Railway Rolling Stock Planning: Robustness Against Large Disruptions

Valentina Cacchiani, Alberto Caprara, Laura Galli, Leo Kroon, Gábor Maróti, Paolo Toth,

Railway Rolling Stock Planning: Robustness Against Large Disruptions

Valentina Cacchiani, Alberto Caprara, Laura Galli
Department of Electronics, Computer Science, and Systems, University of Bologna, 40136 Bologna, Italy
{valentina.cacchiani@unibo.it, alberto.caprara@unibo.it, l.galli@unibo.it}

Leo Kroon
Rotterdam School of Management, Erasmus University Rotterdam, NL-3000 DR Rotterdam, The Netherlands; and
Netherlands Railways, NL-3500 HA Utrecht, The Netherlands, lkroon@rsm.nl

Gábor Maróti
Rotterdam School of Management, Erasmus University Rotterdam, NL-3000 DR Rotterdam,
The Netherlands, gmaroti@rsm.nl

Paolo Toth
Department of Electronics, Computer Science, and Systems, University of Bologna, 40136 Bologna, Italy,
paolo.toth@unibo.it

In this paper we describe a two-stage optimization model for determining robust rolling stock circulations for passenger trains. Here robustness means that the rolling stock circulations can better deal with large disruptions of the railway system. The two-stage optimization model is formulated as a large mixed-integer linear programming (MILP) model. We first use Benders decomposition to determine optimal solutions for the LP-relaxation of this model. Then we use the cuts that were generated by the Benders decomposition for computing heuristic robust solutions for the two-stage optimization model. We call our method *Benders heuristic*. We evaluate our approach on the real-life rolling stock-planning problem of Netherlands Railways, the main operator of passenger trains in the Netherlands. The computational results show that, thanks to Benders decomposition, the LP-relaxation of the two-stage optimization problem can be solved in a short time for a representative number of disruption scenarios. In addition, they demonstrate that the robust rolling stock circulation computed heuristically has total costs that are close to the LP lower bounds. Finally, we discuss the practical effectiveness of the robust rolling stock circulation: When a large number of disruption scenarios were applied to these robust circulations and to the nonrobust optimal circulations, the former appeared to be much more easily recoverable than the latter.

Key words: two-stage optimization; rolling stock planning; integer linear programming

1. Introduction

The optimization of passenger and freight transportation networks has been a fruitful area of applied operations research in the past decades. Numerous papers address and solve real-life problems of airline and railway operators. The first applications concentrated on solving operational scheduling problems with a relatively short planning horizon, and usually considered a static and deterministic problem setting. However, recently a significant amount of effort has been put into dealing with the inherent uncertainties in such networks. This broadened the research area in two directions: On one hand, many papers addressed short-term and real-time rescheduling problems and disruption management; on the other hand, the robustness of the operational schedules has become a research target.

Informally speaking, robust scheduling attempts to find a schedule that is insensitive to irregularities of the operations. Such irregularities may include delays of flights and trains, but also large disruptions where significant parts of the network are temporarily out of order. Robustness can refer to the absorption capacity or to the recoverability of a schedule. On one hand, the *absorption capacity* expresses how a schedule can cope with relatively small disturbances without structural changes—for example, by applying a *push-back* strategy. On the other hand, *recoverability* measures how easily the schedule can be adjusted to cope with a large disruption. In this paper we are particularly interested in recoverability of a railway rolling stock circulation. That is, we focus on the question how the rolling stock circulation can be designed in such a way that it can be recovered relatively easily in case of a large disruption.

1.1. Robust Rolling Stock Planning

In this paper we address the medium-term rolling stock planning problem of *Netherlands Railways* (NS), the main operator of passenger trains in the Netherlands. This planning problem arises two–six months before the actual railway operations, and amounts to assigning the available rolling stock to the trips in a given timetable. The traditional objectives of this rolling stock planning problem are related to service quality, efficiency, and—to a limited extent—robustness. In this *nominal* version, all data are assumed to be known, and disturbances or disruptions are not taken into account. Fioole et al. (2006) describe a *mixed-integer linear programming* (MILP) model for solving this nominal problem. Using commercial MILP software, the solution times on real-life instances of NS are quite low, ranging from a few minutes (for most instances) to a couple of hours (for some particularly complex instances). A software tool based on this model has been in operation within NS since 2004.

The solutions of the nominal rolling stock planning problem are optimal under undisrupted operations only. However, infrastructure failures, rolling stock breakdowns, and accidents are regularly recurring large disruptions of the railway system, causing the nominal solution to become infeasible. In such cases, disruption management must come up with an adjusted timetable, rolling stock circulation, and crew schedules. In this paper, we focus on the recovery of the rolling stock circulation after the timetable has been recovered first. This latter assumption is motivated by the fact that in most European countries infrastructure management is carried out by authorities that are independent of the railway operators.

In recovering the rolling stock circulation, the nominal objective criteria are of marginal importance. Instead, the goal is to quickly find such a feasible recovered solution that fits with the recovered timetable, is close to the original one, and can be implemented easily in practice. As will be explained later, the deviation of a recovered solution from the nominal solution is measured in terms of the number of additionally canceled trips, the number of modified shunting operations, and the number of end-of-day off-balances of the rolling stock inventories at the stations.

Figure 1 shows an example of a disrupted timetable in a time-space diagram. A disruption in this example occurs between the stations Alkmaar (Amr) and Amsterdam (Asd) from 17:30 until 19:30. Some trips of the original schedule between these stations have been canceled, thus requiring the rolling stock circulation to be modified as well.

In this paper, we model the robust optimization of the rolling stock circulation as a two-stage optimization problem. In the first stage, the robust rolling
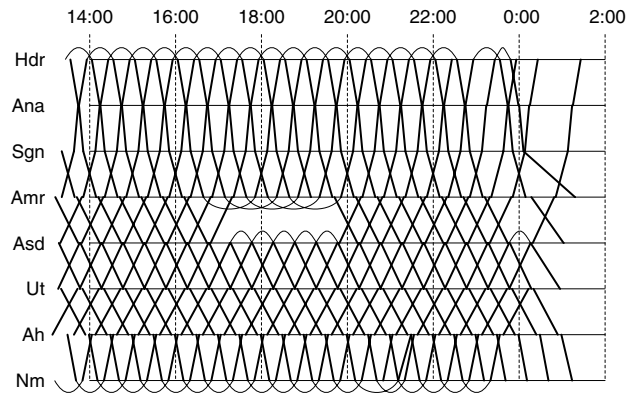

Figure 1   An Example of Disruption Between Stations Alkmaar (Amr) and Amsterdam (Asd) from 17:30 Until 19:30

stock circulation is generated. In the second stage, the optimal recovery actions in response to a finite set of disruption scenarios are represented. The model aims at minimizing the sum of the nominal costs and the maximum of the recovery costs over the disruption scenarios. Under mild assumptions, this two-stage optimization problem can be described as a large MILP model. To solve this model, we propose a Benders decomposition approach for solving the LP-relaxation, leading to a subproblem for each disruption scenario. Then we compute heuristic robust rolling stock circulations based on the cuts that were generated by the Benders decomposition. We call our method *Benders heuristic*. Our computational results on relevant real-life rolling stock planning instances of NS indicate that the Benders decomposition approach takes relatively short computation time, thereby outperforming the straightforward solution of the whole LP-relaxation by a state-of-the-art solver. Moreover, the *Benders heuristic* takes very little additional time after the solution of the LP-relaxation, providing near-optimal solutions.

Due to computation time and space limitations, the generation of the robust rolling stock circulations is based on a limited number of disruption scenarios (about 30). Therefore, we also propose an evaluation framework where solutions to the rolling stock planning problem can be evaluated on a much larger number of disruption scenarios (3,500 in our case). It turns out that the robust rolling stock circulations admit much easier recovery than the nonrobust solutions. However, they have only slightly higher nominal costs.

The current paper is related to Nielsen, Kroon, and Maróti (2009), but has a different focus. Indeed, the earlier paper focuses on the *operational* problem of finding a recovered rolling stock circulation in case of a large disruption of the railway system. The current paper focuses on the *tactical* problem of designing the original rolling stock circulation in such a way that

the recoverability of this circulation in case of a large disruption is increased.

This paper is structured as follows. In the remainder of this section, we give an overview of the literature on robust resource planning in airlines and railway systems. In §2 we describe our robust optimization approach based on Benders decomposition. We also present our Benders heuristic, which is used to calculate robust integer solutions. Section 3 describes the railway rolling stock scheduling problem of NS. Section 4 is devoted to our computational results. We present our evaluation framework and its results in §5. Finally, §6 outlines some conclusions and final remarks.

1.2. Literature

Most robust scheduling applications in the passenger railway context concern the timetabling problem. Kroon et al. (2007) describe a stochastic programming model for improving the absorption robustness of a cyclic timetable. Similar results are presented by Liebchen et al. (2010). Fischetti, Salvagnin, and Zanette (2009) propose the notion of light robustness as an alternative to solving large-scale stochastic programs. Cacchiani, Caprara, and Fischetti (2012) solve the robust noncyclic timetabling problem by adding a simple buffer time measure to a Lagrangian relaxation framework. De Almeida et al. (2008) describe a model for robust rolling stock scheduling in case of relatively small train delays. These papers focus on delay absorption, and consider retiming of trains as the only way to react to the disturbances.

Robustness can be formalized mathematically in many different ways. Some applications consider simple, practice-driven measures, such as time supplements in process times or buffer times between flights or trains. More explicit robust models are based on stochastic programming (see Birge and Louveaux 1997) or on robust optimization (see Ben-Tal and Nemirowski 1998; Ben-Tal, El Ghaoui, and Nemirowski 2009; Bertsimas and Sim 2004). However, stochastic programming requires the use of probabilities for the occurrences of the different disruption scenarios, which may be hard to obtain in practice. Robust optimization may result in conservative solutions, because it aims at finding solutions that are feasible under *all* disruption scenarios. These approaches mainly focus on absorption robustness.

Liebchen et al. (2009) recently introduced the concept of recoverable robustness as a generic framework for modeling robustness with a focus on recoverability. In case of a disruption, they allow a feasible solution to be modified by a recovery algorithm. As a measure of robustness, the authors use the *maximum deviation* of the recovered solutions from the original solution, where the maximum is taken over a set of disruption scenarios. The deviation is an indication of the effort required to modify the nominal solution into a recovered solution. The approach based on the maximum deviation relieves them from the use of probabilities for the occurrences of the different disruption scenarios. In principle, the number of disruption scenarios may be infinite. Liebchen et al. (2009) also define the price of robustness as the relative increase in nominal costs due to the improved recoverability.

Cicerone et al. (2007) consider the recoverable robustness of several railway shunting problems. They analyze the price of robustness for a few concrete recovery algorithms and prove lower and upper bounds. In Cicerone et al. (2009) also timetabling problems are considered. Caprara et al. (2008) propose an exact method for computing recoverable robust solutions by optimally buffering a network for the train platforming problem. Other references to recoverable robustness are Stiller (2008) and Erera, Morales, and Savelsbergh (2009).

In the airline context, robustness has been a research topic for quite some time. Relatively recent work on robust airline-crew scheduling includes Ehrgott and Ryan (2002) and Shebalov and Klabjan (2006). These approaches describe measures to estimate how well a crew schedule is likely to be able to deal with flight delays, the first being related to absorption capacity, the second to recoverability. To that end, Shebalov and Klabjan (2006) introduce the concept of *move-up* crews. Furthermore, Yen and Birge (2006) propose a stochastic programming model for airline crew scheduling, minimizing the sum of the operational costs and the expected recovery costs. The methods of these papers focus primarily on minimizing passenger delays and are tailored to crew-scheduling applications.

With respect to the existing literature, the main contribution of the current paper consists of deriving robust solutions for an important real-world railway planning problem. In fact, we are able to practically prove that our robust solutions obtained on a limited but representative set of scenarios are also robust for a much larger set of scenarios. On one hand, our approach is similar to two-stage stochastic optimization and recoverable robustness, because we consider the rescheduling phase as a second optimization stage. On the other hand, in contrast with stochastic optimization and recoverable robustness, we prefer not to rely on any probability distribution nor to restrict ourselves to a limited set of recovery algorithms.

2. Two-Stage Optimization and Benders Decomposition

In this section we describe a two-stage optimization model to improve the robustness of rolling stock

circulations. The model is a large MILP. We use a solution approach based on Benders decomposition to solve the associated LP-relaxation. Moreover, the cuts generated by the Benders decomposition are used for generating robust integer solutions in a heuristic way.

Our objective is to find a solution to the rolling stock planning problem that is robust under a finite set of disruption scenarios. That is, we want to find a feasible initial rolling stock circulation and at the same time a feasible recovered solution for each disruption scenario. Here we minimize the sum of the cost of the initial rolling stock circulation problem and the maximum recovery cost for transforming this initial solution into a recovered solution, where the maximum is taken over all disruption scenarios.

Given that the notions presented in this section apply to a generic planning problem, we postpone all details about our specific application to §3, mentioning this application here only as a concrete example.

2.1. Definitions
We consider a generic planning problem, shortly, the *nominal problem* (NP), given in the following form:

$$\text{NP} = \min\{c(x) \mid x \in K\}. \tag{1}$$

Here $x \in \mathbb{R}^n$ is the vector of decision variables, $K \subseteq \mathbb{R}^n$ is the feasible region, and $c: K \to \mathbb{R}_+$ is the cost function. In general, some components of the variable vector x are required to be integer valued, which is the case for the nominal rolling stock circulation problem.

Moreover, there is a finite set Σ of *disruption scenarios*, where each disruption scenario $\sigma \in \Sigma$ has its own feasible region K_σ. For example, a disruption scenario may refer to the case of canceling some trains due to infrastructure or rolling stock failures, thereby requiring some kind of recovery action. Furthermore, there is a *recovery algorithm A*. This recovery algorithm takes as input a nominal feasible solution $x \in K$ and a disruption scenario $\sigma \in \Sigma$, and produces a recovered solution $x_\sigma = A(x, \sigma) \in K_\sigma$ that is feasible for disruption scenario σ. Finally, there are functions $\Delta_\sigma: K \times K_\sigma \to \mathbb{R}_+$ measuring the *deviation* $\Delta_\sigma(x, x_\sigma)$ of a recovered solution x_σ for disruption scenario σ from a nominal solution x, and a monotone nondecreasing function $f: \mathbb{R}_+^\Sigma \to \mathbb{R}_+$ penalizing the deviation over all disruption scenarios. The aim is to minimize the deviations from the nominal solution. For a given nominal solution x, the deviations of the recovered solutions x_σ from x are denoted by a vector $z = (z_{\sigma_1}, z_{\sigma_2}, \ldots) \in \mathbb{R}_+^\Sigma$ of auxiliary variables. All deviations z_{σ_i} are nonnegative.

Then the *two-stage optimization problem* (TSOP) that we consider in this paper is formulated as follows:

$$\text{TSOP} = \min\{c(x) + f(z) \mid x \in K, \; x_\sigma = A(x, \sigma) \in K_\sigma,$$
$$z_\sigma = \Delta_\sigma(x, x_\sigma) \; \forall \sigma \in \Sigma\}. \tag{2}$$

In this paper we choose $f(z) = \max_{\sigma \in \Sigma} z_\sigma$. That is, we penalize the *maximum deviation* in the objective function. This is in line with the recoverable robustness approach of Liebchen et al. (2009), and it relieves us from the use of probabilities for the different disruption scenarios. Although this approach may seem to be rather conservative, our computational results show that it allows us to compute robust solutions that are more easily recoverable than the nonrobust solutions, and that do not have much higher nominal costs. A stochastic programming approach would be to use a probability p_σ for each disruption scenario $\sigma \in \Sigma$ and to consider $f(z) = \sum_{\sigma \in \Sigma} p_\sigma z_\sigma$, thereby penalizing the *expected* (or *average*) *deviation*.

2.2. Reformulation as an MILP
Our two-stage optimization model can be applied if the following conditions hold:
- For the recovery costs $f(z)$ we have $f(z) = \max_{\sigma \in \Sigma} z_\sigma$, i.e., the largest deviation from the nominal solution is penalized in the objective function.
- For the nominal objective function $c(x)$ we have $c(x) = c^\mathsf{T} x$ for a given $c \in \mathbb{R}^n$. Furthermore, the set K of feasible solutions to the nominal problem can be expressed as $\{x \mid Ax \geq b, x \text{ integer}\}$ for given $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Thus, the nominal problem can be described as an MILP.
- For each disruption scenario $\sigma \in \Sigma$, the set of feasible recovered solutions K_σ can be expressed as $\{x_\sigma \mid A_\sigma x_\sigma \geq b_\sigma, x_\sigma \text{ integer}\}$ for given $A_\sigma \in \mathbb{R}^{m_\sigma \times n}$ and $b_\sigma \in \mathbb{R}^{m_\sigma}$. Furthermore, the deviation $z_\sigma = \Delta_\sigma(x, x_\sigma)$ of a recovered solution x_σ from a nominal solution x can be expressed as a generic linear function of x and x_σ, as follows: $z_\sigma = d_\sigma^\mathsf{T} x + e_\sigma^\mathsf{T} x_\sigma + k_\sigma$, where $d_\sigma, e_\sigma \in \mathbb{R}^n$ and $k_\sigma \in \mathbb{R}$ are given coefficients that depend on the specific problem. Thus, finding a recovered solution with minimal deviation from a nominal solution can be described as an MILP.

In our rolling stock application described in §3, we show that these assumptions are satisfied indeed, and we indicate how the required coefficients are actually computed. For example, (27) defines the deviation z_σ for this application. Note that (27) is defined in terms of the parameters and the variables of the application, so that the abstract variable z_σ is not mentioned explicitly there.

Given the above assumptions on TSOP, the problem (2) can be formulated as in (3)–(9) below. We call this model the *robust model*. Here λ is an auxiliary variable expressing the recovery cost for the worst-case disruption scenario. We want to point out that an arbitrary feasible nominal solution x with arbitrary feasible recovered solutions x_σ and a sufficiently large λ forms a feasible solution to (3)–(9).

$$\min \quad c^\mathsf{T} x + \lambda \tag{3}$$

$$\text{s.t.} \quad Ax \geq b \tag{4}$$

$$x \geq 0, \text{ integer} \tag{5}$$

$$\lambda \geq 0 \tag{6}$$

$$-d_\sigma^\mathsf{T} x - e_\sigma^\mathsf{T} x_\sigma + \lambda \geq k_\sigma \quad \forall \sigma \in \Sigma, \tag{7}$$

$$A_\sigma x_\sigma \geq b_\sigma \quad \forall \sigma \in \Sigma, \tag{8}$$

$$x_\sigma \geq 0 \text{ integer} \quad \forall \sigma \in \Sigma. \tag{9}$$

In the computational experiments carried out in this paper, we will initially restrict our attention to the case in which the integrality restrictions on x and x_σ are not imposed at all. This yields a lower bound on TSOP. Then, we propose a heuristic method called *Benders heuristic* that uses the cuts generated by the Benders decomposition (which is described in the next section) to find integral solutions for the *robust model* (3)–(9).

2.3. Benders Decomposition for Solving the LP-Relaxation

For solving the LP-relaxation of (3)–(9) one can apply various mathematical programming techniques. In this paper we focus on Benders decomposition, a cutting plane method that exploits the block-diagonal structure of (3)–(9). This method is also known as the L-shaped method (see Benders 1962; Geoffrion 1972). It is an approach that is widely used for solving such problems, for example, for solving stochastic programming problems.

Briefly, the Benders decomposition approach solves the (gradually extended) LP-relaxation of (3)–(6) called the *master problem*. Note that the initial master problem is equivalent to the LP-relaxation of the nominal problem, because $\lambda = 0$ yields the nominal objective function. Based on the optimal solution of the current master problem, the feasibility of the LP-relaxations of the subproblems (7)–(9) is checked. The procedure terminates if the LP-relaxations of the subproblems are feasible, in which case the current optimal solution is optimal also for the LP-relaxation of (3)–(9). In case of infeasibility, inequalities in terms of x and λ are derived and added to the master problem, and the updated master problem is reoptimized.

2.3.1. Benders Heuristic. Here we describe our *Benders heuristic*, which results in a heuristic integer solution to the robust model (3)–(9). We take advantage of the Benders cuts derived when solving the LP-relaxation of (3)–(9), using these cuts to guide the search method rather than dealing with the disruption scenarios explicitly.

More precisely, we first apply the Benders decomposition method to solve the LP-relaxation of (3)–(9). This yields the following Benders cuts:

$$Bx + g\lambda \geq h. \tag{10}$$

Next we solve the following MILP, in which the Benders cuts (10) encode relevant information about the considered disruption scenarios:

$$\min \quad c^\mathsf{T} x + \lambda \tag{11}$$

$$\text{s.t.} \quad Ax \geq b \tag{12}$$

$$x \geq 0, \text{ integer} \tag{13}$$

$$\lambda \geq 0 \tag{14}$$

$$Bx + g\lambda \geq h. \tag{15}$$

Note that this MILP has only variables associated with the nominal problem, which makes it much smaller than (3)–(9). In our application, it is much faster to solve this MILP (11)–(15) than the LP-relaxation of (3)–(9) itself. Representing the scenarios through inequalities involving only the nominal variables is another major advantage of using Benders decomposition, in addition to the smaller LP-relaxation solution times.

Let $(\bar{x}, \bar{\lambda})$ be the optimal solution to (11)–(15). Then $\bar{x}$ is a feasible solution to the nominal problem (3)–(6), but $\bar{\lambda}$ is merely *an estimate* (a lower bound) on the *actual recovery costs* for $\bar{x}$. Nevertheless, one can compute a feasible recovered solution $\bar{x}_\sigma$ and the actual recovery cost for each disruption scenario $\sigma \in \Sigma$ by solving the following MILP:

$$\bar{x}_\sigma = \arg\min\{\Delta_\sigma(\bar{x}, x_\sigma) : x_\sigma \in K_\sigma\}. \tag{16}$$

Then, letting

$$\hat{\lambda} = \max_{\sigma \in \Sigma} \Delta_\sigma(\bar{x}, \bar{x}_\sigma), \tag{17}$$

the variable $\hat{\lambda}$ and the vectors $\bar{x}, \bar{x}_\sigma$ form a feasible solution to the TSOP (3)–(9). This solution is the output of our Benders heuristic.

So, to summarise, our Benders heuristic consists of three phases:

1. Apply Benders decomposition to solve the LP-relaxation of (3)–(9) and hence derive the Benders cuts (10);

2. Solve MILP (11)–(15) and get the estimated recovery costs $\bar{\lambda}$ and the corresponding solution $\bar{x}$;

3. For each disruption scenario $\sigma \in \Sigma$, solve MILP (16) and, finally, compute the actual recovery costs $\hat{\lambda}$ of solution $\bar{x}$ using (17).

To compare the heuristic robust solution to the optimal solution x^* of the nominal problem, we compute the maximal recovery cost λ^* of x^* over all disruption scenarios, similarly to (16) and (17):

$$x^*_\sigma = \arg\min\{\Delta_\sigma(x^*, x_\sigma) : x_\sigma \in K_\sigma\}, \quad \forall \sigma \in \Sigma$$

$$\lambda^* = \max_{\sigma \in \Sigma} \Delta_\sigma(x^*, x^*_\sigma).$$

The comparison of $\hat{\lambda}$ and λ^* indicates how much easier it is to perform recovery for the robust solution

than for the optimal nominal solution. Note that if the robust model (3)–(9) could be solved to optimality, then obviously $\hat{\lambda} \leq \lambda^*$. However, in general the model (11)–(15) will only lead to a heuristic solution for the robust model (3)–(9).

3. Application: Robust Rolling Stock Planning

This section is devoted to the description of the specific real-world case study on which we focus our attention.

3.1. The Nominal Problem

We consider the medium-term railway rolling stock scheduling problem of NS (the main operator of passenger trains in the Netherlands). It arises two–six months before the actual railway operations and has the task of assigning the available rolling stock to the *trips* of a given timetable. In this section we give a brief problem description. Further details about the problem can be found in Fioole et al. (2006) and in Maróti (2006). Moreover, for a literature survey on rolling stock scheduling the reader is referred to Caprara et al. (2007). We note that the purpose of the present paper is not to improve the existing methods for nominal rolling stock scheduling, but to show how our two-stage optimization model applied on top of an existing successful method for solving the nominal rolling stock scheduling problem can lead to more robust solutions.

The rolling stock consists of *train units*. Each unit has driver's seats at both ends and an own engine so that it can operate autonomously in both directions. It is composed of a number of carriages and cannot be split up in everyday operations. Units are available in different *types* and can be combined with each other to form *compositions*. This allows a fine adjustment of the seat capacity to the passenger demand. For example, if there are train types *a* and *b*, then *aab* and *aba* are compositions consisting of three train units. These compositions have the same capacity. However, their *shunting* possibilities are different: from the first composition, the unit of type *b* can be uncoupled easily, which is not the case for the second one. Thus, the order of the train units in the trains is an important issue.

The timetable of NS is quite dense, and the turning times of the trains at the end points of a line are short, often less than 20 minutes. The timetable is given in the form of a number of trips, each one with a start and end station, and with a start and end time. A trip starts and ends at a station where the composition of a train can be modified by a shunting operation. Because of the short turning times, the possibilities to modify a train composition are limited to *coupling* or *uncoupling* of one or two units at the appropriate side of the train. Each trip also has a *successor trip*: the units

that serve in a trip generally go over to the successor trip, although a *composition change* may take place in between a trip and its successor. Similarly, each trip also has a *predecessor trip*. We note that the last trips of the day have a dummy successor trip, whereas the first trips of the day have a dummy predecessor trip. Finally, the end-of-day rolling stock balances must be such that by the end of the day the train units are at the right stations for the next day's operations.

The objective of the nominal problem is threefold. *Service quality* is measured by seat shortage kilometers. It is computed for each trip by comparing the assigned seat capacity to the a priori given anticipated number of passengers; by multiplying the anticipated number of unseated passengers by the length of the trip; and, finally, by summing these values over all trips. *Efficiency* is expressed by the sum over all units of the kilometers traveled, which accounts for the operating costs due to electricity or fuel consumption and maintenance. *Robustness* is taken into account by counting the number of composition changes. Indeed, coupling or uncoupling of units causes additional traffic through the railway nodes, and thereby may lead to delay propagation if some passing trains are late.

We note again that the nominal problem is solved several months before the operations. This leaves enough time to plan the detailed shunting operations at the railway nodes. In particular, shunting drivers are scheduled to carry out the coupling and uncoupling operations.

3.1.1. Model Description. The model for solving the nominal rolling stock circulation problem is very similar to the model described by Fioole et al. (2006). It is basically an integer multicommodity flow model with several additional constraints, which are needed to describe the limited shunting possibilities in the stations. The unique simplification used here with respect to Fioole et al. (2006) concerns the fact that we do not consider the reallocation time that one has to wait before reusing a train unit that was uncoupled. Given that the two models are almost identical, we use here an almost identical notation (by only changing the variable names).

To formulate the nominal rolling stock circulation problem as an MILP, the set of trips is denoted by T, the set of train unit types by M, the set of all possible compositions by P, and the set of stations by S. Let $k_{p,m}$ denote the number of train units of type m in composition p. The time horizon covered by the model is typically one day. Thus, the set T contains all trips to be carried out in a single day. Each trip $t \in T$ has a departure station $d(t)$, an arrival station $a(t)$, and a successor trip $\nu(t)$. For each train unit type $m \in M$, the parameter n_m denotes the number of available train units of type m. Furthermore, we

introduce the set P^s as the set of composition changes $(p \to p')$ from composition p to composition p' involving a shunting operation. That is, at least one train unit is coupled or uncoupled to/from composition p to get composition p'.

The main binary decision variables are the variables $u_{t,p}$, whose value is 1 if composition p is assigned to trip t. Moreover, we have binary variables $v_{t,p,p'}$ whose value is 1 if composition p is assigned to trip t and composition p' is assigned to the successor trip $\nu(t)$ of trip t. These v-variables are only defined for those triples (t, p, p') where the composition change from composition p to composition p' is allowed between trips t and $\nu(t)$. Thus, the local constraints on the composition changes are represented by these v-variables.

The stations are modeled as *inventories* of train units. The inventory of a station at a certain time instant consists of all train units that are temporarily located at that station because they have not been assigned to any trip at that moment. Train units coupled to a train are subtracted from the inventory, and train units uncoupled from a train are added to the inventory. The start-of-day and end-of-day inventories of station s of train units of type m are represented by the variables $y_{s,m}^0$ and $y_{s,m}^\infty$. Whereas the former represent a decision to be taken, the latter are a linear function of the other variables and are introduced only for convenience because the recovery cost takes into account the discrepancies between end-of-day inventories in the nominal and disrupted scenarios.

The parameter $\alpha_{t',p,p',s,m}$ measures the effect on the end-of-day inventory of train unit type m at station s as a result of a shunting operation after trip t' that transforms composition p on trip t' into composition p' on trip $\nu(t')$:

$$\alpha_{t',p,p',s,m} = \begin{cases} k_{p,m} - k_{p',m} & \text{if station } s \text{ is the arrival} \\ & \text{station of trip } t', \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, the parameters $\beta_{t,t',p,p',m}$ indicate how the inventory of train unit type m at the departure station of trip t right after the departure time of trip t increases or decreases as a result of an earlier shunting operation that transforms composition p on trip t' into composition p' on trip $\nu(t')$:

$$\beta_{t,t',p,p',m}$$

$$= \begin{cases} k_{p,m} - k_{p',m} & \text{if the arrival station of trip } t' \text{ is the} \\ & \text{departure station of trip } t, \text{ and if} \\ & \text{trip } t' \text{ arrives earlier at this station} \\ & \text{than the departure time of trip } t, \\ 0 & \text{otherwise.} \end{cases}$$

The objective function coefficients $c_{t,p}^c$ describe aspects such as seat shortages and carriage kilometers; the objective function coefficients $c_{t,p,p'}^s$ describe the complexity and risk of the shunting operations. Then the MILP formulation that we consider here is the following.

$$\min \sum_{t \in T} \sum_{p \in P} c_{t,p}^c \cdot u_{t,p} + \sum_{t \in T} \sum_{(p \to p') \in P^s} c_{t,p,p'}^s \cdot v_{t,p,p'} \quad (18)$$

$$\text{s.t.} \sum_{p \in P} u_{t,p} = 1 \quad \forall t \in T, \quad (19)$$

$$u_{t,p} = \sum_{p' \in P} v_{t,p,p'} \quad \forall t \in T, \ p \in P, \quad (20)$$

$$u_{\nu(t),p'} = \sum_{p \in P} v_{t,p,p'} \quad \forall t \in T, \ p' \in P, \quad (21)$$

$$y_{s,m}^\infty = y_{s,m}^0 + \sum_{t' \in T} \sum_{(p \to p') \in P^s} \alpha_{t',p,p',s,m} \cdot v_{t',p,p'}$$
$$\forall s \in S, \ m \in M, \quad (22)$$

$$y_{d(t),m}^0 + \sum_{t' \in T} \sum_{(p \to p') \in P^s} \beta_{t,t',p,p',m} \cdot v_{t',p,p'} \geq 0$$
$$\forall t \in T, \ m \in M, \quad (23)$$

$$\sum_{s \in S} y_{s,m}^0 = n_m \quad \forall m \in M, \quad (24)$$

$$u_{t,p}, v_{t,p,p'} \text{ binary} \quad \forall t \in T, \ p \in P, \ p' \in P, \quad (25)$$

$$y_{s,m}^0, y_{s,m}^\infty \geq 0, \text{ integer} \quad \forall s \in S, \ m \in M. \quad (26)$$

The objective function (18) takes into account the assignment of compositions to trips and the shunting operations between successive trips. This objective function can incorporate a wide variety of objective criteria related to service quality, efficiency, and robustness. Constraints (19) state that each trip must be assigned exactly one composition. Constraints (20) state that if composition p is assigned to trip t, then the composition p' that is assigned to the successor trip $\nu(t)$ is selected from one of the compositions p' that can be reached from composition p via an allowed shunting operation. Similarly, constraints (21) link the compositions on a trip to compositions on its predecessor trip. Constraints (22) describe that, by the end of the day, the final inventory $y_{s,m}^\infty$ of train units of type m at station s equals the initial inventory $y_{s,m}^0$ plus the increases and decreases of the inventory depending on the local shunting operations, as was explained in the definition of the parameters $\alpha_{t',p,p',s,m}$. In a very similar way, constraints (23) consider the increases and decreases of the inventory of train units of type m at station $d(t)$ right after the departure time of any trip t, imposing that this inventory is always nonnegative. Finally, constraints (24) specify that all available rolling stock is in the inventory of one of the stations by the start of the day.

Experience has shown that for practically meaningful objective coefficients, the LP-relaxation of the model (18)–(26) is very tight, the associated lower bound being always within a few percent of the MILP optimum. As a consequence, for most instances of NS the MILP model (18)–(26) can be solved to optimality within a reasonable computing time with CPLEX (see Fioole et al. 2006).

3.2. The Disruption Scenarios and the Associated Deviations

In our robustness framework, the solutions of the nominal problem are to be operated subject to a number of disruption scenarios. Each disruption scenario is obtained by assuming that a certain part of the network is blocked for a certain time interval of several hours. All trips that interfere directly with the infrastructure blockage are canceled. Such disruptions are quite common in practice; these are the ones that require significant resource rescheduling.

In the Netherlands, in case of a disruption, the timetabling and resource-rescheduling decisions are taken consecutively. The modification of the timetable is determined first. Thereafter, the rolling stock circulation is modified so that it again fits with the timetable. The crew duties are rescheduled in a third step. Therefore, the adjusted timetable that takes care of the disruption is to be considered as input when rescheduling the resources.

We assume that a disruption becomes known at the start of the blockage. Furthermore, we also assume that the timetable has been modified accordingly. The task is then to reschedule the rolling stock from that point on until the end of the day. The solution has to fulfill the same requirements as the nominal problem, the only additional option being the cancellation of trips if there is not enough rolling stock to cover all trips. However, canceling additional trips due to lack of rolling stock is highly undesirable.

In this research we also assume that the exact duration of the disruption is known at its start. Admittedly, on one hand this assumption is rather optimistic for practical purposes. On the other hand, it simplifies the mathematical model, and still enables us to gain insight into the recovery capacity of the rolling stock circulation.

The three main criteria in rescheduling are the following (in decreasing order of importance): (i) minimize the number of canceled trips; (ii) minimize the number of newly introduced shunting operations (units that are coupled or uncoupled in the recovered plan, and not in the original one); and (iii) minimize the deviations from the planned end-of-day rolling stock balances. The first criterion limits the passenger inconvenience. The second criterion aims at keeping the schedule of the shunting drivers intact. The third

criterion tries to restrict the consequences of the disruption to a single day.

3.2.1. Extended Model Description. Although the model (18)–(26) was originally developed for solving the nominal problem, it can be adjusted for solving the recovery problem as well. That is, the feasibility of a recovered solution and the associated recovery costs can be computed as a variant of the nominal model. We note that the real-time rescheduling framework of Nielsen, Kroon, and Maróti (2009) is based on a similar extension to the basic model of Fioole et al. (2006).

Below, we describe how the nominal model (18)–(26) must be extended to represent the complete robust model. Note that a complete description of the robust model is also given in the appendix of this paper.

In addition to the constraints (19)–(26) for the nominal part of the model (nom), these constraints (19)–(26) are to be replicated for each scenario σ with variables $u_{t,p}^\sigma$, $v_{t,p,p'}^\sigma$, and $y_{s,m}^{\infty,\sigma}$. Thus, for example, if $u_{t,p}^\sigma = 1$, then rolling stock composition p is assigned to trip t in the recovered solution for scenario σ. For the scenarios we also allow the assignment of the empty composition $\varnothing$, where $u_{t,\varnothing}^\sigma = 1$ means that trip t is canceled in scenario σ due to lack of rolling stock.

Next, one has to impose that in each scenario the rolling stock circulation cannot be changed until the start of the disruption. To this end, the set of trips for which the assigned composition cannot be changed in scenario σ is called $\bar{T}^\sigma$, and we add the constraints $u_{t,p}^\sigma = u_{t,p}^{\text{nom}}$ for each trip $t \in \bar{T}^\sigma$ and composition $p \in P$ (please refer again to the appendix of this paper).

Finally, the model is extended to express the recovery costs. Altogether, for each disruption scenario σ, the following constraints must be satisfied:

$$\lambda \geq c_1 \sum_{t \in T} u_{t,\varnothing}^\sigma + c_2 \sum_{t \in T} r_t^\sigma + c_3 \sum_{s \in S} \sum_{m \in M} q_{s,m}^\sigma \quad (27)$$

$$q_{s,m}^\sigma \geq y_{s,m}^{\infty,\text{nom}} - y_{s,m}^{\infty,\sigma} \quad \forall s \in S, \ m \in M, \quad (28)$$

$$q_{s,m}^\sigma \geq y_{s,m}^{\infty,\sigma} - y_{s,m}^{\infty,\text{nom}} \quad \forall s \in S, \ m \in M, \quad (29)$$

$$w_t^{\text{nom}} = \sum_{(p \to p') \in P^s} v_{t,p,p'}^{\text{nom}} \quad \forall t \in T, \quad (30)$$

$$w_t^\sigma = \sum_{(p \to p') \in P^s} v_{t,p,p'}^\sigma \quad \forall t \in T, \quad (31)$$

$$r_t^\sigma \geq w_t^\sigma - w_t^{\text{nom}} \quad \forall t \in T, \quad (32)$$

$$y_{s,m}^{\infty,\sigma}, q_{s,m}^\sigma \text{ integer} \geq 0 \quad \forall s \in S, \ m \in M, \quad (33)$$

$$r_t^\sigma, w_t^\sigma \text{ binary} \quad \forall t \in T. \quad (34)$$

The auxiliary variables $q_{s,m}^\sigma$ measure the deviation of the realized end-of-day rolling stock balances for scenario σ from the planned end-of-day rolling stock balances in the nominal solution (constraints (28)–(29)). The value of the auxiliary variable w_t^{nom}

is 1 if the nominal solution has a composition change (i.e., coupling or uncoupling of units) between trips t and $\nu(t)$, and it is 0 otherwise. A similar role is played by the variable w_t^σ in the recovered solution for scenario σ. Thus, the value of the auxiliary variable r_t^σ equals 1 if the recovered solution for scenario σ has a composition change between trips t and $\nu(t)$, and the nominal solution does not, and it is 0 otherwise (constraints (30)–(32)).

The recovery costs (27) include the variables $q_{s,m}^\sigma$ and r_t^σ, as well as all variables $u_{t,\varnothing}^\sigma$ that assign an empty composition to trip t in scenario σ, i.e., trip t has been canceled due to lack of rolling stock. The objective function of the robust model consists of (18) summed to λ.

Note that with the given definitions, each disruption scenario has at least one feasible recovered solution. Indeed, one may cancel all trips from the start of the disruption until the end of the day. Of course, this solution has very high costs, so hopefully cheaper recovered solutions also exist.

4. Computational Results

We implemented the model (3)–(9) for the case of the rolling stock (re-)scheduling problem by using the models described in §3. As indicated earlier, we will refer to this model as the *robust model*.

In our application, we studied the so-called *3,000 line* of NS from Den Helder to Nijmegen; see Figure 2. This is an intercity line with a closed rolling stock
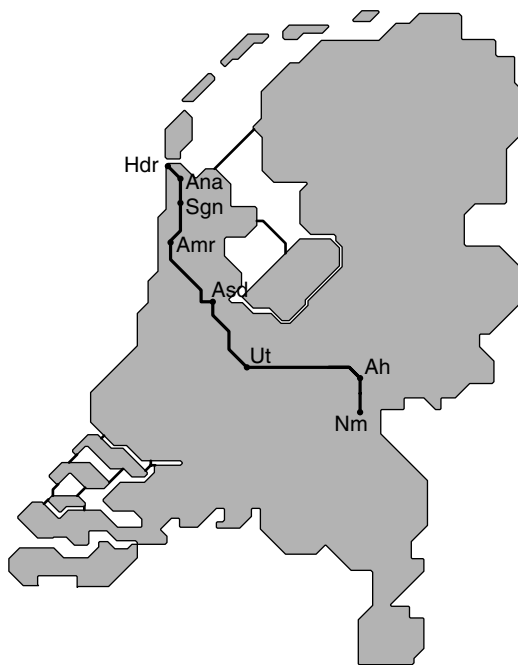

Figure 2   The 3,000 line Connecting Den Helder (Hdr) to Nijmegen (Nm) via Ana Paulowna (Ana), Schagen (Sgn), Alkmaar (Amr), Amsterdam (Asd), Utrecht (Ut), and Arnhem (Ah)

circulation. Because all trains in our instance serve only this intercity line, all trains have the same priority. The instance contains about 400 trips connecting eight stations; that is, the line is divided into seven subtrajectories.

The 3,000 line is served by two types of train units with 11 and 24 units, respectively; restrictions on the train lengths allow eight different compositions for each trip. Denoting the two rolling stock types by a and b, the feasible compositions are $\varnothing$, a, b, aa, ab, ba, bb, and aaa. Composition changes (i.e., coupling or uncoupling) are possible at the terminals Den Helder (Hdr) and Nijmegen (Nm) as well as at the underway stations Alkmaar (Amr) and Arnhem (Ah). The 3,000 line is one of the medium-sized rolling stock instances of NS.

All computations have been carried out on a PC with a Pentium IV 3.2 GHz processor and with 2 GB memory, solving the LPs and MILPs by ILOG CPLEX 10.0. Our computer codes are written in the C language.

The nominal problem is based on the actual one-day timetable of NS. The disruption scenarios have been generated using a program of Nielsen (2008), which also generates the decisions about train cancellations, including the possible changes in the successors of the trips during the disruption. In each disruption, one of the seven subtrajectories of the line is blocked for a certain time period (from one to four hours), and the timetable is modified accordingly.

As mentioned earlier, Figure 1 shows an example of a disrupted timetable in a time-space diagram. A disruption in this example occurs between stations Alkmaar (Amr) and Amsterdam (Asd) from 17:30 until 19:30. Some trips of the original schedule are thus canceled and new successors are defined for some other trips.

Our solution method turns out to be able to deal with a limited number of disruption scenarios. Specifically, for the case study considered and the PC we used, the solution of the LP-relaxation by Benders decomposition runs out of memory for more than 28 scenarios. Accordingly, we selected a set of 28 representative disruption scenarios to conduct our tests, evenly spread throughout the day and among the subtrajectories. Specifically, we generated a large set of scenarios and then selected these 28 by guaranteeing that (i) for each subtrajectory and time instant, there is a scenario in which the subtrajectory is blocked in that instant, (ii) there is a wide variability of the block durations, ranging from one to six hours. The selection of such a restricted number of scenarios is unavoidably a rough choice. To testify a posteriori the representativeness of the selected scenarios, we will also report on the quality of the solutions found

Table 1 Coefficients for the Nominal Objective Function as Well as for the Recovery Costs

Criteria in the nominal problem	
Seat shortage km	100
Traveled km	9
Composition change	5
Criteria for recovery	
Cancelation	1,000,000
Inventory deviation	20,000
New shunting	10,000

with respect to a much larger (two orders of magnitude) number of disruption scenarios.

As for the nominal objective functions and for the recovery costs, we considered the objective coefficients given in Table 1. The reason for this choice is to focus on service quality by heavily penalizing seat shortages and additional trip cancelations. Note that a number of trips must be canceled anyhow due to the disruption of the railway system. The objective function aims at reducing the number of *additionally* canceled trips due to lack of rolling stock at the right time and place. We want to emphasize that although these cost coefficients are realistic for this particular scheduling problem, they might not perfectly reflect the decision makers' preferences. However, finding the best possible balance between the objective criteria goes beyond the purpose of this paper.

4.1. Solving the LPs
As already discussed, the goal of our first set of computational tests is to show how the suggested optimization framework can be effective to compute lower bounds on the optimal TSOP value.

We implemented two solution methods: (i) solving the LP-relaxation of (3)–(9) directly as a single LP; and (ii) applying a canonical Benders decomposition approach as described in §2.3.

The nonrobust nominal model (3)–(6) for our case study has about 14,600 variables, 8,400 constraints and 320,000 nonzeros in the coefficient matrix. The solution approaches have been tested with 1–28 disruption scenarios from the representative set discussed above, implying that the LPs solved by method (i) feature about 44,000–440,000 variables, 27,000–305,000 constraints, and 950,000–9,400,000 nonzeros.

The computational results with the two solution approaches are summarized in Table 2. In column *No. of scenarios* we indicate the number of disruption scenarios considered (for example, "5" means that we are dealing with the master problem and the first five disruption scenarios). Note that the case "0" corresponds to the nonrobust nominal model. Then we present the value of the LP-relaxation obtained with the two approaches and the corresponding computation time (expressed in seconds). The column *Nominal*

objective gives the value of (18) (i.e., the nominal value of the corresponding robust solution), and the column *Recovery cost* the value of λ, equal to the maximum right-hand side of (27) over all considered scenarios. The column *Robust objective* is the sum of the nominal objective and the recovery costs λ.

It turns out that when dealing with a large number of disruption scenarios, the huge LPs in method (i) require a rapidly increasing computation time. The computation time of the Benders decomposition approach in method (ii), on the other hand, increases in a much slower way with the number of disruption scenarios. The presented results prove that, at least for our case study, the LP relaxation can be solved within reasonable time for the considered 28 disruption scenarios.

The lower bounds obtained with the solution of the LP-relaxation of the problem are useful to get an idea of the increase in the costs when taking into account the disruption scenarios. In particular, because the objective value of the LP-relaxation of the nominal problem (with zero disruption scenarios) is 737,198 and the robust objective of the LP-relaxation is 781,296 when taking into account 28 disruption scenarios, we have an increase of up to 6% in the total costs of the solution. The 6% increase is almost entirely due to the recovery costs, whereas the nominal objective value changes by less than 0.6%, namely, from 737,198 to 741,296.

4.2. Solving the MILPs
Having dealt with the LP-relaxations, we tried to compute integer solutions to the robust rolling stock problem by solving the MILP (3)–(9). In Table 3 we show a first attempt of computing integer solutions using CPLEX directly. However, the computation time exceeded one day if more than 10 disruption scenarios were taken into account. Similar to the linear relaxation case, for the cases that could be solved by CPLEX, the robust solution has at most 0.5% higher nominal costs than the optimal nominal solution. Moreover, the gaps with respect to the LP lower-bound value in Table 2 are tiny.

We also tried to solve the robust model (3)–(9) by relaxing the integrality requirement on the x_σ variables. This leads to the same solutions as with integral x_σ variables, but it does not help to decrease the computation times significantly.

To deal with a larger number of disruption scenarios, we applied the Benders heuristic described in §2.3 by solving the MILP model (11)–(15) by CPLEX and then computing the actual recovery costs using (16)–(17). The computational results are given in Table 4. Here the column (11)–(15) *objective* is the objective value (11) of the model (11)–(15), subdivided

Table 2 **Comparison Between the Benders Decomposition Approach and Direct Solution of the LP-Relaxation by CPLEX**

No. of scenarios	Benders					CPLEX	
	Nominal objective	Recovery cost	Robust objective	Number of cuts	CPU time	Robust objective	CPU time
0	737,198	0	737,198	—	12	737,198	12
1	739,611	0	739,611	22	124	739,611	40
2	739,611	0	739,611	22	411	739,611	108
3	740,032	0	740,032	88	1,299	740,032	366
4	740,684	40,000	780,684	177	2,230	780,684	247
5	740,684	40,000	780,684	186	2,174	780,684	944
6	740,684	40,000	780,684	233	3,120	780,684	1,266
7	740,684	40,000	780,684	194	2,403	780,684	1,767
8	740,684	40,000	780,684	233	2,768	780,684	2,090
9	740,684	40,000	780,684	261	2,633	780,684	2,351
10	740,919	40,000	780,919	373	4,044	780,919	2,497
11	740,919	40,000	780,919	286	2,897	780,919	3,650
12	741,090	40,000	781,090	363	4,158	781,090	4,358
13	741,090	40,000	781,090	371	4,023	781,090	4,721
14	741,090	40,000	781,090	396	3,894	781,090	4,918
15	741,090	40,000	781,090	420	3,447	781,090	4,865
16	741,090	40,000	781,090	412	4,066	781,090	7,458
17	741,090	40,000	781,090	472	5,453	781,090	7,710
18	741,090	40,000	781,090	463	5,629	781,090	9,928
19	741,090	40,000	781,090	478	6,811	781,090	9,896
20	741,090	40,000	781,090	511	6,924	781,090	13,100
21	741,090	40,000	781,090	568	7,237	781,090	17,860
22	741,090	40,000	781,090	515	7,591	781,090	16,469
23	741,090	40,000	781,090	531	7,248	781,090	17,964
24	741,090	40,000	781,090	553	8,456	781,090	20,339
25	741,090	40,000	781,090	537	7,618	781,090	24,753
26	741,090	40,000	781,090	551	8,733	781,090	30,350
27	741,090	40,000	781,090	580	8,664	781,090	33,919
28	741,296	40,000	781,296	642	12,408	781,296	32,659

Note. The recovery costs in the CPLEX computations are equal to those in the Benders decomposition approach.

into its two components, *nominal objective* $c^\mathsf{T}\bar{x}$ and *estimated rec. cost* $\bar{\lambda}$, where $(\bar{x}, \bar{\lambda})$ is the optimal solution to (11)–(15). In addition, the column *Actual rec. cost* shows the actual recovery costs $\hat{\lambda}$ as defined in (16)–(17). The column *B-Heur objective* gives the sum of the nominal objective value $c^\mathsf{T}\bar{x}$ and the actual recovery costs $\hat{\lambda}$. The column *B-Heur CPU* represents the computing time for finding the solution to the

Benders heuristic once the Benders cuts have been generated (i.e., the time to solve MILP (11)–(15) and (16)–(17)). It turns out that the solution times for the Benders heuristic are within two minutes for up to 28 disruption scenarios. Finally, the LP gap is given in the column *LP gap %*. This column gives the relative difference between the objective function of the Benders heuristic and the objective function of the LP-relaxation of the robust model (3)–(9).

As one might expect, the MILP (11)–(15) underestimated the actual recovery costs; the difference is rather large in some cases. In most cases, however, the robust solutions are of very good quality, as can be verified by comparing the objective values in Tables 3 and 4. The robust solutions are in fact optimal in the case of $1, 4, 5, 6, 7,$ and 8 disruption scenarios, and very close to optimal in the case of $2, 9,$ and 10 disruption scenarios. For more than 10 disruption scenarios, we cannot compare with the optimal solution value, but only with the lower bound value in Table 2.

Note that there are two cases, namely, for 18 and 21 scenarios, in which the recovery cost is fairly bad, due to the approximation in representing these costs

Table 3 **Solving the Robust Optimization Problem (3)–(9) Directly as an MILP by CPLEX**

No. of scenarios	CPU time	Nominal objective	Recovery cost	Robust objective	LP gap %
1	9	739,611	0	739,611	0.00
2	54	739,611	0	739,611	0.00
3	1,659	740,791	0	740,791	0.10
4	1,059	740,840	40,000	780,840	0.02
5	1,216	740,840	40,000	780,840	0.02
6	1,547	740,840	40,000	780,840	0.02
7	17,116	740,840	40,000	780,840	0.02
8	2,775	740,840	40,000	780,840	0.02
9	16,863	740,840	40,000	780,840	0.02
10	65,090	742,070	40,000	782,070	0.15

Table 4   **Solutions Obtained by the Benders Heuristic**

No. of scenarios	(11)–(15) objective	Nominal objective	Estimated rec. cost	Actual rec. cost	B-Heur objective	B-Heur CPU	LP gap %
0	737,490	737,490	0	0	737,490	1.50	0.04
1	739,611	739,611	0	0	739,611	1.13	0.00
2	740,385	740,385	0	0	740,385	7.73	0.10
3	780,840	740,840	40,000	40,000	780,840	5.88	5.23
4	780,840	740,840	40,000	40,000	780,840	4.83	0.02
5	780,840	740,840	40,000	40,000	780,840	6.14	0.02
6	780,840	740,840	40,000	40,000	780,840	5.68	0.02
7	780,840	740,840	40,000	40,000	780,840	6.53	0.02
8	780,840	740,840	40,000	40,000	780,840	7.00	0.02
9	782,042	742,042	40,000	50,000	792,042	38.08	1.43
10	782,042	742,042	40,000	50,000	792,042	72.65	1.40
11	782,042	742,042	40,000	60,000	802,042	42.63	2.63
12	781,614	741,614	40,000	50,000	791,614	12.79	1.33
13	782,224	742,224	40,000	70,000	812,224	34.26	3.83
14	782,224	742,224	40,000	70,000	812,224	29.55	3.83
15	781,614	741,614	40,000	50,000	791,614	19.45	1.33
16	781,614	741,614	40,000	50,000	791,614	25.75	1.33
17	782,224	742,224	40,000	50,000	792,224	43.78	1.41
18	782,224	742,224	40,000	2,050,000	2,792,224	63.29	72.03
19	782,252	742,252	40,000	70,000	812,252	108.39	3.84
20	782,224	742,224	40,000	70,000	812,224	69.01	3.83
21	782,252	742,252	40,000	1,010,000	1,752,252	63.62	55.42
22	782,224	742,224	40,000	70,000	812,224	67.41	3.83
23	782,224	742,224	40,000	70,000	812,224	80.94	3.83
24	782,224	742,224	40,000	50,000	792,224	79.60	1.41
25	782,252	742,252	40,000	70,000	812,252	64.87	3.84
26	781,614	741,614	40,000	50,000	791,614	35.51	1.33
27	782,252	742,252	40,000	70,000	812,252	92.90	3.84
28	782,414	742,414	40,000	70,000	812,414	65.31	3.83

by the Benders constraints. These two solutions are clearly dominated, e.g., by the solution for 26 scenarios. This also shows that solutions with the same nominal objective value are not the same, otherwise, e.g., the recovery cost for 20 scenarios would be at least as large as the one for 18 scenarios. Not counting the two "unlucky" instances, for 10 or more disruption scenarios the gap between the robust objective value of the heuristic MILP solution in Table 4 and the LP solution value in Table 2 is 2.7% on average, and 5.2% at most for the case with 28 scenarios.

Let us now consider the heuristic integer solution for 28 disruption scenarios; in what follows, it is referred to as the *robust solution*. Then, the increase in the nominal objective value in going from 0 to 28 disruption scenarios is only 0.7%, namely, from 737,490 to 742,414.

5. Evaluation of the Solution
The results of the heuristic two-stage optimization algorithm reported in §4 are based on 28 disruption scenarios. It is thus a natural question whether the *robust solution* (i.e., the solution when heuristically optimizing over 28 disruption scenarios) admits significantly lower recovery costs than the *optimal nominal solution* (i.e., the optimal solution to the nominal

problem) also on much *larger* sets of disruption scenarios. Clearly, it depends on how representative the 28 disruption scenarios are for the larger sets of disruption scenarios.

To answer this question, we implemented a framework for evaluating the robustness of an arbitrary rolling stock circulation: we randomly generated 3,500 disruption scenarios, and we analyzed how well the involved rolling stock circulation could handle these. The location of the disruption is uniformly distributed along the 3,000 line; the start time of the disruption is uniformly distributed on the time interval from 8:00 to 20:00; finally, the duration of the disruption is uniformly distributed between one hour and four hours. Each disruption scenario leads to a modified timetable as described in §4 and as shown in Figure 1. Then the robustness of the rolling stock circulation is evaluated by computing the recovery costs of each of the 3,500 disruption scenarios. Note that this does not pose any computational difficulty; the recovery costs of each disruption scenario can be determined within 2–3 seconds because now the disruption scenarios are independent of each other.

We applied the evaluation sketched above for the optimal nominal solution as well as for the robust solution that was determined in §4.2. In what follows,
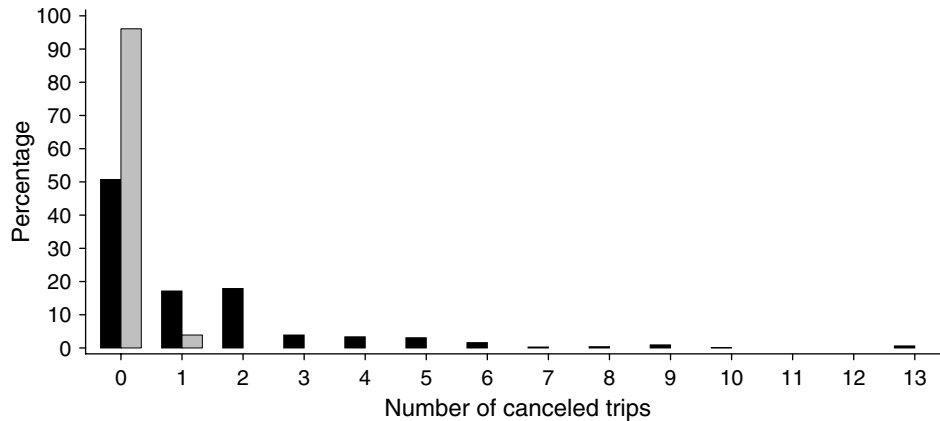
Figure 3 Distribution of the Number of Additionally Canceled Trips When Evaluating the Optimal Nominal Solution (Black Columns) and the Robust Solution (Gray Columns) on 3,500 Disruption Scenarios

we split the recovery costs into two terms: (i) the trip cancelation costs for additionally canceled trips due to lack of rolling stock, and (ii) the costs for new shunting operations and for end-of-day off-balances of the rolling stock inventories at the stations. Recall that the first term is a sum of penalties of 1,000,000 each, whereas the second term is a sum of values 10,000 and 20,000.

5.1. Additionally Canceled Trips
First, we consider the number of additionally canceled trips due to lack of rolling stock. The evaluation revealed that the recovery requires an average of about 1.5 additionally canceled trips (ranging from 0 to 13) for the optimal nominal solution, and an average of 0.039 additionally canceled trips (ranging from 0 to 1) for the robust solution. The distribution of the number of canceled trips is shown in Figure 3. The details of the results are given in Table A.1 in the appendix of this paper. In fact, for the robust solution,

96% of the disruption scenarios can be solved without any additionally canceled trip.

5.2. New Shunting Operations and Off-Balances
Second, we consider the recovery costs for new shunting operations and for end-of-day off-balances of the rolling stock inventories at the stations. The average contribution of these terms is about 269,000 (ranging from 0 to 670,000) for the optimal nominal solution, and about 29,000 (ranging from 0 to 180,000) for the robust solution. The distribution of these values is shown in Figure 4. The details of the results are shown in Table A.2 in the appendix.

We can conclude that the robust solution allows much lower recovery costs also in the case of a much larger number of disruption scenarios than the 28 disruption scenarios that were used for its generation. This also gives empirical evidence that the initially selected 28 disruption scenarios are indeed representative of the vast majority of the potential disruption scenarios.
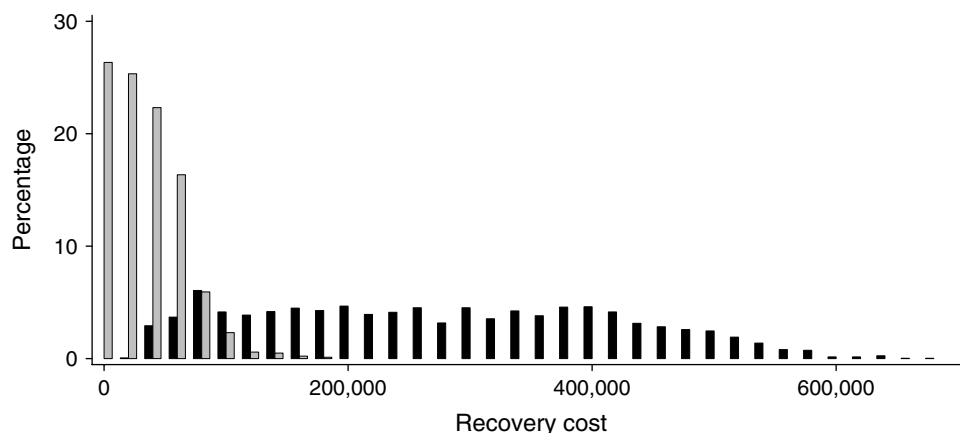

Figure 4 Recovery Cost of New Shunting Operations and of End-of-Day Off-Balances When Evaluating the Optimal Nominal Solution (Black Columns) and the Robust Solution (Gray Columns) on 3,500 Disruption Scenarios
Note. Each column represents a value range of 20,000.

5.3. Structure of the Robust Solution

It is interesting to compare the structures of the robust and of the optimal nominal solutions. First, the robust solution tries to balance the rolling stock between the stations. This is particularly true for the reserve units: the units whose daily duty does not contain any trip. The optimal nominal solution happens to assign all these reserve units at the Southern end of the 3,000 line. The worst-case recovery costs become particularly high: the disruption cuts the 3,000 line into two parts in such a way that the Northern part gets an insufficient number of units, which may lead to the inevitable cancellation of up to 13 trips. The robust solution avoids this trap by keeping reserve units at several locations along the line.

The better allocation of the reserve units does not explain the improvement in new shunting operations and in inventory deviations. In fact, we evaluated another rolling stock circulation, called the *half-robust solution*, which is obtained from the optimal nominal solution by reallocating the reserve units in the same way as in the robust solution. The evaluation reveals that the half-robust solution almost perfectly coincides with the gray columns (i.e., with the robust solution) in Figure 3, and with the black columns (i.e., with the optimal nominal solution) in Figure 4. This indicates that the improvement that is clearly visible in Figure 4 arises from a better composition assignment and from a better shunting pattern.

When performing recovery for any feasible solution, the recovery costs include penalties for introducing new shunting operations, but no penalties for canceling a planned shunting operation. Therefore, one could expect that the robust solution has many more shunting operations than the optimal nominal solution: more planned shunting operations give more flexibility in the recovery. Somewhat surprisingly, the robust solution has only two more shunting operations than the optimal nominal solution. Thus, the spatial and temporal distribution of the shunting operations is what helps, not their sheer number. In fact, the robust solution tends to place the shunting operations later in the day, which is certainly helpful to fix the end-of-day balances after a disruption. Nevertheless, it is far from straightforward—even for expert practitioners—to judge how easy or difficult the recovery for a given schedule is.

One may wonder how good or bad the recovery costs for the robust solution would be in practice. Experts of NS have the opinion that these recovery costs would be an excellent achievement in real-life railway disruption management. One should keep in mind, however, that our model does not take each aspect of the railway process into account. Even more importantly, this paper has the strong assumption that the duration of the disruption is a priori known; that

is, we do not deal with the uncertainty of future events. This may be difficult to handle in the robust optimization model. However, it will be possible to handle this in the evaluation model. This is a subject for further research.

6. Conclusions

In this paper we presented a two-stage optimization model to improve the recoverability of the rolling stock circulations of Netherlands Railways, the main operator of passenger trains in the Netherlands. This model is used to design a rolling stock circulation that can be recovered more easily than the nominal solution in case of a disruption. The model that we describe gives a practical implementation of the concept of recoverable robustness, as defined by Liebchen et al. (2009).

To generate a robust solution to the LP-relaxation of the large involved MILP model, we used a Benders decomposition approach. The latter turns out to be an effective approach to handle this type of large models. Based on the obtained Benders cuts, we also heuristically generated a robust rolling stock circulation for our case study. This approach is called the Benders heuristic.

The computational results indicate that, for the considered case study, the problem is tractable with up to 28 disruption scenarios in the two-stage optimization model. An evaluation based on 3,500 randomly generated disruptions shows that for the robust rolling stock circulation fewer trains have to be canceled than for the optimal nominal solution, and also that the other recovery costs are significantly lower. Moreover, the nominal costs of the robust solution are only slightly higher than those of the optimal nominal solution.

Appendix

Complete Model for the Two-Stage Optimization Problem

In the following, we give the complete model for the two-stage optimization problem model (3)–(9) for the specific case of the rolling stock (re-)scheduling problem. We use

the notation of §3. The index $\sigma \in \Sigma \cup \{\text{nom}\}$ is the scenario index, where $\sigma = \text{nom}$ is the nominal case.

$$\min \sum_{t \in T} \sum_{p \in P} c_{t,p}^c \cdot u_{t,p}^{\text{nom}} + \sum_{t \in T} \sum_{(p \to p') \in P^s} c_{t,p,p'}^s \cdot v_{t,p,p'}^{\text{nom}} + \lambda$$

$$\sum_{p \in P} u_{t,p}^\sigma = 1 \quad \forall t \in T,\ \sigma \in \Sigma \cup \{\text{nom}\},$$

$$u_{t,p}^\sigma = u_{t,p}^{\text{nom}} \quad \forall t \in \bar{T}^\sigma,\ p \in P,\ \sigma \in \Sigma,$$

$$u_{t,p}^\sigma = \sum_{p' \in P} v_{t,p,p'}^\sigma \quad \forall t \in T,\ p \in P,\ \sigma \in \Sigma \cup \{\text{nom}\},$$

$$u_{\nu(t),p'}^\sigma = \sum_{p \in P} v_{t,p,p'}^\sigma \quad \forall t \in T,\ p' \in P,\ \sigma \in \Sigma \cup \{\text{nom}\},$$

$$y_{s,m}^{\infty,\sigma} = y_{s,m}^0 + \sum_{t' \in T} \sum_{(p \to p') \in P^s} \alpha_{t',p,p',s,m} \cdot v_{t',p,p'}^\sigma$$
$$\forall s \in S,\ m \in M,\ \sigma \in \Sigma \cup \{\text{nom}\},$$

$$y_{d(t),m}^0 + \sum_{t' \in T} \sum_{(p \to p') \in P^s} \beta_{t,t',p,p',m} \cdot v_{t',p,p'}^\sigma \geq 0$$
$$\forall t \in T,\ m \in M,\ \sigma \in \Sigma \cup \{\text{nom}\},$$

$$\sum_{s \in S} y_{s,m}^0 = n_m \quad \forall m \in M,$$

$$\lambda \geq c_1 \sum_{t \in T} u_{t,\varnothing}^\sigma + c_2 \sum_{t \in T} r_t^\sigma + c_3 \sum_{s \in S} \sum_{m \in M} q_{s,m}^\sigma \quad \forall \sigma \in \Sigma,$$

$$q_{s,m}^\sigma \geq y_{s,m}^{\infty,\text{nom}} - y_{s,m}^{\infty,\sigma} \quad \forall s \in S,\ m \in M,\ \sigma \in \Sigma,$$

$$q_{s,m}^\sigma \geq y_{s,m}^{\infty,\sigma} - y_{s,m}^{\infty,\text{nom}} \quad \forall s \in S,\ m \in M,\ \sigma \in \Sigma,$$

$$w_t^\sigma = \sum_{(p \to p') \in P^s} v_{t,p,p'}^\sigma \quad \forall t \in T,\ \sigma \in \Sigma \cup \{\text{nom}\},$$

$$r_t^\sigma \geq w_t^\sigma - w_t^{\text{nom}} \quad \forall t \in T,\ \sigma \in \Sigma,$$

$$u_{t,p}^\sigma,\ v_{t,p,p'}^\sigma \text{ binary} \quad \forall t \in T,\ p,p' \in P,\ \sigma \in \Sigma \cup \{\text{nom}\},$$

$$y_{s,m}^{\infty,\sigma},\ q_{s,m}^\sigma \text{ integer} \geq 0 \quad \forall s \in S,\ m \in M,\ \sigma \in \Sigma,$$

$$r_t^\sigma,\ w_t^\sigma \text{ binary} \quad \forall t \in T,\ \sigma \in \Sigma \cup \{\text{nom}\},$$

$$y_{s,m}^0 \text{ integer} \geq 0 \quad \forall s \in S,\ m \in M.$$

Table A.1 Distribution of the Number of Additionally Canceled Trips When Evaluating the Optimal Nominal Solution and the Robust Solution on 3,500 Disruption Scenarios

Number of cancelations	Nominal (%)	Robust (%)
0	50.72	96.10
1	17.15	3.90
2	17.89	—
3	3.90	—
4	3.35	—
5	3.07	—
6	1.60	—
7	0.28	—
8	0.37	—
9	0.92	—
10	0.12	—
11	—	—
12	—	—
13	0.61	—

Table A.2 Distribution of the Recovery Cost of New Shunting Operations and of End-of-Day Off-Balances When Evaluating the Optimal Nominal Solution and the Robust Solution on 3,500 Disruption Scenarios

Recovery cost value	Nominal (%)	Robust (%)	Recovery cost value	Nominal (%)	Robust (%)
0	0.00	26.34	360,000	3.81	—
20,000	0.06	25.33	380,000	4.58	—
40,000	2.92	22.32	400,000	4.61	—
60,000	3.69	16.35	420,000	4.15	—
80,000	6.06	5.93	440,000	3.14	—
100,000	4.15	2.31	460,000	2.83	—
120,000	3.87	0.58	480,000	2.58	—
140,000	4.18	0.49	500,000	2.46	—
160,000	4.49	0.22	520,000	1.91	—
180,000	4.27	0.12	540,000	1.38	—
200,000	4.67	—	560,000	0.80	—
220,000	3.93	—	580,000	0.74	—
240,000	4.12	—	600,000	0.15	—
260,000	4.52	—	620,000	0.15	—
280,000	3.17	—	640,000	0.25	—
300,000	4.52	—	660,000	0.03	—
320,000	3.54	—	680,000	0.03	—
340,000	4.24	—	700,000	—	—

In Table A.1 we show the distribution of the number of canceled trips. Table A.2 shows the distribution of the recovery cost of new shunting operations and of end-of-day off-balances when evaluating the optimal nominal solution and the robust solution on 3,500 disruption scenarios.

References

Benders, J. 1962. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik* **4**(1) 238–252.

Ben-Tal, A., A. Nemirowski. 1998. Robust convex optimization. *Math. Oper. Res.* **23**(4) 769–805.

Ben-Tal, A., L. El Ghaoui, A. Nemirowski. 2009. *Robust Optimization*. Princeton University Press, Princeton, NJ.

Bertsimas, D., M. Sim. 2004. The price of robustness. *Oper. Res.* **52**(1) 35–53.

Birge, J., F. Louveaux. 1997. *Introduction to Stochastic Programming*. Springer, New York.

Cacchiani, V., A. Caprara, M. Fischetti. 2012. A Lagrangian heuristic for robustness, with an application to train timetabling. *Transportation Sci.* **46**(1).

Caprara, A., L. Galli, S. Stiller, P. Toth. 2008. Recoverable-robust platforming by network buffering. Technical Report ARRIVAL-TR-0157, ARRIVAL Project, http://www.cti.gr/arrival.

Caprara, A., L. Kroon, M. Monaci, M. Peeters, P. Toth. 2007. Passenger railway optimization. C. Barnhart, G. Laporte, eds. *Handbooks in Operations Research and Management Science*, Vol. 14. Elsevier, Amsterdam, 129–187.

Cicerone, S., G. D'Angelo, G. Di Stefano, D. Frigioni, A. Navarra. 2007. Robust algorithms and price of robustness in shunting problems. C. Liebchen, R. K. Ahuja, J. A. Mesa, eds. *Proc. 7th Workshop on Algorithmic Approaches for Transportation Model., Optim., Systems (ATMOS)*, Schloss Dagstuhl, Dagstuhl, Germany.

Cicerone, S., G. D'Angelo, G. Di Stefano, D. Frigioni, A. Navarra, M. Schachtebeck, A. Schöbel. 2009. Recoverable robustness in shunting and timetabling. R. Ahuja, R. Möhring, C. Zaroliagis, eds. *Robust and On-Line Large Scale Optimization*. Lecture Notes in Computer Science, Vol. 5868. Springer-Verlag, Berlin, 28–60.

De Almeida, D., N. Marcos, S. Vianey, C. Weber. 2008. Robustness of rolling stock schedules. Technical Report ARRIVAL-TR-0129, ARRIVAL project, http://www.cti.gr/arrival.

Ehrgott, M., D. Ryan. 2002. Constructing robust crew schedules with bicriteria optimization. *J. Multi-Criteria Decision Anal.* **11**(3) 139–150.

Erera, A., J. Morales, M. Savelsbergh. 2009. Robust optimization for empty repositioning problems. *Oper. Res.* **57**(2) 468–483.

Fioole, P., L. Kroon, G. Maróti, A. Schrijver. 2006. A rolling stock circulation model for combining and splitting of passenger trains. *Eur. J. Oper. Res.* **174**(2) 1281–1297.

Fischetti, M., D. Salvagnin, A. Zanette. 2009. Fast approaches to improve the robustness of a railway timetable. *Transportation Sci.* **43**(3) 321–335.

Geoffrion, A. 1972. Generalized Benders decomposition. *J. Optim. Theory Appl.* **10**(4) 237–260.

Kroon, L., G. Maróti, M. Retel Helmrich, M. Vromans, R. Dekker. 2007. Stochastic improvement of cyclic railway timetables. *Transportation Res. Part B* **42**(6) 553–570.

Liebchen, C., M. Lübbecke, R. Möhring, S. Stiller. 2009. The concept of recoverable robustness, linear programming recovery, and railway applications. R. Ahuja, R. Möhring, C. Zaroliagis, eds. *Robust and Online Large-Scale Optimization.* Lecture Notes in Computer Science, Vol. 5868. Springer-Verlag, Berlin, Heidelberg, 1–27.

Liebchen, C., M. Schachtebeck, A. Schöbel, S. Stiller, A. Prigge. 2010. Computing delay resistant railway timetables. *Comput. Oper. Res.* **37**(5) 857–868.

Maróti, G. 2006. Operations research models for railway rolling stock planning. Ph.D. thesis, Eindhoven University of Technology, Eindhoven, The Netherlands.

Nielsen, L. 2008. Disruption generator for railway rolling stock re-scheduling, Software, Erasmus University Rotterdam, The Netherlands.

Nielsen, L., L. Kroon, G. Maróti. 2009. A rolling horizon based framework for rolling stock rescheduling. Technical Report ARRIVAL-TR-0205, ARRIVAL Project, http://www.cti.gr/arrival.

Shebalov, S., D. Klabjan. 2006. Robust airline crew planning: Move-up crews. *Transportation Sci.* **40**(3) 300–312.

Stiller, S. 2008. Extending concepts of reliability, network creation games, real-time scheduling, and robust optimization. Ph.D. thesis, Technische Universität Berlin, Berlin.

Yen, J., J. Birge. 2006. A stochastic programming approach to the airline crew scheduling problem. *Transportation Sci.* **40**(1) 3–14.