018). To do so, DB employs about 1,300 locomotives, 5,500 coaches (i.e., passenger wagons), and 4,200 train sets (i.e., permanently coupled sets of self-powered units), including 274 ICE high-speed units, which are the flagships of DB's operations (Figure 2). DB also transports 700,000 tons of freight each day using 2,900 freight trains (in Germany), 2,700 locomotives, and 83,000 freight wagons. As the Federal Ministry of Transport and Digital Infrastructure (2020) reports, in 2018, these numbers corresponded to steadily increasing shares in the German market of 8.3% for passenger

Figure 2. (Color online) Deutsche Bahn Currently Operates Four Generations of ICE High-Speed Train Sets, Which We Show from Left to Right: ICE1 (Oldest), ICE2, ICE3, and ICE4 (Newest)

transport (p. 221) and 19.1% for freight transport (p. 247). Rail transport also has a positive impact on our environment (Federal Environment Agency 2012) because it saves up to two-thirds of energy (p. 14) and 75% of CO_2 (p. 16) in comparison with road transport.

A key asset of every railway company is its train sets, locomotives, and wagons. These units are referred to individually as vehicles, and together as rolling stock, to distinguish them from the network infrastructure of stations, tracks, switches, signals, and control centers that are fixed in place. Putting the rolling stock to efficient use is essential for every optimization strategy. Although it would seem logical for railway companies to use operations research methods, as airlines have used them to schedule their aircraft and public transit companies their buses since the 1980s, this was not the case at DB until recently. The main reason is mathematical. Unlike in air transportation or public transit, solving a railway vehicle routing problem to compute sequences of trips for individual vehicles (i.e., rotations) is not a satisfactory solution. Scheduling railway rolling stock also requires handling the composition of vehicles into trains. One option is to decompose the problem into train composition and train routing steps, which are solved in sequence. In rail freight transport, for example, cars are assembled into blocks in step 1, trains are routed in step 2, and locomotives are scheduled in step 3 (Ireland et al. 2004, Ahuja et al. 2005), often using network flow techniques. Approaches using partial integration include Ahuja et al. (2005), who consider penalizing the breakup of consists (i.e., locomotive compositions) by using consist busting, and Cacchiani et al. (2019), who use capacity constraints on train compositions. Kroon et al. (2009) present an integrated approach, but it is designed for daily settings with fixed turns on the spot (i.e., unique, immediate follow-on trips). This model is suboptimal for operations that rely on repositioning vehicles over long distances and over several days, such as in long-distance passenger transport or in European cargo transport. Therefore, no mathematical methods were available to simultaneously handle both aspects of vehicle scheduling and train composition for DB's complex and large-scale rolling stock rotation scheduling problems.

DB launched a project to develop a new rolling stock rotation planning (RSRP) system in 2008. Its name FEO is an acronym for "Fahr- und Einsatz-Optimierung," which can be translated as fleet employment optimization. It is based on new mathematical methods of a concept we call "algorithmic hypergraph theory."

This paper describes the RSRP problem at DB and our operations research approach to modeling and solving it. We discuss how we implemented our approach and how FEO helped DB to improve its

business processes. We then report on success stories for the three DB divisions (DB Fernverkehr AG, DB Regio AG, and DB Cargo) that are responsible for long-distance passenger transport, regional transport, and cargo transport, respectively, and conclude with a summary of the benefits accrued.

The Rolling Stock Rotation Planning Problem

The objective of rolling stock rotation planning is to determine how a railway uses its rolling stock to compose trains and to implement a given timetable subject to a number of objectives and constraints. In this section, we provide background information.

Train Composition

Rolling stock refers to a railway company's set of vehicles. Traditionally, the vehicles have been (and to a large extent, still are) locomotives and wagons; however, today vehicles can also be train sets (i.e., self-powered, fixed concatenations of several units). Such train sets look like trains, but they cannot be easily disassembled or reassembled, at least not by a coupling operation. For our purposes, we treat them as indivisible vehicles (of a larger size). Figure 3 shows a typical train set.

Similar to locomotives and wagons, train sets can be coupled to form trains if they feature clutches. Rather than using the term train, we use the term train composition or composition to highlight the combinatorial choices that are made in forming a train (we use the term train in the title of the paper because it is a customary word in everyday speech). ICE2 and ICE3 high-speed train sets have special Scharfenberg clutches—clutches that allow them to be coupled to run in double traction (i.e., two self-powered units; in this case, train sets that are coupled together) (Figure 4); theoretically, they can also run in triple traction, but because platform length in Germany is limited to

Figure 3. (Color online) The Vast Majority of DB's Train Sets Are Used in Regional Transport Commuter Trains

Figure 4. (Color online) ICE2 and ICE3 Train Sets Have Clutches and Can Be Coupled to Run as Double-Traction Trains

400 meters, triple-traction ICEs are not allowed. Although the clutches of ICE2 and ICE3 train sets fit together mechanically, the electronic and pneumatic systems are not compatible; therefore, forming trains by coupling ICE train sets of different vehicle types (in this case, ICE generations) is not permitted. ICE vehicles also feature a variety of technical equipment that the neighboring European countries of Austria, Belgium, France, the Netherlands, and Switzerland require. The presence or absence of such equipment results in many vehicle subtypes and complex rules that address which vehicles can be coupled and how this coupling can be done. These rules can be expressed in terms of numbers of vehicles, types of vehicles, or positions of vehicles within a train composition.

A less obvious train composition rule refers to the orientation of a vehicle (i.e., whether it is positioned to drive forward or backward). For example, an ICE1 train set consists of five first-class wagons, followed by a restaurant and seven second-class wagons, and has an overall length of 390 meters. Positioning the train at the correct side of the platform is important to facilitate transfers of passengers between trains that are at the same platform; to ensure first-class wagons are closer to the station, thus minimizing the distance first-class passengers must walk; and to optimize crew relief. If an ICE train set goes forward with the first-class wagons in front, it is said to be in a "tick" orientation; if the second-class wagons are in front, it is called a "tock" orientation. The orientation must also be considered for train composition because of technical constraints. For example, consider two ICE2 train sets that are coupled in double traction as tock-tick, such that the two cockpits at the first-class side of the train sets are face to face in the middle of the train. Such a composition can only run at a maximum speed of 200 km/h. Because the timetable is designed for a train speed of 250 km/h, this is a severe restriction.

For these reasons, technical, organizational, and marketing considerations give rise to many train composition rules that have significant combinatorial complexity. We provide an illustration in Figure 5.

All train compositions are rated to assess their desirability and to reflect, for example, the number of vehicles used. Multiplied by travel distances, these weighted factors provide the primary (number of vehicles used), secondary (monetary), and tertiary (operational convenience) objectives of a rolling stock rotation planning problem.

In one sense, the positions and orientations of vehicles within a train composition are less critical than the number of vehicles and their types. Positions and orientations will often end up correctly, or they can be adjusted in a subsequent scheduling step (but not in operation) at no extra cost. This observation can be exploited algorithmically. Therefore, introducing the term "configuration" to refer to the numbers and types of vehicles in a train composition is useful. The train configuration represents a "coarsening" of the composition; that is, the positions and orientations are ignored. The result is significantly fewer configurations than compositions. In the example in Figure 5, the 16 compositions are coarsened to three configurations: (1) two white framed, (2) one white framed and one gray framed, and (3) two gray framed.

Turns, Deadhead Trips, and Regularity
Determining the compositions of the timetabled trips is an important aspect of rolling stock rotation planning, but other factors must also be considered. The timetabled trips must be connected. This involves turns and deadhead trips. The simplest types of turns are direct through-station (i.e., a station at which a train can stop and then continue with its composition unchanged) turns and direct turns in a dead-end station (i.e., a station at which the track ends such that the train reverses its orientation). To avoid a reversal, a turn with a (hopefully short) turnaround trip can be used,

provided the infrastructure allows the reversal. Sometimes, however, a train might have to be reoriented or repositioned using an unproductive empty trip without payload over a larger distance; such a trip is called a deadhead trip. A significant problem involving shunting operations, whose feasibility and duration depend on the track geometry, can arise in scheduling coupling and uncoupling processes on turns and deadhead trips. We will show later in the paper that such (un-)coupling rules can be handled by setting train composition rules for turns and deadhead trips, such that we do not have to delve into the (complicated) subject of railway station operations.

Fortunately, we can also set rules to address the requirement for regularity. Regularity refers to the similarity of operations on different days of the week to provide railways and other industries with the ability to set operational routines. In their package delivery optimization approach, Holland et al. (2017) emphasize the value of regular patterns for their UPS delivery routes, whereas Ahuja et al. (2005) stress the significance of consistent (i.e., regular) locomotive compositions in rail freight transport for trips that repeat in five or more of the seven days of the week. Scheduling similar activities each day is advantageous in every complex system. For this reason, railway timetables are regular (or periodic) and should also be operated using regular train compositions and vehicle rotations. We explain that the mathematical approach we use to address train compositions can also be used to address this requirement.

Summarizing the Rolling Stock Rotation Planning Problem
There are two degrees of freedom in rolling stock rotation planning: selecting the train compositions for the timetabled trips, the turns, and the deadhead trips and selecting the turns and deadhead trips. The number of potential turns and deadhead trips exceeds the number of timetabled trips by orders of magnitude. As both degrees of freedom multiply, the treatment of turns and deadheads becomes the main algorithmic difficulty. Turns, deadhead trips, and the compositions of timetabled trips, turns, and deadhead trips must be chosen in such a way that feasible routes (i.e., rotations) are generated for the vehicles involved. Rotations and train compositions must be synchronized to ensure a railroad can cover all timetabled trips at the minimum cost—this is the rolling stock rotation planning problem.

Rolling stock rotation planning has additional constraints: in particular, maintenance rules and station and parking requirements. The FEO system handles these constraints; however, although they are important, we do not discuss them in this paper.

Figure 5. Coupling Two Train Sets of Different Types (e.g., White Framed and Gray Framed) at Two Positions Using Two Orientations (Tick = White Filled, Tock = Gray Filled) Results in 16 Possible Train Compositions
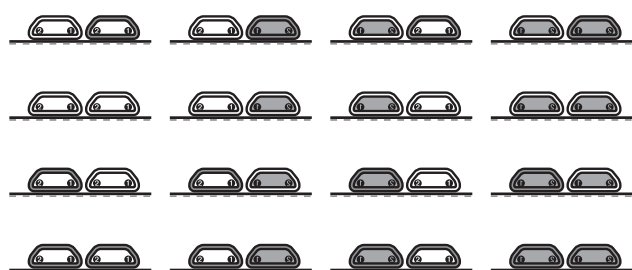

Note. The driving direction is from left to right.

Uses for Rolling Stock Rotation Planning

The rolling stock rotation planning problem is relevant to several DB business processes. Rolling stock rotations are first considered in developing the strategic plan, which is devised 10–20 years prior to its implementation. The strategic plan includes the creation of a prototypical timetable for a standard week (i.e., a week that repeats indefinitely in a cyclic fashion). In a cyclic timetable, the vehicle routes are cycles that extend over one or more weeks (hence, the term rotation); we illustrate this in Figure 6. The standard week is useful in evaluating the feasibility or cost of a timetable with respect to an existing fleet or fleet procurement. The tactical plan, which is developed approximately five years before its implementation, includes the development and differentiation of summer and winter timetables and establishes the transition between them. It also considers long-term construction sites in the railway track network.

In the operational plan, which is developed one year (or sometimes less than one year) prior to its implementation, holidays, important events such as major exhibitions, and most importantly, construction sites are considered on a fully dated basis (i.e., for specific days in a calendar period) and in increasing detail, as additional information becomes available. For this purpose, the standard week is rolled out over an acyclic (i.e., noncyclic) planning horizon with a start date and an end date (e.g., December 1, 2021, to December 31, 2021), adding information for specific days (e.g., extra trips during the Christmas season);

see Figure 7 for an illustration. A fully dated scenario involves more timetable data. It also shifts attention from the consideration of exchangeable vehicles (or vehicle types) to the planning of individual vehicles with specific characteristics, such as start and end positions, maintenance states, and home depots. Using our approach, an individual vehicle can be modeled as a unique and distinct vehicle type; however, this may come with a drawback. For example, the resulting increase in the number of trips and the number of vehicle types initially makes the fully dated problem appear to be much harder to solve than its cyclic counterpart.

However, fully dated problems are not optimized from scratch. The purpose of strategic and tactical planning in FEO is to construct a set of base schedules that are modified only in the fully dated scenarios, thus allowing DB to provide operational stability. The purpose of planning one year ahead or less than one year ahead is to adapt the schedule to a previously generated base plan. Planners should follow the base plan whenever the planning department considers doing so to be feasible and cost effective; a planner who deviates from it should return to following it as soon as possible. This approach is similar to the use of

Figure 6. (Color online) Rotations Through a Cyclic Timetable of Seven Days in a Standard Week Can Be Arranged in a Torus
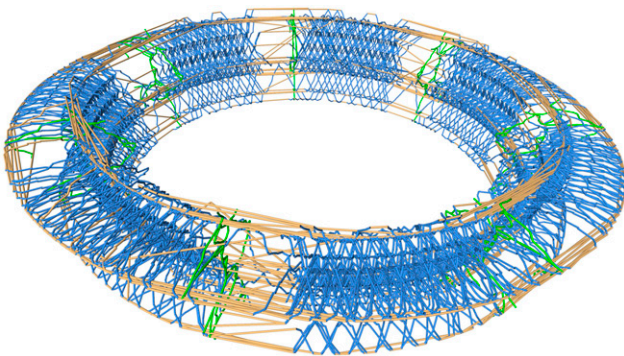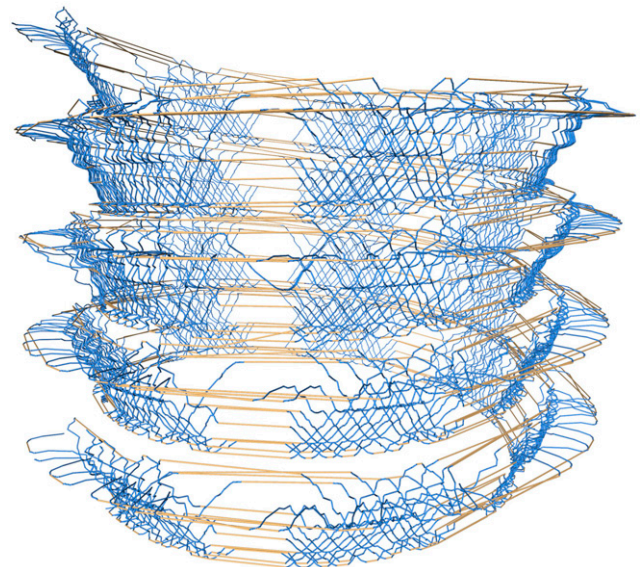

Notes. This three-dimensional diagram uses cylindrical coordinates in which time goes counterclockwise (a full cycle represents one week); the other two coordinates describe positions in a map of Germany. The seven denser sections correspond to the daytime operations on the seven days of the week, and the seven sparse sections correspond to overnight train parking. The edges represent timetabled trips and deadhead trips in cylindrical space and time. The diagram visualizes the rotations of the entire ICE1 fleet for one standard week; the rotations correspond to a set of cycles that cover all timetabled trips. The train compositions are not visible (i.e., no hyperarcs are visualized).

Figure 7. (Color online) A Fully Dated (i.e., Concrete Calendar Period) Scenario of Four Weeks Can Be Arranged in a Helix

Notes. This three-dimensional diagram also uses cylindrical coordinates, but now, time goes counterclockwise and upward: that is, time does not repeat indefinitely; instead, there is a start at the bottom and an end at the top. Each of the four windings corresponds to one week of operation: that is, the diagram visualizes a calendar period of four weeks. The vehicle rotations are no longer cycles but become paths. They are still called rotations because vehicles eventually return to their home depots. Comparing the rotations at different windings highlights irregularities (differences), which result from events such as holidays or blockages such as construction sites.

base routes in the UPS package delivery approach (Holland et al. 2017).

Such a reoptimization setting requires, as an additional input, a base plan that is used as a scheduling template to penalize deviations. This base plan is usually infeasible for the current scenario; otherwise, we would not have to adapt it. Nevertheless, a base plan provides enough guidance to compensate for the increase in size, such that fully dated reoptimization scenarios are different but not substantially harder to solve than an optimization developed from scratch in the base case. In reality, the computationally hardest problems are found in strategic settings in which one has to determine the feasibility of a scenario in a fleet minimization run; we discuss such applications in section Tendering Regional Passenger Services and in the procurement cases we address in the section ICE Procurement and Construction Site Management.

The Hypergraph Model for Rolling Stock Rotations

At the heart of our operations research approach to rolling stock rotation planning is a novel model of the problem using hypergraphs. This model provides a mathematically elegant and unified treatment of all train composition and regularity requirements and for all optimization and reoptimization uses discussed in the previous section. It also provides the basis for a generalization of the theory of network flows and more generally, a generalization of algorithmic graph theory to "algorithmic hypergraph theory."

Figure 8 illustrates this concept. Figure 8(a) shows two (standard) arcs that connect two pairs of nodes. Each node represents an arrival or departure event of a vehicle at a station and the state of the vehicle with respect to type, position, and orientation. The arcs that connect these state-event nodes represent the movement of a vehicle and the retention or change of its position and orientation (the type must stay the same). An arc from a departure to an arrival node represents the movement of a vehicle in a timetabled

trip; an arc from an arrival to a departure event represents a movement (or just waiting) in a turn or a deadhead trip. Adding all possible state-event nodes and all possible transition arcs results in a vehicle scheduling graph. Flow on an arc reflects the transition from one node to another (e.g., a vehicle moving from one station to another). Two flows on two parallel arcs that represent the same timetabled trip indicate the use of two vehicles to service this trip; this can be used to model a double traction of two vehicles, as we show in Figure 8(a). Flow conservation results in vehicle rotations. The problem with this naïve flow model is that it provides no control of train compositions because each vehicle can move as if it was operated individually.

Figure 8(b) shows how to obtain this control by merging compatible standard arcs into hyperarcs. Such hyperarcs model feasible state-event transitions of train compositions. Adding (not adding) appropriate hyperarcs allows (rules out) train compositions on timetabled trips, turns, and deadhead trips. Figure 9 shows a more

Figure 9. A Hyperflow in a State-Event Hypergraph Represents Feasible Train Compositions and Vehicle Rotations
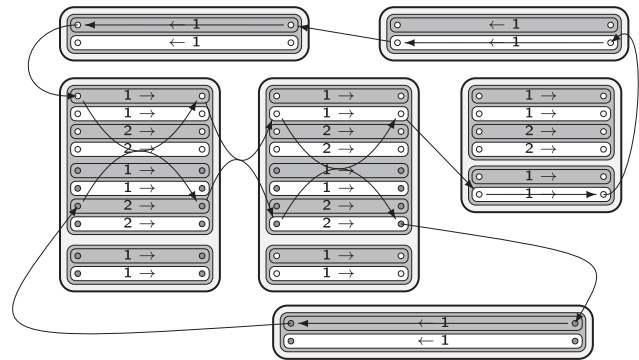
Notes. In this example, the six framed boxes represent six timetabled trips. The nodes shown as small circles model possible departure-arrival events in terms of the vehicle types (i.e., white-filled and gray-filled nodes), positions (1 or 2 with reference to the driving direction →), and orientations (white subbox: tick; gray subbox: tock). The two trips at the top can only be serviced by a white-filled vehicle in single traction in either tick or tock orientation. Likewise, the trip at the bottom requires a gray-filled vehicle. The three trips in the middle can be run in single traction (bottom of the box) and double traction (top of the box). The most complicated one is on the left. It can be operated by a gray-filled single-traction train composition in either orientation (lower two subboxes) or by a double-traction train composition, including an additional white-filled vehicle, in any position and orientation (upper subboxes). The hyperarc drawn in this trip encodes the composition 1:white:tock, 2:gray:tock. The succeeding hyperarc represents a turn to 1:gray:tick, 2:white:tick in which the entire composition is turned around with reference to its driving direction; this is typical for a turn in a dead-end station. At the turn after the next trip, the train composition is split into two single-traction train compositions. After the return trips, the single-traction train compositions are recoupled into a double-traction train composition to start over. Any other kind of regime can be encoded by choosing corresponding hyperarcs—and vice versa.

Figure 8. (a) The Simultaneous Use of Several (Here: Two) Standard Arcs Can Be Represented by a Hyperarc to (b) Conveniently Model Train Composition Options Using Hypergraphs
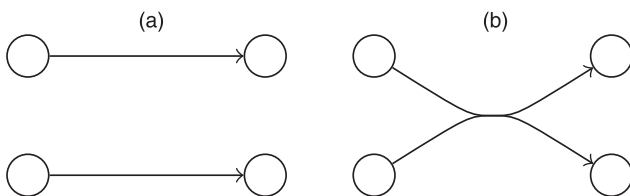

Notes. The example shows two vehicles combined into a train. Because the nodes are not labeled by positions and orientations, the example looks as if it is at the configuration level. However, the same principle applies to compositions.

elaborate example with six trips, turns, and deadhead trips. As in a standard network flow model, flow conservation at every node results in a hyperflow that can be decomposed into rotations of the individual vehicles. The trips are covered because of flow constraints that enforce the use of exactly one standard arc or hyperarc for every timetabled trip. Rolling stock rotation planning thus amounts to solving a minimum cost hyperflow problem. This can in turn be translated into a binary integer program in which each hyperarc and standard arc becomes a binary decision variable. Because many train compositions are possible, many hyperarcs that reflect the underlying combinatorial complexity are also possible. In a full model, there are, for every trip, turn, and deadhead trip, as many hyperarcs as there are possible train compositions. However, this impairs neither the conceptual simplicity of the model nor its usefulness because hyperarcs can be generated dynamically using a column-generation approach.

The hypergraph approach can also handle regularity requirements. The ability to model simultaneous movements of several vehicles applies to both one trip and to multiple trips: that is, all repetitions of a generic timetabled trip on several days of a planning horizon. Running all such trips in the same way gives rise to regularity hyperarcs. Figure 10 shows an example of a regular turn from the everyday arrival of a set of regular timetabled trips on the left to the departures of another set of regular timetabled trips on the right. This regular turn can be modeled by using a hyperarc, as the example shows. Giving this hyperarc a reward in the objective function configures the model to prefer regularity. Other regularity requirements can be expressed similarly via suitable hyperarcs (which could become "hyperarcs of hyperarcs").

Figure 10. A Hyperarc Can Model a Regular Turn (i.e., a Turn that Is Identical on Each Day of the Week) Between Regular Arrivals and Regular Departures, Which Also Repeat Each Day
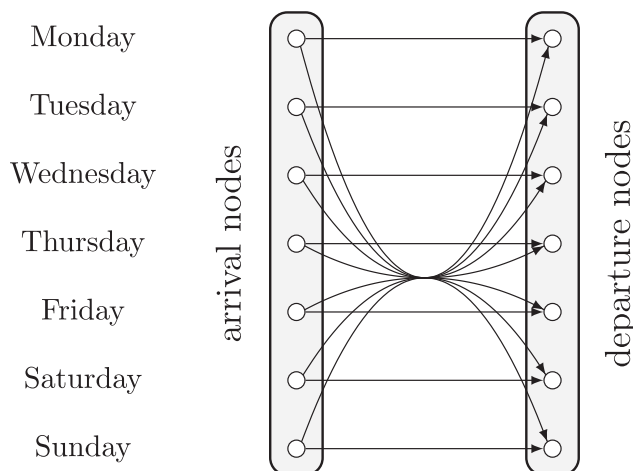
Hypergraphs have thus far had limited popularity in the optimization community because they can result in large models that have little structure and are hard to solve. The hypergraphs present in rolling stock rotation planning, however, have a special structure; that is, all hyperarcs are unions of standard arcs, which we refer to as graph-based hypergraphs. The node-hyperarc incidence matrices of graph-based hypergraphs have only $0/ \pm 1$ coefficients, and all columns add up to zero, like standard network matrices (see the mixed integer programming (MIP) formulation in the appendix for details). Although hyperflow problems are provably hard (i.e., APX-hard) for even the simplest imaginable settings (Borndörfer and Heismann 2015), we found that instances arising from rolling stock rotation planning problems are computationally well behaved and provide strong lower bounds on the optimal objective value (Reuther 2017, chapter 7). Their structure can also be exploited to design a combinatorial hyperflow network simplex algorithm for the solution of the linear programming relaxation (Beckenbach 2018), which is currently in an early theoretical stage and not yet used for computations in FEO. These results indicate the potential to establish an algorithmic hypergraph theory based on suitable structures, such as those that are graph based. A number of theoretical results for related classes of hypergraphs, including a generalized Hall theorem for normal hypergraphs (Beckenbach and Borndörfer 2018) and a tight-cut decomposition algorithm for matching-covered uniformizable hypergraphs (Beckenbach 2019), corroborate this assumption. Hyperflow models are also potentially useful for other applications that involve a synchronization of paths, such as scheduling work for teams, truck-and-drone scheduling, or modelling chemical reaction networks.

Maintenance requirements, which are handled by using an additional model component of resource flows and consume resources along vehicle rotations (Giacco et al. 2014), cutting planes (Grimm et al. 2019), or path-based models (Lusby et al. 2017), have also been investigated. The mathematical details of the complete model are published in several articles, including Borndörfer et al. (2011, 2015, 2018, chapter 10) and extensively in Reuther (2017).

The Coarse-to-Fine Method for Rolling Stock Rotation Optimization

Hyperflow problems arising from actual business cases at DB cannot be solved by setting up an integer program and entering it into a standard solver. The major problem is the size of the hypergraph, which can have 10^8 hyperarcs, each of which gives rise to a binary variable. Although direct hyperarc column-generation

approaches provide better results, they are not sufficient in our situation. We found, however, that a hierarchical column-generation approach, which makes the major decisions on a coarse (and hence, computable) level and checks the position and orientation details on a fine level, is a workable approach. In the algorithm we implemented in FEO, we use three levels (i.e., layers), which we define in Figure 11. The resulting hierarchical pricing algorithm is called the coarse-to-fine (C2F) methods (Borndörfer et al. 2014, 2018, chapter 10). The C2F approach is not limited to rolling stock rotation optimization but is applicable to large-scale linear programming in general. Reuther (2017) includes an in-depth presentation.

The C2F method is inspired by the expertise of DB planners and can be seen as a mathematical formalization of a procedure that has been used successfully over a long period. DB planners knew that tackling a dedicated scenario often requires making a relatively small number of key decisions. For example, suppose we know the number of vehicles and have information about some essential deadhead trips with crucial repositioning. Then, the remaining direct turns are often obvious, and further details (e.g., orientations and positions of the vehicles in the train compositions) are tightly restricted by the timetable and by technical constraints such that they frequently fall easily into place. The C2F method mirrors this approach in a mathematically precise and algorithmically powerful way.

The concept is to represent coarse and fine decisions in terms of hierarchical layers, as we illustrate in Figure 11. The finest layer is called the composition layer; it contains all details of the full hypergraph model. This is the model that we want to solve; the other two layers are relaxations (i.e., approximate models that have more solutions). The configuration layer is derived from the composition layer by ignoring the positions and orientations of the vehicles in a train composition. This layer represents the key decisions on which types of vehicles are combined and whether coupling and uncoupling operations must be performed. The configuration layer is a coarse hyperflow model of its own. Going one final step higher in the hierarchy, we also ignore vehicle types (e.g., white-filled and gray-filled types) and coupling requirements and arrive at the vehicle layer. This layer simplifies the problem to the computation of a standard single-commodity minimum cost flow. It reflects essential decisions on the number of vehicles needed and which deadhead trips are used. The layered construction has the property that the nodes and hyperarcs of the detailed layers have canonical representatives in the coarser layers, and vice versa, coarse nodes and hyperarcs correspond to sets of nodes and hyperarcs arcs in the finer layers. In this

way, compositions can be mapped to configurations, configurations to individual vehicle movements, and vice versa; see Borndörfer et al. (2011, 2015) for details. The particular mappings between compositions and configurations and between configurations and standard flow arcs are inspired by the problem structure and the planning expertise of DB experts.

The layers are used to implement a hierarchical column-generation algorithm that yields efficient pricing rules and strong, exact pruning. Briefly, the algorithm prices on the coarsest layer and checks the implementation of the resulting coarse solution on the induced part (and only on the induced part) of the next-finer layer and recursively on the finest layer (i.e., overall over two layers). Then, the induced part of the finest layer is used to compute shadow prices (i.e., a dual solution), and the method iterates. In this way, the algorithm can exploit the simplicity and reduced size of the coarse layers and yet check detailed and complex rules on the fine layers. To compute the shadow prices and to guarantee that the procedure finds the fine linear programming optimum, an algebraic linear programming (and finally, integer programming) implementation of the C2F method must be derived. Its mathematical core is a "coarsening lemma," which ensures that the reduced costs in coarse layers always overestimate the attractivity of arcs or hyperarcs in the finer layers. That is, the coarsening lemma proves that coarsening produces relaxations. This ensures the theoretical convergence of the C2F method. In practice, the integer programming optimum is often reached to a persistent duality gap that is small enough that one can prove the optimality of the computed fleet size by means of reduced cost arguments. The FEO implementation of the C2F method is fully parallelized in various ways. Run times vary between a few minutes for easy reoptimization scenarios to several days for large-scale scenarios that may be infeasible.

The FEO System

The FEO decision support system allows planners to edit, manage, visualize, optimize, and analyze rolling stock rotations using advanced functionality. FEO is a multiple-tier application featuring a fat windows client, which interacts with a central server component in the enterprise cloud. The server is connected to a single database and can distribute computing tasks to multiple optimizer instances in a flexible and scalable way. The FEO graphical user interface (GUI) gives access to all relevant data through configurable lists and freely scalable Gantt charts, including tool tips that provide detailed information (Figure 12). The GUI is client specific; that is, different versions are available for the various DB divisions. The FEO

Figure 11. The C2F Algorithm in FEO Uses Three Layers
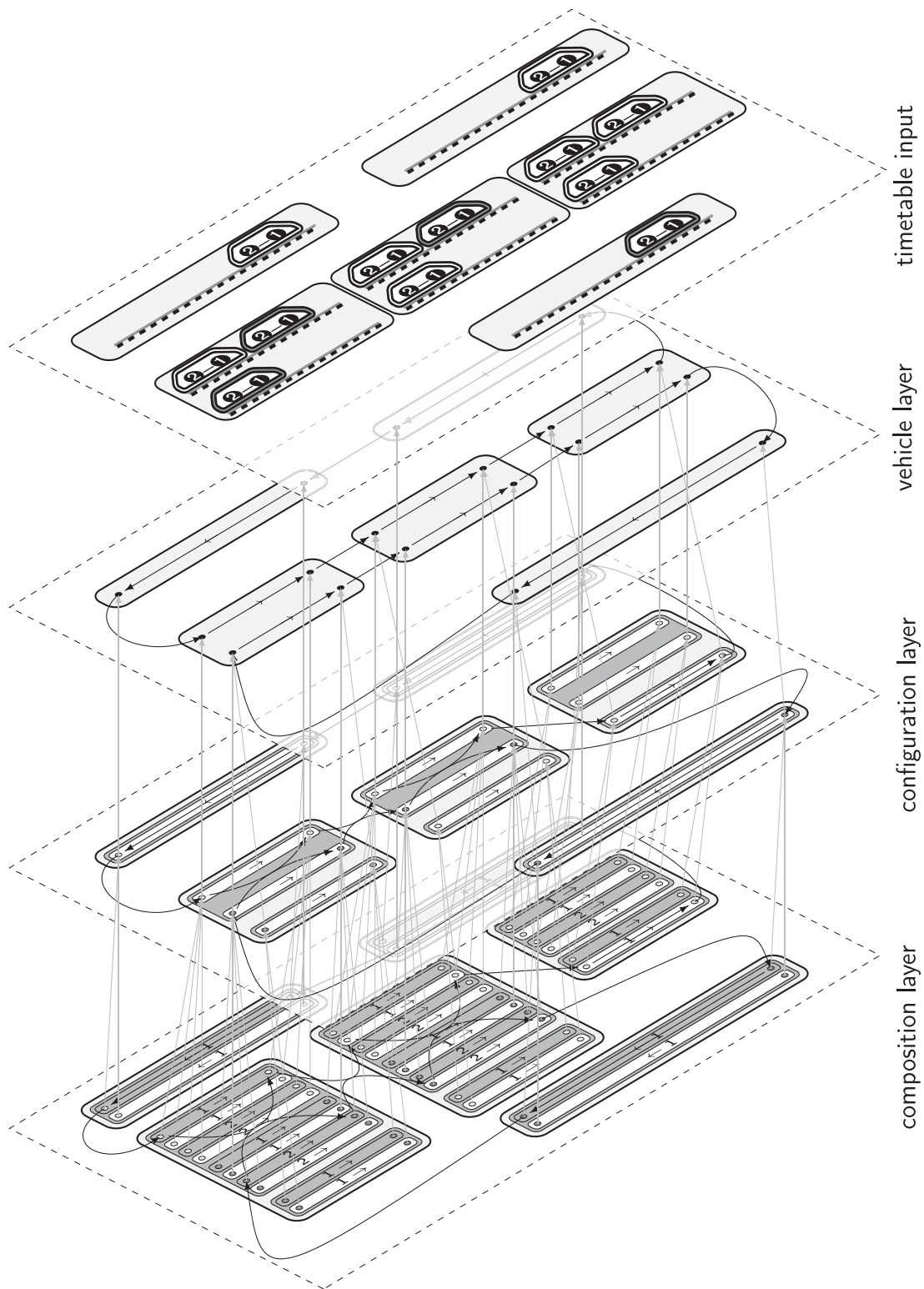
Notes. In this example, the timetable with possible vehicle compositions that suit the example in Figure 9 is arranged at the top. The three layers that the FEO coarse-to-fine algorithm uses are (from top to bottom) the vehicle layer, the configuration layer, and the composition layer; the composition layer coincides with Figure 9. The cone-shaped light gray lines indicate the coarsening projection (from bottom to top), respectively, and the refinement of nodes and (hyper-)arcs (from top to bottom). The level of detail and the complexity of the model increase as we view the diagram from top to bottom.

Figure 12. (Color online) The GUI of the FEO Long-Distance Passenger Transports Version Displays Useful Information on Vehicle Rotations
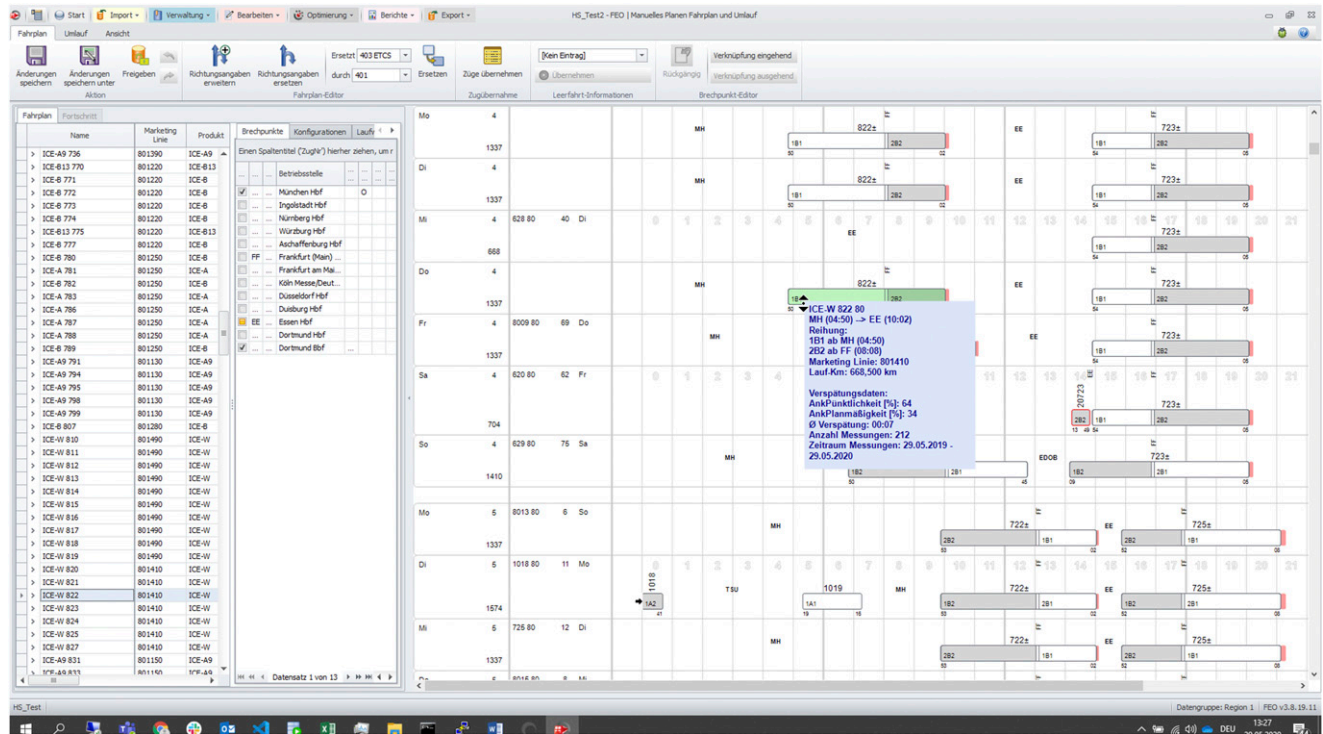
Notes. Data can be inspected and edited using list views (left), Gantt charts (right), and sometimes maps. The small bars on the extreme right (i.e., to the right of the white or gray rectangles) visually show delay averages.

versions are configured automatically based on the role of the software user.

Setting Up a Scenario

The availability of high-quality data, which is a conditio sine qua non for optimization, and fixing faulty data are major problems that can seriously impair user acceptance; indeed, planning experts consider them as dumb data-scrubbing activities that keep them from their creative planning work. The development of a working database is therefore a top priority. Despite all efforts, however, the data delivery systems may not be of sufficient quality, or necessary data may be missing. FEO mitigates this problem by detecting and reporting issues as lists of errors or warnings that describe the problem found (*what*) and provide a reference to the input (*where*) using railway terminology, such as train numbers and station names. FEO includes powerful editors that allow the postprocessing of complex data sets.

Some real-world requirements may not be covered fully or may be covered only by making a disproportionate effort. For example, including rarely applied exceptions, such as scheduling turns without the typical slack time, would require unnecessary effort. Such exceptions are not only allowed, they are sometimes necessary. When they are applied, they

require only some small but well-specified adaptions to allow FEO to find a much better solution. For such cases, FEO offers interactive planning options that are implemented via a drag-and-drop user interface, including filtering, and embedded reporting functionality; examples include the number of kilometers since the last maintenance event and statistical data regarding historical delays (Figure 12).

Working with FEO

Business use cases typically require more than one call to FEO's optimization algorithm. An optimizer will exploit errors (e.g., a driving or turning time of zero is attractive for the optimizer but is probably wrong) to construct solutions that cannot be implemented. When such flaws become apparent, they must be fixed in a subsequent optimization. In addition, a scenario may be infeasible, or computed turns may be incomprehensible (*why did the optimizer do this?*). In such cases, additional side constraints, objective function adjustments, and data fixes must be identified to clarify the situation. The FEO design includes several functions to support these optimization-evaluation issues. For example, in addition to consistency and plausibility checks, FEO categorizes the input data into hard data that cannot be changed and soft data that can be changed if the user confirms the

changes. FEO also provides an editor for exceptions to overrule soft data. If the optimization results are not as expected, FEO can suggest adaptations (i.e., small changes to soft input data) to obtain better results. After an initial phase of data cleaning, parameterization, and the development of FEO profiles for typical use cases, this process is now smooth. Where a single solution was constructed formerly, FEO now usually offers four to five solutions, which a user can compare, analyze, and then choose the best one.

Reoptimization. One of FEO's most important optimization functions is the reoptimization of rolling stock rotations. For this, an existing rotation (i.e., a *reference rotation*) can be passed to the algorithm. The reference rotation was typically created for a different scenario and might be suboptimal, infeasible, or even incomplete. The algorithm automatically configures its objective function to compute a feasible (optimal) solution that reproduces as many details of the reference rotation as possible. This is useful in a wide range of applications. One is to deal with operational disruptions (e.g., by fixing rotations that cannot be implemented because of construction sites or timetable changes). Another is to impound regularity by deriving fully dated rotations from a base plan. A third is scenario exploration. A particularly important aspect of reoptimization is that it can make optimization results transparent to the planner. In reoptimization, the algorithm tries to keep as much as possible of a reference rotation, which the planner can create manually. This simplifies the planner's focus on the differences. Thus, many obscurities and doubts can be quickly resolved. Reoptimization is algorithmically supported by warm starting the optimizer, which guarantees short run times. The similarity of rotations can be controlled via configurable templates; see Borndörfer et al. (2017) for details. Such a reoptimization is usually much faster than a computation made from scratch.

Timetable Improvement. Another popular feature is FEO's ability to identify advantageous shifts of timetabled trips. For this purpose, variables and constraints reflecting the timetable modifications are added to the basic model. For each timetabled trip, variants (shifted copies of the original trip) are added within some time interval, as are constraints stipulating that FEO must choose exactly one variant and penalties that represent the cost of the associated timetable change (the more extreme the shift, the higher the cost). Because the original trip has no change costs, FEO tries to keep as many trips as possible in their original time slots. If FEO finds a better overall solution (i.e., the benefits outweigh the change costs), it reports a suggested solution to the user.

Handout Optimization. FEO includes a feature that improves the look of rolling stock rotations to make them more readable for planners by enabling them to visualize the regularity of rotations, as Figure 13 illustrates. FEO optimizes the rotation handout using a clustering approach; see Borndörfer et al. (2019) for details.

Simulation. FEO can simulate future operations up to 52 weeks in advance; see Öztürk (2015) for details. The purpose is to detect several months in advance if and when additional vehicles will be required because of additional services or timetable changes resulting from construction work. The challenge here is not only to ensure that the trains will still run (additional vehicles can be rented to service additional trips) but to also satisfy crew planning and dispatching requirements. In particular, previously planned rotations should not be assigned to different types of vehicles because drivers are frequently qualified for only specific vehicle types. Early detection of such

Figure 13. The Graphic Shows Two Handouts of the Same Rolling Stock Rotation; the Handout on the Left Is Easier to Read

Notes. The lines represent days of work of a vehicle. Each line is assigned to a rotation day, which we define as a sequence of trips that one vehicle can perform on the seven days of one week. In this example, the rotation day is the "1" or "2" in the first column. The last column defines the rotation day of the succeeding line on the next day. A line of the first rotation day, which continues into the same rotation day, has a "standard turn." Otherwise, it has a "divergent turn." The handout on the left features regular rotation days with mostly repeating days of work. It can easily be read line by line with only two exceptional divergent turns (both on Sunday). The handout on the right shows irregular days of work and only divergent turns. Reading it requires permanent skipping between the two rotation days. In this example, the vehicle starts in rotation day 1 on Monday with trips 374 and 1,061. The "2" at the end of the line indicates that the next line is the one for Tuesday in rotation day 2, which contains trips 374 and 1,061 again. The "1" at the end of the line indicates that we must switch back to the Wednesday line of rotation day 1 and so on. Although standard turns are not always possible, planners clearly prefer them because they result in rotations that are more regular and readable and have more structure.

situations enables foresight in crew planning (e.g., vacation planning) to ensure the availability of drivers with the right qualifications.

Reporting and Scenario Management. An FEO solution can be inspected using various reporting options, and the user is encouraged to experiment with different settings. To analyze and compare different schedule variants, FEO supports sophisticated scenario management. A scenario combines specific data sets in various domains into a complete set of data as required by the optimizer. FEO allows temporarily inconsistent settings while the user works on the data, thus enabling that user to quickly make small changes. When the optimizer starts, it does a validity check to ensure that all data references can be resolved. With many users and use cases, the set of scenarios can become confusing. FEO allows users to see only data sets and scenarios that are relevant to them while enabling them to also work on the same consistent set of core data.

FEO Development
The FEO system was developed by DB Analytics, DB's internal service provider for analytics, forecasts, optimization, and simulation. The project began in 2008 and followed the agile software development philosophy. Throughout the endeavor, users from the three transport divisions of DB worked closely with the software development team to ensure the project's success. User integration was a key success factor for FEO in the early use cases. The mathematical component required research, which was carried out in cooperation with Zuse Institute Berlin (ZIB), a research institute in applied mathematics. The cooperative agreement was in place for almost 11 years and encompassed several phases. During these phases, the research was partially supported by the German Federal Ministry for Research and Education (BMBF), the MATHEON Research Center, and the BMBF MODAL Research Campus. Since the successful implementation and integration of the optimizer, the ZIB spin-off company, LBW Optimization GmbH, has provided additional development, maintenance, and support.

FEO Success Stories
FEO was implemented in direct response to the 2007 liberalization of the regional and cargo railway markets in Europe. It immediately became clear that tendering for regional transport was a question of implementing a service using the minimum number of vehicles. Similarly, crossborder operations required different locomotives for each country and thus, resulted in high fixed costs for freight traffic. Optimization was therefore the key to winning or losing business. DB's long-distance transport division was under pressure to implement the timetable subject to an increasing number of

construction sites, as we explain. These presented, and still present, considerable challenges—but also, significant opportunities. Railroad transportation volumes have increased steadily in the past two decades (Federal Ministry of Transport and Digital Infrastructure 2020). Between 2001 and 2018, the numbers of long-distance and regional railway passengers in Germany rose by almost 10% and 46%, respectively (Federal Ministry of Transport and Digital Infrastructure 2020, pp. 216–217), and contrary to public perception, freight traffic tonnage rose by almost 35% (Federal Ministry of Transport and Digital Infrastructure 2020, pp. 240–241). The growth in passenger and ton kilometers is even more significant as people and goods are being transported over increasing distances. Long-distance and regional passenger kilometers have increased by 22% and 37%, respectively, (Federal Ministry of Transport and Digital Infrastructure 2020, pp. 218–219), and ton kilometers have increased by 64% (Federal Ministry of Transport and Digital Infrastructure 2020, pp. 244–245). To meet this rising demand, the German government and DB have established a long-term investment program valued at 156 billion euro (€) through 2030, with an objective of substantially expanding the capacity of the network. FEO was designed to seize these opportunities and to address the challenges through smarter planning of rolling stock rotations.

ICE Procurement and Construction Site Management
DB's long-distance passenger transport division, DB Fernverkehr AG, uses FEO for both strategic and operational planning purposes. It achieved its first major benefit in 2009 when it used the optimizer in its strategic planning department to support the procurement of the new ICE4 fleet. Its planners performed extensive scenario analyses to determine the number of train sets to procure, their sizes (i.e., numbers of coaches), and whether the units should feature clutches to allow for double traction (like the ICE2) or should not (like the ICE1), and they made a decision to order 130 train sets of 7 cars and 90 train sets of 10 cars, all without clutches.

Since the rollout of FEO in 2013, its main use has been in construction site management. Construction sites force detours (which cost time and mileage) and sometimes result in service splits. For example, a line A-B-C-D might be split into two lines A-B and C-D if a construction site at segment B-C cannot be bypassed; as a result, passengers and vehicles cannot proceed from B to C. Detours and service splits can (theoretically) be addressed by employing additional vehicles, which are not part of the vehicle fleet available in practice, or by service cancellations, which result in loss of revenue and customer dissatisfaction—consequences that DB

wishes to avoid. Because of the mentioned network capacity expansion program, DB Fernverkehr must contend with a record high number of construction sites—currently 850 per day. Thus, scheduling trains through bottlenecks and around temporarily closed track segments, while also maintaining a sufficient level of regularity, is increasingly difficult. Maintenance is a particular complication. To avoid expensive deadhead trips over long distances, the rolling stock rotations would ideally automatically pass by the proper maintenance facilities, further adding to the complexity of the scheduling task. In Figure 14, we illustrate how we used FEO to manage the disruption that construction sites can cause.

In addition to the example in Figure 14(b), we used four dedicated cases from the fourth quarter of 2019. In each, we analyzed FEO solutions in detail to compare its final optimized solutions with alternative manual ones and assess the impact. For the construction sites shown in Figure 14(a) (Gelnhausen (construction duration 3 days, start of October; marked as G), Offenbach (3 days, middle of October; marked as O), Dollbergen-Fallersleben (6 days, start of November; marked as D), and Ulm (38 days, November/December; marked as U)), FEO recommendations were used to prevent service cancellations of approximately 28,200, 21,300, 8,100, and 304,000 kilometers (km), respectively, (i.e., 360,600 total train km). Extrapolating this to an annual value, FEO allows DB Fernverkehr to drive roughly 1 million additional train km with the existing fleet. This equals the annual mileage of two ICE double-traction train sets and would generate additional fare revenue of approximately 24 million € per year or equate to the transport of 784,000 additional travelers. Each train journey saves around 44 kilograms of CO_2 in comparison with the amount that would have been expended had each passenger driven an automobile. Thus, FEO saves approximately 34,000 tons of CO_2 emissions per year.

The construction sites pose a serious problem because of the significant planning effort that must be made for each of them and for their combinations in the face of constant time shifts. Considering 850 construction sites per day, avoiding business losses related to managing around these sites is extremely important, and FEO has been of considerable help in preventing these losses.

Tendering Regional Passenger Services

All regional passenger traffic in Germany is tendered by the states of Germany. Train operating companies (TOCs) such as DB's regional transport division, DB Regio AG, can apply in these tenders; the TOC that requires the lowest public subsidy in its offer usually wins the contract for a period of up to 15 years. The long length of the terms reflects the refinancing

periods of the vehicle procurement costs. These costs are enormous and can range from 50,000,000 to 400,000,000 € for a single order. To win a tender, a TOC must determine the size and type of fleet that results in the most favorable operational plan. For this purpose, a TOC must study a large number of fleet variants (often 5–30 variants) in a short time (e.g., two to three weeks). Each variant requires studying different scenarios and operational procedures, performing sensitivity analyses, and forecasting delays.

We describe three documented cases in which FEO was instrumental in winning a tender for the state of Baden-Württemberg in the southwest of Germany. These are the tenders Netz 3b from Stuttgart to Crailsheim (2.1 million train km per year) (Ministry of Transport Baden-Württemberg 2015), Netz 4 (4.8 million km per year) (Ministry of Transport Baden-Württemberg 2017), and Netz 7b (4.5 million km per year) (Ministry of Transport Baden-Württemberg 2019), which together cover a mileage of roughly 11 million train km per year; see Figure 15.

DB Regio won the Netz 3b tender by implementing all timetable and vehicle reserve requirements using one vehicle *fewer than the intrinsic minimum fleet size*: that is, 16 BR 442 Bombardier Talent 2 one-story electric railcars. Undercutting the minimum fleet size enabled DB Regio to offer services at a price lower than its competitors. The tender was designed to permit small service tasks (e.g., cleaning) on reserve vehicles. The vehicles could be repositioned for this purpose, and such a repositioning would not interrupt the required reserve time window of 24 hours. However, the terms also permitted a deadhead trip of a reserve vehicle between two stations within the reserve time. One of DB Regio's planning experts noticed this and realized that FEO could exploit it to construct a solution that saved one vehicle. Although this solution was considered to be unorthodox and the crucial deadhead option was disallowed for future tenders, it was instrumental in helping DB Regio to win the Netz 3b tender.

Extensive coupling/uncoupling was the key to winning the Netz 4 tender. The tender consisted of two lots. Lot 1 was for fast regional express trains between Karlsruhe, Offenburg, and Basel. Lot 2 was for the "Breisgau S-Bahn" commuter trains, which stop at every station and need strong acceleration. Demand was high around the three regional centers of Karlsruhe, Freiburg, and Basel but low in between these centers. The tender therefore stipulated capacities that varied between 800 and 400 seats on the respective segments. These requirements could be satisfied by always using large vehicles or by alternating between single, double, and triple tractions of small vehicles. The latter concept requires permanent coupling and uncoupling, which is much more

Figure 14. The Graphics Show That the Cologne-Rhine/Main Construction Site Caused a Disruption in a Major Link in the German ICE Network
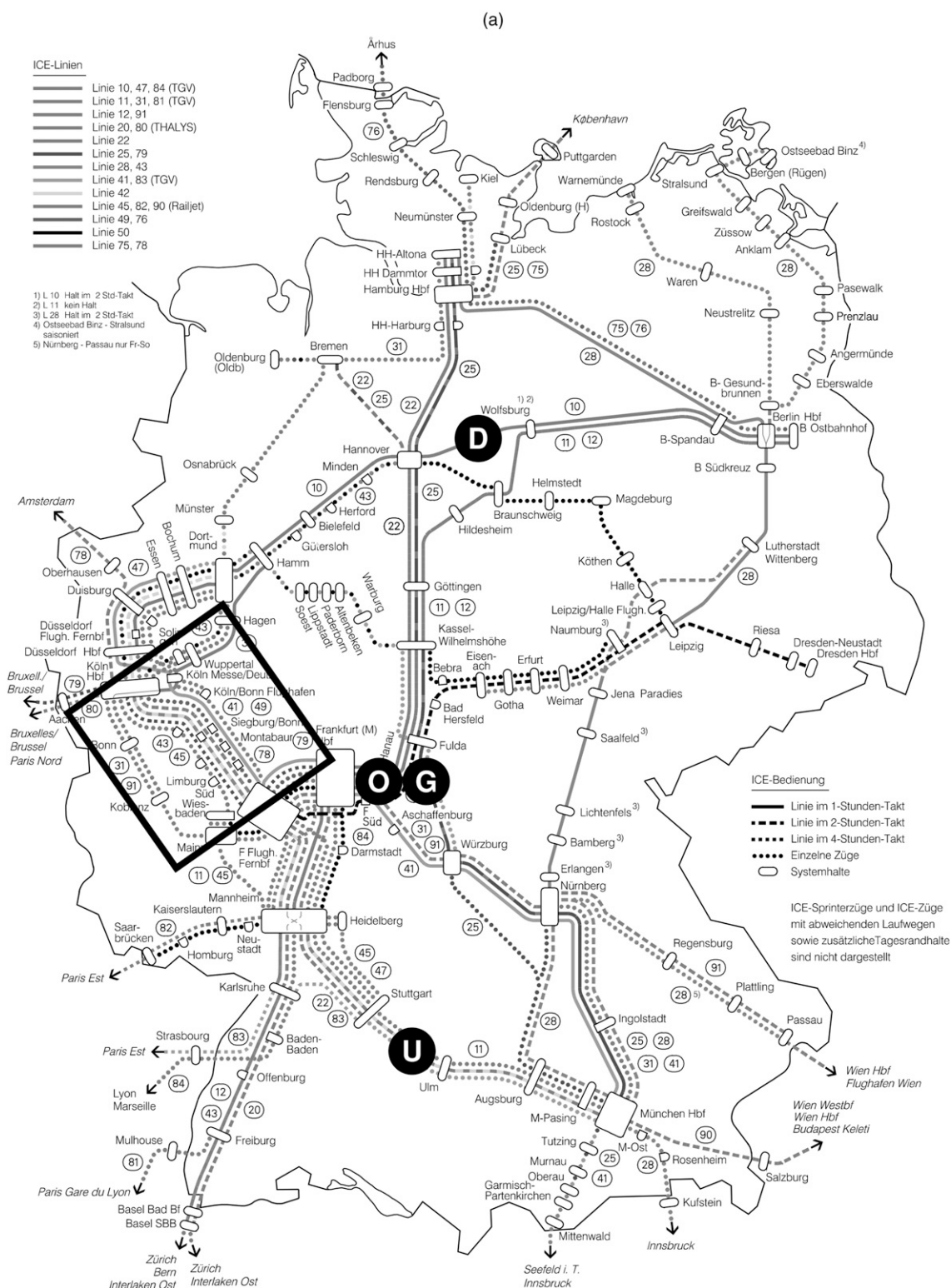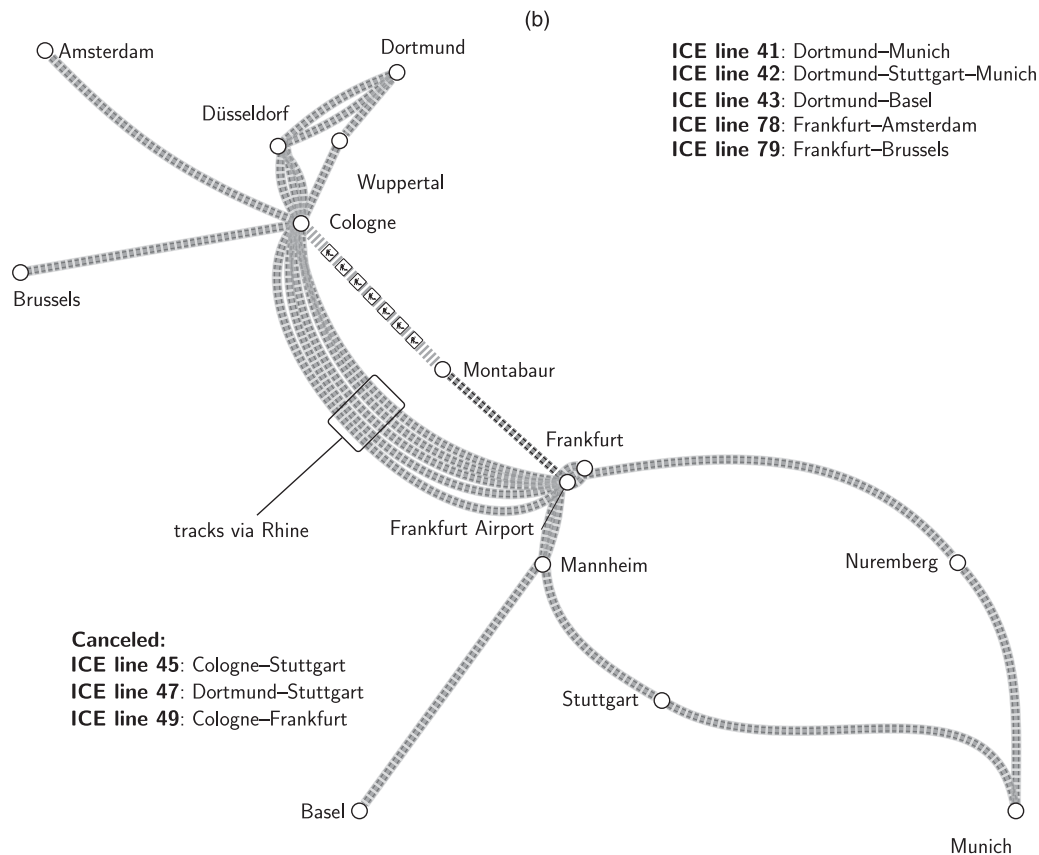

(a)

Figure 14. (Continued)

(b)

ICE line 41: Dortmund–Munich
ICE line 42: Dortmund–Stuttgart–Munich
ICE line 43: Dortmund–Basel
ICE line 78: Frankfurt–Amsterdam
ICE line 79: Frankfurt–Brussels

Canceled:
ICE line 45: Cologne–Stuttgart
ICE line 47: Dortmund–Stuttgart
ICE line 49: Cologne–Frankfurt

Notes. Construction work on 117 km of the high-speed link between Cologne and Frankfurt (see panel (b) and the square inset in panel (a)) on four successive weekends affected almost the entire ICE1 fleet, which was forced to take a detour along the scenic, but slow, old Rhine valley route. The driving time increase of more than 60 minutes was too high to allow a shift of connections by only one timetable period, which is 60 minutes. Using FEO, a "minimally invasive" solution was constructed that left one line (82 to Paris) unaffected. This was important because the vehicles that travel on this line require special electric power systems and train control equipment to meet the varying safety standards of the individual countries in Europe. The solution involved canceling the weekend trips of three lines (45, 47, and 49) but operating the remaining five lines (41, 42, 43, 78, and 79) on their usual frequencies, complete with the transitions into and out of the four weekends on which construction work was occurring.

complicated to operate and requires a second engineer; however, it reduces mileage and maintenance costs and is the most efficient solution. DB Regio won both lots. The key to this success was to use vehicles with excellent driving dynamics: in this case, 15 BR 462 Siemens Desiro HC electric railcars with 410 seats for Lot 1 and 24 BR 463 Siemens Mireo electric railcars with 215 seats per vehicle for Lot 2. These vehicles could be operated in such a way that specific arrival times could be advanced by one to two minutes. These timetable changes in turn enabled coupling and uncoupling operations at stations, at which (un-)coupling was not possible using the "traditional" timetable included in the tender. FEO provided the analytics to verify the feasibility of this complex operation and to evaluate the financial impact of this novel and superior concept.

DB Regio won the Netz 7b tender in essentially the opposite way. The tender consisted of two lots.

The Complex Lot 1 was for the "Karlsruher Netz." It seemed to require a mixed operation of single and double tractions to serve varying capacities between 140 and 280 seats similar to Netz 4. Likewise, Lot 2, the "Nordbaden Express," stipulated capacities between 200 and 400 seats. Again, coupling is important, but the timetable in this case favored the use of large vehicles, although their use required significant and counterintuitive overcapacities on some segments. Over the tendering period of 13 years, the higher procurement costs of large vehicles are outweighed by smaller operational costs; that is, replacing couplings by overcapacities is cost effective. To win this tender, DB Regio acquired 19 BR 440 Alstom Coradia Continental electric railcars (6 large ones with five units and 13 small ones with three units) for Lot 1 and 7 BR 463 Siemens Mireo electric railcars with three units for Lot 1.

To distinguish between cases such as the Netz 4 and Netz 7b tenders, which initially seemed similar, precise operational costs had to be computed over a period of more than a decade. These calculations would have been impossible without FEO. However, the cases also testify to the importance of human creativity. The successful offers relied on innovative ideas. In all three cases, the same expert planner identified pivotal optimization opportunities and used the FEO decision support system with its powerful optimizer to explore, evaluate, and finally implement these concepts. This represents DB Regio's best practice in tendering.

Crossborder Freight Rail Operations

DB's freight transport division, DB Cargo AG, uses FEO for strategic and operational locomotive scheduling. In contrast to passenger transport, the rail freight business is highly demand driven. Transports are volatile, and operations often do not follow synchronized timetables on standardized train routes. The main challenge is crossborder operations, which constitute more than 60% of DB Cargo's transports (Figure 16). Because the European railway infrastructure has not yet been harmonized (e.g., the electrical power systems, the signaling/train

Figure 15. Network 3b (Black; East) Covers the Route from Stuttgart to Crailsheim, Network 7b (Dark Gray and Light Gray; Center and North, Respectively) Contains Services in the Karlsruhe Area, and Network 4 (Medium Gray; West) Is the Connection from Karlsruhe to Basel
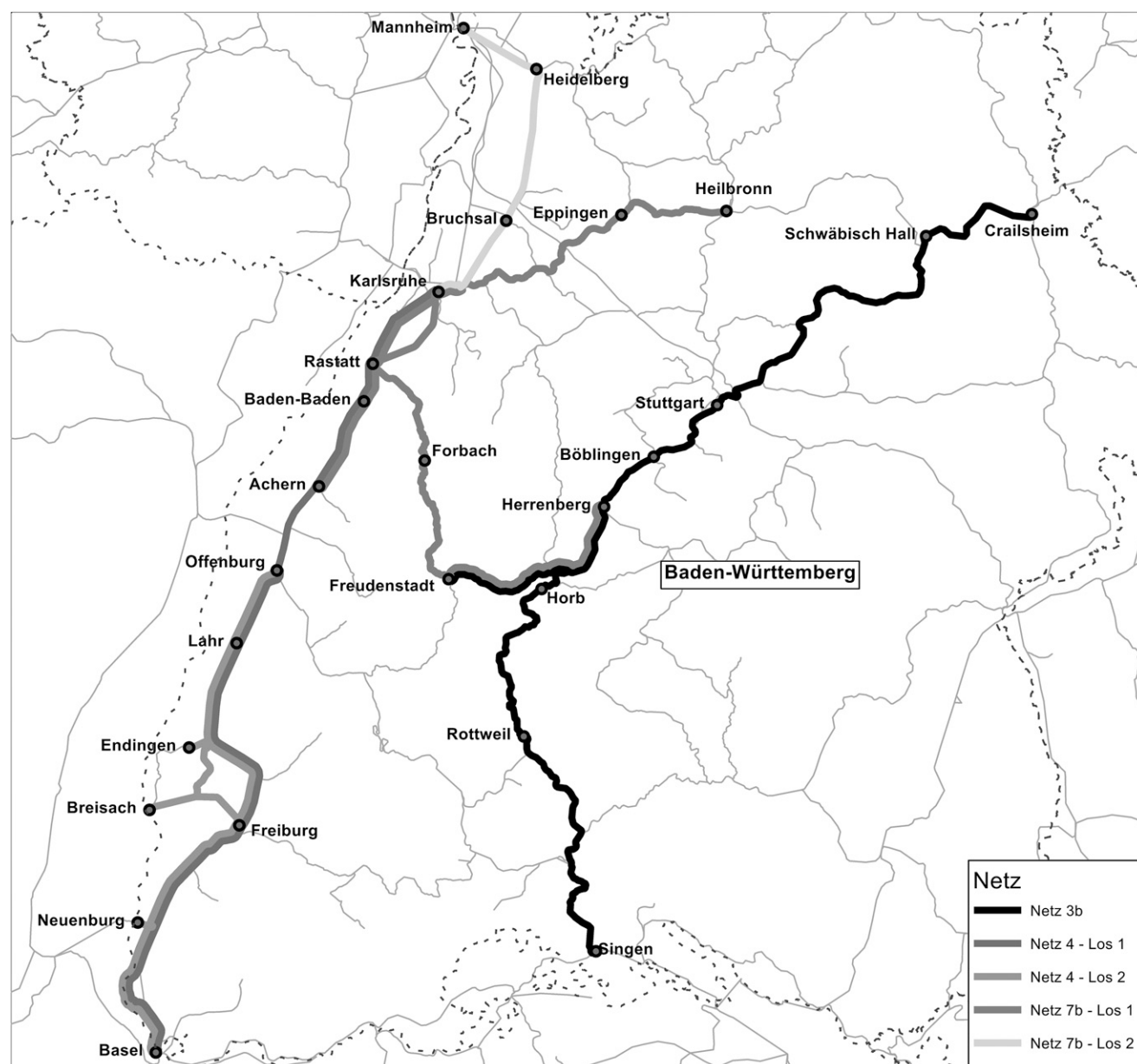
control systems, and the safety standards differ), engines must be changed if they are not interoperable; this incurs coupling work (Figure 17). Traffic from the large North Sea harbors of Rotterdam and Hamburg along the Rhine and across the Alps to the industrial centers of northern Italy crosses several countries and hence, can require several locomotive changes. Because the price of a locomotive ranges between 2 and 5 million €, efficiently employing these resources is critical.

DB Cargo started to use FEO in 2010 to optimize the locomotive operations in Belgium and France. At that time, more than 10 locomotive types were in use, many of them from the first generation of electric locomotives from the 1960s (locomotives can be used for several decades). The old locomotives were long depreciated and so, were inexpensive from that financial perspective. However, they used considerable energy, acquiring spare parts was a problem, and they required qualified drivers. A FEO analysis showed that putting 10–15 old locomotives out of service would be cost effective, necessarily increasing

the mileage of the modern ones. Implementing this solution was a forerunner of later locomotive standardization at DB Cargo.

The Belgium Hinterland traffic was further optimized in 2011. It consists of oil, gas, automotive, and intermodal (i.e., container) transports, which are often time critical. Using FEO's timetable improvement feature, 3 trains (of over 400) were identified, such that shifting them by as little as one hour would require two fewer locomotives. Implementing the shifts involved negotiations with the customer and the infrastructure provider. The timetable improvement was a huge success and has subsequently become a primary FEO application at DB Cargo.

Alpine transit operations were optimized in 2016 in connection with the opening of the Gotthard basis tunnel. These operations, which feature heavy trains loaded with containers, steel, scrap, minerals, and tiles, were operated using a diverse locomotive fleet. By an integrated optimization of these transports, including a timetable improvement, the number of locomotives required was reduced from 30 to 25.

Figure 16. (Color online) DB Cargo Uses Interoperable Locomotives in Its Extensive European Crossborder Operations
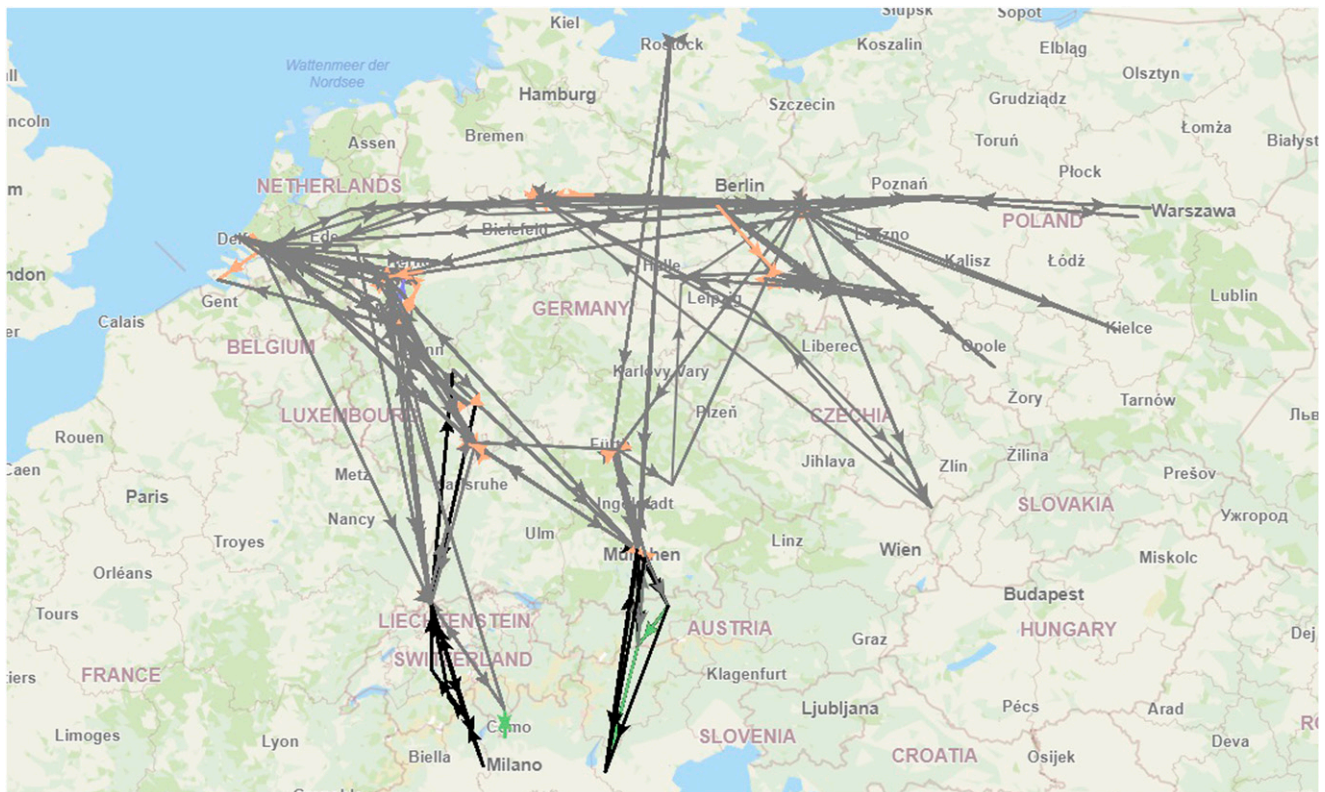
Notes. The Alps can be crossed using the Gotthard or the Lötschberg/Simplon passes in Switzerland ("over the Alps") or the Lötschberg/Simplon and the new Gotthard Basis tunnels ("under the Alps"). The tunnels charge higher track fees, but they provide an almost level route without special traction requirements, and the routes are much faster than the passes; going over the Alps is a necessity to circumnavigate construction sites; however, doing so requires more traction force (e.g., a stronger, a second, or even a third locomotive). Additional locomotives can be used to push or pull. The feasible train compositions depend on the inclination of the track, the weight and the length of the train, and the delivery deadline.

Figure 17. (Color online) Arduous Coupling Work Is a Major Problem in Crossborder Operations

All of these results testified to the significant benefits that could be derived by overcoming assignments of dedicated locomotives to individual services. However, this traditional concept is not devoid of advantages. Such transports are extremely reliable because they are not impacted by delays caused by other customers, which is a serious problem in rail freight transport. Using dedicated locomotives ensures a high quality of service and a high level of customer satisfaction. Would it be possible to establish a more integrated and more resource-efficient operation with a reasonable level of quality? This key question could, for the first time, be analyzed with FEO and was answered in the affirmative. One consequence was the introduction of a hub-and-spoke system in which the expensive intermodal locomotives circle between long-distance hubs and cheaper locomotives implement national distribution in the spokes. These concepts were successful and achieved an average increase in efficiency of between 3% and 5%. Extrapolated to the entire fleet of more than 2,700 locomotives, this amounts to annual cost savings of 27 million €.

DB Cargo is now moving toward integrated operations with standardized locomotives. This new strategy became palpable in DB Cargo's biggest strategic investment during the past decades: that is, the procurement of a new fleet of 104 interoperable third-generation BR 193 "Vectron" locomotives for standardized crossborder operations. The support of this procurement was one of the biggest successes attributable

to FEO. As compared with the ICE4 procurement at DB Fernverkehr, the problem involved not only determining the number of locomotives and their types (a locomotive can be equipped for only five to six European countries, not for all) but also, redesigning the crossborder operations per se. Using FEO, DB Cargo was able to address this problem. The first of the new engines has been in use since spring 2019 and has already improved on-the-job safety by eliminating 600 coupling operations per year.

The increasingly integrated and standardized production is revolutionary for DB Cargo's operations. It requires a constant and careful balancing of resource integration and customer satisfaction subject to complex operational constraints in a volatile market. FEO showed that this strategy is viable and helped DB Cargo to implement it.

Summary of Benefits

The optimization of railway systems taps significant, but previously unused, potential to increase the efficiency and the quality of rail transport services. These stem from an optimization of existing production processes as well as from the ability to explore new concepts. Mastering the combinatorial complexity, the integrated use of resources, and improved scheduling processes are key components. We have shown how the use of advanced mathematical optimization methods in combination with human creativity resulted in the following benefits, which can be attributed directly to FEO: (1) annual savings/revenue increases of 74 million € (i.e., 24 million € at DB Fernverkehr, 23 million € at DB Regio, and 27 million € at DB Cargo); (2) savings in CO_2 emissions of 34,000 tons per year; (3) elimination of 600 coupling operations per year; (4) increased regularity of operations when managing construction site delays; (5) greatly increased flexibility, ease, and speed of planning processes; (6) new company-wide planning practices; (7) evaluation and implementation of new and superior production concepts; and (8) progress in railway optimization and the implementation of algorithmic hypergraph theory.

FEO's hypergraph-based operations research approach addresses the key problem of rolling stock rotation optimization and provides DB's planners with a powerful tool that helps them to reinvent the rail transport of the twenty-first century. Of all means of transportation, railways have the greatest potential to become the comfortable, fast, efficient, and environmentally friendly conveyance of the future. Operations research provides the methods to allow companies to make smarter planning and control decisions. Rail transport is on the right track—for the people, the economy, and the environment.

Appendix. The Hyperflow Model
We show the MIP model that FEO uses. We list the terminology.

(V, H) Hypergraph with nodes V and hyperarcs H; hyperarc h carries cost $c_h \in \mathbb{R}$;

(V, A) Underlying standard graph with nodes V and arcs A;

(V, C) Trunk graph with nodes V and arcs C;

T Set of (cyclic) timetabled trips;

R Set of resources, visiting node v in hyperarc h consumes $\omega_{vh}^r \in \mathbb{R}_{\geq 0}$ units of resource r with upper bound $u^r \in \mathbb{R}_{\geq 0}$;

Sr Set of service stations (e.g., maintenance facilities) for resource $r \in R$;

F Set of fleets (i.e., vehicle types); fleet f has maximum size $u^f \in \mathbb{N}_0$;

P Set of parking facilities; parking p has capacity $u^p \in \mathbb{N}_0$;

K Set of "trunks" (i.e., trip sequences that must be serviced by a common "trunk" vehicle to provide a direct passenger connection); trunk k must be serviced by $\ell^k \in \mathbb{N}_0$ vehicles.

For a node v and a set of arcs $E \in \{H, A, C\}$, we denote by $\delta_E^+(v)$ and $\delta_E^-(v)$ the arcs from E that go out of and into node v, respectively, and we denote by $E(*)$ the set of arcs of type $E \in \{H, A, C\}$ that are associated with the entity $*$; for example, $H(t)$ denotes all hyerarcs that cover trip t. Introducing binary variables x for the hyperflow, real variables w for a resource flow, and binary variables y for a trunk flow, the MIP model reads as follows:

$$\min_{x(H(t))} c^t x = 1 \quad \forall t \in T \quad \text{trip covering}$$

$$x(\delta_H^+(v)) = x(\delta_H^-(v)) \quad \forall v \in V \quad \text{vehicle flow}$$

$$w_a^r \leq u^r x(H(a)) \quad \forall a \in A, \ r \in R$$
$$\text{resource coupling}$$

$$w^r(\delta_A^+(v)) - \sum_{h \in \delta_H^+(v)} \omega_{vh}^t x_h = w^r(\delta_A^-(v)) \forall v \in V, r \in R \quad \text{resource flow}$$

$$w^r(\delta_A^+(v)) - w_v^r(\delta_H^+(v)) = 0 \quad \forall r \in R, \ v \in S^r \quad \text{resource reset}$$

$$x^f(\delta_A^+(\text{Mo } 0:00)) \leq u^f \quad \forall f \in F \quad \text{fleet size}$$

$$x(H(p)) \leq u^p \quad \forall p \in p \quad \text{parking capacity}$$

$$y_a \leq x(H(a)) \quad \forall a \in C \quad \text{trunk coupling}$$

$$y(\delta_C^+(u)) = y(\delta_C^-(u)) \quad \forall u \in U \quad \text{trunk flow}$$

$$y(C(k)) \geq \ell_k \quad \forall k \in K \quad \text{trunk start}$$

$$x \in \{0,1\}^H \quad \text{hyperflow integrality}$$

$$w \in \mathbb{R}_{\geq 0}^{A \times R} \quad \text{resource nonneg.}$$

$$y \in \{0,1\}^E \quad \text{trunk integrality.}$$

The objective minimizes the total cost of all hyperarcs. The covering constraints ensure that every timetabled trip is covered by exactly one hyperarc. The vehicle flow constraints produce a hyperflow (actually a hypercirculation). Together with the integrality requirements for the hyperarc variables x, these constraints constitute the combinatorial part of the model, as we describe in this paper. The remainder deals with resource and trunk constraints. The resource variables w sum up resource consumptions along vehicle rotations according to the resource flow constraints but only on standard arcs underlying the hyperflow, as enforced by the resource coupling constraints. These also address upper bounds on resource consumptions. The resources are reset at the service nodes via the resource reset constraints. The trunk graph consists of a set of disjoint paths of standard arcs, which model a contiguous service by the same vehicle in terms of the y variables. The trunk start constraints initiate such paths, the trunk flow constraints extend them, and the trunk coupling constraints enforce suitable train compositions to support the trunk paths. Reuther (2017) includes a detailed discussion of the model.

References
Ahuja RK, Liu J, Orlin JB, Sharma D, Shughart LA (2005) Solving real-life locomotive scheduling problems. *Transportation Sci.* 39(4): 503–517.

Beckenbach I (2018) A hypergraph network simplex algorithm. Kliewer N, Ehmke JF, Borndörfer R, eds. *Operations Research Proceedings 2017* (Springer International Publishing, Cham, Switzerland), 309–315.

Beckenbach I (2019) Matchings and flows in hypergraphs. Doctoral dissertation, Freie Universität Berlin, Berlin.

Beckenbach I, Borndörfer R (2018) Hall's and Kőnig's theorem in graphs and hypergraphs. *Discrete Math.* 341(10):2753–2761.

Borndörfer R, Heismann O (2015) The hypergraph assignment problem. *Discrete Optim.* 15(February):15–25.

Borndörfer R, Reuther M, Schlechte T (2014) A coarse-to-fine approach to the railway rolling stock rotation problem. Funke S, Mihalák M, eds. *Proc. 14th Workshop Algorithmic Approaches Transportation Model Optim. Systems OpenAccess Series Informatics (OASIcs)*, vol. 42 (Schloss, Dagstuhl, Germany), 79–91.

Borndörfer R, Grimm B, Reuther M, Schlechte T (2017) Template-based re-optimization of rolling stock rotations. *Public Transport* 9(1–2):365–383.

Borndörfer R, Grimm B, Reuther M, Schlechte T (2019) Optimization of handouts for rolling stock rotations. *J. Rail Transport Planning Management* 10(August):1–8.

Borndörfer R, Reuther M, Schlechte T, Weider S (2011) A hypergraph model for railway vehicle rotation planning. Caprara A, Kontogiannis S, eds. *Proc. 11th Workshop Algorithmic Approaches Transportation Model Optim. Systems OpenAccess Ser. Informatics (OASIcs)*, vol. 20 (Schloss, Dagstuhl, Germany), 146–155.

Borndörfer R, Reuther M, Schlechte T, Waas K, Weider S (2015) Integrated optimization of rolling stock rotations for intercity railways. *Transportation Sci.* 50(3):863–877.

Borndörfer R, Klug T, Lamorgese L, Mannino C, Reuther M, Schlechte T, eds. (2018) *Handbook of Optimization in the Railway Industry* (Springer International Publishing, Cham, Switzerland).

Cacchiani V, Caprara A, Toth P (2019) An effective peak period heuristic for railway rolling stock. *Transportation Sci.* 53(3):746–762.

Federal Environment Agency (2012) Daten zum Verkehr. Accessed August 25, 2020, https://www.umweltbundesamt.de/sites/default/files/medien/publikation/long/4364.pdf.

Federal Ministry of Transport and Digital Infrastructure (2020) Verkehr in Zahlen 2019/2020. Accessed August 3, 2020, https://www.bmvi.de/SharedDocs/DE/Publikationen/G/verkehr-in-zahlen-2019-pdf.

Giacco GL, D'Ariano A, Pacciarelli D (2014) Rolling stock rostering optimization under maintenance constraints. *J. Intelligent Transportation Systems Tech. Planning Oper.* 18(1):95–105.

Grimm B, Borndörfer R, Reuther M, Schlechte T (2019) A cut separation approach for the rolling stock rotation problem with vehicle maintenance. Cacchiani V, Marchetti-Spaccamela A, eds. *Proc. 19th Symposium Algorithmic Approaches Transportation Model Optim. Systems OpenAccess Ser. Informatics (OASIcs)*, vol. 75 (Schloss, Dagstuhl, Germany), 1:1–1:2.

Holland C, Levis J, Nuggehalli R, Santill B, Winters J (2017) UPS optimizes delivery routes. *Interfaces* 47(1):8–23.

Ireland P, Case R, Fallis J, Van Dyke C, Kuehn J, Meketon M (2004) The Canadian Pacific Railway transforms operations by using models to develop its operating plans. *Interfaces* 34(1):5–14.

Kroon L, Huisman D, Abbink E, Fioole P-J, Fischetti M, Maróti G, Schrijver A, Steenbeek A, Ybema R (2009) The new Dutch timetable: The OR revolution. *Interfaces* 39(1):6–17.

Lusby RM, Haahr JT, Larsen J, Pisinger D (2017) A branch-and-price algorithm for railway rolling stock rescheduling. *Transportation Res. Part B: Methodological* 99(May):228–250.

Ministry of Transport Baden-Württemberg (2015) DB is awarded the contract for local rail transport in network 3b between Crailsheim and Konstanz/Freudenstadt. Accessed August 4, 2020, https://vm.baden-wuerttemberg.de/de/service/presse/pressemitteilung/pid/db-erhaelt-den-zuschlag-fuer-spnv-im-netz-3b-zwischen-crailsheim-und-konstanzfreudenstadt/.

Ministry of Transport Baden-Württemberg (2017) Local traffic on the Rhine will be expanded significantly and faster. Accessed August 4, 2020, https://vm.baden-wuerttemberg.de/de/service/presse/pressemitteilung/pid/nahverkehr-auf-der-rheinschiene-wird-deutlich-ausgeweitet-und-schneller/.

Ministry of Transport Baden-Württemberg (2019) Karlsruher Netze: DB Regio is to receive a contract for transport services. Accessed August 4, 2020, https://vm.baden-wuerttemberg.de/de/service/presse/pressemitteilung/pid/karlsruher-netze-db-regio-soll-zuschlag-fuer-verkehrsleistungen-erhalten/.

Öztürk A (2015) Implementierung eines rolling horizon ansatzeszur anwendung bei derlokomotiveinsatzoptimierung. Bachelor's thesis, Helmut-Schmidt-Universität Hamburg, Hamburg, Germany.

Reuther M (2017) Mathematical optimization of rolling stock rotations. Doctoral dissertation, Technische Universität Berlin, Berlin.

Ralf Borndörfer is a professor of discrete mathematics at Freie Universität Berlin, head of the optimization department at Zuse Institute Berlin, and a cofounder of LBW Optimization GmbH. He holds a master's degree in mathematics with economics from Augsburg University and graduated and habilitated at Technische Universität Berlin. His research interests include combinatorial optimization, integer programming, and their applications: in particular, in transportation and traffic.

Thomas Eßer is responsible for the operation and use of information technology solutions at Deutsche Bahn (DB) Cargo's resource planning department. He was also project manager for the implementation of the Lok-Einsatz-Optimierung (Locomotive Optimization) (LEO). He joined DB Cargo's European Service Design division in 2008 and worked there in different roles. At Katholische Universität Eichstätt–Ingolstadt, he studied business administration with a focus on operations research and service design.

Patrick Frankenberger is a software developer and software architect at DB Analytics at Deutsche Bahn. He is responsible for the development of several decision support systems that integrate optimization algorithms into end-user software. He also is the tech lead of the fleet employment optimization (FEO) development team. He is a graduate of computer science at Technische Universität Darmstadt.

Andreas Huck is a mathematician at DB Analytics. He is developing optimization algorithms especially for solving transportation problems. Before joining Deutsche Bahn AG, he was an assistant professor of mathematics at the University of Hannover (Germany), where he also obtained his doctoral degree and his habilitation. His main research interest was graph theory. He is author of several papers and he made contributions to some famous graph theoretical conjectures such as the cycle double cover conjecture.

Christoph Jobmann joined the department of DB Analytics at Deutsche Bahn in 2007. He has dealt with software testing, customer communication, end-user support, and software development. The main focus was always on the software that is nowadays known as FEO. Before joining DB Analytics, he studied mathematics and applied computer science at the University of Göttingen, graduating with a diploma and MSc, respectively.

Boris Krostitz is the responsible lead of DB Analytics, which is the department for transport modelling, OR, data science, and simulation within Deutsche Bahn Group. In this role, he focuses on the development of algorithmic approaches to get a better understanding of the transport market and to suggest algorithmic based decisions to raise the modal split and economic success of the German railway system. He earned his PhD at the University of Stuttgart in the modeling of metropolitan transport systems.

Karsten Kuchenbecker is the responsible manager of the transport modelling and simulation system group of DB Analytics. His models are used to understand and optimize processes in a wide field of projects, starting from passenger flow analysis within railway stations to nationwide railway simulations. He earned his PhD at the Karlsruhe Institute of Technology with the focus on the combination of classic strategic transport modelling approaches with a feedback-loop-oriented system dynamic modelling.

Kai Mohrhagen studied mathematics at the Ruprecht-Karls-University of Heidelberg with focus on graph theory and hyper graphs. He earned his PhD on his work in neurophysiology at the Max-Delbrück-Centrum Berlin. Later on, he became a professional software developer and now leads software development teams at Deutsche Bahn as software architect. He specializes in the creation of software implementing complex mathematical algorithms and methods including but not limited to operations research.

Philipp Nagl is chief operating officer (COO) of Deutsche Bahn (DB) Fernverkehr AG. He is responsible for the operations of almost 900 InterCity Express high-speed and intercity trains in Europe per day. Before that, he spent a number of

years in various leadership functions at DB and Österreichische Bundesbahnen (ÖBB). He holds a doctoral degree in economics from Vienna University of Economics and Business.

Michael Peterson is the chief executive officer (CEO) of Deutsche Bahn Fernverkehr AG. He joined DB Fernverkehr in 2014 to build up the product management department and became chief marketing officer (CMO) in 2016 and CEO in 2019. Prior to joining DB, he was a partner at Booz & Company. He is a business engineer, having graduated from the Karlsruhe Institute of Technology and he holds a doctorate degree in business administration from the University of Oldenburg.

Markus Reuther is an mathematical optimization expert at LBW Optimization. After receiving his master's degree in mathematics, he has graduated on novel algorithmic models and methods for complex railway applications at Technische Universität Berlin. His PhD thesis is the mathematical core of the Rotation Optimization for Railways (RotOR) software. He initially programmed RotOR at Zuse Institute Berlin and is now responsible for its further development and maintenance at LBW Optimization.

Thilo Schang is a team leader in the principles of capacity management division (Grundsätze Kapazitätsmanagement) at the German railway infrastructure operator DB NETZ AG. He began his career as an operations research specialist in the DB Analytics division at Deutsche Bahn AG after graduating with a diploma in mathematics at the Technical University of Darmstadt.

Michael Schoch is responsible for software development and IT infrastructure at the DB Analytics department of Deutsche Bahn AG (German railway). He studied computer science at the University Karlsruhe and obtained his PhD in economics on the assessment of European transport corridors. After he joined Deutsche Bahn, he managed several software development projects with a focus on resource planning based on mathematical optimization.

Hanno Schülldorf works as project manager and consultant at the DB Analytics department at Deutsche Bahn. He has worked on several operations research projects for all divisions within DB and is product owner of the software

FEO. Moreover, he is research coordinator and supervises research and development projects with several universities. He is a graduate in mathematics at Technische Universität Darmstadt.

Peter Schütz is a business engineer, having graduated from Technical University of Darmstadt with a master's degree in 1994. Since 1996, he has held several leading positions in passenger transportation at Deutsche Bahn. From 2005 to 2013, he was responsible for timetabling and vehicle scheduling for the entire fleet of long-distance trains in Germany. Since 2014, he has held the position of Chief Experience Officer of traveler information at Deutsche Bahn.

Tobias Therolf is a senior project manager at Deutsche Bahn (DB) Regio. He is responsible for the creation and calculation of competitive offers in public procurement procedures for regional transport contracts in the southwest of Germany. Since 2014, he has been responsible for the operational components of the offers, in particular for vehicle and crew planning. He and his team have often been instrumental in determining the best vehicle concept through the use of fleet employment optimization.

Kerstin Waas is an information technology project manager at Deutsche Bahn AG. She graduated with a PhD in mathematics from the University of Bielefeld. At DB, she has been managing IT projects for several different divisions, working on projects such as timetable planning, vehicle rotation planning, and rail traffic control. She was jointly responsible for the development, implementation, and rollout of the fleet employment optimization system for long-distance passenger railways.

Steffen Weider is cofounder and manager of LBW Optimization GmbH. He obtained his master's degree in mathematics with economics from Technical University Berlin. He wrote his PhD thesis about integrating vehicle and duty scheduling for public transit. These methods are now part of a commercial planning suite that is used in public transit companies all over the world. He was part of the research team that developed the Rotation Optimization for Railways (RotOR) software.