



Innovative Applications of O.R.

## A rolling horizon approach for disruption management of railway rolling stock

Lars Kjær Nielsen<sup>a</sup>, Leo Kroon<sup>b,c</sup>, Gábor Maróti<sup>c,\*</sup><sup>a</sup> Quintiq Applications, 's Hertogenbosch, The Netherlands<sup>b</sup> Rotterdam School of Management, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands<sup>c</sup> NS Reizigers, Process Quality and Innovation, P.O. Box 2025, 3500 HA Utrecht, The Netherlands

## ARTICLE INFO

## Article history:

Received 17 May 2010

Accepted 20 January 2012

Available online 28 January 2012

## Keywords:

Transportation

Railway optimization

Rolling stock

Disruption management

Rolling horizon

Integer programming

## ABSTRACT

This paper deals with real-time disruption management of rolling stock in passenger railway transportation. We describe a generic framework for dealing with disruptions of railway rolling stock schedules. The framework is presented as an online combinatorial decision problem, where the uncertainty of a disruption is modeled by a sequence of information updates. To decompose the problem and to reduce the computation time, we propose a rolling horizon approach: rolling stock decisions are only considered if they are within a certain time horizon from the time of rescheduling. The schedules are then revised as time progresses and new information becomes available. We extend an existing model for rolling stock scheduling to the specific requirements of the real-time situation, and we apply it in the rolling horizon framework. We perform computational tests on instances constructed from real-life cases of Netherlands Railways (NS), the main operator of passenger trains in the Netherlands. We explore the consequences of different settings of the approach for the trade-off between solution quality and computation time.

© 2012 Elsevier B.V. All rights reserved.

## 1. Introduction

The planning and operation of a busy passenger railway network is a complex task. Rolling stock and crew have to be scheduled to serve the timetable with ever growing demand for capacity. This has led to extensive research on optimizing the utilization of railway resources such as the infrastructure, the rolling stock and the crew, see [Caprara et al. \(2007\)](#) and [Huisman et al. \(2005\)](#). The developed methods have resulted in resource schedules that are highly efficient when operations run as planned, see [Kroon et al. \(2009\)](#). However, when operating a dense timetable, operations occasionally have to deviate from the plans. Currently, such deviations are handled manually under large time pressure. As a consequence, there is a demand for decision support systems for effectively dealing with the challenges posed by the real-time operations of a passenger railway system. However, to this date only limited research has been conducted on the real-time rescheduling of railway resources.

Irregularities of the railway operations are more or less inevitable and require real-time rescheduling of the system. Practitioners distinguish minor and major incidents; minor incidents are called *disturbances*, while major incidents are called *disruptions*. The distinction is not exactly defined, but depends on the impact on the operations. As a rule of thumb, disruptions are incidents that require significant changes of the pre-set resource schedules.

Disruptions may be caused by various internal or external factors such as a faulty switch on a busy track, broken down rolling stock, or damaged overhead wires. In a disrupted situation, the planned resource schedules are no longer feasible and will have to be updated to take the actual situation into account. Disturbances, on the other hand, only need simple recovery measures. An example of a disturbance is a delay caused when the boarding of passengers at a station takes unexpectedly long. Disturbances of such kind are either absorbed by the slack in the system or can be handled by small changes in the timetable.

Disruption management of railway resources often relies on an updated timetable. In fact, the utilization of the railway infrastructure of many European countries is managed by an independent authority, while the operators are responsible for the resource schedules. In the Netherlands, hundreds of so called *handling scenarios* have been defined, in order to speed up the timetable updating process. A handling scenario is a set of rules telling which timetable services are to be canceled, rerouted or delayed in case of a particular disruption. An example of a handling scenario is described in [Appendix A](#).

The challenges in real-time decision making differ from the static planning tasks in a number of aspects. The most important difference is that real-time rescheduling needs to deal with uncertain information as the environment evolves, see [Séguin et al. \(1997\)](#), [Grötschel et al. \(2001\)](#) and [Yu and Qi \(2004\)](#). In addition, real-time decisions must be made within a tight time frame. Finally, a highly optimized plan already exists, and any changes to that plan will have to be communicated to the involved parties. Decision support

\* Corresponding author. Tel.: +31 30 235 7763.

E-mail addresses: [lars.nielsen@quintiq.nl](mailto:lars.nielsen@quintiq.nl) (L.K. Nielsen), [lkroon@rsm.nl](mailto:lkroon@rsm.nl) (L. Kroon), [gabor.maroti@ns.nl](mailto:gabor.maroti@ns.nl), [gmaroti@rsm.nl](mailto:gmaroti@rsm.nl) (G. Maróti).

methods for real-time rescheduling need to address these aspects adequately.

### 1.1. Contribution of this paper

This paper deals with the real-time rescheduling of rolling stock in case of a disruption of the railway system. We aim at contributing to the central (i.e. network level) rolling stock rescheduling process while taking into account several aspects of the local (i.e. station level) operational details.

We first formulate a generic framework for the Rolling Stock Rescheduling Problem (RSRP) as a static decision making problem, and then we define its *online* variant in order to be able to deal with uncertain information. We propose a *rolling horizon approach* to solve the online RSRP. That is, the solution method amounts to rescheduling the rolling stock within a certain rolling horizon of limited length, and then periodically updating the plans as time progresses and more information becomes available. Rolling horizons are often used in various areas of operations management; we refer to Chand et al. (2002) for an extensive overview of such applications. Also, we note that our proposed framework and solution approach are generic in that they can be applied to any rolling stock rescheduling model.

The rolling horizon approach is a heuristic problem decomposition in order to pursue the overall goals of the online RSRP. The problem specification of online RSRP involves global criteria that cannot be taken into account directly when considering a single rolling horizon iteration. In this paper we propose a novel heuristic way to deal with such a criterion: to guide the rolling horizon solution process towards the desirable end-of-day rolling stock distribution.

An essential contribution of this paper is that we apply the generic framework to the specific realistic rolling stock rescheduling problems of Netherlands Railways (NS), the major passenger railway operator in the Netherlands. The method is based on an extension of the existing rolling stock scheduling model of Fioole et al. (2006) for the special needs of real-time rescheduling. The model of Fioole et al. (2006) is a well-tested integer linear programming model which has successfully been used by NS for medium-term planning since 2004.

The computational results indicate that the rolling horizon approach is capable to deal with fairly large problem instances. The outcome of the rolling horizon approach in an online setting turns out to be reasonably close to the off-line optimum. This also shows the effectiveness of the heuristic guidance for the desired end-of-day rolling stock distribution.

This paper is structured as follows. In Section 2 we set the terminology we use and we give a literature overview. Section 3 describes the rolling stock rescheduling process from a practical point of view. In Section 4 we describe the generic framework for the Rolling Stock Rescheduling Problem and for its online variant, and we propose the rolling horizon approach. In Section 5 we specialize the generic framework to real-life problems of NS. Section 6 is devoted to the computational results on a particular set of disruptions. Finally, we draw some conclusions in Section 7.

The paper is supplied with two appendices. In Appendix A we give a detailed example of a disruption on a train line of NS. In Appendix B we report further computational results on another set of realistic disruption instances of NS.

## 2. Terminology and literature

In this section we briefly describe the notions of the rolling stock operations at NS. The concrete models as well as the computational experiments are based on these assumptions. For more

details we refer to Fioole et al. (2006). Having set the terminology, we give a brief literature overview on rolling stock scheduling.

### 2.1. Railway rolling stock terminology

The passenger train services at NS are primarily operated by electrically powered *rolling stock units*. A unit is self-propelled and consists of a fixed number of carriages. We point out that this paper focuses on such units; locomotives and hauled carriages lead to substantially different (re) scheduling problems.

Units are divided into *rolling stock types* depending on their characteristics, and units of the same type are considered as interchangeable during planning. In the operations a few units may be singled out due to maintenance requirements, but the remaining ones are treated as interchangeable as well.

Rolling stock units can be combined to form *compositions*. This way capacity can better meet demand during peak hours. An important feature of NS is the fact that, due to efficiency reasons, train compositions are adapted during operations by uncoupling units from or coupling units to trains. We emphasize that the order of the units in a composition is essential because of the shunting possibilities. For example, if *a* and *b* are different rolling stock types, then *aab* and *aba* are different compositions, but with the same capacity.

A *trip* is a train service from a departure station to an arrival station at a specific time. A timetable is a set of trips. At NS, the timetable is cyclic, which means that in principle the same one hour timetable is repeated several times.

Each trip has a rolling stock connection to its *successor* trip. This means that, when a composition on a trip arrives at a station, it is specified by the timetable which trip is to be served next by these units, possibly after adjusting the composition. The rolling stock connections exist both along the line and at stations at the end of the line; the latter is known as *turning* the train. Two trips that form a rolling stock connection are called *consecutive* trips.

Note that real-life problem instances may contain trips with two successor trips, which means that one train is split into two trains. It may also happen that two trains have the same successor trip, that is, two trains are combined into one train. These situations can be handled with an extension of the model described in Section 5 of this paper (for more details we refer to Fioole et al., 2006). In fact, our computational tests use the extended model. However, we want to focus in this paper on the rescheduling framework, therefore we omit the splitting and combining of trains from the problem- and model formulations.

In this paper we only consider two kinds of *shunting operations* that are relevant for rolling stock planning, since they change the composition of a train: *uncoupling* of units from trains and *coupling* of units to trains. These are the operations that can be performed between consecutive trips; if none of them is applied, then the trip and its successor trip have the same composition. Usually only certain types of shunting operations are allowed between two trips: they mainly depend on the layout of the station. For example, the standard allowed operations are to couple train units to the front end of a train and to uncouple train units from the rear end of a train.

The details of the shunting operations are locally monitored and rescheduled by local dispatchers. This means that the full information on the track occupancy inside stations and shunting yards is not globally available during operations to the central dispatchers. As a consequence, any changes to the planned shunting operations by the central dispatchers must first be evaluated and accepted by local dispatchers to ensure local feasibility. From the central dispatchers' point of view, and thus throughout this paper, shunting is simplified to coupling and uncoupling operations.

The *duty* of a rolling stock unit is the set of trips that the unit will serve along with the position of the unit in the composition of each trip. The assignment of all rolling stock units to duties is called the *rolling stock circulation*. The number of available units of each type at a station at any given time is called the *inventory* of the station at that time. The inventory at the end of the day is called the *end-of-day rolling stock balance*. The end-of-day balance connects the rolling stock circulation to that of the next day.

The scope of the rolling stock rescheduling process usually lasts until the end of the day. When the operations deviate from the plan, some units may end up at different stations than planned. This may have undesirable consequences for the operations of the next day or it may lead to necessary repositioning of train units during night time. A deviation of one unit from the planned number of units of a specific type ending up at a station is known as an *off-balance*.

## 2.2. Literature overview

There is a rich list of publications on rolling stock problems in the early planning stages where resource schedules are created from scratch. The time horizon ranges from a couple of weeks to a few months. Caprara et al. (2007) provide a comprehensive overview of these papers. We mention here Alfieri et al. (2006), Fioole et al. (2006) and Peeters and Kroon (2008). The problems considered in these three papers have the same underlying structure and assumptions as the rescheduling problems in this paper.

There are several applications of short-term rescheduling where a schedule has to be adjusted in a deterministic setting with a planning horizon of a few days. These models, however, do not take into account the stochastic nature of the real-time operations. Budai et al. (2010) consider the rolling stock balancing problem of NS where the dominant rescheduling objective is to minimize the end-of-day off-balances. Lingaya et al. (2002) describe models for rescheduling locomotive hauled carriages at VIA Rail Canada. Ben-Khedher et al. (1998) combine short-term rolling stock rescheduling with revenue management at SNCF.

The literature on real-time rescheduling of rolling stock is scarce. Jespersen-Groth and Clausen (2006) consider the problem of reinserting canceled train units after a disruption. Jespersen-Groth et al. (2007) offer a comprehensive overview of the practical aspects of disruption management in passenger railway transportation. Further, a number of papers address the problem of routing trains through a dense railway network, the primary focus being on dispatching of timetable services and on management of delays. See Törnquist (2005) for a recent overview of this topic.

Operational crew rescheduling in the passenger railway industry is a field that has recently attracted increasing attention. Potthoff et al. (2010) describe a column generation algorithm for rescheduling the crews when a disruption occurs in a passenger railway system. Rezanova and Ryan (2010) present a fast set partitioning heuristic for the train driver recovery problem and test it on real-life instances of the Danish railway operator DSB S-Tog.

Walker et al. (2005) present an integrated model for train and crew recovery in the event of a disruption. The approach incorporates the complex crew constraints into the timetable modification decisions and minimizes the train idle times. The model does not explicitly consider rolling stock, as it does not contain possibilities for changing the assignment of rolling stock to the trains.

Disruption management in the context of airlines is a well studied field. For an overview of the research on aircraft rescheduling we refer to Clausen et al. (2010) and Kohl et al. (2007). However, there are several key differences between airline systems and passenger railway systems.

First of all, the airline industry uses seat reservation systems, so that the destinations of the individual passengers in the system are

known. Thus, after a disruption, the recovery actions can be based on this information, which may help to reroute the individual passengers to their destinations. In a railway system without a seat reservation system, information on the destinations of the individual passengers is not known. Thus, after a disruption, the disrupted passenger flows and the destinations of the passengers can only be estimated.

Furthermore, airplanes are single units with a fixed capacity, operating individually. In contrast, trains may be composed of a number of coupled train units: train capacities can be adapted by coupling or uncoupling train units, and there are usually complex rules describing which changes in the train compositions are allowed. As a consequence, results from aircraft (re) scheduling cannot be directly carried over to the railway setting.

## 3. Real time control in passenger railways

In this section we describe the practical aspects of the real-time monitoring and rescheduling of rolling stock. We discuss the options that are open to dispatchers for rescheduling the rolling stock during operations. For a comprehensive overview of the disruption management process in passenger railways we refer to Jespersen-Groth et al. (2007).

### 3.1. The real-time control process

The resource plans that have been constructed through the early planning stages are optimized towards a number of managerial goals, such as service and efficiency. The execution of the plans involves monitoring the positions and movements of all resources, and reacting to any deviations from the plans.

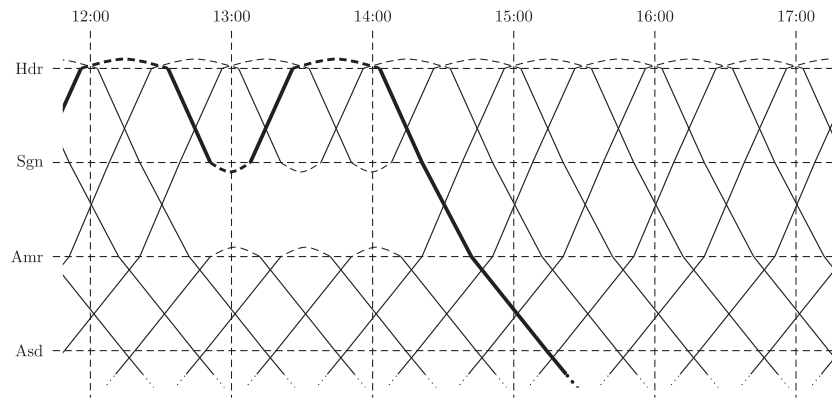
In case of a disruption, the disruption management process starts with updating the timetable based on certain pre-determined scenarios and on an estimate of the duration of the disruption. These scenarios describe which trips will be canceled, rerouted, or short-turned inside the stations. They also determine the sequence of trips to be carried out by each train, as is illustrated in the example shown in Fig. 1.

Fig. 1 shows the effect of a scenario, involving a disruption between Schagen (Sgn) and Alkmaar (Amr) starting at 12:45 on the so-called 3000-line of NS. The estimated length of the disruption is one and a half hour, during which no trains can be operated between Schagen and Alkmaar. The figure shows the stations in the vertical dimension and time in the horizontal dimension. For a more extensive description of the involved example we refer to Appendix A.

After the timetable has been adjusted according to the scenario, the rolling stock is rescheduled, which means that an appropriate rolling stock composition is assigned to each trip. The compositions on successive trips are such that they fit with each other. Finally, the crew is rescheduled according to the updated timetable and rolling stock circulation.

Real-time rolling stock rescheduling is carried out by central and local rolling stock dispatchers. Central dispatchers handle the rolling stock circulation on a network level: they assign a rolling stock composition to each timetabled trip. Local dispatchers handle the details of the shunting processes in the stations.

The central dispatchers can apply various actions for changing the rolling stock circulation. First, the lengths of the trains may be adjusted by modifying the uncoupling-coupling pattern of the train units. Second, the dispatchers may decide not to restore the originally planned end-of-day rolling stock balance on the current day; then the off-balance is resolved either by repositioning train rides during the night or by postponing the recovery actions to the following day. Last, train services may be canceled entirely if



**Fig. 1.** Time-space diagram showing the effect of a scenario in case of a disruption between Schagen (Sgn) and Alkmaar (Amr) starting at 12:45. The estimated length of the disruption is one and a half hour. The bold line indicates a train whose path differs from the original timetable.

no appropriate train units can be found, but this is only an option if all else failed.

A complicating issue in the process is the inherent uncertainty related to large-scale disruptions: it is usually not known at the occurrence of a disruption how long the involved resources will be unavailable – at best, a forecast is available. This means that the results of any step in the rescheduling process can turn out to be infeasible when more information on the resource availability becomes known.

Currently, the primary goal in the real-time operations is to keep the circulation feasible. In addition to that, central dispatchers measure the quality of the rescheduling process by several objective criteria. The following three criteria concern the rescheduling process itself.

1. Cancellations of trains form a major inconvenience for the passengers, thus they are penalized. This criterion has a very high priority in practice.
2. The end-of-day rolling stock off-balance is to be minimized, in order to prevent the effect of the disruption to propagate to the next day.
3. The changes in the shunting plans need to be minimized, in order to enhance the chance that the local dispatchers approve the modified plan.

Furthermore, service quality and efficiency are relevant issues in railway practice, leading to the following two additional criteria.

4. The allocated seat capacity of a train should be sufficient for the anticipated seat demand. This is measured by means of *seat shortage kilometers*.
5. The efficiency of the plan is taken into account by minimizing the operational costs. This is measured by means of *carriage kilometers*.

Some of these objectives are mutually conflicting, and some are difficult to quantify due to lacking real-time information. For example, the precise numbers of passengers on trains and their destinations are not available in real-time – at best a forecast based on historical data is available.

Currently the process related criteria are considered most important in practice. Indeed, the idea is that the original plan is optimized based on the last two criteria, so that efficiency and service in the real-time operations are handled by not deviating too much from the original plan. In our computational experiments we therefore mainly focus on the process related criteria. This also keeps the number of experiments limited. However, in [Appendix B](#)

we also explore the trade-off between the process related criteria and service and efficiency.

### 3.2. Interaction between central and local dispatchers

The rescheduling tool proposed in this paper is meant to be used by the central dispatchers to re-allocate rolling stock compositions to each trip in the modified timetable. The workflow assumes that the timetable has already been adjusted based on the latest available information and according to the selected handling scenario. Also, the original rolling stock schedule is known.

Modifications of the rolling stock circulation directly give rise to changes in the local shunting plans inside the stations. Thus rolling stock rescheduling heavily depends on the ability to evaluate the proposed changes in the shunting plans, which is a complex task, see [Kroon et al. \(2008\)](#). Therefore, the central dispatchers communicate intensively with the local dispatchers about the proposed changes in the shunting plans, either directly or by phone.

Different measures on the network level have different likelihoods of being locally implementable, depending on their impact on the existing shunting plans: If no shunting movement was originally planned between two consecutive trips, then introducing a new shunting movement there requires a major modification of the shunting plan. Indeed, one needs to find an available shunting driver. Canceling a planned shunting movement only requires communication between the local dispatchers and the shunting drivers. Therefore, the cancelation of a planned shunting movement is always preferable over the introduction of a new one. Changing the type of a shunting movement (coupling becomes uncoupling or vice versa) has a complexity in between.

After the rescheduling model has been used to reschedule the rolling stock, the proposed tool graphically shows the implied modifications in the shunting plans per station to the local dispatchers, so that they can evaluate these modifications. If a shunting plan has undesirable features according to local dispatchers, the rescheduling model can be run again with modified parameters or constraints. For example, a new shunting movement at a certain station at a certain time can be penalized or forbidden by appropriately setting the corresponding objective function coefficients; or, in contrast, a new shunting movement can be forced to take place at a certain station and time. This process is repeated until a satisfactory solution is found. For such a system to be accepted by practitioners, the running time must be low – preferably a solution must be available within seconds.



#### 4. The Rolling Stock Rescheduling Problem

The Rolling Stock Rescheduling Problem (RSRP) amounts to adjusting the current circulation of the rolling stock to a changed timetable. RSRP can be considered as an off-line optimization problem for disruption management without uncertainty. Later in this section we use RSRP as a building block of the online variant of RSRP in order to deal with uncertainty.

An instance of RSRP is a 4-tuple  $\langle \mathcal{T}_0, C_0, \tau, \mathcal{T}' \rangle$ . Here  $\mathcal{T}_0$  is the original timetable which consists of a set of trips including the rolling stock connections between them.  $C_0$  is the original rolling stock circulation which is a feasible assignment of rolling stock compositions to the trips in  $\mathcal{T}_0$ . Finally,  $\tau$  is the time instant at which the updated timetable  $\mathcal{T}'$  becomes known. The updated timetable is to be served from time instant  $\tau$  till the end of the day.

Then RSRP amounts to modifying  $C_0$  to a new circulation  $C'$  in order to serve timetable  $\mathcal{T}'$  using the available rolling stock. The circulation  $C'$  must fulfill several technical and market constraints. Clearly, the rolling stock compositions assigned to trips in  $\mathcal{T}_0$  that depart before time  $\tau$  cannot be changed.

The objective of RSRP measures the deviation of the rescheduled circulation  $C'$  from the original circulation  $C_0$  by weighing the objective criteria for canceled trips, changes to the shunting processes, and off-balances.

##### 4.1. Online RSRP

In case of a disruption, the information becomes gradually available to the dispatchers. Therefore it is natural to model the uncertainty as a sequence of timetable updates which require to adjust the last computed circulation to the changed circumstances by solving an instance of the deterministic RSRP.

More formally, an instance of online RSRP consists of the pair  $\langle \mathcal{T}_0, C_0 \rangle$  as well as of a list  $\langle t_1, \mathcal{T}_1 \rangle, \dots, \langle t_n, \mathcal{T}_n \rangle$  of changes to the timetable. Again,  $C_0$  is the original circulation for the original timetable  $\mathcal{T}_0$ . An element  $\langle t_i, \mathcal{T}_i \rangle$  is a pair consisting of a time instant  $t_i$  at which the updated timetable  $\mathcal{T}_i$  is revealed. We assume that  $t_1 < \dots < t_n$ . The  $t_i$ 's are for example the time instants when new information about the duration of the disruption becomes available. The list of timetable changes represents the uncertainty of the real-time operations in that there is no knowledge available at  $t_i$  of future timetables  $\mathcal{T}_{i+1}, \dots, \mathcal{T}_n$ , nor of the time instants  $t_{i+1}, \dots, t_n$ , nor of the number of timetable updates  $n$ .

The task is then at time instant  $t_i$  to solve the RSRP instance  $\langle \mathcal{T}_{i-1}, C_{i-1}, t_i, \mathcal{T}_i \rangle$ . This gives rise to circulation  $C_i$  which is feasible for the latest updated timetable  $\mathcal{T}_i$ . We point out again that at time instant  $t_i$ , the rolling stock assigned to trips in  $\mathcal{T}_i$  departing before  $t_i$  is fixed. The objective of online RSRP measures the deviation of the final circulation  $C_n$  from the original circulation  $C_0$  by weighing the objective criteria for canceled trips, changes to the shunting plans and off-balances.

This way of modeling the progression of a disrupted situation is known in practice as a wait-and-see approach, see Wets (2002). In this approach, no assumptions are made on the probability distribution of the length of the availability of the resources.

It is easy to prove that no algorithm for online RSRP can achieve a constant competitive ratio, see Nielsen (2011). Therefore we do not investigate the theoretical properties of this problem but focus on a solution approach that is satisfactory from a practical point of view.

##### 4.2. Rescheduling with rolling horizon

In this section we propose a rolling horizon approach for solving online RSRP. During a disruption dispatchers only plan a certain time ahead and then revise the solution as time progresses and more information about the disruption becomes available. This is possible since, in practice, only the most immediate decisions are executed and the remaining ones can be rescheduled later. This approach is known as planning with a rolling horizon. We note that Törnquist (2007) uses a similar approach for railway traffic disturbance management. For an extensive overview of several other logistic areas and applications where a rolling horizon is used, we refer to Chand et al. (2002).

The rolling horizon approach works by considering only those trips that are within a certain time horizon from the time at which rescheduling takes place. Whenever new information on the timetable and rolling stock becomes available, the circulation is rescheduled for a further time horizon.

More formally, let  $h$  be a horizon parameter that expresses how far ahead we wish to take the current information into account. At time instant  $t_i$  when the information  $\langle t_i, \mathcal{T}_i \rangle$  becomes known, trips in  $\mathcal{T}_i$  that depart no later than  $t_i + h$  are taken into account and the remaining trips are ignored. If a new timetable update  $\langle t_{i+1}, \mathcal{T}_{i+1} \rangle$  arrives, then the circulation is rescheduled accordingly for the time interval from  $t_{i+1}$  to  $t_{i+1} + h$ . Note that the applied approach does

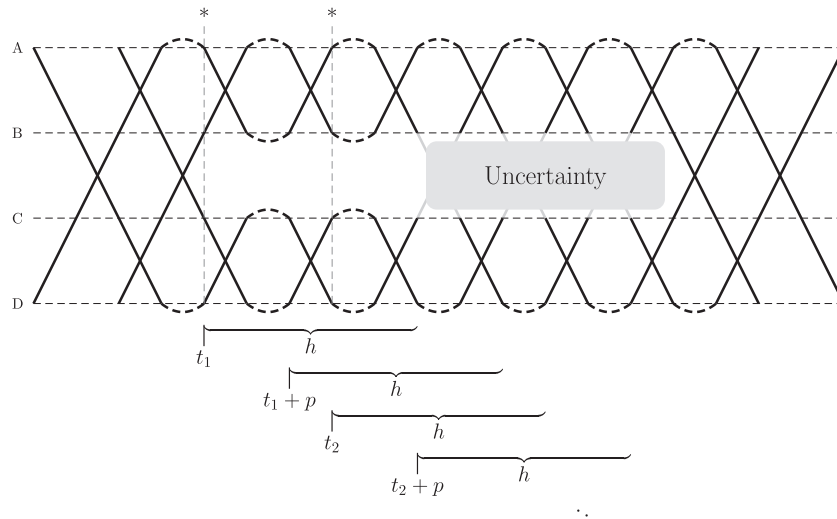


Fig. 2. Rescheduling using a rolling horizon. When new information is available, the circulation is rescheduled  $h$  time steps ahead. (\*) marks the time instants where new information arrives.

not include an *inertia interval*, during which earlier made decisions cannot be changed. However, such an inertia interval could be implemented easily.

For a time interval without any timetable update, the current timetable is the best available forecast for the development of the disruption. Still, a feasible plan only exists from the last update point until the end of the horizon. Therefore we introduce a new parameter  $p$ , called the *period*, which denotes how often the circulation should be updated when the available information does not change. Then, if no new information arrives for  $p$  minutes, we create an artificial information update  $\langle t_i + p, T_i \rangle$  and reschedule accordingly. Clearly, the update should take place before the end of the horizon, so  $p < h$ .

The idea of rolling horizon rescheduling is shown in Fig. 2. A time–space diagram is shown for a timetable with trips between stations A to D. At time  $t_1$  a disruption occurs which leads to some cancellations of trips. The modified turning pattern of the trains in the updated timetable is indicated by dashed arcs in the figure. The circulation is rescheduled to serve the updated timetable until  $t_1 + h$ . Since no new information arrives for some time, rescheduling is performed again at time  $t_1 + p$  with a horizon  $t_1 + p + h$ . However shortly later, at time  $t_2$ , new information becomes available and again the circulation is rescheduled until  $t_2 + h$ . Then no new information arrives for some time, which means that rescheduling is performed at time  $t_2 + p$  and so on.

Realistic values of  $h$  are around two to three hours as that allows dispatchers to react to the most immediate conflicts. Parameter  $p$  should be sufficiently less than  $h$  to allow for a smooth roll of the horizon with sufficient overlap. It could be for example one hour if  $h$  is two hours.

Beside its ability to deal with uncertainty, the rolling horizon approach has an additional strength. For real-life rolling stock scheduling problems, the solution time of state-of-the-art optimization methods tends to increase rapidly with the problem size. This easily leads to computation times of several minutes or even hours, which is unacceptably long for real-time decision making. The rolling horizon method decomposes the rescheduling problem into a number of moderately sized problem instances, making it possible to find solutions quickly, in particular for the urgent trips that have to be carried out soon. The drawback of the rolling horizon approach is of course its heuristic nature.

#### 4.3. The objective in the rolling horizon approach

As was described in Section 3.1, the overall objective of online RSRP contains penalty terms of three kinds: (i) train cancellations, (ii) changed shunting plans and (iii) end-of-day off-balances. Terms (i) and (ii) can directly be minimized in each iteration of the rolling horizon approach. However, the end-of-day off-balances are a fundamentally different matter. In a generic iteration, the rolling stock circulation model is applied for the time interval  $[t_i; t_i + h]$ . Therefore the end-of-day off-balances cannot be taken explicitly into account during the day with a rolling

horizon approach. This motivates our heuristic approach to account for off-balances, which is described next.

Consider iteration  $i$  when the circulation is rescheduled at time  $t_i$ . The RSRP instance to be solved uses penalties for deviations from the rolling stock balance at the end of the horizon, i.e., at  $t_i + h$ . The question is which *target* inventories should the deviation be measured from. We propose to use the original (undisrupted) rolling stock circulation  $C_0$  at  $t_i$  as a guideline.

We consider the intermediate inventories  $i_{s,m}^t$  according to  $C_0$ : these numbers describe how many units of type  $m$  are located at station  $s$  at time instant  $t$ . The intermediate inventories are computed simply by restricting the circulation  $C_0$  to the trips that depart before  $t_i$ . Then we declare the values  $i_{s,m}^{t_i+h}$  to be the target inventories when planning for the time interval  $[t_i; t_i + h]$ .

This guideline may be fairly inaccurate for early time instants  $t_i$ , and intuitively it becomes more and more precise as  $t_i$  approaches the end of the day. Therefore the importance of the off-balances in the objective function should increase with  $t_i$ . This is achieved by multiplying the penalties for off-balances by a factor  $q(t_i)$ , depending on the rescheduling time  $t_i$ . That is, the objective function contains the term

$$q(t_i) \cdot \beta \cdot \sum_{s \in S} \sum_{m \in M} D_{s,m}. \quad (1)$$

Here  $D_{s,m}$  is a decision variable in the rolling stock rescheduling model that measures the deviation of the inventories at the end of the horizon at station  $s$  for rolling stock type  $m$ , while  $\beta$  is the penalty for one single off-balance.

In our application we used a scaled logistic function to increase the relative importance  $q(t)$  of the intermediate inventories from 0 to 1 over time, see Fig. 3. We introduce two parameters  $a$  and  $b$  that guide the relative importance of the deviations from the target intermediate inventories. The function is scaled so that it maps the interval  $[a, b]$  onto the interval  $[0, 1]$ . We used this scaled logistic function because it gives a smooth transition from zero to one.

We note that the exact shape of the function does not play a crucial role. Indeed, the function  $q$  is evaluated a relatively small number of times in the course of the algorithm. We expect that a similarly shaped function will give similar results.

#### 5. A concrete RSRP application

In Section 4 we described a generic framework for RSRP and on-line RSRP. In this section we extend an existing rolling stock scheduling model to be able to solve real-life RSRP instances of NS. The computational results reported in Section 6 are all based on this extended model.

##### 5.1. A model for rolling stock scheduling at NS

Fioole et al. (2006) and Maróti (2006) describe a model for rolling stock scheduling that is used by NS to set up generic rolling stock circulations at an early planning stage with planning hori-

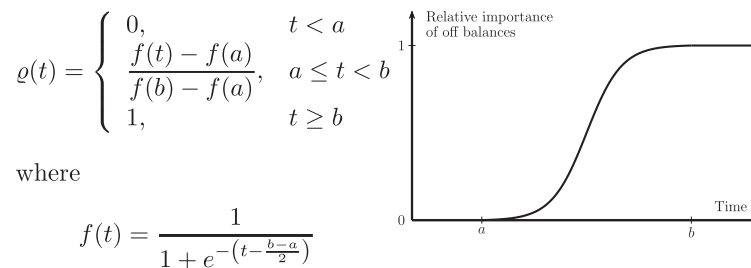
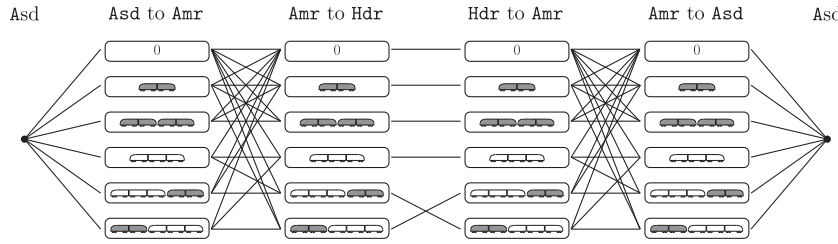


Fig. 3. The relative importance of off-balances increases over time according to a scaled logistic function.



**Fig. 4.** Transition graph for a chain of trips. The chain starts in Amsterdam (Asd). In Alkmaar (Amr) it is possible to couple a unit to the front of the train or to uncouple a unit from the rear of the train. The train is turned in Den Helder (Hdr). An empty composition, represented by 0, has been added to each trip to represent the possibility of canceling the trip.

zons of a few weeks or months. Here we give a simplified description of this model.

Let  $R$  be the set of trips, let  $M$  be the set of rolling stock types, let  $C$  be the set of allowed compositions, and let  $S$  be the set of stations. Let  $T \subseteq R \times R$  be the set of pairs of consecutive trips, also called transitions. As was mentioned before, we assume for the sake of simpler notations that each trip has at most one successor trip and at most one predecessor trip.

The units in the inventories at the stations can be coupled to departing trains, while uncoupled units are added to the local inventories. The inventories thus form a simple model of the shunting processes. At each station, specific end-of-day inventory targets must be achieved in order to seamlessly fit the rolling stock circulations of two subsequent days to one another.

The model is a constrained multi-commodity flow model in a graph constructed from the time-space diagram of the timetable. The core of the model decides which composition is assigned to each trip. The compositions on consecutive trips are linked through the possible shunting operations at the stations and to the inventory there. The assignment of compositions to trips for a train follows a path in a transition graph (see Fig. 4 for an example). Here possible compositions are represented by nodes in the graph while the arcs indicate how the compositions can change between a trip and its successor trip.

The model contains the following variables.

- Binary decision variables  $X_{r,c}$  (corresponding to the nodes in the transition graph) describe whether composition  $c \in C$  is used on trip  $r \in R$ .
- Let  $(r, r') \in T$  be a pair of consecutive trips and let  $(c, c') \in C \times C$  be a pair of compositions. Then binary variable  $Z_{r,r',c,c'}$  (corresponding to the arcs in the transition graph) describe whether composition  $c$  on trip  $r$  is changed to composition  $c'$  on trip  $r'$ .

Using these variables, the objective and the constraints of the model can be expressed as follows.

$$\min F(X, Z) \quad (2)$$

$$\text{subject to } \sum_{c \in C} X_{r,c} = 1 \quad \forall r \in R \quad (3)$$

$$X_{r,c} = \sum_{c' \in C} Z_{r,r',c,c'} \quad \forall (r, r') \in T, c \in C \quad (4)$$

$$X_{r',c'} = \sum_{c \in C} Z_{r,r',c,c'} \quad \forall (r, r') \in T, c' \in C \quad (5)$$

$$0 \leq i_{s,m}^0 + \sum_{r \in R} \sum_{c \in C} \alpha_{r,s,t} \cdot n_{c,m} \cdot X_{r,c} \quad \forall t, s \in S, m \in M \quad (6)$$

$$i_{s,m}^\infty = i_{s,m}^0 + \sum_{r \in R} \sum_{c \in C} \alpha_{r,s}^\infty \cdot n_{c,m} \cdot X_{r,c} \quad \forall s \in S, m \in M \quad (7)$$

Here  $i_{s,m}^0$  and  $i_{s,m}^\infty$  are the a priori given start-of-day and end-of-day inventory of type  $m$  at station  $s$ , and  $n_{c,m}$  is the number of units of type  $m$  in composition  $c$ . Moreover,  $\alpha_{r,s,t} = 1$  if  $s$  is the arrival station

of trip  $r$  and if trip  $r$  arrives before time instant  $t$ ;  $\alpha_{r,s,t} = -1$  if  $s$  is the departure station of  $r$  and if  $r$  departs before time  $t$ , and  $\alpha_{r,s,t} = 0$  otherwise. Parameter  $\alpha_{r,s}^\infty$  is defined similarly, but is related to the end of the day.

The objective (2) is to minimize a linear function of the composition variables  $X$  and the transition variables  $Z$ . The costs related to the composition variables  $X$  may include carriage kilometer costs and seat shortage kilometer costs. The costs related to the transition variables  $Z$  express the complexities of the implied shunting movements. According to (3), each trip is assigned exactly one composition. Constraints (4) and (5) express the flow conservation in the transition graph. Constraints (6) prescribe that the inventories at all stations must always be non-negative. Constraints (7) require the end-of-day inventories at the stations to be equal to the prescribed target values.

We note that the model does not take the duties of the rolling stock units explicitly into account. Instead it works with the units at an aggregated level and divides them into types of interchangeable units. However, it is always possible to convert the compositions resulting from the model to duties, as an integer flow can always be decomposed into unit valued path flows, see Ahuja et al. (1993). Experience has shown that the duties created from the compositions can indeed be implemented in practice.

The model (2)–(7) can be solved by a commercial MIP solver (e.g. CPLEX) for realistic instances in reasonable time. Fiiole et al. (2006) apply it for medium term planning problems of NS. The running times range from a few seconds on smaller instances to several minutes or even hours on larger instances, also depending on the parameters of the objective function. For these larger instances the running time is unacceptably long for use in a real-time setting. Thus for real-time RSRP we use the rolling horizon approach, which reduces the computation time for each subproblem to be solved to a few seconds.

## 5.2. Extending the basic model for RSRP

Having seen the basic model, we extend it to the rescheduling context by adding a number of features. First, the original rolling stock circulation has to be taken into account. We assume that timetable  $T_0$  and circulation  $C_0$  are known; in particular, we know which rolling stock composition is assigned to each trip in the original circulation, as well as the original set of shunting movements.

The rescheduling takes place at time  $\tau$  and  $T'$  is the new timetable. Since the past cannot be changed, all compositions on trips departing before time  $\tau$  are fixed in the model, so that these trips get the same composition as in  $C_0$ . For the remainder of the day, compositions may be modified.

In order to accommodate the possibility of canceling trips, a new binary variable  $X_{r,0}$  is added for each trip  $r \in R$  that departs after time  $\tau$ . The variable denotes the assignment of a special *empty composition* to the trip which is the equivalent of assigning zero rolling stock units to the trip. In other words, the trip is canceled.

In Fig. 4 the composition changes are adapted to the possibility of canceling trips.

For measuring the off-balances in the end-of-day inventories, some additional variables are needed. For each rolling stock type  $m \in M$  and for each station  $s \in S$  the variable  $I_{s,m}^\infty$  denotes the final inventory, and the variable  $D_{s,m}$  denotes the deviation from the target value  $i_{s,m}^\infty$ . Recall that these variables  $D_{s,m}$  are used, in a generic context, for the heuristic objective function (1).

The final inventory  $I_{s,m}^\infty$  is computed by replacing the term  $i_{s,m}^\infty$  in (7) by  $I_{s,m}^\infty$ . The deviation  $D_{s,m}$  is measured by adding the following constraints:

$$D_{s,m} \geq i_{s,m}^\infty - I_{s,m}^\infty \quad \forall s \in S, m \in M$$

$$D_{s,m} \geq I_{s,m}^\infty - i_{s,m}^\infty \quad \forall s \in S, m \in M$$

In the rescheduling model, the objective function depends on the composition variables  $X$ , the transition variables  $Z$ , and the deviation variables  $D$ . In particular, it is composed of the following terms.

$$\sum_{r \in R} \sum_{c \in C} w_{r,c} X_{r,c} + \sum_{r,r' \in T} \sum_{c,c' \in C \times C} \gamma_{r,r',c,c'} Z_{r,r',c,c'} + \beta \cdot \sum_{s \in S} \sum_{m \in M} D_{s,m} \quad (8)$$

The first term incurs a penalty of  $w_{r,c}$  for assigning composition  $c$  to trip  $r$ . This may also include a penalty for assigning a composition  $c$  to trip  $r$  that differs from the composition on trip  $r$  in the original circulation  $C_0$ . The penalty for canceling trip  $r$  can be incorporated in  $w_{r,0}$ . If needed, then penalties for seat shortage kilometers and carriage kilometers can also be incorporated in  $w_{r,c}$ , depending on the composition  $c$  and on the expected number of passengers and the length of trip  $r$ .

The second term incurs a cost of  $\gamma_{r,r',c,c'}$  for the shunting operation from composition  $c$  to composition  $c'$  between consecutive trips  $r$  and  $r'$ . The value  $\gamma_{r,r',c,c'}$  depends on the deviation from the shunting operation that was originally planned between the trips, as was explained in Section 3.1. The third term penalizes a deviation of rolling stock type  $m$  in the end-of-day balance at station  $s$ ; the coefficient  $\beta$  is the penalty for an off-balance of one unit.

## 6. Computational results

In this section we present our computational results for online RSRP. In particular, we explore the relationship between the horizon parameters, and the solution quality and characteristics using a set of test instances. We first describe the test instances and then discuss the results.

### 6.1. Test instances

To analyze the relationship between parameter settings and solutions, we use a realistic set of instances of NS. The instances are divided into four groups, each concerning a particular disruption in the network. In this section we describe one group of instances to analyze the parameters, while the results of the remaining groups are presented in Appendix B.

The group of instances described in this section involves a disruption on the so-called Noord-Oost lines, a system of interconnected lines with a closed rolling stock circulation (see Fig. 5). The Noord-Oost lines form the most challenging cases for rolling stock scheduling at NS, as they have a complicated structure involving underway splitting or combining of trains.

The disruption occurs at 12:00 between Utrecht (Ut) and Amersfoort (Amf). As a consequence, no trains can run between these two stations, and all trains are turned in Utrecht and Amersfoort according to the handling scenario. The disruption turns out to last two and a half hours. However, the actual length of the disruption is not known at the occurrence of the disruption, but only an estimated length is available. An instance consists of the original time-

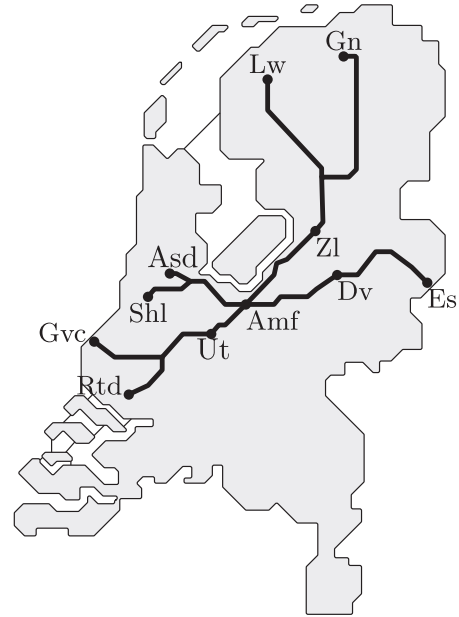


Fig. 5. The Noord-Oost lines.

table with the original circulation and a list of timetable updates that become available at certain times. Table 1 shows the timetable updates in each instance. The timetable  $\mathcal{T}_{t_1-t_2}$  denotes the timetable for the Noord-Oost lines with all trips canceled between Utrecht and Amersfoort in the time interval from  $t_1$  to  $t_2$ . For example, in instance #6, the disruption is first estimated to last one and a half hours, but at time 13:00 the estimated length of the disruption is changed to two and a half hours in total.

The difference between the instances in this group is the time that additional information becomes available and the accuracy of that information. All instances in the group have the same optimal off-line solution.

#### 6.1.1. Objective parameters

As described in Section 4.3, the overall objective of online RSRP contains penalty terms for train cancellations, for changed shunting plans, and for end-of-day off-balances.

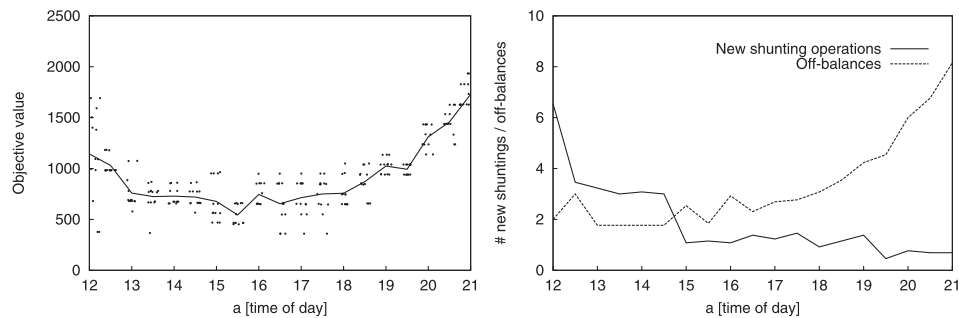
Based on discussions with dispatchers, we chose the following penalty values. The cancellation of a trip due to lack of rolling stock costs 10,000 which outweighs all other objective parameters. This reflects the practitioners' preference to use train cancellations only as a very last option. Each end-of-day off-balance costs 200, and each new shunting operation costs 100. Minor changes to already planned shunting operations cost either 5, 2 or 1, depending on

Table 1

Instances with a disruption on the Noord-Oost lines between Utrecht and Amersfoort.

Instance	Timetable updates
#1	$\langle 12 : 00, \mathcal{T}_{12:00-13:00} \rangle, \langle 12 : 30, \mathcal{T}_{12:00-14:30} \rangle$
#2	$\langle 12 : 00, \mathcal{T}_{12:00-13:00} \rangle, \langle 13 : 00, \mathcal{T}_{12:00-13:30} \rangle, \langle 13 : 30, \mathcal{T}_{12:00-14:30} \rangle$
#3	$\langle 12 : 00, \mathcal{T}_{12:00-13:00} \rangle, \langle 13 : 00, \mathcal{T}_{12:00-14:00} \rangle, \langle 13 : 30, \mathcal{T}_{12:00-14:30} \rangle$
#4	$\langle 12 : 00, \mathcal{T}_{12:00-13:00} \rangle, \langle 13 : 00, \mathcal{T}_{12:00-14:00} \rangle, \langle 14 : 00, \mathcal{T}_{12:00-14:30} \rangle$
#5	$\langle 12 : 00, \mathcal{T}_{12:00-13:00} \rangle, \langle 13 : 00, \mathcal{T}_{12:00-14:30} \rangle$
#6	$\langle 12 : 00, \mathcal{T}_{12:00-13:30} \rangle, \langle 13 : 00, \mathcal{T}_{12:00-14:30} \rangle$
#7	$\langle 12 : 00, \mathcal{T}_{12:00-13:30} \rangle, \langle 13 : 00, \mathcal{T}_{12:00-14:00} \rangle, \langle 13 : 30, \mathcal{T}_{12:00-14:30} \rangle$
#8	$\langle 12 : 00, \mathcal{T}_{12:00-13:30} \rangle, \langle 13 : 00, \mathcal{T}_{12:00-14:00} \rangle, \langle 14 : 00, \mathcal{T}_{12:00-14:30} \rangle$
#9	$\langle 12 : 00, \mathcal{T}_{12:00-13:30} \rangle, \langle 13 : 30, \mathcal{T}_{12:00-14:30} \rangle$
#10	$\langle 12 : 00, \mathcal{T}_{12:00-14:00} \rangle, \langle 13 : 00, \mathcal{T}_{12:00-14:30} \rangle$
#11	$\langle 12 : 00, \mathcal{T}_{12:00-14:00} \rangle, \langle 13 : 30, \mathcal{T}_{12:00-14:30} \rangle$
#12	$\langle 12 : 00, \mathcal{T}_{12:00-14:00} \rangle, \langle 14 : 00, \mathcal{T}_{12:00-14:30} \rangle$
#13	$\langle 12 : 00, \mathcal{T}_{12:00-14:30} \rangle$





**Fig. 6.** Left: Objective function of each instance as a function of parameter  $a$ . The line represents the average. Right: Average number of new shunting operations and off-balances.

their nature and their likelihood to succeed: new, modified, or canceled shunting movements, see Section 3.2). The same objective function is used for each test instance.

## 6.2. Exploring the horizon parameters

In this section we investigate different settings of the parameters of the rolling horizon approach.

### 6.2.1. Parameters for the intermediate inventories

As mentioned in Section 4.3, off-balances are handled heuristically by increasing the importance of balancing the intermediate inventories. The parameters  $a$  and  $b$  determine when the intermediate inventories start to be taken into account and when they are taken into account with full weight, respectively.

In the first set of experiments, we set the horizon to  $h = 3$  hour and the update parameter to  $p = 1$  hour. We then applied the rolling horizon framework with  $a$  at different values between 12:00 and 21:00 with 30 minutes between values. The parameter  $b$  is set to 6 hour after  $a$  in all tests.

The left diagram of Fig. 6 shows the objective values of all instances as a function of the parameter  $a$ . Each dot represents the outcome of one instance with the given parameters. The average outcome over all instances is plotted as a line in the diagram as well. It turned out that none of the obtained solutions requires cancelations of trips. In the right diagram of Fig. 6 the average numbers of new shunting operations and off-balances are shown, as they are the major contributors to the objective function.

When the balancing is initiated rather late, more off-balances remain in the solutions. However, when balancing is initiated rather early, more new shunting operations are introduced without resolving more off-balances. For the involved instances, the best balanced solutions seem to be found with  $a$  between 15:00 and 17:00. Since  $h = 3$  hours, this means that the intermediate inventories are taken into account once the end of the horizon reaches 18:00.

The additional tests in Appendix B show that the best choice of the parameter  $a$  depends on the structure of the involved lines and on the time interval of the disruption. If possible,  $a$  should be chosen in such a way that there is still sufficient time to get a train unit from any station along the line to any other station. If there is insufficient time, then the balancing should start immediately after the disruption. Thus for shorter lines the balancing can be initiated later.

The characteristics of all results of the first group of instances with a specific choice of parameters are shown in Table 2. The parameters are  $h = 3$  hour,  $p = 1$  hour,  $a = 15:30$  and  $b = 21:30$ . In all instances 1 or 2 off-balances occur and up to 2 new shunting operations are introduced. In the optimal off-line solution it is possible to reschedule the rolling stock in such a way that no

**Table 2**

Results of all instances with parameters  $h = 3$  hour,  $p = 1$  hour,  $a = 15:30$  and  $b = 21:30$ . The figures in the column headings indicate the (weighted) objective coefficients. (\*) shows the results of all instances when rescheduling with the same horizon parameters except that off-balances are not taken into account.

Instance	New shunting operations (200)	Off-balances (100)	Weighted number of minor changes to shunting plans (1)	Objective
1	2	2	60	660
2	2	2	57	657
3	2	1	52	552
4	2	0	65	465
5	2	0	65	465
6	2	0	65	465
7	1	2	53	453
8	2	0	65	465
9	1	2	53	453
10	2	2	60	660
11	2	2	52	652
12	2	0	65	465
13	2	2	60	660
(*)	11–13	0–2	22–26	–
off-line	0	0	37	37

off-balances occur and no new shunting operations are introduced: only some minor changes of the shunting operations are needed.

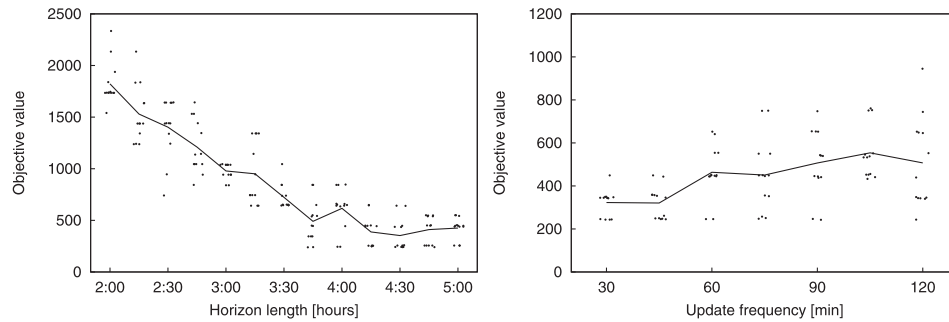
When the balancing heuristic is not used and off-balances are *not* taken into account when rescheduling with a rolling horizon (as indicated by (\*) in the table), it results in 11–13 off-balances. The large reduction in off-balances comes at a cost of some minor changes to the existing shunting operations. This suggests that the heuristic guidance of the rolling stock to the target inventories works well in these instances. Note that Table 2 does not show a clear relationship between the obtained results and the way the disruption is revealed.

### 6.2.2. Horizon length

A set of experiments have been conducted to explore the relationship between the length of the horizon and the solution quality. The length of the horizon represents how far ahead in the current timetable trips are taken into account.

The instances have been tested with values for the horizon parameter  $h$  between 2 and 5 hour with 15 minutes between values. Parameter  $a$  is set to 17:00 and  $b$  is set to 23:00. The update parameter  $p$  is set to 1 hour. The objective values have been plotted as a function of the horizon length in Fig. 7(left).

In principle, a longer horizon gives better solutions. However, this comes at a cost of computation time, and there is no guarantee to obtain better solutions due to the heuristic nature of the approach. The latter explains the slightly upward trend in Fig. 7(left) after a horizon length of 4:30 hour. Note that there



**Fig. 7.** Left: objective function of each instance as a function of parameter  $h$ . Right: objective function of each instance as a function of parameter  $p$ . In both cases, the line represents the average.

are barely improvements for horizons longer than 3:30 hour, which is in line with experience from practice.

### 6.2.3. Update parameter

The parameter  $p$  determines how often the circulation is rescheduled when no information updates arrive. This way it controls the roll of the horizon. We have tested the instances for values of  $p$  from 30 to 120 minutes with 15 minutes between values. The remaining parameters are fixed at  $h = 4$  hour,  $a = 17:00$  and  $b = 23:00$ . The objective values are plotted as a function of the parameter  $p$  in Fig. 7(right). It shows that updating more often in principle yields better solutions, as non-executed decisions can be revised when later trips come into the horizon. The objective function does not strictly increase with the value of  $p$  though, again due to the heuristic nature of the approach.

### 6.3. Computation times

The tests have been conducted on a Pentium 4 2.8 GHz desktop with 1 GB of RAM using CPLEX 10.1. As discussed earlier, the size of each instance depends on the horizon length, so the relationship between computation time and horizon length is explored. The computation times per iteration are presented in Fig. 8 as a function of the horizon length. Values of  $h$  from 2 hour to 5 hour were tested with 15 minutes between values.

There is a significant correlation between the computation times and the length of the horizon. When using a horizon of less than three hours, the computation times are always within a few seconds. When using a horizon of more than three and a half hours, longer computation times may occur, and also the variability in the computation times increases. We note that when setting  $h$  large enough to reschedule the rest of the day, computation times of up to 10 minutes occurred, which is too long for the purpose of

rescheduling in real-time. This shows the value of the rolling horizon approach with lower values of  $h$ .

## 7. Conclusions

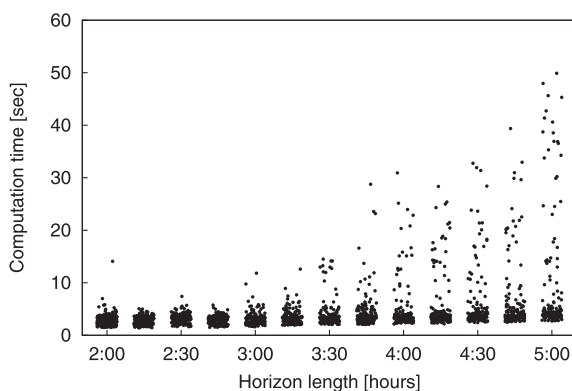
This paper deals with real-time disruption management of railway rolling stock. We defined the Rolling Stock Rescheduling Problem (RSRP). The main assumption is that timetable updates are explicitly given via the handling scenarios. The goal is to adjust the original rolling stock schedules for the updated timetables, taking various objectives into account.

We also defined the online variant of RSRP where the uncertainty about the duration of the disruption is modeled by a sequence of timetable updates. In order to deal with such uncertainties and to reduce the computation times, we proposed a rolling horizon framework as a solution approach. In this framework we only consider rolling stock decisions within a certain horizon of the time of rescheduling. The schedules are then revised as the situation progresses and more accurate information becomes available.

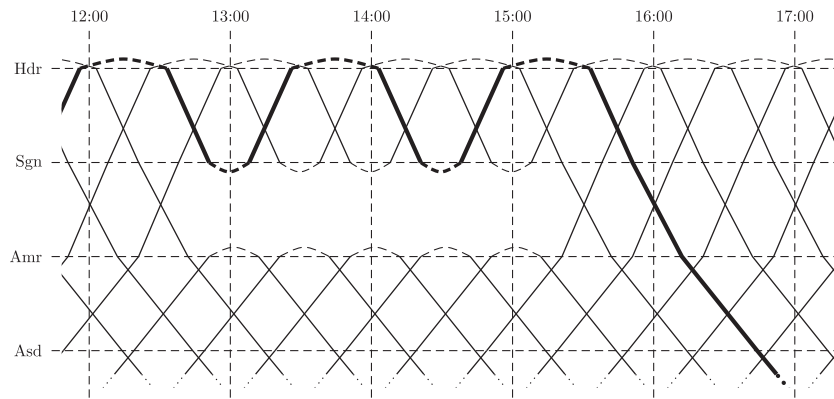
The (online) RSRP is an abstract framework which needs to be adapted for the concrete specifications of real-life railway scheduling problems. In this paper we extended an existing rolling stock scheduling model that has been applied successfully by the major Dutch passenger railway operator NS for its medium term planning since 2004. We mainly penalize cancellations of trains, modifications of the shunting processes, and end-of-day rolling stock off-balances.

The rolling horizon approach takes the off-balances into account in a heuristic way. Based on the undisrupted rolling stock circulation, we define target inventories during the day, and we minimize the deviations from these targets. The heuristic target becomes more accurate as the current horizon approaches the end of the day. Therefore the objective function coefficients of the off-balances increase as the horizon shifts ahead in time. Although this is a heuristic approach for minimizing the end-of-day off-balances, it turned out to work well for our test instances.

The performance of the rolling horizon approach and the heuristic for balancing the end-of-day inventories depends on a number of parameters. The test results indicate that starting the balancing too early results in many changes to the shunting plans without much effect on the final balances. However, starting too late leaves little time for balancing the rolling stock and leads to many off-balances. The best time to start the balancing process depends on the length of the involved lines and on the time interval of the disruption. If possible, the balancing process should start at such a time that there is still sufficient time to get a train unit from any station along the line to any other station. If the disruption takes place in the late afternoon or in the evening, then balancing should start as early as possible after the disruption.



**Fig. 8.** Computation time of each step of all instances as a function of parameter  $h$ .



**Fig. 9.** Time-space diagram for part of the timetable for the 3000 line. At 13:30 it turns out that the disruption lasts two and a half hour. The bold line is a train whose path differs from the first updated timetable.

The length of the horizon offers a trade-off between solution quality and computation time, though longer horizons do not offer strictly better solutions, because the rolling horizon approach is a heuristic. The tests did not reveal a clear relationship between the objective function and the way the disruption is revealed (i.e. in one step, or in several steps).

The tests show that the method can be used to reschedule the rolling stock during a disruption with minor effects for the shunting plans. At the same time, the number of off-balances can be reduced by a few changes to the planned shunting operations. This, together with the short computation times, indicates that the approach is a good candidate for being implemented as the core of a decision support system for rolling stock rescheduling.

The currently implemented model leaves a number of practically relevant issues out of account, such as maintenance of rolling stock. These issues require further extensions of the model. Also, further discussions with dispatchers are needed to refine the weights in the objective function in order to reach a better match with their preferences, for example related to the relative weights of the different types of changes in the shunting plans.

The current paper focused on the operational problem of rescheduling the rolling stock in case of a disruption. In a forthcoming paper, we will also address the tactical question whether the original rolling stock circulation can be made more robust against disruptions.

A further important feature of real-life disruptions is that rerouting of passengers may lead to a substantially altered seat demand on the detour routes. Such demand changes are not yet considered in this paper. We are currently working on a rolling stock rescheduling model that can optimize both the passenger delays and the process related rescheduling costs.

## Acknowledgments

We would like to thank the dispatchers at the Rolling Stock Managing Center of NS for their input on the practical details of rolling stock operations.

This work was partially supported by the Future and Emerging Technologies Unit of EC (IST priority – 6th FP), under contract No. FP6-021235-2 (project ARRIVAL).

## Appendix A. An example of a disruption

In this appendix we describe an example of the rolling stock rescheduling problem. The example involves the so-called 3000 line of NS.

### A.1. The 3000 line

The 3000 line is an intercity line of NS from Den Helder to Nijmegen. The trains operating the line call at a number of stations on the line including Den Helder (Hdr), Schagen (Sgn), Alkmaar (Amr), Amsterdam (Asd), Utrecht (Ut), Ede (Ed), Arnhem (Ah) and Nijmegen (Nm). The line is operated with double-deck train units that are available in lengths of 4 and 6 carriages, which can be coupled with each other to form longer trains. The timetable is cyclic and has trains in both directions every half hour.

### A.2. Disruption on the 3000 line

In this example, a disruption occurs at 12:45. The tracks cannot be used in either direction between Schagen and Alkmaar. At the occurrence, the infrastructure is estimated to be unavailable for one and a half hour. The trips between Schagen and Alkmaar are canceled and the timetable is updated according to a handling scenario. In the updated timetable trains are turned at each side of the disrupted region (i.e. in Schagen and Alkmaar). A *time-space diagram* of the updated timetable for the estimated duration of the disruption of one and a half hour was shown already in Fig. 1.

However, at time 13:30 it becomes clear that the necessary repairs of the infrastructure will take an hour longer than first estimated. This leads to a new updated timetable with two more canceled trips in each direction, see Fig. 9. The new update of the timetable is a forecast of the development of the disruption, and it may turn out to need further adjustment later on.

### A.3. Challenges for rescheduling

The original rolling stock circulation determines the composition of each train, and also where and when units are coupled to or uncoupled from trains. However, the original circulation is clearly not feasible for the updated timetable.

Originally, the train departing from Den Helder at 12:33 is supposed to arrive in Alkmaar at 13:09. But due to the disruption, it is turned in Schagen. In Fig. 1 the path of the train is marked with a bold line and, as can be seen, the train now serves different trips in the timetable. At the same time, another train takes over the original trips of this train. Since this train may have a different composition than the original one, some of the couplings and uncouplings meant for the original train must be changed or canceled.

The rolling stock circulation is rescheduled according to the updated timetable, yet it is still uncertain if the disruption will last shorter or longer than first estimated. When it becomes clear at 13:30 that the repairs of the infrastructure will last longer than ex-

pected, the recently updated circulation must be changed again. In Fig. 9 the path of the before mentioned train was changed again.

#### A.4. The disruption example as an instance of RSRP

The described disruption of the 3000 line can be modeled as an instance of online RSRP in the following way. The original timetable  $\mathcal{T}_0$  for the 3000 line is given, as well as the rolling stock circulation  $\mathcal{C}_0$ , using 11 units of length 4 and 20 units of length 6.

The disruption between Alkmaar and Schagen starts at 12:45 and the information on the duration of the disruption changes at 13:30. This is modeled as a list of changes to the resource availability containing the following two elements:

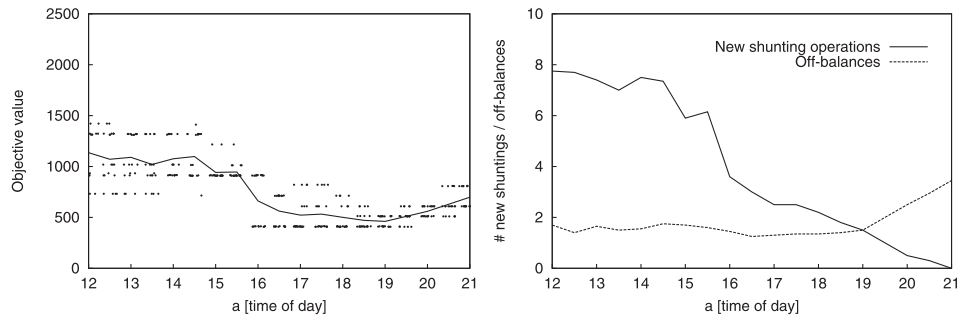
- The first element is  $\langle t_1, \mathcal{T}_1 \rangle$  where  $t_1 = 12:45$ ,  $\mathcal{T}_1$  is equal to  $\mathcal{T}_0$  except from the canceled trips according to Fig. 1.
- The second element is  $\langle t_2, \mathcal{T}_2 \rangle$  where  $t_2 = 13:30$ ,  $\mathcal{T}_2$  is equal to  $\mathcal{T}_1$  except from the additionally canceled trips according to Fig. 9.

The task is then to reschedule the rolling stock at time  $t_1$  to serve timetable  $\mathcal{T}_1$  without the knowledge of the later update  $\langle t_2, \mathcal{T}_2 \rangle$ . That information becomes available at time  $t_2$ , where the rolling stock must be rescheduled to serve timetable  $\mathcal{T}_2$ .

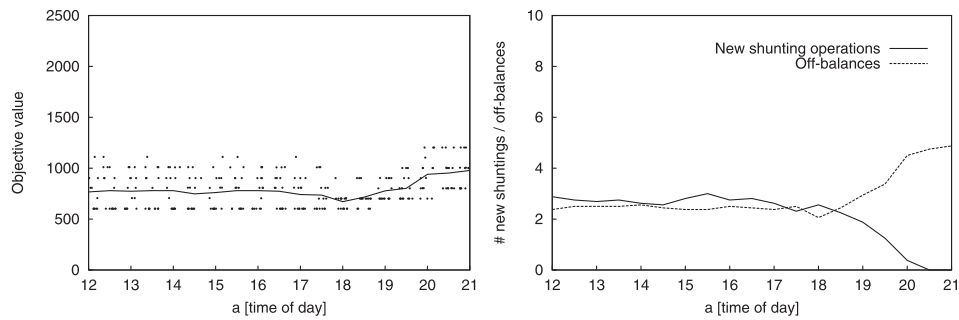
#### Appendix B. Further computational results

This appendix contains the computational results for the other instances that were mentioned in Section 6 but could not be reported there. Also, it contains a short study of the trade-off between service quality and efficiency.

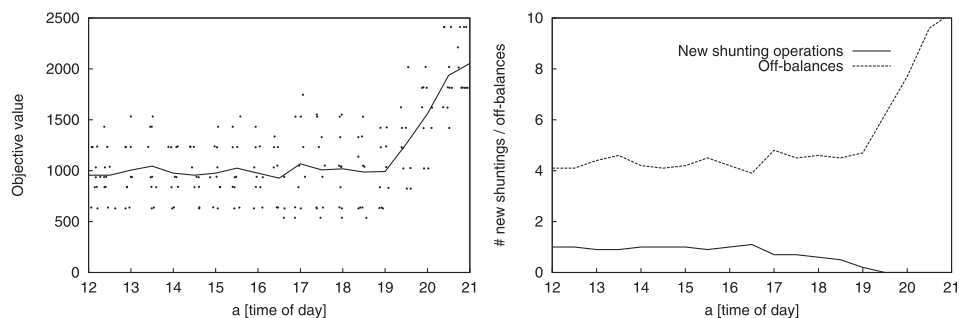
The first set of additional instances involves the 3000 line (see Appendix A) with a disruption between Schagen and Alkmaar at time 12:00. The disruption lasts 4 hour. The set has 20 instances. The second set involves the 3000 line with a disruption between Utrecht and Arnhem at time 19:00. The disruption lasts 3 hour. The set has 16 instances. The third set involves the Noord-Oost



**Fig. 10.** Left: objective function of each instance for the disruption at 12:00 on the 3000 line as a function of parameter  $a$ . The line represents the average. Right: average number of new shunting operations and off-balances.

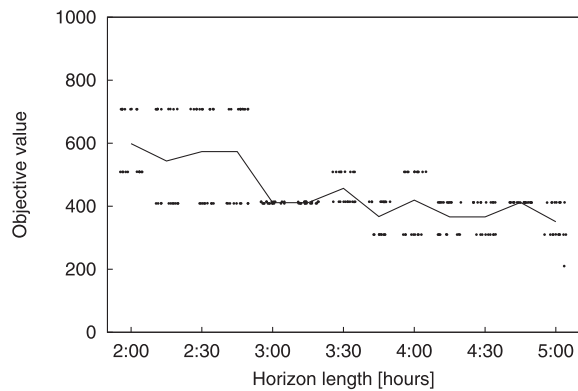


**Fig. 11.** Left: objective function of each instance for the disruption at 19:00 on the 3000 line as a function of parameter  $a$ . The line represents the average. Right: average number of new shunting operations and off-balances.



**Fig. 12.** Left: objective function of each instance for the disruption at 19:30 on the Noord-Oost lines as a function of parameter  $a$ . The line represents the average. Right: average number of new shunting operations and off-balances.





**Fig. 13.** Objective function of each instance for the disruption at 12:00 on the 3000 line as a function of parameter  $h$ . The line represents the average.

**Table 3**  
Set of objective parameters.

Objective	Cost
Cancel a trip	10,000
Off-balance	200
New shunting operation	100
Canceled shunting operation	1
Different type of train unit shunted	2
Coupling replaced by uncoupling (or vice versa)	5
Carriage kilometer	0.01
Seat shortage kilometer	Varied between 0.0004 and 0.0025

lines with a disruption between Utrecht and Amersfoort at time 19:30. The disruption lasts 2 hour. The set has 10 instances.

### B.1. Horizon parameters

Figs. 10–12 show the objective values of the instances and the trade-off between off-balances and introducing new shunting operations. We observe that setting  $a$  to around 18:00 to 19:00 yields the best trade-off for the 3000 line. Note that the 3000 line is shorter than the Noord-Oost lines. Therefore, on this line the balancing process can start later than in the earlier described case on the Noord-Oost lines. For the disruptions that take place in the evening there is no point in starting the balancing process earlier since there are no off-balances in the intermediate inventories before the disruption.

Fig. 13 shows the relationship between the objective value and the horizon length  $h$ . Generally, longer horizons result in better solutions, although for horizons longer than 3:30 hour the added value diminishes.

### B.2. Impact on service quality and efficiency

To further explore the effects of rescheduling the Noord-Oost lines on service quality and efficiency, we varied the seat shortage kilometer costs between 0.0004 and 0.0025, while all other cost coefficients are held constant according to the weights in Table 3. In particular, the costs related to the modified shunting operations are fixed: a new shunting operation costs 100, and minor changes to the shunting operations cost 1, 2 or 5 depending on their type (as indicated in the table). Note that in this experiment *only* the costs of the seat shortage kilometers are varied.

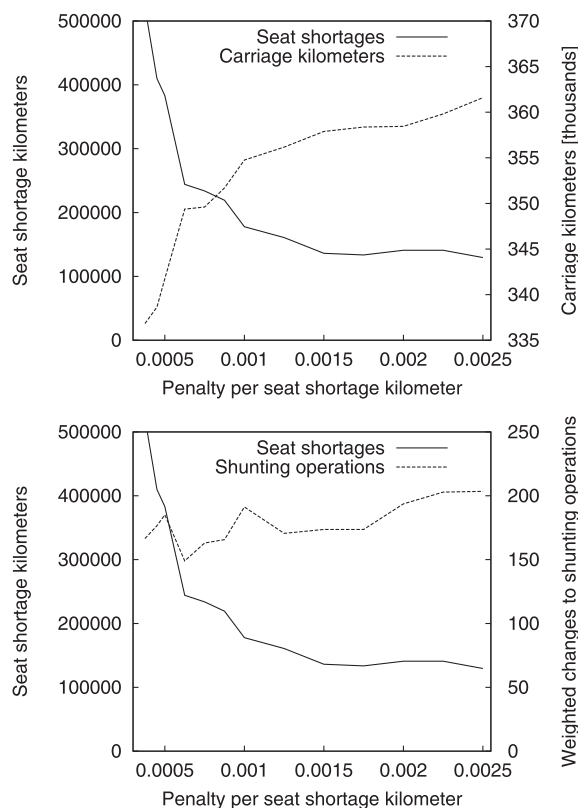
We conducted experiments with a horizon of  $h = 3$  hour with rescheduling every  $p = 1$  hour. The top diagram in Fig. 14 shows the relationship between the resulting number of seat shortage kilometers and carriage kilometers averaged over all instances. We note a clear trade-off between carriage kilometers and seat shortage kilometers. The originally planned rolling stock circulation used 369,751 carriage kilometers of which 4,255 are canceled directly due to the disruption. At the same time the original plan had 52,553 seat shortage kilometers. Depending on the parameter settings for the seat shortage kilometers, the rescheduled plan achieves a somewhat lower number of carriage kilometers with at least the double number of seat shortage kilometers.

The additional seat shortage kilometers are due to the heuristic nature of the rolling horizon approach, whereas the original plan is an optimal one. Indeed, in order to cover a certain seat demand on a certain train, one may have to take the right decisions already long time in advance. This may be overlooked by the rolling horizon approach, resulting in an increased number of seat shortages.

The bottom diagram in Fig. 14 shows the trade-off between seat shortage kilometers and another characteristic of the solution: the changes to the shunting plans. Again we note that providing better service, in the sense of limiting seat shortages, comes at a cost of somewhat more changes to the shunting plans.

## References

- Ahuja, R., Magnanti, T., Orlin, J., 1993. Network Flows: Theory Algorithms and Applications. Prentice-Hall, Englewood Cliffs, NJ.
- Alfieri, A., Groot, R., Kroon, L., Schrijver, A., 2006. Efficient circulation of railway rolling stock. Transportation Science 40 (3), 378.
- Ben-Khedher, N., Kintanar, J., Queille, C., Stripling, W., 1998. Schedule optimization at SNCF: From conception to day of departure. Interfaces 28, 6–23.
- Budai, G., Maróti, G., Dekker, R., Huisman, D., Kroon, L., 2010. Re-scheduling in railways: The rolling stock balancing problem. Journal of Scheduling 13, 281–297.



**Fig. 14.** Top: seat shortage kilometers and carriage kilometers for different values of the seat shortage penalty. Bottom: seat shortage kilometers and the weighted number of changes to the shunting operations.

- Caprara, A., Kroon, L., Monaci, M., Peeters, M., Toth, P., 2007. Passenger Railway Optimization. In: Barnhart, C., Laporte, G. (Eds.), *Handbook in OR and MS*, Vol. 14. Elsevier, pp. 129–187.
- Chand, S., Hsu, V., Sethi, S., 2002. Forecast, solution and rolling-horizons in operations management problems: A classified bibliography. *Manufacturing and Service Operations Management* 4 (1), 25–43.
- Clausen, J., Larsen, A., Larsen, J., Rezanova, N.J., 2010. Disruption management in the airline industry – Concepts, models and methods. *Computers and Operations Research* 37, 809–821.
- Fioole, P.J., Kroon, L., Maróti, G., Schrijver, A., 2006. A rolling stock circulation model for combining and splitting of passenger trains. *European Journal of Operational Research* 174, 1281–1297.
- Grötschel, M., Krumke, S.O., Rambau, J., Winter, T., Zimmermann, U.T., 2001. Combinatorial Online Optimization in Real Time. In: Grötschel, M., Krumke, S.O., Rambau (Eds.), *Online Optimization of Large Scale Systems*. Springer, pp. 679–704.
- Huisman, D., Kroon, L.G., Lentink, R.M., Vromans, M., 2005. Operations research in passenger railway transportation. *Statistica Neerlandica* 59 (4), 467–497.
- Jespersen-Groth, J., Clausen, J., 2006. Optimal Reinsertion of Cancelled Train Lines. Technical report, Informatics and Mathematical Modelling, Technical University of Denmark.
- Jespersen-Groth, J., Potthoff, D., Clausen, J., Huisman, D., Kroon, L.G., Maróti, G., Nielsen, M.N., 2007. Disruption Management in Passenger Railway Transportation. Technical report, Econometric Institute, Erasmus University Rotterdam.
- Kohl, N., Larsen, A., Larsen, J., Ross, A., Tiourine, S., 2007. Airline disruption management – Perspectives, experiences and outlook. *Journal of Air Transport Management* 13 (3), 149–162.
- Kroon, L.G., Lentink, R.M., Schrijver, A., 2008. Shunting of passenger train units: An integrated approach. *Transportation Science* 42 (4), 436–449.
- 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.
- Lingaya, N., Cordeau, J.F., Desaulniers, G., Desrosiers, J., Soumis, F., 2002. Operational car assignment at VIA Rail Canada. *Transportation Research B* 36, 755–778.
- Maróti, G., 2006. Operations Research Models for Railway Rolling Stock Planning. PhD thesis, Eindhoven University of Technology, Eindhoven, The Netherlands.
- Nielsen, L.K., 2011. Rolling Stock Rescheduling in Passenger Railways. PhD thesis, Erasmus University, Rotterdam, The Netherlands.
- Peeters, M., Kroon, L.G., 2008. Circulation of railway rolling stock: A branch-and-price approach. *Computers and Operations Research* 35 (2), 538–556.
- Potthoff, D., Huisman, D., Desaulniers, G., 2010. Column generation with dynamic duty selection for railway crew rescheduling. *Transportation Science* 44, 281–297.
- Rezanova, N.J., Ryan, D.M., 2010. The train driver recovery problem – A set partitioning based model and solution method. *Computers and Operations Research* 37, 845–856.
- Séguin, R., Potvin, J.Y., Gendreau, M., Crainic, T.G., Marcotte, P., 1997. Real-time decision problems: An operational research perspective. *Journal of the Operational Research Society* 48 (2), 162–174.
- Törnquist, J., 2005. Computer-based decision support for railway traffic scheduling and dispatching: A review of models and algorithms. *ATMOS2005 (Algorithmic Methods and Models for Optimization of RailwayS)*, Palma de Mallorca, Spain.
- Törnquist, J., 2007. Railway traffic disturbance management – An experimental analysis of disturbance complexity, management objectives and limitations in planning horizon. *Transportation Research Part A* 41 (3), 249–266.
- Walker, C.G., Snowdon, J.N., Ryan, D.M., 2005. Simultaneous disruption recovery of a train timetable and crew roster in real time. *Computers and Operations Research* 32 (8), 2077–2094.
- Wets, R.J.-B., 2002. Stochastic programming models: Wait-and-see versus here-and-now. In: Greengard, C., Ruszczyński, A. (Eds.), *Decision Making Under Uncertainty*. Springer.
- Yu, G., Qi, X., 2004. *Disruption Management: Framework, Models and Applications*. World Scientific.