

Maintenance Appointments in Railway Rolling Stock Rescheduling

Joris Wagenaar, Leo Kroon, and Marie Schmidt

Rotterdam School of Management, Erasmus University
Technology and Operations Management
Burgemeester Oudlaan 50, 3062 PA, Rotterdam The Netherlands
e-mail: jwagenaar@rsm.nl, lkroon@rsm.nl, schmidt2@rsm.nl

January 7, 2016

Abstract

This paper addresses the Rolling Stock Rescheduling Problem (RSRP), while taking maintenance appointments into account. After a disruption, the rolling stock of the disrupted passenger trains has to be rescheduled in order to restore a feasible rolling stock circulation. Usually, a number of train units have a scheduled maintenance appointment during the day: these appointments must be taken into account while rescheduling the rolling stock. In this paper we propose three **Mixed Integer Programming (MIP)** models for this purpose. All models are extensions of the Composition model from literature, **which does not distinguish individual train units**. The Extra Unit Type model adds an additional rolling stock type for each train unit that requires maintenance. The Shadow-Account model keeps track of a shadow account for each train unit that requires maintenance. The Job-Composition model creates a path **for each train unit** such that the train units that require maintenance are on time for their maintenance appointments. All models are tested on instances of Netherlands Railways (NS). The results show that especially the Shadow-Account model and the Job-Composition model are effectively able to take maintenance appointments into account during real-time rescheduling. It depends on the characteristics of an instance whether the Shadow-Account model or the Job-Composition model performs best.

1 Introduction and Contributions

1.1 Introduction

In passenger railway transportation, an extensive planning process is used to develop a satisfying rolling stock circulation. First, in the *strategic* planning phase, the purchase of rolling stock takes place and specific train series are developed. A train series represents a line between stations A and B and back again, with possible intermediate stations. In the *tactical* planning phase, the timetable is created and train units are assigned to all trips within the timetable. This results in a rolling stock circulation, **usually involving anonymous rolling stock duties, i.e. no physical train units have been assigned to the rolling stock duties yet.**

In the *operational* planning phase, **physical train units are assigned to the anonymous duties. Furthermore, the rolling stock circulation is modified by taking into account specific operational aspects, such as the short-term maintenance that is required by certain physical train units. A train unit requires maintenance after a certain number of kilometers or a certain amount of time since its previous maintenance appointment.** A train unit requiring maintenance gets a fixed maintenance appointment, **assigned by the maintenance company**, at a given time and location. In the operational planning phase, the rolling stock circulation obtained in the tactical planning phase is modified in such a way that the maintenance appointments are met by the corresponding **physical** train units. The latter means that they arrive at the appropriate locations on time, see for instance the maintenance routing models of Maróti and Kroon [2005, 2007].

In the *real-time* phase, the railway network inevitably experiences disruptions and therefore fast rescheduling is required. There are three major resource schedules which need to be rescheduled due to a disruption: the timetable, the rolling stock circulation, and the crew schedule. In the Netherlands, the timetable is usually rescheduled based on a predefined contingency plan. Then, with the rescheduled timetable as input, the rolling stock is rescheduled, and, finally, with both the rescheduled timetable and rolling stock circulation as input, the crew is rescheduled.

In this paper the focus is on the second step: the Rolling Stock Rescheduling Problem(RSRP). We assume that the timetable has been rescheduled already based on a contingency plan. Given the rescheduled timetable, the RSRP aims to find a new feasible rolling stock circulation that upholds as much of the passenger service as possible. It is required that the fixed maintenance appointments are taken into account directly in the rescheduling process.

However, current rolling stock rescheduling models, see for example Nielsen [2011], assign anonymous train units to the trips during a disruption. They assume that all train units of the same type are interchangeable. That means, for instance, that there is no distinction between train units that require maintenance and train units that do not. As a result, if maintenance appointments are not taken into account, then the train units scheduled for maintenance will probably not be in time for their maintenance appointments. Thus, when rescheduling the rolling stock, the maintenance appointments of the train units must be considered.

In this paper, three MIP formulations for solving additionally constrained multi-commodity flow problems are used as solution method for this problem. The maintenance appointments that have been scheduled in the operational planning phase are taken into account in these model formulations. The developed models are able to reschedule the rolling stock in real-time such that the maintenance appointments are still met by the corresponding train units as much as possible.

1.2 Contributions and structure of the paper

Although there exist models for rescheduling the rolling stock circulation in the operational phase including maintenance appointments, the current paper is, to the best of our knowledge, the first to include maintenance appointments in the real-time rescheduling phase. By including these maintenance appointments in the RSRP models, the models are able to guide the maintenance units to their scheduled maintenance appointments after the occurrence of a disruption.

The main contribution of this paper is the development and comparison of three MIP models which are able to handle the complicating factor that physical train units of the same type are not fully interchangeable due to their maintenance appointments.

The contributions of the current paper can be summarized as follows:

- We take scheduled maintenance appointments into account while rescheduling the rolling stock, with the rescheduled timetable as input.
- We describe one straightforward extension of an existing model and introduce two new models.
- We provide an experimental comparison of the three models.

The paper begins in Section 2 with explaining the maintenance problem in detail. Then a literature overview is given in Section 3. Thereafter the Composition model from Fiore et al. [2006] and Nielsen [2011] for rescheduling the rolling stock without maintenance appointments is presented. This model is used as the base model for all three models that take maintenance into account. The notation used for describing the maintenance aspects is explained in Section 5.

Following, three approaches for including maintenance in rolling stock rescheduling models are given. First, the Extra Unit Type model is discussed in Section 6. Secondly, the Shadow-Account model is presented in Section 7. Finally the Job-Composition model is proposed in Section 8. Then, in Section 9, all models are tested on real life instances of Netherlands Railways (NS), the main operator of passenger trains in the Netherlands. All models use the same objective function. Therefore, we mainly compare the models with respect to their computation time and the number of times a proven optimal solution is found within a certain time limit. In Section 10 conclusions and topics for further research are given.

2 Maintenance problem

In this section the maintenance problem is explained in detail. We start with some general remarks on rolling stock scheduling. Thereafter we present an example of the maintenance problem. Finally we discuss the assumptions that are taken into account in this paper.

Rolling stock units of different types are available for passenger transportation. There exist large differences between the different types. First, there exist self-propelled train units and carriages hauled by a locomotive. In this paper we focus on self-propelled train units. Furthermore, there exist train units with two floors (called double-deck) and train units with a single floor. The main difference between the types we consider in this paper is the number of carriages of which they consist (e.g. a VIRM6 unit consists of 6 carriages and a VIRM4 unit consists of 4 carriages). See, for instance, Figure 1 for a train unit of type VIRM4.



Figure 1: Train unit of type VIRM4

Train units can be coupled onto each other before being used on a trip. Such an ordered combination of train units is called a *composition*. Compositions are used to assign sufficient capacity for the passenger demand on a trip. *Composition changes* can possibly take place at stations in order to increase or decrease the capacity assigned to the next trip. This is done by either coupling or uncoupling train units to and from the train. In the Netherlands it is predefined in the station rules whether a train unit is (un)coupled at the front or at the rear of a train at a station. In addition, the middle unit of a composition cannot be uncoupled. As a consequence, the order of the train units within a composition is important.

Figure 2 shows a time-space diagram of the scheduled rolling stock circulation on the 3000 series between the stations Nijmegen (Nm) and Den Helder (Hdr). Time is displayed on the horizontal axis and the stations are shown on the vertical axis. Every line between two stations represents a train unit assigned to the corresponding trip. Two (or more) lines close to each other form a composition of two (or more) train units.

This circulation is infeasible due to a disruption between Utrecht (Ut) and Amsterdam (Asd) from 09:00-11:00 indicated by the black rectangle. There are in total 25 train units available, where two train units require maintenance: one that starts in Alkmaar (Amr) with an appointment at 16:00 in station Nijmegen (indicated by a black line), and one that starts in Den Helder with an appointment at 22:00 in station Nijmegen (indicated by a dark grey line). Both units have a maintenance appointment that lasts for two hours, thereafter they are available for usage again. The maintenance appointments are visualized by a black and a grey triangle, representing their location and timing. Due to the disruption, the circulation needs to be rescheduled for the remainder of the day, such that the maintenance appointments are still met by the corresponding train units.

Figure 3 shows the solution after rescheduling. The modified timetable, rescheduled based on a contingency plan, is given as input while rescheduling the rolling stock. As can be seen in the figure, in the new timetable the trains turn in Amsterdam and Utrecht during the disruption, as is specified in the contingency plan. Furthermore, in the rescheduled rolling stock circulation, both units are still on time for their maintenance appointment. The models we discuss in this paper are able to reschedule the rolling stock in this way.

The problem is thus to guide certain individual train units in time to their maintenance appointment while rescheduling the rolling stock. This requires extensions of the existing rolling stock rescheduling models, since these models do not distinguish individual train units.

The assumptions that are taken into account in the developed models are the following:

- (1.) The timetable has been rescheduled already before the rolling stock is to be rescheduled.
- (2.) The maintenance appointments of the train units have been fixed in the operational planning phase, and cannot be modified in the real-time rescheduling phase.
- (3.) In the Netherlands, usually less than 5% of the train units have a maintenance appointment. Due to this fact, there is never more than one maintenance unit in a composition in real-life. Therefore

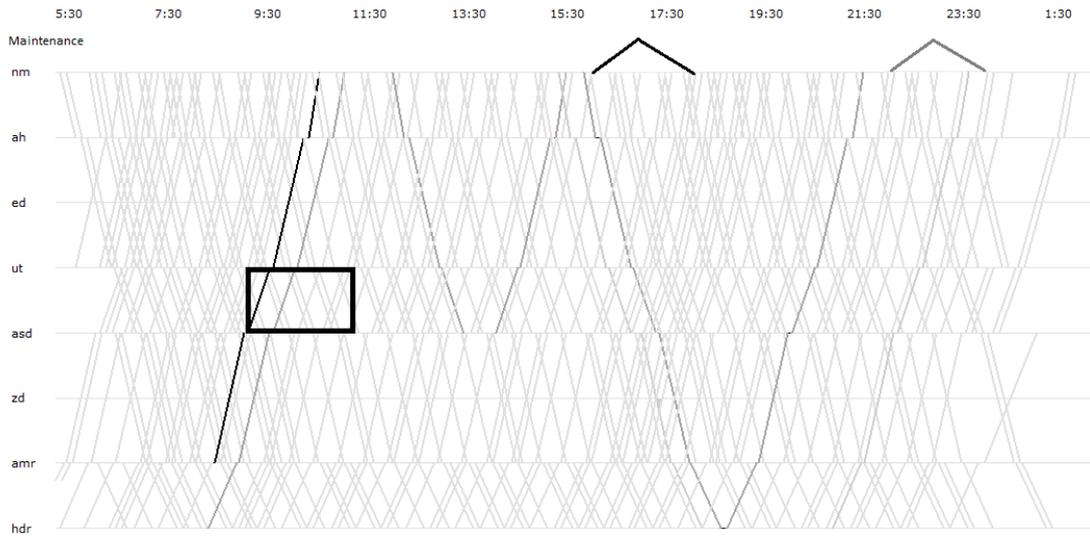


Figure 2: Time space diagram with a disruption

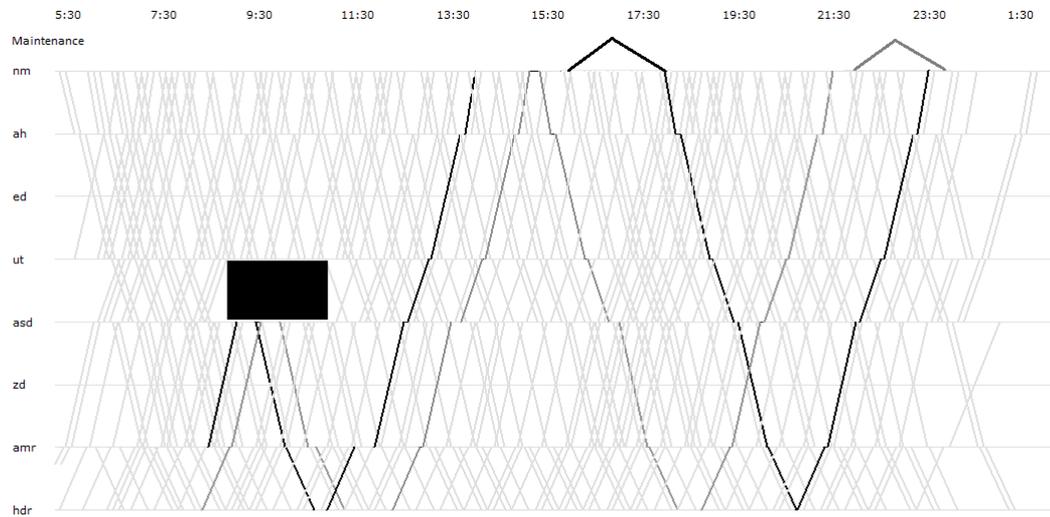


Figure 3: Rescheduled time space diagram

we make the realistic assumption that at most one train unit requiring maintenance is present in each composition.

- (4.) In the Netherlands, it is very unusual to both couple and uncouple train units from and to an arriving train. Usually, either a coupling activity, or an uncoupling activity, or no shunting activity at all takes place when a train arrives at a station. Therefore we assume that coupling and uncoupling at a station at the same time is not allowed. Note, however, that this assumption can be relaxed by slightly adjusting the models.

3 Literature

Table 1 gives an overview of the literature related to rolling stock (re)scheduling. These papers can be classified based on two characteristics: models developed for either scheduling or rescheduling, and models where maintenance is included or not.

¹This paper

	<i>Scheduling</i>	<i>Real-time rescheduling</i>
<i>No maintenance</i>	<ul style="list-style-type: none"> • Fioole et al. [2006] • Cordeau et al. [2001] • Lingaya et al. [2002] • Brucker et al. [2003] • Mellouli and Suhl [2007] • Peeters and Kroon [2008] 	<ul style="list-style-type: none"> • Nielsen [2011] • Nielsen et al. [2012] • Kroon et al. [2014] • Sato et al. [2009] • Sato and Fukumura [2012]
<i>Maintenance</i>	<ul style="list-style-type: none"> • Maróti and Kroon [2005] • Maróti and Kroon [2007] • Borndörfer et al. [2015] 	<ul style="list-style-type: none"> • Wagenaar et al. [2015]¹

Table 1: Overview railway literature

3.1 Scheduling, no maintenance included

Fioole et al. [2006] formulate a MIP model to assign rolling stock to the timetable in the tactical planning phase. The model is called the *Composition model* and is an integer multi-commodity flow model with additional constraints. This model is solved by CPLEX. The model can handle complicated line structures, such as combining and splitting of trains. NS has been using this model to generate rolling stock schedules since 2004. The model takes the order of the train units in each composition into account. However, maintenance routing is out of its scope. Peeters and Kroon [2008] consider the same problem, but they describe a Branch & Cut approach as solution method.

Cordeau et al. [2001] describe the rolling stock scheduling problem as the routing of locomotives and carriages through a railway network. The locomotives and carriages have to be combined to form a train group which has to be routed through the network. Their problem focusses on the tactical planning phase. They do not take the order of carriages into account. The problem is modelled as an integer multi-commodity flow model and is solved with CPLEX. Similar problems are considered by Brucker et al. [2003] and Mellouli and Suhl [2007].

Lingaya et al. [2002] also study the problem of scheduling locomotives and carriages in the tactical planning phase. However, they do take the order of the carriages in a train into account. They consider a train as a Last-In-First-Out (LIFO) stack, where carriages can be coupled or uncoupled from the rear part of the train in LIFO order only.

3.2 Scheduling, maintenance included

There are several papers that take maintenance into account in the operational phase. First of all, the problem was tackled in the airline industry before it was considered in the railway industry. For instance, Barnhart et al. [1998], Talluri [1998], and Clarke et al. [1997] propose models to solve the routing of maintenance for aircraft. Their models cannot be directly translated to models for scheduling rolling stock with maintenance constraints in the railway industry due to practical complications, such as the order of the train units in a composition.

One of the first to include maintenance routing in the operational phase of the railway industry were Maróti and Kroon [2005, 2007]. They propose two different MIP formulations for maintenance routing of rolling stock for passenger trains: the “Transition Model” and the “Interchange Model”. Both models use the scheduled rolling stock circulation as input, and exchange train unit duties such that maintenance

requirements are met. Both models are designed for the operational phase.

Giacco et al. [2014] develop a MIP formulation for integrating maintenance planning in the rolling stock planning problem in the operational phase. **Their formulation does not consider the order of the train units in a composition. Train units have to undergo maintenance after a certain time or a certain number of kilometers since their previous maintenance appointment. These maintenance appointments are not fixed, but determined by the model, in combination with the rolling stock circulation.** A commercial MIP solver is used to find efficient solutions in short time. The model is tested on real-world instances from the main Italian railway company Trenitalia.

Recently, Borndörfer et al. [2015] introduced a hypergraph formulation to create a rolling stock circulation for a generic week in a long distance railway network. The hypergraph formulation is used for the tactical planning phase. In this model several practical requirements are taken into account, such as scheduling of the maintenance for the train units. As in Giacco et al. [2014], the maintenance appointments are not fixed, but determined by the model. The model is tested on real life instances of the German railway company Deutsche Bahn. Circulations are found in between 10 minutes and 4 days of computation time. The model is not applicable in real-time for rescheduling due to time limitations.

3.3 Real-time phase, no maintenance included

All of the above models are applicable in the tactical or operational planning phase of the railway process. Maintenance requirements are taken into account to schedule maintenance appointments for certain train units. During a disruption the rolling stock circulation becomes infeasible, but the train units requiring maintenance still have their appointment. Furthermore, during a disruption less time is available, and, as a result, fast models are required for rescheduling.

Cacchiani et al. [2014] give an extensive literature overview on recent research within passenger railway disruption management. Papers on rescheduling the timetable on microscopic and macroscopic level, rescheduling the rolling stock, and rescheduling the crew are discussed. We refer to this paper for all literature on timetable and crew rescheduling. In the current paper the focus is on rescheduling the rolling stock, so the remainder of the discussed literature is on rolling stock.

Nielsen [2011] extends the model of Fioole et al. [2006] to cope with rescheduling. He formulates a MIP model with the adjusted timetable and the original rolling stock schedule as input, and an adjusted rolling stock schedule as output. This model will be used as base model in the current paper and is referred to as the *Composition model*. Subsequently, Nielsen et al. [2012] propose a rolling horizon to solve the RSRP. The idea behind the rolling horizon is that at the beginning of the disruption not all information about the duration of the disruption is known: this information becomes gradually available. The rescheduling is periodically performed within a limited rolling horizon length, possibly taking new information into account. At each time instant where an updated timetable becomes available, or when a certain amount of time has passed without any update, the MIP model is solved for the rolling horizon time window. This model is tested on instances of NS. Solutions with small deviations from the original plan are found in a short time.

Kroon et al. [2014] consider real-time rescheduling of rolling stock during large disruptions while taking dynamic passenger flows into account. They use the rescheduled timetable as input. Then they apply a two-stage feedback loop, where in the first stage the rolling stock allocation is rescheduled by using the model of Nielsen [2011] and in the second stage the effect of the rolling stock allocation on the passenger flows is determined by means of simulation. This passenger simulation provides feedback in terms of passenger delays due to limited capacity of the assigned rolling stock. The feedback is then used in the optimization model to reallocate the rolling stock again, in such a way that the total passenger delay is reduced. Given the reallocation of the rolling stock, the passenger simulation is performed again and feedback is given to the optimization model. This process continues for a number of iterations. **Results demonstrate that passenger delays can be reduced significantly.**

Sato et al. [2009] give a formulation to reallocate resources in a railway network in case of a disruption. Resources may refer either to rolling stock or to crew. The resources are reallocated to trips in such a way that the resource allocation differs as little as possible from the ones in the original plan. They use two phases to solve the problem: in the first phase conflicts created by the disruption are resolved through changes in the resource duties. The second phase is a local search heuristic which attempts to iteratively improve the rescheduled resource duties. The algorithm is tested on one line of the Japanese railway network.

In a subsequent paper, Sato and Fukumura [2012] consider the problem of reassigning locomotives to tasks in the case of a disruption in the railway network. A task consists of hauling a number of carriages from one station to another. They first enumerate possible sequences of tasks, to determine the corresponding costs for each sequence. A MIP model based on set-partitioning is used in order to assign locomotives to sequences of tasks with minimum cost, and a column generation technique is proposed as a solution approach. Based on the solutions found for instances of the Japan Freight Railway Company between Kuroiso and Shimonoseki in Japan, the authors conclude that locomotive reassignments can be found within a practical amount of time.

3.4 Real-time phase, maintenance included

None of the above rescheduling models includes maintenance appointments. These appointments should be taken into account during rescheduling, otherwise the train units will most likely miss their appointments. To the best of our knowledge no rescheduling models exist that take maintenance appointments into account. In this paper we fill that gap in the existing literature.

4 Composition model

We start with introducing the base model. This is the Composition model developed by Fioule et al. [2006] and Nielsen [2011]. As shown in Nielsen [2011], this model is fast enough to be used during rescheduling. However the model **does not distinguish between train units of the same type. Therefore, the model** does not include maintenance appointments of the rolling stock.

Let T be the set of trips in the timetable and S the set of stations. A trip is defined as a train driving from one station until the next station at a fixed point in time. Only stations where the composition of a train may be changed are taken into account. Denote $s_t^{dep}(s_t^{arr})$ as the station where trip $t \in T$ starts (ends) and define $\tau_t^{dep}(\tau_t^{arr})$ as the departure (arrival) time of trip $t \in T$.

In many countries, such as the Netherlands, trips are part of a predefined route. That means either that a trip has a predefined successor trip, or that the route ends after the trip. Take Figure 4 as an example of a predefined route between stations A, B, and C.

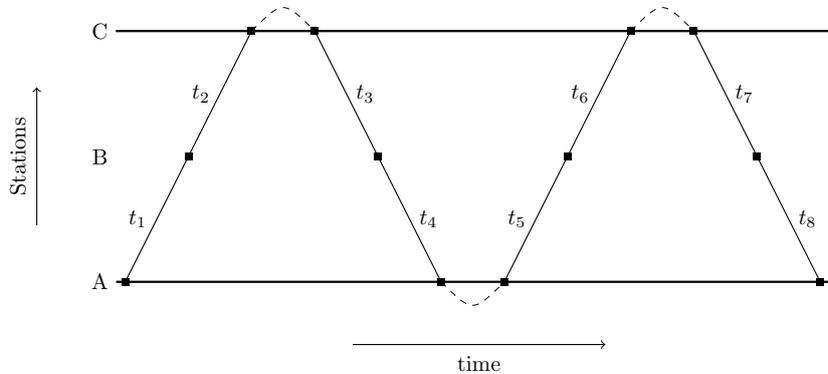


Figure 4: Predefined route

There are two trips between stations A and B, two trips between stations B and C, two trips between stations C and B, and two trips between stations B and A. Trip t_1 is the first trip of the route, so t_1 is not a successor trip of any other trip. Thereafter, we have that trip t_2 succeeds t_1 , t_3 succeeds t_2 , and so on. Trip t_8 does not have a successor trip, so the route ends after trip t_8 . If no coupling or uncoupling takes place, it means that the successor trip uses exactly the same train units as its predecessor. If a trip does not have a successor trip, then this means that all train units are moved to the shunting yard.

As can be seen in Figure 4, also trips at an end station of a line can have a successor trip. For instance, trip t_3 succeeds trip t_2 . In that case, the train units that will be used on the successor trip wait at the platform track where the trip arrives until the successor trip starts. A dashed line represents a turning between two succeeding trips.

We define $\sigma(t)$ as the successor trip of trip t , R as the set of routes, and $r := (t_1, \dots, t_n)$ is a route consisting of a sequence of trips, such that t_1 does not have a predecessor, $\sigma(t_i) = t_{i+1}$ for all $i = 1, \dots, n - 1$ and $\sigma(t_n) = \emptyset$. Then, $r(t)$ is the (uniquely defined) route trip t belongs to.

Now, let M be the set of rolling stock types. We denote P as the set of possible compositions, where a composition is an ordered combination of train units that can be used on a trip. For example, in Figure 5 the composition ab is assigned to trip t and the composition a is assigned to trip $\sigma(t)$. For each trip $t \in T$, $P(t)$ denotes the set of allowed compositions on the trip. Note that for each trip the empty composition is an element of $P(t)$, meaning that each trip may be cancelled. However, cancelling a trip is highly undesirable.

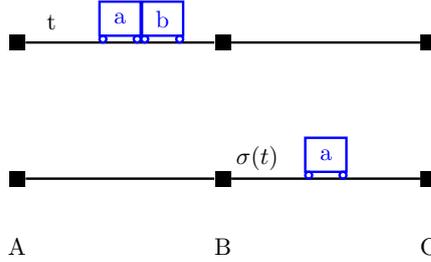


Figure 5: The successor trip of trip t

At the end of a trip, the composition of a train can be changed, depending on the shunting rules at the station, before the train departs on its successor trip. Recall that coupling and uncoupling takes place at either the front or the rear side of the train, this is defined by the station rules. A composition change denotes the composition of the incoming trip, the composition of the outgoing trip, and which train units are coupled or uncoupled during the composition change. To that end, let $\rho(t)$ be the set of possible composition changes at the end of trip $t \in T$, p_q the incoming composition of a trip when composition change $q \in \rho(t)$ is used, and o_q the outgoing composition when composition change q is used. For a given composition change $q \in \rho(t)$, $\alpha_{q,m}$ denotes the number of uncoupled train units of type $m \in M$ and $\beta_{q,m}$ denotes the number of coupled train units of type $m \in M$.

For instance, the composition change $ab \rightarrow a$ takes place at station B in Figure 5, so $p_q = ab$ and $o_q = a$. Furthermore, $\alpha_{q,a} = 0$, $\alpha_{q,b} = 1$, and $\beta_{q,m} = 0$ for both $m = a$ and $m = b$.

The time at which coupling takes place just before the start of trip $t \in T$ is denoted by τ_t^+ and the time at which an uncoupled unit is available after uncoupling after trip $t \in T$ is denoted by τ_t^- .

The available number of train units of type $m \in M$ at station $s \in S$ at the beginning of the planning period is denoted by $i_{s,m}^0$ and the desired number of available train units of type $m \in M$ at station $s \in S$ at the end of the planning period is given by the parameter $i_{s,m}^\infty$. This is usually the end of the day.

Besides the defined parameters, the model uses the following decision variables:

- $X_{t,p} \in \{0, 1\}$ denotes whether composition $p \in P(t)$ is used on trip $t \in T$.
- $Z_{t,q} \in \{0, 1\}$ denotes whether composition change $q \in \rho(t)$ is used at the end of trip $t \in T$.
- $I_{t,m} \in \mathbb{Z}_0^+$ denotes the number of train units of type $m \in M$ in the inventory at station s_t^{dep} immediately after time τ_t^+ .
- $C_{t,m}$ and $U_{t,m} \in \mathbb{Z}_0^+$ denote the number of train units $m \in M$ that are coupled and uncoupled at the start and end of trip $t \in T$, respectively.
- $D_{s,m} \in \mathbb{Z}$ denotes the deviation from the desired end-of-day balance at station $s \in S$ for rolling stock type $m \in M$.

Model:

$$\min f(X, Z, D) \quad (4.1)$$

subject to:

$$\sum_{p \in P(t)} X_{t,p} = 1 \quad \forall t \in T \quad (4.2)$$

$$X_{t,p} = \sum_{q \in \rho(t): p_q = p} Z_{t,q} \quad \forall t \in T, p \in P(t) \quad (4.3)$$

$$X_{\sigma(t),p} = \sum_{q \in \rho(t): o_q = p} Z_{t,q} \quad \forall t \in T, p \in P(\sigma(t)) \quad (4.4)$$

$$C_{\sigma(t),m} = \sum_{q \in \rho(t)} \beta_{q,m} Z_{t,q} \quad \forall t \in T, m \in M \quad (4.5)$$

$$U_{t,m} = \sum_{q \in \rho(t)} \alpha_{q,m} Z_{t,q} \quad \forall t \in T, m \in M \quad (4.6)$$

$$i_{s,m}^\infty + D_{s,m} = i_{s,m}^0 - \sum_{t \in T, s_t^{dep} = s} C_{t,m} + \sum_{t \in T, s_t^{arr} = s} U_{t,m} \quad \forall s \in S, m \in M \quad (4.7)$$

$$I_{t,m} = i_{s,m}^0 - \sum_{t' \in A_t} C_{t',m} + \sum_{t' \in B_t} U_{t',m} \quad \forall t \in T, m \in M, s \in S : s = s_t^{dep} \quad (4.8)$$

$$X_{t,p} \in \{0, 1\} \quad \forall t \in T, p \in P(t) \quad (4.9)$$

$$C_{t,m}, U_{t,m}, I_{t,m} \in \mathbb{R}_+ \quad \forall t \in T, m \in M \quad (4.10)$$

$$D_{s,m} \in \mathbb{R}_+ \quad \forall s \in S, m \in M \quad (4.11)$$

$$Z_{t,q} \in \mathbb{R}_+ \quad \forall t \in T, q \in \rho(t) \quad (4.12)$$

The objective function (4.1) is a linear function that depends on the assigned compositions (X), this includes the number of cancelled trips, the capacity shortages, and the number of carriage kilometers. The objective also depends on the shunting movements (Z): modified shunting movements, with respect to the original plan, are penalized. Finally, the total deviation from the end-of-day balance (D) is penalized.

Constraints (4.2) specify that to each trip exactly one composition is assigned, this composition is in the set of allowed compositions, $P(t)$, of that trip. Note that the compositions of the trips before and at the start of the disruption are fixed, because these trips are already underway. For those trips the set of allowed compositions consists of only a single composition. Constraints (4.3) state that if composition $p \in P(t)$ is assigned to trip $t \in T$, then only a composition that can originate from composition p can be assigned to the succeeding trip $\sigma(t)$. Constraints (4.4) state that if composition $p \in P(\sigma(t))$ is assigned to the succeeding trip $\sigma(t)$, then only a composition that fits with composition p can be assigned to trip $t \in T$.

Constraints (4.5) specify the number of coupled train units at the beginning of a trip and Constraints (4.6) specify the number of uncoupled train units at the end of a trip. Constraints (4.7) specify the end-of-day balance at a station plus the total deviation from the scheduled end-of-day balance. Their sum equals the initial inventory at the station ($i_{s,m}^0$), minus all units that have been coupled onto a train and plus all units that have been uncoupled from a train during the day. Constraints (4.8) keep track of the inventory of rolling stock type $m \in M$ at station s_t^{dep} immediately after the coupling time τ_t^+ . The sets A_t and B_t are explained below in detail. This inventory equals the initial inventory at the station, minus all train units that have been coupled onto a departing train before time τ_t^+ (all train units that have been coupled at the start of the trips in the set A_t), and plus all train units that have been uncoupled from an arriving train before time τ_t^+ (all train units that have been uncoupled at the end of the trips in the set B_t). Finally, Constraints (4.9), (4.10), (4.11), and (4.12) specify the character of the decision variables. Since $X_{t,p}$ is binary, all other variables can be defined as continuous variables, see Maróti [2006].

In Constraints (4.8), the subsets A_t and B_t are defined as:

1. $A_t = \{t' \in T : s_{t'}^{dep} = s_t^{dep}, \tau_{t'}^+ \leq \tau_t^+\}$

$$2. B_t = \{t' \in T : s_{t'}^{arr} = s_t^{dep}, \tau_{t'}^- \leq \tau_t^+\}$$

The set A_t contains all trips which depart from station s_t^{dep} before time τ_t^+ . This is the set of trips to which train units may have been coupled from station s_t^{dep} up to (and including) the departure time of trip t . We take as an example Figure 5 showing 8 different trips $\{t_1, t_2, t_3, \dots, t_8\}$. To explain the set A_t , we focus on trip t_8 . Trip t_8 departs from station B. All trips that have departed from station B before and including trip t_8 are the trips t_2, t_4, t_6 , and t_8 . So, $A_{t_8} = \{t_2, t_4, t_6, t_8\}$.

Furthermore, the set B_t contains all trips which have arrived at station s_t^{dep} before time τ_t^+ . This is the set of trips from which train units may have been uncoupled to the inventory of station s_t^{dep} up to the departure time of trip t . If we look at Figure 5 again, then all trips arriving at station B before the departure time of trip t_8 are the trips t_1, t_3, t_5 , and t_7 . So, $B_{t_8} = \{t_1, t_3, t_5, t_7\}$.

The output of the Composition model is a list of trips with compositions assigned to them. Note that these compositions can be decomposed into individual duties for train units in a postprocessing step, because an integer flow can always be decomposed into train unit valued path flows, see Ahuja et al. [1993]. However, this does not guarantee that there exists a feasible individual duty for train units that have a maintenance appointment. This is because the Composition model assumes all train units of the same type $m \in M$ to be interchangeable. As a consequence, no distinction can be made between train units requiring maintenance and train units that do not require maintenance. Thus, individual maintenance constraints cannot be imposed on the train units requiring maintenance.

Therefore we describe in Sections 6, 7, and 8 three extensions of the Composition model that allow guiding individual train units to their maintenance appointments.

5 Maintenance notation

The notation for maintenance units we use throughout this paper is the following. Let M' be the set of train units that require maintenance. Denote h_m as the time that train unit $m \in M'$ has its maintenance appointment, g_m as the duration of the appointment, and f_m as the location of the appointment. Furthermore, all maintenance units still belong to their regular rolling stock type (e.g. a train unit of type a that requires maintenance is still a train unit of type a). To that end, let $b_m \in M$ be the corresponding regular rolling stock type of train unit $m \in M'$. Finally, train units with a maintenance appointment at the same time, at the same location, and with the same corresponding regular type can have the same maintenance type $m \in M'$. Then a_m denotes the number of train units with the specific maintenance appointment.

Train units that require maintenance need to be in inventory at the right maintenance location and in time for their appointment. The inventory is measured immediately after the coupling time, τ_t^+ , of every trip $t \in T$, see Constraints (4.8). Thus, to be able to measure the inventory at the maintenance station at relevant points in time, we introduce for each maintenance unit $m \in M'$ an additional set of trips T_m . This set T_m contains the following (dummy) trips:

- (i): One trip t' with parameters: $\tau_{t'}^{dep} = \tau_{t'}^{arr} = \tau_{t'}^+ := h_m$ and $s_{t'}^{dep} = s_{t'}^{arr} := f_m$.
- (ii): One trip t'' with parameters: $\tau_{t''}^{dep} = \tau_{t''}^{arr} = \tau_{t''}^+ := h_m + g_m$ and $s_{t''}^{dep} = s_{t''}^{arr} := f_m$.
- (iii): For each trip $t \in T$ with $s_t^{arr} = f_m$ and $h_m \leq \tau_t^{arr} \leq h_m + g_m$, the set T_m contains one trip t^* with parameters: $\tau_{t^*}^{dep} = \tau_{t^*}^{arr} = \tau_t^{arr}$, and $s_{t^*}^{dep} = s_{t^*}^{arr} := f_m$.

These trips are used to trigger the measurement of the inventory (i) just after the start of the maintenance appointment, (ii) just after the end of the maintenance appointment, and (iii) just after the arrival of a trip at the maintenance station at a time instant in between. The trips t^* may bring a maintenance unit at a too late point in time to its maintenance location. Since these trips are used only to measure the inventory, no composition may be assigned to these trips, so the set of allowed compositions $P(t)$ for $t \in T_m$ consists of only the empty composition.

6 Extra Unit Type model

The Extra Unit Type (EUT) model is an extension of the Composition model. By adding an additional rolling stock type for every train unit that has a maintenance appointment, maintenance constraints can be **imposed** on such a train unit.

Consider the same example as in Section 2. There are 25 train units, this time 10 train units of type a and 15 train units of type b . There are again 2 train units that require maintenance, one of type a starting in Alkmaar with an appointment in Nijmegen at 16:00 and one of type b starting in Den Helder with an appointment in Nijmegen at 22:00. That means that the following rolling stock types are used in the **EUT** model: a (9 train units), b (14 train units), a' (1 train unit) and b' (1 train unit). So, two additional rolling stock types have been added to the model.

The train units that require maintenance are **used to redefine the set M by adding the train units requiring maintenance: $M := M \cup M'$** . Furthermore, we introduce the decision variable $A'_{t,m}$, to denote the number of train units of type $m \in M'$ that are not present at their maintenance location immediately after the coupling time τ_t^+ of a trip $t \in T_m$. Then, Constraints (6.1) denote that maintenance units need to be in inventory at the time of their appointment and during their appointment. Otherwise the decision variable $A'_{t,m}$ is equal to the number of train units of type $m \in M'$ that are not at their appointment immediately after time τ_t^+ for $t \in T_m$.

$$I_{t,m} + A'_{t,m} \geq a_m \quad \forall m \in M', t \in T_m \quad (6.1)$$

As a result, a penalty value θ_t can be set upon train units missing their appointment completely **and on being late for their appointment**. The objective function (4.1) is extended with the variables $A'_{t,m}$ for missing maintenance appointments. Together with Constraints (6.1) and Constraints (4.2)-(4.7) this forms the **EUT model**.

A drawback of this approach is that, by taking additional rolling stock types into account, the number of possible compositions increases rapidly. As a result of **Assumption (3.) from Section 2** that a composition contains at most one maintenance unit, we have that the increase in the number of compositions by adding an additional rolling stock type due to maintenance appointments only depends on the number of regular rolling stock types and **on the allowed composition length**. Indeed, a composition of length k , measured in the number of train units, consists of at most 1 train unit that requires maintenance and at least $k - 1$ regular train units that do not require maintenance. Denote n as the number of available regular rolling stock types. Adding one additional type leads thus to $n^{k-1} \cdot k$ new compositions of size k . The maximum length of a composition, measured in the number of train units, is denoted by c . **Then, adding one additional train unit that requires maintenance leads to a maximum total increase in the number of compositions that is equal to:**

$$\begin{aligned} \sum_{k=1}^c (n^{k-1} \cdot k) &= \frac{c \cdot n^{c+1} - (c+1) \cdot n^c + 1}{(n-1)^2} \\ &= \frac{(c(n-1) - 1) \cdot n^c + 1}{(n-1)^2} \end{aligned}$$

This is polynomial in n , since c is fixed. In the Netherlands c is usually not larger than 5 and n not larger than 3, without taking maintenance appointments into account. **Figure 6 shows a 3D surface plot of the above formula for c up to 5 and n up to 3. As can be seen, the number of additionally required compositions grows rapidly in c and n .**

7 Shadow-Account model

To overcome the problem that the EUT model quickly grows when taking additional rolling stock types into account for every train unit that requires maintenance, we introduce the Shadow-Account (SA) model. We start with an introduction to the model in **Section 7.1**. Thereafter, in **Section 7.2**, we explain the shadow account part. **Finally, in Section 7.3**, we explain the linking part.

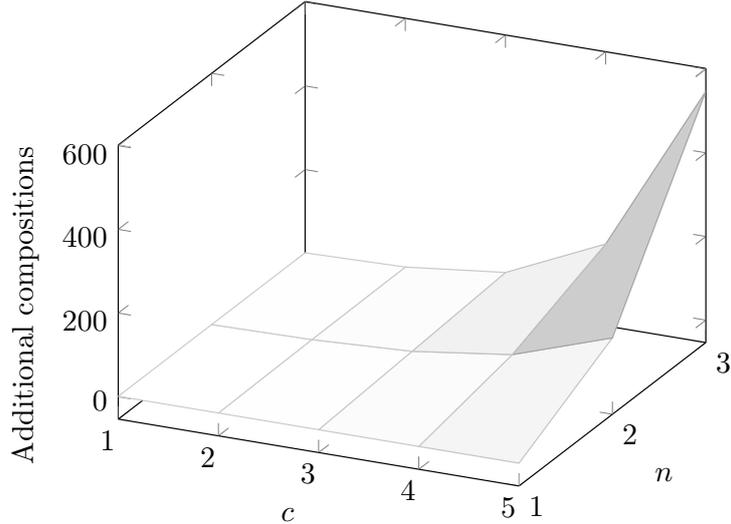


Figure 6: 3d plot of the number of additional compositions required

7.1 Introduction

The second approach to include maintenance in the rolling stock rescheduling problem is to create two parallel rolling stock circulations. The first circulation takes care of covering all trips with train units in the same way as in the Composition model. Thereby it does not make a distinction between train units requiring maintenance and train units that do not. The second circulation takes care of the maintenance appointments. To this end, a so called “shadow train unit” is created for each regular train unit. A shadow train unit is not denoted by a rolling stock type, (e.g. a, b, \dots), but by a ‘Shadow Account’ type (‘SA’ type), (e.g. $0, a', b', \dots$), representing maintenance appointments. A train unit with SA type 0 stands for a train unit that does not require maintenance and a train unit with SA type a', b', \dots stands for a train unit having a maintenance appointment. So, most train units are of SA type 0 and just a few train units have a different SA type. A train unit and its corresponding shadow train unit are synchronized, resulting in matching regular and shadow rolling stock circulations. For this reason the model is called the Shadow-Account model.

For instance, consider the same situation as in the previous section. There are 10 train units of type a and 15 train units of type b . The same 2 train units require maintenance, one of type a starting in Alkmaar with an appointment at 16:00 in Nijmegen and one of type b starting in Den Helder with an appointment at 22:00 in Nijmegen. In the Composition part of the model there are still 10 train units of type a and 15 train units of type b , however, in the shadow account there are 23 train units of SA type 0, one train unit of SA type a' , and one train unit of SA type b' .

See Figure 7 for the corresponding composition circulation of the train units of the example. As can be seen, the composition circulation represents only the regular train units, it is not clear which train units have a maintenance appointment and which train units do not. On the contrary, in Figure 8 the SA circulation is visualized. In this circulation there is no distinction between train units that do not require maintenance. They are all represented by light gray lines. However, there is a distinction between train units that require maintenance (dark gray and black lines), so this circulation is specifically used to create paths to the maintenance appointments.

The two circulations have to match in, among other factors, terms of the lengths of the assigned rolling stock compositions, otherwise the maintenance paths cannot be used. The precise definition of the matching of the two circulations will be presented in Section 7.3.

As in the previous section, we denote M as the set of different rolling stock types, and the maximum size of a composition is still denoted by c . Assuming that every SA composition contains at most one train unit that requires maintenance, one can verify that the total number of additionally required SA

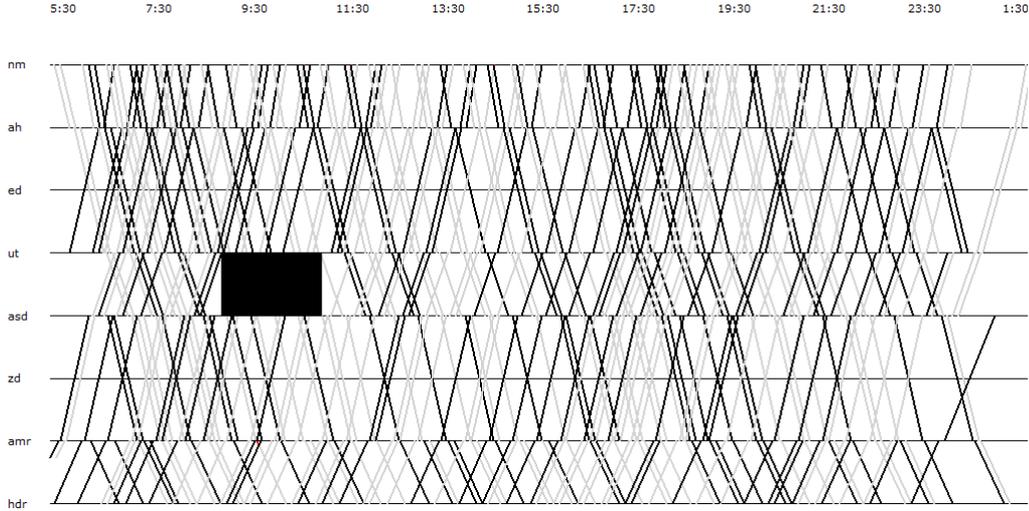


Figure 7: A solution for the **Composition** part, gray line = type a , black line = type b

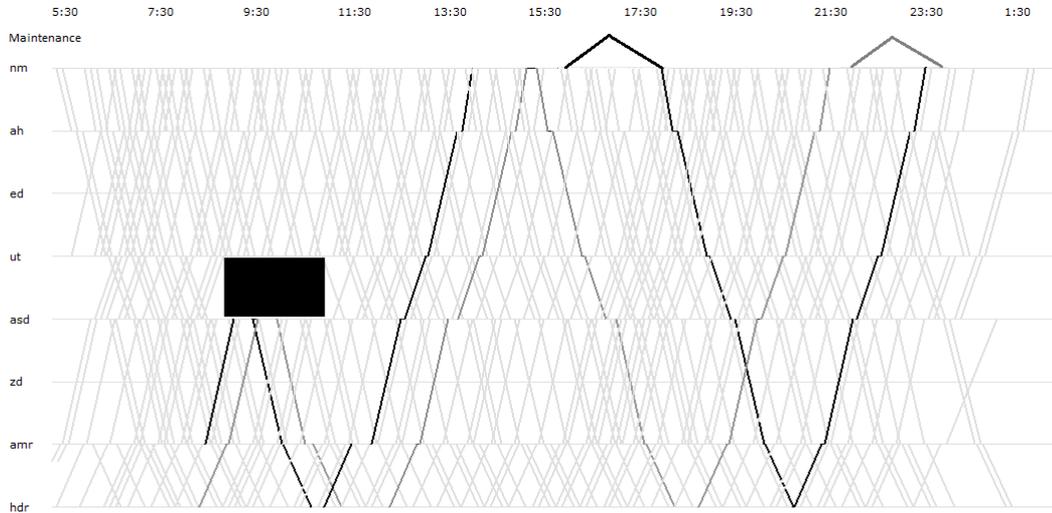


Figure 8: A matching solution for the SA part, light gray line = SA type 0, black line = SA type a' , dark gray line = SA type b'

compositions after adding one train unit that requires maintenance equals $\sum_{k=1}^c k = \frac{1}{2}c(c+1)$. For example, a SA composition of length three (000) leads to three new SA compositions ($a'00, 0a'0, 00a'$). The increase is quadratic in c and **does not** longer depend on n . For instance, with $c = 5$, there are only 15 additional SA compositions required, while there are 547 additional compositions required with the EUT model if $n = 3$ and $c = 5$.

The constraints in the complete SA model can be decomposed into three different parts: The **Composition** part, the SA part, and the Linking part.

All constraints in the **Composition** part are exactly the same as the constraints in the **Composition** model described in Section 4. So, the **Composition** part consists of constraints (4.2)-(4.7). The other parts will be discussed in the coming subsections.

7.2 Shadow-Account part

The SA part creates a rolling stock circulation for the shadow train units. To that end, define the set M' to be the set of SA types $\{0, a', b', \dots\}$ and P' as the set of possible SA compositions. The SA part of the MIP model consists of a second copy of Constraints (4.2)-(4.7), for the shadow types M' instead

of the regular rolling stock types. The same kind of variables are used as well, e.g. the variable $X'_{t,p}$ states whether composition $p \in P'$ is used on trip $t \in T$.

We note that for the start inventory $i'_{s,m}$ of the shadow types $m \in M'$ it holds by definition that the total number of SA train units in inventory at the start of the day is equal to the total number of regular train units in inventory at the start of the day.

$$\sum_{m \in M'} i'_{s,m} = \sum_{m \in M} i_{s,m} \quad \forall s \in S \quad (7.1)$$

This is not a constraint, but a condition that is to be satisfied by the data.

In a similar way as in the EUT model, we introduce the decision variable $A'_{t,m}$ denoting the number of train units of type $m \in M'$ that are not in inventory at their corresponding maintenance location at time τ_t^+ for $t \in T_m$. Then, Constraints (7.2) specify that either a maintenance unit is present at the station where its maintenance appointment is scheduled at the time of the appointment for the duration of the appointment, or the train unit is too late or misses its appointment completely. Constraints (7.2) are only needed for maintenance units $m \in M'$. **As a consequence, the restriction $m \neq 0$ is used in the constraint set. The objective function (4.1) is then extended with the variable $A'_{t,m}$ and a penalty θ_t , just as in the EUT model.**

$$I'_{t,m} + A'_{t,m} \geq a_m \quad \forall m \in M' : m \neq 0, t \in T_m \quad (7.2)$$

Constraints (7.2) are added to the SA copy of Constraints (4.2)-(4.7) to form the SA part.

7.3 Linking part

The **Composition** part creates a rolling stock circulation for the regular train unit types (e.g. a, b, \dots). The SA part creates a second rolling stock circulation for the SA train unit types (e.g. $0, a', b', \dots$). The SA part should give a shadow account of the **Composition** part. That means that the SA part should be linked to the **Composition** part.

Before introducing what we exactly mean by linking the SA and the **Composition** part, we first introduce the parameter N_p as the total number of train units in composition $p \in P$. Then the SA part is said to be linked to the **Composition** part if the following conditions are satisfied:

- (1.) For each trip $t \in T$, the lengths of the compositions assigned to trip t in the **Composition** part and in the SA part are the same: $\sum_{p \in P(t): N_p = v} X_{t,p} = \sum_{p \in P'(t): N_p = v} X'_{t,p} \quad \forall t \in T, v \in \{0, 1, \dots, c\}$
- (2.) For each trip $t \in T$, the numbers of uncoupled train units at the end of trip t are the same in the **Composition** part and in the SA part: $\sum_{m \in M} U_{t,m} = \sum_{m \in M'} U'_{t,m} \quad \forall t \in T$
- (3.) For each trip $t \in T$, the numbers of coupled train units at the start of trip t are the same in the **Composition** part and in the SA part: $\sum_{m \in M} C_{t,m} = \sum_{m \in M'} C'_{t,m} \quad \forall t \in T$
- (4.) At each point in time, the numbers of regular and SA train units in inventory are the same in the **Composition** part and in the SA part: $\sum_{m \in M} I_{t,m} = \sum_{m \in M'} I'_{t,m} \quad \forall t \in T$
- (5.) If a train unit of SA type $m \in M' : m \neq 0$ is assigned to trip $t \in T$ in the SA part, then its corresponding regular type $b_m \in M$ is assigned to trip t in the **Composition** part.
- (6.) If at some point in time a train unit of SA type $m \in M' : m \neq 0$ is in inventory in the SA part, then its corresponding regular type $b_m \in M$ is in inventory in the **Composition** part.

Theorem 7.1 *After including Constraints (7.3) the variables $U_{t,m}$ & $U'_{t,m}$, $C_{t,m}$ & $C'_{t,m}$, and $I_{t,m}$ & $I'_{t,m}$ are linked in the way as stated in conditions (1.), (2.), (3.), and (4.).*

$$\sum_{p \in P(t): N_p = v} X_{t,p} - \sum_{p \in P'(t): N_p = v} X'_{t,p} = 0 \quad \forall t \in T, v \in \{0, 1, \dots, c\} \quad (7.3)$$

Proof: Constraints (7.3) are the same as condition (1.). Next we will prove in steps that conditions (2.), (3.) and (4.) are true as well if Constraints (7.3) are satisfied.

- $U_{t,m}$ and $U'_{t,m}$. We will show that $\sum_{m \in M} U_{t,m} = \sum_{m \in M'} U'_{t,m}$ for all $t \in T$ by contradiction. Assume that $\sum_{m \in M} U_{t,m} > \sum_{m \in M'} U'_{t,m}$ for at least one trip $t \in T$. This means that at the end of trip t more train units are uncoupled in the **Composition part** than in the SA part. By Constraints (7.3) the lengths of the compositions **in the Composition part and in the SA part are the same for trip t . This holds for trip $\sigma(t)$ as well.** It is assumed that more train units are uncoupled at the end of trip t in the **Composition part**. This is only possible if also more train units are coupled there, otherwise the lengths of the compositions assigned to trip $\sigma(t)$ **in the Composition part and in the SA part** are not equal anymore. However, it is not allowed to both couple and uncouple train units at the end of a trip. This leads to a contradiction, and so $\sum_{m \in M} U_{t,m} \leq \sum_{m \in M'} U'_{t,m}$.

The same proof holds in the other direction, thus $\sum_{m \in M'} U'_{t,m} \leq \sum_{m \in M} U_{t,m}$. We can conclude that

$$\sum_{m \in M'} U'_{t,m} = \sum_{m \in M} U_{t,m} \quad \forall t \in T$$

- $C_{t,m}$ and $C'_{t,m}$. We can use the same proof as for $U_{t,m}$ and $U'_{t,m}$ to find that:

$$\sum_{m \in M'} C'_{t,m} = \sum_{m \in M} C_{t,m} \quad \forall t \in T$$

- $I_{t,m}$ and $I'_{t,m}$. Assume that $\sum_{m \in M} I_{t,m} > \sum_{m \in M'} I'_{t,m}$ immediately after time τ_t^+ of at least one trip $t \in T$. The inventories at the start of the day are by definition equal **on each station**, see Equation (7.1), so a difference between $I_{t,m}$ and $I'_{t,m}$ arises during the operations. Note that, **from Constraints (4.8), we have that:**

$$\sum_{m \in M} I_{t,m} = \sum_{m \in M} \left(i_{s,m}^0 - \sum_{t' \in A_t} C_{t',m} + \sum_{t' \in B_t} U_{t',m} \right) = \sum_{m \in M} i_{s,m}^0 - \sum_{t' \in A_t} \sum_{m \in M} C_{t',m} + \sum_{t' \in B_t} \sum_{m \in M} U_{t',m}$$

and

$$\sum_{m \in M'} I'_{t,m} = \sum_{m \in M'} \left(i'_{s,m}{}^0 - \sum_{t' \in A_t} C'_{t',m} + \sum_{t' \in B_t} U'_{t',m} \right) = \sum_{m \in M'} i'_{s,m}{}^0 - \sum_{t' \in A_t} \sum_{m \in M'} C'_{t',m} + \sum_{t' \in B_t} \sum_{m \in M'} U'_{t',m}$$

This means that a difference between $I_{t,m}$ and $I'_{t,m}$ can only be caused by **a difference in** either the start inventory, $C_{t,m}$ or $U_{t,m}$, but we just showed that $\sum_{m \in M} C_{t,m} = \sum_{m \in M'} C'_{t,m}$ and $\sum_{m \in M} U_{t,m} =$

$\sum_{m \in M'} U'_{t,m}$. So, it holds that

$$\sum_{m \in M} I_{t,m} = \sum_{m \in M'} I'_{t,m} \quad \forall t \in T$$

Theorem 7.3 is used for the first four linking conditions. We now introduce the second set of linking constraints for Condition (5.). These are Constraints (7.4). To this end, denote $w_{i,p}$ (or $w'_{i,p}$) as the rolling stock type assigned to position $i \in \{1, \dots, c\}$ in composition $p \in P$ (or $p \in P'$). Constraints (7.4) then state that when a train unit of SA type $m \in M'$ with $m \neq 0$ resides in a SA composition on position

i , then a corresponding regular train unit $b_m \in M$ must reside on position i in the corresponding regular composition as well.

$$\sum_{\substack{p \in P' \\ w'_{i,p} = m}} X'_{t,p} \leq \sum_{\substack{p \in P \\ w_{i,p} = b_m}} X_{t,p} \quad \forall t \in T, i \in \{1, \dots, c\}, m \in M' : m \neq 0 \quad (7.4)$$

Finally, for condition (6.), we use Constraints (7.5): if a train unit of type $m \in M' : m \neq 0$ is in inventory in the SA part, then a train unit of type $b_m \in M$ must be in the regular inventory.

$$I'_{t,m} \leq I_{t,b_m} \quad \forall t \in T, m \in M' : m \neq 0 \quad (7.5)$$

Constraints (7.3)-(7.5) take care of synchronizing the two circulations (the Composition part and the SA part). The complete SA model is hence given by the Composition part, the SA part, the Linking part, and the objective function.

8 Job-Composition model

In this section the third model to take maintenance into account is introduced. This model is called the Job-Composition (JC) model. This model is based on the concept of *jobs*. At the beginning of a day, all train units are in inventory. During the day, each train unit is assigned to a certain departing trip and fulfills a number of successor trips until the train unit is uncoupled and becomes part of the inventory again. A job is such a sequence of succeeding trips between coupling and uncoupling. So, a job starts **when** a train unit is coupled to a trip, and the job ends **when** the train unit is uncoupled from a trip. **Note that a train unit may carry out more than one job per day.**

The problem now becomes to appoint both regular train units and maintenance units to jobs, while synchronizing the movements of the maintenance units with those of the corresponding regular train units, just as in the SA model. In this way no additional compositions have to be taken into account for every train unit having a maintenance appointment.

A complicating factor is that trains must be considered, to a large extent, as *double sided stacks*. That means that, at both sides of a train, train units can be coupled or uncoupled in principle only in Last-In-First-Out (LIFO) order. As a consequence, if a pair of jobs does not correspond with a correct order of couplings and uncouplings per side of the train, then one assigned train unit will block the other one when the latter is to be uncoupled. In such a case, the two jobs are called *incompatible*. In Section 8.1 we characterize the pairs of incompatible jobs, and in Section 8.3 we present constraints that prevent two incompatible jobs to be selected both.

8.1 Jobs

In a preprocessing step, we create a list of all possible jobs during the day, and denote J as this set of possible jobs. Let $T(j)$ be the set of trips covered by job $j \in J$. Every job $j \in J$ has a start (and final) trip denoted by λ_j (γ_j). For all trips $t_1, \dots, t_n \in T(j)$ we have that $\lambda_j = t_1$, $\sigma(t_i) = t_{i+1}$, and $t_n = \gamma_j$. As can be seen from this notation, every job $j \in J$ takes place on a route $r \in R$, **where R denotes the set of routes as defined in Section 4.** Recall that the length of a predefined route depends on the shunting rules at a station and on the maximum turnaround time, as was explained in Section 4. The longer the maximum turnaround time, the longer the route, **and the more possible jobs exist.** As a consequence, the JC model contains more decision variables then.

Along each route $r \in R$ runs a train v_r , which consists of the actual train units that are used on the trips within the route. Each physical train has two sides. For further convenience, from now on these sides are called the *A-side* and the *B-side* of the train. **We define the A-side of train v_r to be the front side of the train at the first trip of its route $r \in R$.** Then we denote ζ_t as the side of train v_r that is **the front side during trip $t \in T$ in route $r \in R$, and ζ_t^{-1} as the rear side.**

Between two succeeding trips on a route $r \in R$, turnings can take place. In Figure 4 a route consisting of trips t_1, \dots, t_8 with 3 turnings is shown. **When train v_r turns, its front and rear side are exchanged.** **Note that turnings only take place at stations between two trips, not during a trip itself.** So, in order to keep track of which side is **the front side during trip $t \in T$, we need to keep track of the number of**

turnings taking place in route $r(t)$ up to **the start of** trip t . To that end, denote h_t^{turn} as the number of turnings taking place in route $r(t)$ up to the start of trip $t \in T$. Then, ζ_t can be **determined** as in Equation (8.1).

$$\zeta_t = \begin{cases} A & \text{If } h_t^{turn} \text{ is even or } 0 \\ B & \text{Otherwise} \end{cases} \quad (8.1)$$

Coupling can take place before the start of trip $t \in T$. **Recall that in the Netherlands** it is predefined in the station rules whether a train unit is coupled to the rear or to the front of the train. Let η_t denote whether a train unit has to be coupled to the front ($\eta_t = 0$) or to the rear ($\eta_t = 1$) of the outgoing train on trip t . Either the A -side or the B -side can be **the front side** of the train, this depends on the number of turnings taking place up to **the start of** trip t . We define the coupling side $\omega_j \in \{A, B\}$ as the side where coupling takes place before the start of job $j \in J$. The coupling side is determined as in Equation (8.2).

$$\omega_j = \begin{cases} \zeta_{\lambda_j} & \text{If } \eta_{\lambda_j} = 0 \\ \zeta_{\lambda_j}^{-1} & \text{Otherwise} \end{cases} \quad (8.2)$$

At the end of trip t , a train unit can be uncoupled from the composition. Just as with coupling, it is predefined in the station rules, whether a train unit is uncoupled from the rear or **from** the front of the train. To this end, let η'_t denote whether a train unit has to be uncoupled from the front ($\eta'_t = 0$) or the rear ($\eta'_t = 1$) of the incoming train. **Again, this** can be the A -side or the B -side of the train, this depends on the number of turnings taking place up to trip t . We denote π_j as the side where uncoupling takes place at the end of job $j \in J$, called the uncoupling side. **Turnings can not take place during a trip, so the number of turnings until the end of a trip is equal to the number of turnings until the start of the trip.** Therefore, π_j is defined as in Equation (8.3).

$$\pi_j = \begin{cases} \zeta_{\gamma_j} & \text{If } \eta'_{\gamma_j} = 0 \\ \zeta_{\gamma_j}^{-1} & \text{Otherwise} \end{cases} \quad (8.3)$$

Note that in case the coupling and uncoupling sides are not predefined by the station rules, we can adjust the model by taking them into account as variables in the model.

Definition 1 A set of jobs $J' \subset J$ on route $r \in R$ is said to be compatible, if for every job $j \in J'$, the train unit assigned to job j , that is coupled at the start of trip $\lambda_j \in T$ with coupling side ω_j , can be uncoupled **from** its uncoupling side π_j after trip $\gamma_j \in T$ without being blocked by any other train unit assigned to a job $j' \in J'$.

Lemma 8.1 A set of jobs $J' \subset J$ is compatible if and only if for each pair of jobs j and $j' \in J'$ the following two conditions hold:

1. If $\tau_{\lambda_j}^{dep} < \tau_{\lambda_{j'}}^{dep} < \tau_{\gamma_j}^{arr} < \tau_{\gamma_{j'}}^{arr}$, then $\omega_{j'} \neq \pi_j$
2. If $\tau_{\lambda_{j'}}^{dep} < \tau_{\lambda_j}^{dep} < \tau_{\gamma_j}^{arr} < \tau_{\gamma_{j'}}^{arr}$, then $\omega_j = \pi_j$

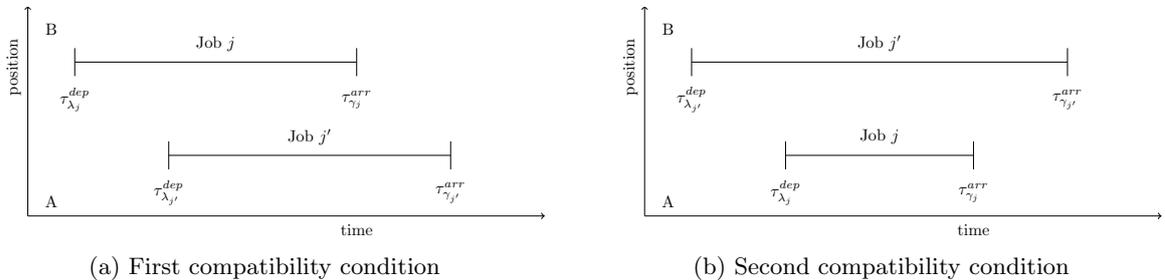


Figure 9: **Compatibility conditions**

Before proving Lemma 8.1, we first visualize the lemma's conditions in Figures 9a and 9b. Here time is displayed on the horizontal axis and the position of a train unit in the composition on the vertical axis.

For convenience, we define the upper side of the figure as the *B*-side and the bottom side of the figure as the *A*-side of the train, **assuming that no turnings take place**. In Figure 9a the train unit assigned to job j' is coupled to the train unit assigned to job j at the *A*-side. Thereafter, the train unit assigned to job j is uncoupled from the train before the train unit assigned to job j' is uncoupled. The train unit assigned to job j can not be uncoupled from the *A*-side, because the train unit assigned to job j' is still there, **so it must be uncoupled from the *B*-side**. In Figure 9b the train unit assigned to job j is coupled to the train unit assigned to job j' and later uncoupled from the train unit assigned to job j' again. This is only possible if the train unit assigned to job j is uncoupled from the same side as where it was coupled.

Proof of Lemma 8.1: Assume that the set of jobs J' is compatible. Let j and j' be a pair of jobs in J' . First, suppose $\tau_{\lambda_j}^{dep} < \tau_{\lambda_{j'}}^{dep} < \tau_{\gamma_j}^{arr} < \tau_{\gamma_{j'}}^{arr}$. Since, by assumption, the uncoupling side of job j is not blocked by job j' , we have that $\omega_{j'} \neq \pi_j$. Otherwise job j is being blocked. Second, suppose $\tau_{\lambda_{j'}}^{dep} < \tau_{\lambda_j}^{dep} < \tau_{\gamma_j}^{arr} < \tau_{\gamma_{j'}}^{arr}$. Again, the uncoupling of job j is not blocked by job j' , so we must have that $\omega_j = \pi_j$. This completes the proof of the “only-if”-part of the lemma.

Next, suppose that each pair of jobs j and $j' \in J'$ satisfies the two conditions, and that the set J' is *not* compatible. Then, by definition there is at least one job $j \in J$, whose uncoupling after trip γ_j from its uncoupling side π_j is blocked by another job $j' \in J'$. Clearly, $T(j) \cap T(j') \neq \emptyset$ and $\tau_j^{arr} < \tau_{j'}^{arr}$. The latter follows from the fact that if the end times of the jobs are the same, then also their uncoupling sides would be the same. Thus job j' would not be blocking the uncoupling of job j in that case.

Furthermore, if $\tau_{\lambda_j}^{dep} = \tau_{\lambda_{j'}}^{dep}$, then, without loss of generality, we may assume that job j' is not blocking the uncoupling of job j . Otherwise the positions of jobs j and j' in the train could have been interchanged just before coupling. Thus we may assume that $\tau_{\lambda_j}^{dep} \neq \tau_{\lambda_{j'}}^{dep}$. That leaves us with the cases $\tau_{\lambda_j}^{dep} < \tau_{\lambda_{j'}}^{dep}$ and $\tau_{\lambda_{j'}}^{dep} < \tau_{\lambda_j}^{dep}$.

If $\tau_{\lambda_j}^{dep} < \tau_{\lambda_{j'}}^{dep}$, then we have $\tau_{\lambda_j}^{dep} < \tau_{\lambda_{j'}}^{dep} < \tau_{\gamma_j}^{arr} < \tau_{\gamma_{j'}}^{arr}$. Thus, by assumption, we have that $\omega_{j'} \neq \pi_j$. In addition, if $\tau_{\lambda_{j'}}^{dep} < \tau_{\lambda_j}^{dep}$, then we have that $\tau_{\lambda_{j'}}^{dep} < \tau_{\lambda_j}^{dep} < \tau_{\gamma_j}^{arr} < \tau_{\gamma_{j'}}^{arr}$. Thus, by assumption, we have that $\omega_j = \pi_j$. However, it is clear that in both cases job j' does not block the uncoupling of job j . This contradiction completes the proof of the lemma.

A set of jobs is not compatible if it contains a pair of jobs not fulfilling one of the two conditions in Lemma 8.1. As a result, we can add constraints to the model guaranteeing that there is no such pair of jobs selected by the model. So, no pair of jobs of the **conflicting** sets CJ^1 and CJ^2 , described by Equations (8.4) and (8.5), may be chosen.

$$CJ^1 = \{(j, j') \in J \times J : T(j) \cap T(j') \neq \emptyset \wedge \tau_{\lambda_j}^{dep} < \tau_{\lambda_{j'}}^{dep} < \tau_{\gamma_j}^{arr} < \tau_{\gamma_{j'}}^{arr} \wedge \omega_{j'} = \pi_j\} \quad (8.4)$$

$$CJ^2 = \{(j, j') \in J \times J : T(j) \cap T(j') \neq \emptyset \wedge \tau_{\lambda_{j'}}^{dep} < \tau_{\lambda_j}^{dep} < \tau_{\gamma_j}^{arr} < \tau_{\gamma_{j'}}^{arr} \wedge \omega_j \neq \pi_j\} \quad (8.5)$$

8.2 Further notation

During the whole day jobs are carried out by train units. At the moment a disruption occurs, there are jobs already being carried out by train units. Compositions of trips that have already departed at the start of the disruption cannot be changed. However, jobs can be changed, as long as the compositions assigned to the trips before the start of the disruption do not change. Denote the set of trips that have departed before the start of the disruption and that are still underway when the disruption starts by $T^< \subset T$ and set the parameter $G_{t,p}$ equal to 1 if composition $p \in P$ is assigned to trip $t \in T^<$.

Finally, the following additional decision variables are necessary in the JC model:

- $K_t \in \{0, 1\}$ denotes whether trip $t \in T$ is cancelled or not.
- $W_j \in \{0, 1\}$ denotes whether job $j \in J$ is **selected** or not.
- $Y_{j,m} \in \mathbb{Z}^+$ denotes the number of train units $m \in M$ assigned to job $j \in J$.
- $Q_{j,m} \in \{0, 1\}$ denotes the number of *maintenance* units $m \in M'$ assigned to job $j \in J$.
- $I_{t,m} \in \mathbb{R}^+$ denotes the inventory of maintenance units $m \in M'$ at station s_t^{dep} , just after the departure of trip $t \in T$.

- $A'_{t,m} \in \mathbb{Z}^+$ denotes the number of maintenance units of type $m \in M'$ that are not available at their maintenance location immediately after time τ_t^+ for $t \in T_m$.

8.3 Model

Constraints (8.7)-(8.19), together with the objective function (8.6), form the JC model.

$$\min f(Y, D, A') \quad (8.6)$$

$$\sum_{j \in J: T(j) \ni t} W_j + K_t \geq 1 \quad \forall t \in T \quad (8.7)$$

$$\sum_{m \in M} Y_{j,m} - W_j \geq 0 \quad \forall j \in J \quad (8.8)$$

$$W_j + W_{j'} \leq 1 \quad \forall (j, j') \in CJ^1 \cup CJ^2 \quad (8.9)$$

$$I_{t,m} = i_{s,m}^0 - \sum_{t' \in A_t} \sum_{\substack{j \in J \\ \lambda_j = t'}} Y_{j,m} + \sum_{t' \in B_t} \sum_{\substack{j \in J \\ \gamma_j = t'}} Y_{j,m} \quad \forall t \in T, m \in M, s \in S : s = s_t^{dep} \quad (8.10)$$

$$I_{t,m} = i_{s,m}^0 - \sum_{t' \in A_t} \sum_{\substack{j \in J \\ \lambda_j = t'}} Q_{j,m} + \sum_{t' \in B_t} \sum_{\substack{j \in J \\ \gamma_j = t'}} Q_{j,m} \quad \forall t \in T, m \in M', s \in S : s = s_t^{dep} \quad (8.11)$$

$$i_{s,m}^\infty + D_{s,m} = i_{s,m}^0 - \sum_{\substack{j \in J \\ s_{\lambda_j}^{dep} = s}} Y_{j,m} + \sum_{\substack{j \in J \\ s_{\gamma_j}^{arr} = s}} Y_{j,m} \quad \forall s \in S, m \in M \quad (8.12)$$

$$I_{t,m} + A'_{t,m} \geq a_m \quad \forall m \in M', t \in T_m \quad (8.13)$$

$$Q_{j,m} - Y_{j,b_m} \leq 0 \quad \forall j \in J, m \in M' \quad (8.14)$$

$$I_{t,m} - I_{t,b_m} \leq 0 \quad \forall t \in T, m \in M' \quad (8.15)$$

$$W_j \in \{0, 1\} \quad \forall j \in J \quad (8.16)$$

$$Y_{j,m} \in \mathbb{Z}^+ \quad \forall j \in J, m \in M \quad (8.17)$$

$$I_{t,m} \in \mathbb{R}^+ \quad \forall t \in T, m \in M' \quad (8.18)$$

$$Q_{j,m} \in \{0, 1\} \quad \forall j \in J, m \in M' \quad (8.19)$$

The objective (8.6) is the same as in the previous models. Note that the variable Y is now used instead of the variables X and Z that were used in the previous models. This variable Y penalizes cancelling a trip, seat shortages, carriage kilometers, and shunting activities, just as this is done in the previous models by the variables X and Z .

Constraints (8.7) state that at least one job covers trip $t \in T$ or else the trip is cancelled. Every chosen job has to be performed by at least one rolling stock type $m \in M$, see Constraints (8.8). At most one job of each pair of jobs in the sets CJ^1 and CJ^2 can be chosen to be performed. This is modelled by Constraints (8.9).

Constraints (8.10) keep track of the inventory of all train units $m \in M$ just after the departure of trip $t \in T$, and Constraints (8.11) keep track of the inventory of *maintenance* train units $m \in M'$ just after the departure of trip $t \in T$. Then, Constraints (8.12) determine the end-of-day inventory of train units $m \in M$ at station $s \in S$. Furthermore, Constraints (8.13) state that every maintenance unit must be in inventory for the duration of its appointment and at the right location, or else the train unit was either too late or missed its appointment completely. Just as in the SA model, linking constraints are required between the maintenance units and the corresponding regular train units. If a maintenance unit $m \in M'$ is assigned to job $j \in J$, then its corresponding regular type $b_m \in M$ must also be assigned to job $j \in J$, see Constraints (8.14). The same holds for the inventory: if a maintenance unit of type $m \in M'$ is in inventory, then at least one of its corresponding train units of type $b_m \in M$ must also be in inventory, as is required by Constraints (8.15). Finally, Constraints (8.16)-(8.19) specify the domains of the variables.

Note that Assumption (4.) from Section 2 that multiple maintenance units cannot occur in the same composition does not longer influence the size of the JC model. This is because the number of possible compositions does not depend on the number of maintenance units. However, we do not relax this assumption here, because we want to have comparable results for all three models.

8.4 Composition part

The computational results demonstrated that including the constraints of the Composition model (Constraints (4.2)-(4.7)) speeds up the computation time of the JC model significantly. A possible explanation is that earlier research [Fioole et al., 2006] has shown that the Composition model is a tight model formulation leading to strong LP-bounds. Therefore including this part in the JC model speeds up the computation. Besides improving the computation time, using the constraints in the Composition model makes it easy to fix compositions on trips that have departed before the disruption starts. Therefore, we add Constraints (4.2)-(4.7) to the JC model, together with Constraints (8.20)-(8.22) to link the two parts to each other.

$$X_{t,p} - G_{t,p} = 0 \quad \forall t \in T^<, p \in P(t) \quad (8.20)$$

$$C_{t,m} - \sum_{j \in J: \lambda_j = t} Y_{j,m} = 0 \quad \forall t \in T, m \in M \quad (8.21)$$

$$U_{t,m} - \sum_{j \in J: \gamma_j = t} Y_{j,m} = 0 \quad \forall t \in T, m \in M \quad (8.22)$$

All trips $t \in T^<$ that have departed before the start of the disruption should have the same composition as originally assigned, as is modelled by Constraints (8.20).

Constraints (8.21) state that the number of coupled train units at the start of a trip is equal to the number of train units that start their job at the trip. The number of uncoupled train units at the end of a trip is equal to the number of train units that finish their job at the end of the trip, as is modelled by Constraints (8.22). Note that Constraints (8.21) and (8.22) are required to link the Job part to the Composition part of the JC model. This is due to the fact that the start or end of a job in the Job part leads to a composition change in the Composition part.

8.5 Strengthening the formulation

In the JC model as described in the previous section, there are only constraints forbidding *pairs* of jobs to be chosen at the same time. However, these constraints can be tightened by forbidding *sets* of jobs, instead of pairs, to be chosen at the same time.

To that end, we define the undirected graph $G_r = (V_r, E_r)$, where the jobs $j \in J$ present in route $r \in R$ form the set of vertices V_r . There is an edge e between every two jobs j and j' , if and only if $(j, j') \in CJ_1 \cup CJ_2$. This means that every pair of adjacent jobs is not compatible. We call this graph the conflict graph of route r .

A clique is a subset of vertices $cl \subset V_r$ such that for every two vertices in cl there exists an edge connecting the two. Bron and Kerbosch [1973] present a more thorough explanation of cliques and a heuristic to find (maximum) cliques in a graph. So, every clique of jobs, $cl \subseteq V_r$ within the conflict graph G_r is a set of pairwise incompatible jobs. Hence we can strengthen our formulation by **replacing Constraints (8.9) with Constraints (8.23) for all cliques $cl \subset V_r$ for all $r \in R$.**

$$\sum_{j \in cl} W_j \leq 1 \quad \forall r \in R, cl \subset V_r : cl \text{ clique} \quad (8.23)$$

Finding and adding all (maximum) cliques may increase the size of the MIP model and the overall solution time drastically, since there may be an exponential number of cliques. For this reason, we add only some easy to find cliques. We use two such types of cliques, as described below.

Both types of cliques contain jobs $j^1, j^2, \dots, j^n \in J$. All jobs in a clique are related to the same route and have at least one common trip. Furthermore, for both types it holds that the train unit assigned to job j^i , for $i \in \{1, \dots, n-1\}$, is coupled to the composition earlier than the train unit assigned to job j^{i+1} .

The first type of cliques (CJ^1) in our conflict graph is constructed such that in each clique job j^1 is uncoupled first, then j^2 and so on. Furthermore, for all $i = 1, \dots, n$ the uncoupling side of job j^i is equal to the coupling side of jobs $j^{i+1}, j^{i+2}, \dots, j^n$ ($\pi_{j^i} = \omega_{j^{i+1}} = \omega_{j^{i+2}} = \dots = \omega_{j^n}$). This is not feasible, because the train units assigned to jobs j^{i+1}, \dots, j^n are **blocking the uncoupling of job j^i** . So, all tuples

of jobs within CCJ^1 are pairwise incompatible.

$$\begin{aligned}
CCJ^1 := & \{(j^1, j^2, \dots, j^n) \in J^n : T(j^1) \cap T(j^2) \cap \dots \cap T(j^n) \neq \emptyset \\
& \wedge \tau_{j^1}^{dep} < \tau_{j^2}^{dep} < \dots < \tau_{j^n}^{dep} < \tau_{j^1}^{arr} < \tau_{j^2}^{arr} < \dots < \tau_{j^n}^{arr} \\
& \wedge \pi_{j^i} = \omega_{j^{i+1}} = \dots = \omega_{j^n} \quad \forall i = 1, 2, \dots, n-1\}
\end{aligned} \tag{8.24}$$

The second type of cliques within our conflict graph (CCJ^2) consists of sets of jobs $\{j^1, \dots, j^n\}$ such that the train unit assigned to job j^{i+1} is uncoupled before the train unit assigned to job j^i is uncoupled. Furthermore, the uncoupling side of job j^{i+1} is different from the side where it was coupled. This is not allowed, because the train units assigned to job j^1, j^2, \dots, j^i are still there (see the set CJ^2 as an example of a single job **blocking the uncoupling of job j^{i+1}**). So, all tuples of jobs within CCJ^2 are pairwise incompatible. **Thus the set CCJ^2 can be described as follows:**

$$\begin{aligned}
CCJ^2 := & \{(j^1, j^2, \dots, j^n) \in J^n : T(j^1) \cap T(j^2) \cap \dots \cap T(j^n) \neq \emptyset \\
& \wedge \tau_{j^1}^{dep} < \tau_{j^2}^{dep} < \dots < \tau_{j^n}^{dep} < \tau_{j^n}^{arr} < \tau_{j^{n-1}}^{arr} < \dots < \tau_{j^1}^{arr} \\
& \wedge \pi_{j^i} \neq \omega_{j^i} \quad \forall i = 1, 2, \dots, n\}
\end{aligned} \tag{8.25}$$

There is a final constraint that can strengthen the model formulation. Recall that in the Netherlands it is not allowed to both couple and uncouple between two succeeding trips. That means that it is not allowed **that two selected jobs end and start directly after each other at the same station**. So, Constraints (8.26) can be added to the formulation as valid inequality.

$$W_j + W_{j'} \leq 1 \quad \forall (j, j') \in J \times J : \lambda_j = \sigma(\gamma_{j'}) \tag{8.26}$$

9 Results

In this section we discuss the results of applying the EUT model, the SA model, and the JC model on different instances of the main Dutch passenger railway operator NS. **We start in Section 9.1 with a description of the instances and the used parameters. Thereafter, in Sections 9.2 and 9.3 we present the results with a maximum turnaround time of 10 minutes and with a maximum turnaround time of 30 minutes, respectively. Finally, in Section 9.4 we give an overview of the objective function components for each model.**

All computations described in this section are ran with CPLEX 12.5.1 on an Intel (R) Core (ITM) i5-3210M processor with 2.50 GHz and 8GB RAM. The maximum computation time is set to 500 seconds per instance and the allowed gap size is set to 0%.

9.1 Instances and Parameters

In this section we first describe the instances and thereafter the parameters. We ran different experiments on trips of the 2200, 2800, and 3000 line in the Netherlands. Here trains are travelling from Breda (Bd) to Amsterdam (Asd) (2200 line), from Rotterdam (Rtd) to Deventer (Dv) (2800 line), and from Den Helder (Hdr) to Nijmegen (Nm) (3000 line). These lines lead to a total of 1095 trips per day. See Figure 10 for a visual representation, **where the 2200 line is represented by black edges, the 2800 line by black dotted edges, and the 3000 line by dark grey edges.**

Table 2 gives an overview of the instances on which the models have been tested. Here, “#RS types” denotes the number of regular rolling stock types used. As can be seen, this is either two or three. In the instances with two rolling stock types, the train units consist of either three or four carriages, while in the instances with three different rolling stock types they consist of either three, four or five carriages. The maximum number of carriages in a train equals 15, so in total there are 31 compositions and 356 composition changes possible when using two different rolling stock types, and 72 compositions and 884 composition changes are possible when using three different rolling stock types.

The column “**Maximum** turnaround time” denotes the maximum time a train is allowed to wait for its succeeding return trip in an end station. This maximum time equals 10 minutes in half of the instances and 30 minutes in the rest. The **actual** turnaround time denotes the time between the arrival time of an incoming trip and the departure time of the first return trip in an end station.

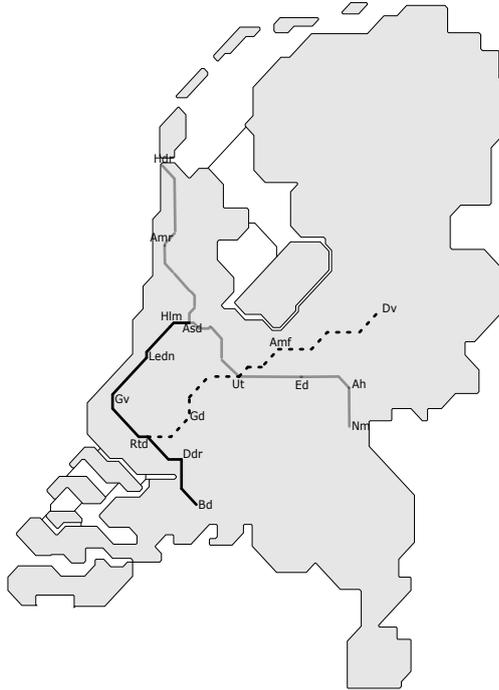


Figure 10: Case lines

In the instances with a maximum turnaround time of 30 minutes most of the incoming trips have a succeeding return trip in an end station, since the lines are operated with a frequency of 2 trains per hour. As a consequence, in these instances the routes are long, and thus for each route $r \in R$ there are many jobs $j \in J$ with $j \subset r$. This leads to a large number of possible jobs in the JC model, which increases the computation time.

If in an end station the time between an incoming trip and the first return trip exceeds the maximum turnaround time, then the train units are assumed to be transferred to the corresponding shunting yard in between the trips. In that case, the incoming trip and the return trip do not belong to the same route anymore.

The column “Disrupted area” describes the location where the disruption takes place. A disruption takes place **either** between the stations The Hague (Gv) and Rotterdam (Rtd) or between the stations Utrecht (Ut) and Amsterdam (Asd). In order to test whether the start time of the disruption has any

Case number	#RS types	Maximum turnaround time	Disrupted area
1a	2	10	Ut - Asd
1b	2	10	Gv - Rtd
2a	3	10	Ut - Asd
2b	3	10	Gv - Rtd
3a	2	30	Ut - Asd
3b	2	30	Gv - Rtd
4a	3	30	Ut - Asd
4b	3	30	Gv - Rtd

Table 2: Different instances

influence on the computation time, we let disruptions take place between 07:00-09:00, 07:03-09:03, 07:06-09:06, ..., and 07:57 - 09:57, so in total **during** 20 different time slots. Furthermore, we experiment with a number of train units requiring maintenance **varying** between one and six. All instances in Table 2 are solved for the different time slots and for the different numbers of train units requiring maintenance. As a result, there are in total $8 * 20 * 6 = 960$ instances which we solve with each of the three models.

Parameter	Penalty
Cancelling	10000
EOD deviation	100
Capacity shortage kilometer	1
Carriage kilometer	1
Deviation original plan	50
Missing maintenance	300

Table 3: Objective function penalties

All models use the *same* objective function based on the objective coefficients mentioned in Table 3. The complete objective function is a weighted sum of the objective coefficients multiplied with their accompanying decision variables. Here “Cancelling” denotes the penalty for cancelling a trip. Since we consider cancelling a trip as the worst thing that can happen, the penalty for cancelling a trip is higher than any other penalty. “EOD deviation” means the penalty for deviating from the scheduled end-of-day balance. For each negative difference, a dead-heading trip must be scheduled during the night to rebalance. This is expensive, so we want to keep the deviation small.

“Capacity shortage kilometer” stands for the penalty on the number of passengers that do not fit in an assigned composition, measured per kilometer. Note that **we use the original passenger demand for each trip as provided by NS**. So, the demand on the trips that are operated is assumed to be unchanged during the disruption. Taking dynamic passenger demand directly into account during the disruption is outside the scope of this paper. We refer to Kroon et al. [2014] for a paper that does take dynamic passenger demand into account (but no maintenance appointments). A similar approach could have been applied in the current paper as well.

“Carriage kilometer” is the penalty on the number of carriages assigned per kilometer. A trade-off exists between minimizing the number of carriage kilometers and minimizing the seat-shortages for passengers, because both objectives are conflicting. **Minimizing seat-shortages will lead to appointing large compositions to trips, while minimizing the number of carriage kilometers will lead to appointing small compositions to trips.**

“Deviation original plan” stands for the penalty on the difference between the original and the rescheduled plan in terms of the numbers of couplings and uncouplings taking place. Each additional shunting movement requires an additional crew task for which a crew member must be found. This takes time, and in a real-time situation not much time is available. As a consequence, we want to keep the number of additional shunting movements low.

Finally, “Missing maintenance” stands for the penalty on the number of train units that miss their scheduled maintenance appointment. As explained before, it is undesirable that a train unit misses its maintenance appointment, so we set a large penalty on this. **The applied penalties come from existing literature or from discussions with dispatching experts of NS.**

It is important to note that that, in case the models are able to prove optimality, then they find optimal solutions with the same optimal objective value . The models were able to prove optimality for most of the instances. This does not necessarily mean that the models produce exactly the same rolling stock circulation. However, the circulations are equally good with respect to the objective function. Therefore, we will mainly compare the models in terms of the computation time and in terms of the number of times an optimal solution was found.

9.2 Maximum turnaround time 10 minutes

In this subsection we compare the results obtained by the three models for all instances with a maximum turnaround time of 10 minutes. We show average computation times for all problem instances with the same number of rolling stock types on the same initial locations, with the same maximum turnaround

time, and with the same number of train units requiring maintenance. In other words, the average is taken over the 20 different disruption time slots while all other instance parameters remain fixed.

In all tables presenting the results, the first column (M) denotes the number of train units that require maintenance, the second column (Model) denotes which model was used to solve the instances, the third column (Time) represents the average computation time required to solve the problem instances, the fourth column ($\#NO$) presents the total number of times no **proven** optimal solution was found, the fifth column ($\#C$) **presents** the average number of constraints in the model, and the sixth column ($\#V$) **presents** the average number of variables used to solve the instances.

First, we show the results for the instances with a maximum turnaround time of 10 minutes. The results of using two regular rolling stock types are shown in Table 4 and Figure 11. As can be seen, the JC model performs significantly better than the SA and the EUT model: both the number of times a proven optimal solution was found and the average computation time are better in the JC model than in the EUT and the SA model. Furthermore, in some of the instances when there are many train units requiring maintenance, the EUT model is not able to find a proven optimal solution within 500 seconds. The JC model finds a proven optimal solution for all instances, and the SA model for all instances but two.

M	Model	Time	$\#NO$	$\#C$	$\#V$
1	EUT	58	0	165339	1063531
	JC	38	0	83992	421309
	SA	64	0	132646	645108
2	EUT	95	0	274219	1823416
	JC	66	0	87004	438030
	SA	87	0	168126	855405
3	EUT	133	0	383100	2562576
	JC	74	0	90017	439832
	SA	110	0	203607	1065572
4	EUT	172	0	474373	3216358
	JC	83	0	92309	441043
	SA	161	0	229304	1205559
5	EUT	268	0	568219	3931422
	JC	97	0	94641	452064
	SA	213	0	260132	1443211
6	EUT	380	6	666201	4803231
	JC	112	0	97032	459216
	SA	256	1	291240	1603491

(a) Case 1a Ut - Asd

M	Model	Time	$\#NO$	$\#C$	$\#V$
1	EUT	51	0	165444	1064512
	JC	50	0	84115	436677
	SA	57	0	132762	645703
2	EUT	97	0	274498	1825098
	JC	60	0	87104	438490
	SA	106	0	168275	856194
3	EUT	138	0	383490	2585684
	JC	92	0	90166	440303
	SA	128	0	203789	1066685
4	EUT	227	0	474856	3219325
	JC	76	0	92461	441531
	SA	199	0	229509	1206651
5	EUT	328	2	569021	3942681
	JC	105	0	96721	452497
	SA	239	0	260405	1443981
6	EUT	417	8	667091	4805197
	JC	118	0	97333	460134
	SA	279	1	291865	1604754

(b) Case 1b Rtd - Gv

Table 4: Results with 2 regular types and 10 minutes maximum turnaround time

The results of applying the models on instances with three regular rolling stock types are shown in Table 5 and Figure 12. The JC model performs again best, both in terms of the computation **time** and in terms of the number of proven optimal solutions. The SA model is second and the EUT model performs worst, having difficulty to solve instances with many maintenance appointments.

The JC model performs better because it does not need any additional compositions for an additional train unit that requires maintenance. The more additional train units require maintenance, the more beneficial this becomes. As can be seen, it results in less variables and constraints than the SA and EUT model require. We conclude that the JC model performs best if the maximum turnaround time is 10 minutes.

9.3 Maximum turnaround time 30 minutes

In contrast with the results with a maximum turnaround time of 10 minutes, the SA model outperforms both the EUT and the JC model when the maximum turnaround time equals 30 minutes. Due to the larger maximum turnaround time, **the jobs are now longer and** there are many more possible jobs. **This makes it a harder problem to solve.**

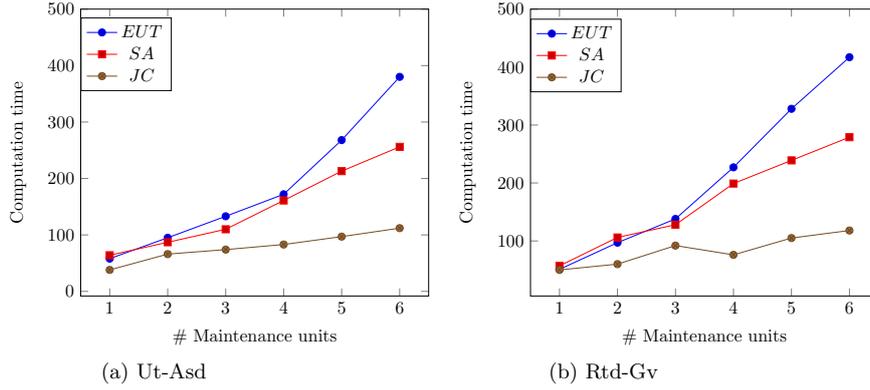


Figure 11: Computation times with 2 regular types and 10 minutes maximum turnaround time

M	Model	Time	#NO	#C	#V
1	EUT	112	0	305545	2048938
	JC	141	0	169642	1013684
	SA	101	0	219016	1221847
2	EUT	191	0	471149	3217491
	JC	173	0	172654	1015486
	SA	180	0	257384	1432144
3	EUT	276	5	572205	3921008
	JC	200	0	175666	1017288
	SA	227	0	285972	1571981
4	EUT	358	10	737810	5089561
	JC	220	0	178679	1019090
	SA	270	0	324341	1782278
5	EUT	404	14	885807	6131286
	JC	251	0	181688	1020892
	SA	304	2	352926	1922115
6	EUT	462	16	986867	6834803
	JC	297	1	184704	1022694
	SA	359	4	381618	2061952

(a) Case 2a Ut - Asd

M	Model	Time	#NO	#C	#V
1	EUT	88	0	305856	2050828
	JC	89	0	169852	1014675
	SA	73	0	219217	1222974
2	EUT	131	0	471629	3220459
	JC	121	0	172877	1016488
	SA	125	0	257618	1433465
3	EUT	221	1	572788	3924625
	JC	190	0	175902	1018301
	SA	160	0	286229	1573431
4	EUT	304	3	738562	5094256
	JC	183	0	178928	1020114
	SA	200	0	324631	1783922
5	EUT	421	11	886710	6136942
	JC	226	0	181950	1021927
	SA	227	0	353239	1923888
6	EUT	456	15	987873	6841108
	JC	241	0	184979	1023740
	SA	294	2	381854	2063854

(b) Case 2b Rtd - Gv

Table 5: Results with 3 regular types and 10 minutes maximum turnaround time

For the instances with two regular rolling stock types, the results are shown in Tables 6 and Figure 13. All models tend to be slower than in the instances with a maximum turnaround time of 10 minutes. The SA model outperforms the EUT model **in terms of the computation time and in terms of the number of proven optimal solutions**. The JC model performs worst. As explained before, this is due to the fact that the number of possible jobs has increased. On the other hand, the computation time for the JC model does not increase as quickly as for the other models when more train units require maintenance. In the end, the JC model performs even better than the EUT model when 6 train units have a maintenance appointment.

In conclusion, the SA model performs better than both the JC and the EUT model when using two regular types and a maximum turnaround time of 30 minutes.

The results when using three regular rolling stock types are shown in Tables 7 and Figure 14. As can be seen, the SA model outperforms both the JC and the EUT model in terms of the computation time and in terms of the number of proven optimal solutions again. The model is able to find proven optimal solutions for most of the instances within the time limit, while the other models have more problems with finding proven optimal solutions.

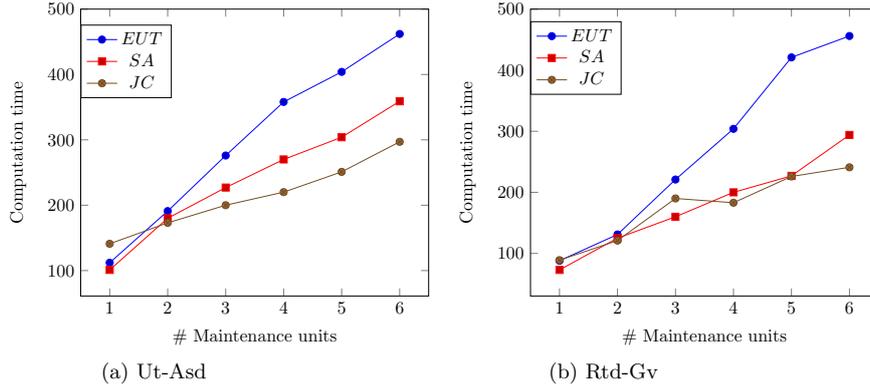


Figure 12: Computation times with 3 regular types and 10 minutes maximum turnaround time

M	Model	Time	#NO	#C	#V
1	EUT	45	0	175221	1063525
	JC	190	0	148931	443767
	SA	39	0	142730	645102
2	EUT	103	0	291243	1823410
	JC	238	0	153389	447078
	SA	81	0	183456	855399
3	EUT	203	0	407249	2583295
	JC	270	0	157845	450389
	SA	156	0	224180	1065696
4	EUT	235	0	504462	3216352
	JC	315	1	162300	453700
	SA	232	0	252641	1205533
5	EUT	361	2	602319	3904671
	JC	369	3	167027	456918
	SA	270	0	280120	1360012
6	EUT	430	7	700232	4700129
	JC	425	8	171872	499812
	SA	320	2	301321	1421208

(a) Case 3a Ut - Asd

M	Model	Time	#NO	#C	#V
1	EUT	41	0	175340	1064506
	JC	182	0	137231	443456
	SA	37	0	141586	645697
2	EUT	84	0	291522	1825092
	JC	190	0	141613	446626
	SA	71	0	181967	856188
3	EUT	172	0	407639	2585678
	JC	210	0	145993	449796
	SA	123	0	222346	1066679
4	EUT	231	0	504882	3219319
	JC	220	0	150372	452966
	SA	188	0	253262	1206645
5	EUT	290	2	607123	3924021
	JC	296	1	154871	456723
	SA	211	1	282432	1399864
6	EUT	410	7	710023	4647910
	JC	398	5	158120	459102
	SA	280	1	310212	1581904

(b) Case 3b Rtd - Gv

Table 6: Results with 2 regular types and 30 minutes maximum turnaround time

Just as with two regular rolling stock types, it takes longer to find an optimal solution for the JC model from the start, but the computation time does not increase quickly when more train units require maintenance. However, the JC model is not able to find a proven optimal solution for any of the instances with 6 maintenance units. Note that the model was able to find feasible solutions for most cases with 6 maintenance units, but it was not able to prove optimality within the time window of 500 seconds.

To conclude, both the SA and the EUT model are not influenced heavily by the maximum turnaround time. The computation times differ little between having a maximum turnaround time of 10 minutes or having one of 30 minutes. On the other hand, the maximum turnaround **time** has a significant influence on the JC model. With a maximum turnaround time of 30 minutes, the computation times increase drastically. Furthermore, the SA model outperforms both the JC and the EUT model in terms of the computation time and in terms of the number of proven optimal solutions found when there are three regular types and a maximum turnaround time of 30 minutes.

9.4 Evaluating the objective components

As mentioned before, the three models give the same optimal objective value **on the same instance**. This

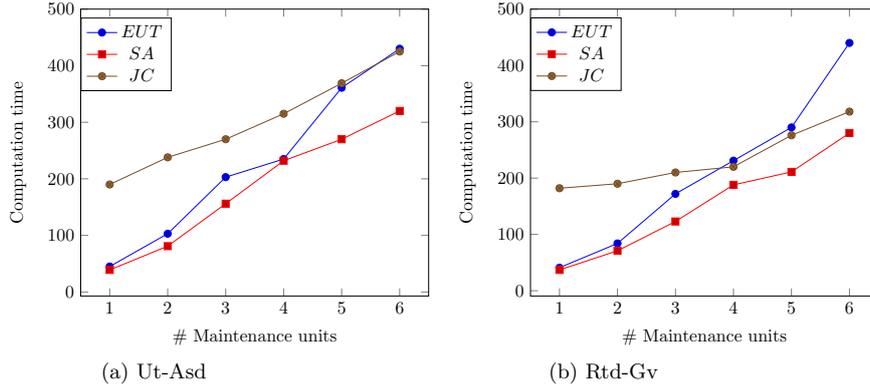


Figure 13: Computation times with 2 regular types and 30 minutes maximum turnaround time

M	Model	Time	#NO	#C	#V
1	EUT	104	0	324633	2049838
	JC	366	3	252781	1026777
	SA	116	0	232064	1221847
2	EUT	152	0	482393	3069888
	JC	410	9	257343	1030192
	SA	151	0	257159	1361684
3	EUT	194	0	590055	3794180
	JC	430	10	258106	1030312
	SA	180	0	282257	1501521
4	EUT	262	4	747896	4835905
	JC	460	12	262743	1033742
	SA	210	1	307352	1641358
5	EUT	341	9	855553	5508888
	JC	470	15	267380	1037172
	SA	222	2	332447	1781195
6	EUT	427	14	963208	6242939
	JC	-	20	272003	1040602
	SA	241	2	357582	1921032

(a) Case 4a Ut - Asd

M	Model	Time	#NO	#C	#V
1	EUT	75	0	324628	2067133
	JC	387	4	235246	1023441
	SA	152	0	232229	1229186
2	EUT	166	0	482867	3093514
	JC	388	5	239729	1026715
	SA	179	0	257335	1370023
3	EUT	281	4	590112	3810032
	JC	420	7	244215	1029989
	SA	240	1	282444	1581189
4	EUT	300	8	747882	4839291
	JC	440	10	248698	1033263
	SA	260	3	307550	1670246
5	EUT	363	11	855538	5539422
	JC	460	11	253181	1036537
	SA	290	4	332656	1813218
6	EUT	482	14	963196	6268116
	JC	-	20	257664	1039811
	SA	340	8	357782	1972034

(b) Case 4b Rtd - Gv

Table 7: Results with 3 regular types and 30 minutes maximum turnaround time

does not necessarily mean that the solutions are the same. In order to investigate whether the three models lead to structurally different solutions, Figure 15 shows a pie chart where the average percentage contribution of each objective aspect is shown per model.

As can be seen, the contribution of cancelling a trip is the same for the three models. This is because the penalty for cancelling a trip is by far the largest penalty. Hence all models cancel as few additional trips as possible. Furthermore, the three models differ a little with respect to capacity shortages, carriage kilometers, end-of-day balance deviations, met maintenance appointments, and deviations from the original plan. There are small differences between the percentages, but the contribution of each aspect is almost the same for all models. Only in the EUT model maintenance appointments are actually missed, but this happened in just two instances. From the pie chart we can conclude that there are no structural differences between the results of the three models.

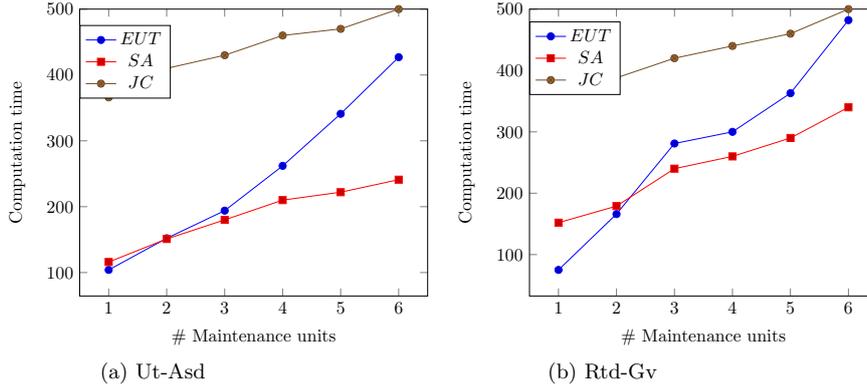


Figure 14: Computation times with 3 regular types and 30 minutes maximum turnaround time

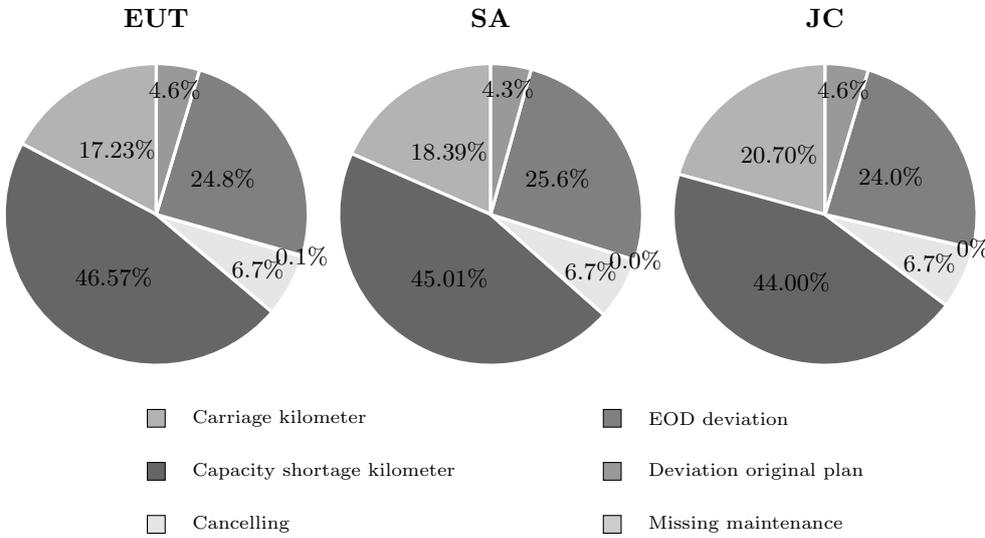


Figure 15: Objective function contributions per model

10 Conclusions and further research

In this paper, three models are presented and compared for rescheduling the rolling stock of passenger trains during large disruptions, while taking scheduled maintenance appointments into account. The considered problem is an extension of the Rolling Stock Rescheduling Problem (RSRP). All models extend the Composition model of [Fiole et al. \[2006\]](#) and [Nielsen \[2011\]](#), which is known for rescheduling rolling stock without maintenance appointments.

The Extra Unit Type (EUT) model uses an additional rolling stock type for each train unit that require maintenance. In this way constraints can be **imposed** on them. This extension has the drawback that adding additional rolling stock types leads to a rapid increase in the number of possible compositions and composition changes. As a result, the EUT model tends to require more computation time than the other models when more train units require maintenance.

The second model is the Shadow-Account (SA) model. Within the SA model a shadow account for all train units is maintained, in particular for the maintenance units. In this way, maintenance constraints can be **imposed** on the train units that require maintenance.

The third model is the Job-Composition (JC) model. This model assigns train units to jobs. As a result, a path is created for each train unit. Specific paths can be created for train units that require maintenance leading to the corresponding maintenance locations.

The three models have been tested on a large number of instances of NS, the main Dutch operator

of passenger trains. The models use the same objective function. Therefore, if the models find a proven optimal solution for an instance, then the obtained optimal objective function values are the same. Therefore we compared the models on their computation time and on the number of times they found a proven optimal solution.

The results show that the SA and the EUT model are rather insensitive to the maximum turnaround time. Their computation times differ little between instances with a maximum turnaround time of 10 minutes and instances with a maximum turnaround time of 30 minutes that are otherwise fully comparable. This is in contrast with the JC model: this model performs best on instances with a maximum turnaround time of 10 minutes. However, it is considerably slower on instances with a longer maximum turnaround time. This is due to the increase in the number of possible jobs in case of a longer maximum turnaround time.

As a consequence, the EUT model is inferior to the SA model and to the JC model. Whether the SA or the JC model performs better depends on the maximum turnaround time. The SA model performs better than the JC model on instances with a maximum turnaround time of 30 minutes. The JC model performs better than the SA model on instances with a maximum turnaround time of 10 minutes. An additional advantage of the JC model is that it is not necessary to assume that at most one train unit requires maintenance in a composition, as is necessary in the other models.

There are several directions for further research. First, a dynamic Branch & Price & Cut approach may be used to solve the JC model. In this way it may be possible to solve the instances with a longer maximum turnaround time faster. This can be incorporated with a column generation technique. The main challenge when using a column generation technique is that the order of the train units in the compositions is important. As a result, the columns are highly dependent on each other. This is in contrast with, for instance, crew rescheduling where the different crew members (driver, conductor) per train are rather independent of each other, unless it is specified that they should operate as much as possible as a *team*.

Finally, other practical aspects are important to be included in the RSRP. Especially the integration of accurate dynamic passenger demand with the maintenance appointments is an interesting topic for further research. Furthermore, station routing and robustness should be incorporated in the disruption management models.

References

- R. Ahuja, T. Magnanti, and J. Orlin. *Network Flows: Theory, Algorithms and Applications*. Prentice-Hall Englewood Cliffs, 1993.
- C. Barnhart, N.L. Boland, L.W. Clarke, E. Johnson, G.L. Nemhauser, and R.G. Shenoi. Flight string models for aircraft fleetings and routing. *Transportation Science*, 32(3):208–220, 1998.
- R. Borndörfer, M. Reuther, T. Schlechte, K. Waas, and S. Weider. Integrated optimization of rolling stock rotations for intercity railways. *Transportation Science*, 2015. doi: <http://dx.doi.org/10.1287/trs.2015.0633>.
- C. Bron and J. Kerbosch. Algorithm 457: finding all cliques of an undirected graph. *Communications of the ACM*, 16(9):575–577, 1973.
- P. Brucker, J. Hurink, and T. Rolfes. Routing of railway carriages. *Journal of Global Optimization*, 27(2-3):313–332, 2003.
- C. Cacchiani, D. Huisman, M. Kidd, L.G. Kroon, P. Toth, L. Veelenturf, and J. Wagenaar. An overview of recovery models and algorithms for real-time railway rescheduling. *Transportation Research Part B: Methodological*, 63:15–37, 2014.
- L. Clarke, E. Johnson, and G. Nemhauser. The aircraft rotation problem. *Annals of Operations Research*, 69:139–151, 1997.
- J. Cordeau, G. Desaulniers, N. Lingaya, F. Soumis, and J. Desrosiers. Simultaneous locomotive and car assignment at VIA Rail Canada. *Transportation Research Part B: Methodological*, 35(8):767–787, 2001.

- P. Fioole, L.G. Kroon, G. Maróti, and A. Schrijver. A rolling stock circulation model for combining and splitting of passenger trains. *European Journal of Operational Research*, 174:1281–1297, 2006.
- G.L. Giacco, A. D’Ariano, and D. Pacciarelli. Rolling stock rostering optimization under maintenance constraints. *Journal of Intelligent Transportation Systems*, 18(1):95–105, 2014.
- L.G. Kroon, G. Maróti, and L.K. Nielsen. Rescheduling of railway rolling stock with dynamic passenger flows. *Transportation Science*, 49(2):165–184, 2014.
- N. Lingaya, J. Cordeau, G. Desaulniers, J. Desrosiers, and F. Soumis. Operational car assignment at VIA Rail Canada. *Transportation Research Part B: Methodological*, 36(9):755–778, 2002.
- G. Maróti. Operations research models for railway rolling stock planning. *PhD thesis, Eindhoven University of Technology, The Netherlands*, 2006.
- G. Maróti and L.G. Kroon. Maintenance routing for train units: the transition model. *Transportation Science*, 39(4):518–525, 2005.
- G. Maróti and L.G. Kroon. Maintenance routing for train units: the interchange model. *Computers & Operations Research*, 34(4):1121–1140, 2007.
- T. Mellouli and L. Suhl. Rotation planning of locomotive and carriage groups with shared capacities. In *Algorithmic Methods for Railway Optimization*, pages 276–294. Springer, 2007.
- L.K. Nielsen. Rolling stock rescheduling in passenger railways. *Ph.D thesis, Erasmus University, Rotterdam, The Netherlands*, 2011.
- L.K. Nielsen, L.G. Kroon, and G. Maróti. A rolling horizon approach for disruption management of railway rolling stock. *European Journal of Operational Research*, 220:496–509, 2012.
- M. Peeters and L.G. Kroon. Circulation of railway rolling stock: a branch-and-price approach. *Computers and Operations Research*, 35(2):538 – 556, 2008.
- K. Sato and N. Fukumura. Real-time freight locomotive rescheduling and uncovered train detection during disruptions. *European Journal of Operational Research*, 221:636–648, 2012.
- K. Sato, S. Sakikawa, T. Morita, N. Ueki, and T. Murata. Crew and vehicle rescheduling based on a network flow model and its application to a railway train operation. *IAENG International Journal of Applied Mathematics*, 39(3), 2009.
- K.T. Talluri. The four day aircraft maintenance routing problem. *Transportation Science*, 32(1):43–53, 1998.
- J.C. Wagenaar, L.G. Kroon, and M. Schmidt. Maintenance appointments in railway rolling stock rescheduling. *ERIM Report Series Reference No. ERS-2015-002-LIS*, 2015.

ERIM Report Series <i>Research in Management</i>	
ERIM Report Series reference number	ERS-2016-001-LIS
Date of publication	2016-01-07
Version	07-01-2016
Number of pages	31
Persistent URL for paper	http://hdl.handle.net/1765/79441
Email address corresponding author	jwagenaar@rsm.nl
Address	Erasmus Research Institute of Management (ERIM) RSM Erasmus University / Erasmus School of Economics Erasmus University Rotterdam PO Box 1738 3000 DR Rotterdam, The Netherlands Phone: +31104081182 Fax: +31104089640 Email: info@erim.eur.nl Internet: http://www.erim.eur.nl
Availability	The ERIM Report Series is distributed through the following platforms: RePub, the EUR institutional repository Social Science Research Network (SSRN) Research Papers in Economics (RePEc)