



Coordinating maintenance windows and train traffic: a case study

Tomas Lidén^{1,2}

Accepted: 12 February 2020 / Published online: 26 April 2020
© The Author(s) 2020

Abstract

This paper concerns a case study for optimal planning and coordination of railway maintenance windows and train traffic. The purpose is to validate a previously presented optimization model on a demanding real-life problem instance and to obtain results that apply in similar planning situations. A mixed integer linear programming model is used for a 913 km long, single-track railway line through the northern part of Sweden, with traffic consisting of 82 trains per day, most of which are freight trains. Cyclic 1-day schedules are produced, which show that 2 h long maintenance windows can be scheduled with small adjustments of the train traffic. The sensitivity for cost changes is studied, which shows that the train costs must increase by more than 30% in order to change the structure of the window solutions. Resource efficient window schedules are obtained by assigning maintenance teams to all windows while respecting crew work and rest time restrictions. A comparison with manually constructed plans from the Swedish Transport Administration indicates that larger window volumes can be scheduled at a lower cost and with solution structures which are deemed reasonable and useful as guidance for constructing the real window patterns. Finally, we estimate that using an integrated planning approach (where maintenance and trains are jointly planned) instead of a sequential approach (where a train timetable has precedence over the maintenance windows), will give maintenance cost savings of 11–17%, without incurring any large cost increases for the train traffic. The paper also presents a method for achieving cyclic schedules without any period-deciding variables, and discusses the consequences of the aggregated capacity usage model that has been adopted.

Keywords Railway scheduling · Maintenance planning · Case study

✉ Tomas Lidén
tomas.liden@liu.se

¹ Linköping University, Department of Science and Technology, 601 74 Norrköping, Sweden

² Swedish National Road and Transport Research Institute, Traffic Analysis and Logistics,
114 28 Stockholm, Sweden

1 Introduction and motivation

Railway infrastructure maintenance consumes large budgets, is complicated to organize and has numerous challenging planning problems (Lidén 2015). In monetary volumes, the European countries have an average spending for maintenance and renewals of 70,000 EUR per km track and year (EIM-EFRTC-CER Working Group 2012). Statistically, the maintenance spendings are in the order of 40% of the gross value added for the railway sector in Europe, with similar figures found in Sweden (Lidén 2016).

Train services and maintenance tasks should ideally be planned together, but are usually treated by different organisations. Historically, practice and research about railway scheduling has focused mainly on train operations and timetabling, while maintenance planning has received less attention. Although the monetary volumes cited above indicate a large potential for a more efficient use of maintenance spendings if the coordination with train traffic can be improved, there are few research publications that treat such joint scheduling problems. Lidén and Joborn (2017) presented a macroscopic optimization model for the integrated scheduling of maintenance windows and train services, which was improved and extended with maintenance crew resource considerations in Lidén et al. (2018). The methods developed in these previous publications were tested on various theoretical test instances, but were not applied on any practical case, which is the aim of the current work.

In this paper, we present a case study for a real-life problem instance concerning the main railway line through the northern part of Sweden, which is over 900 km long, mostly single-track, and has over 80 trains per day, predominantly freight trains. The purpose of the case study is to (a) validate the model on a real and demanding planning problem, and to (b) obtain results that apply for similar railway lines. For this study, it was decided to produce 1-day periodic production plans which can be repeated every week day. Since the longest train runs take more than 10 h, the train traffic must run continuously over the day and with considerable overlap between consecutive days. Hence a cyclic schedule is necessary, and we adapt the optimization model so as to handle cyclicity as well as variable maintenance window lengths.

The contributions of this paper are: (1) verification that optimal schedules can be obtained for real-life one-day instances of practical interest; (2) results showing that the obtained schedules are stable for relatively large cost uncertainties (up to 30%) and that consideration of crew resources give more efficient solutions for the contractors. Furthermore: (3) it is validated that the obtained solution structures are reasonable and useful as guidance for constructing the real window patterns; and (4) that integrated planning of maintenance and trains give estimated maintenance cost savings of 11–17%, as compared to sequential planning (where an existing or new train timetable has precedence over the maintenance windows), without incurring any large cost increases for the train traffic. Finally, a model approach for cyclic scheduling that does not require any period-deciding variables is presented and the consequences of the aggregated capacity usage model are discussed.

The paper is organized as follows: Sect. 2 presents research literature close to our problem setting. Section 3 describes the planning problem and features considered. The mathematical model is presented and discussed in Sect. 4, followed by the case study and results in Sect. 5. Finally, some conclusions and future directions are given in Sect. 6.

2 Literature review

This section describes research literature that is close to our problem setting, i.e. tactical coordination of train traffic and infrastructure maintenance, while handling resource limitations and cyclic (periodic) scheduling. We focus on how cyclic aspects have been treated, and refer to Lidén et al. (2018) regarding handling of resource considerations. The publications are divided into three groups: those that schedule infrastructure maintenance (in Sect. 2.1), those that schedule trains or adjust timetables to make room for maintenance (2.2) and those where maintenance and trains are jointly scheduled (2.3).

A neighboring research field, which we do not cover here is maintenance handling for rolling stock problems [see e.g. Giacco et al. (2014), Lai et al. (2015)] and robustness of such rolling stock plans during disruptions [see e.g. Cacchiani et al. (2012), Cadarso and Marín (2011)]. The basic difference is that rolling stock maintenance is individually scheduled for each vehicle and directly coupled to the train timetable, while infrastructure maintenance is scheduled at a network level, not as part of the train schedules.

2.1 Maintenance scheduling

We use the same categorization as proposed in Lidén (2015, 2016), where three classes of tactical maintenance scheduling problems have been identified: (1) deterioration-based maintenance scheduling, (2) maintenance vehicle routing and team scheduling, and (3) possession scheduling.

The first class concerns the scheduling of preventive and corrective maintenance, sometimes also renewals, while considering the deterioration of the track. The deterioration aspects make these scheduling problems non-cyclic. The majority of papers focus on tamping, such as Vale et al. (2012), Vale and Ribeiro (2014), Gustavsson (2014), Wen et al. (2016), but the coordination with train traffic is seldom included. One exception is Su et al. (2017) who study grinding of tracks to handle rail cracks and squats. A three-level decision model is used where the middle level finds suitable time slots for the maintenance actions such that the interferences with train traffic volumes are minimized. This paper considers the repeated (monthly) replanning that will take place and how the system will evolve over a longer time period (several years) as new measurements and maintenance actions are performed. Thus, the planning is non-cyclic.

The second class concerns the assignment and scheduling of a given set of maintenance tasks to different maintenance teams, which can vary regarding capabilities,

equipment and home location. The problem is sometimes labelled curfew planning and if train traffic is considered, it is done by imposing constraints or costs on how and which tasks that can be scheduled simultaneously. Some examples are Peng and Ouyang (2012), Borraz-Sánchez and Klabjan (2012), Camci (2015), but none of these treat cyclic schedules.

The third class concerns the scheduling of work possessions. In these problems, the traffic impact must be considered and various approaches have been used for doing so. Higgins (1998) minimizes the expected interference delays (train delays due to late ending job as well as job delays due to late trains) while Budai et al. (2006) cluster short routine tasks of preventive maintenance and larger projects together, in order to minimize possession time and maintenance cost. Boland et al. (2013) adjust a given maintenance plan for a complete transportation chain so as to maximize the transported throughput (where traffic is treated as flows of trains), and Savelsbergh et al. (2015) use a similar approach for optimizing a transportation plan. All these models are non-cyclic.

The only reference known to us, that treats a tactical cyclic maintenance scheduling problem is van Zante-de Fokkert et al. (2007). The paper describes how the maintenance work planning was reorganized in The Netherlands by dividing the track into work zones (den Hertog et al. 2005), constructing regular work possession patterns (called single-track grids), and schedule them in a cyclic four-week plan so as to give access to every part of the network at least once every fourth week. Consideration to train traffic is handled manually when constructing the single-track grids but no method for dimensioning these patterns is given.

2.2 Train scheduling (with fixed maintenance closures)

This research field has an abundance of literature and there are also several surveys which cover different problem types and aspects, for example Caprara et al. (2007, 2011), Cacchiani and Toth (2012), Cacchiani et al. (2014), Corman and Meng (2014). Cyclic planning is common and one of the earliest models was introduced by Serafini and Ukovich (1989), labelled the Periodic Event Scheduling Problem (PESP). As discussed in Caprara et al. (2007) the PESP model needs decision variables to determine the periodic separation of events and these variables make the problem considerably harder to solve than the non-cyclic scheduling problem.

Despite the many train scheduling publications, there are relatively few publications that consider infrastructure maintenance. As shown in Caprara et al. (2006), fixed maintenance closures that are known beforehand can be handled in a timetabling problem by reducing the train scheduling possibilities in the graph representation. For an existing timetable, the replanning or timetable adjustment problem due to given and fixed track closures or maintenance activities has been studied in: Brucker et al. (2005) (scheduling of single-track traffic past a working site on a line section); Vansteenwegen et al. (2015) (robust rescheduling due to planned track closures on large stations and junctions); Veelenturf et al. (2016) and Louwerse and Huisman (2014) (rescheduling of timetables, rolling stock and crew during major disruptions in operational dispatching); Van Aken et al. (2017a, b) (timetable

adjustment for multi-day infrastructure possessions); Arenas et al. (2018) (short-term and detailed adjustment of regular trains as well as introduction of maintenance trains due to planned track works); and Zhu et al. (2018) (combined rescheduling, reordering and adjustment of train turning patterns during simultaneous disruptions). Only Van Aken et al. consider a cyclic problem, since the work possessions in that problem are substantially longer (1–2 days) than the timetable period (30–60 min). The other publications consider shorter possessions and how to adjust the train scheduling before, during and after the track closures. Louwerse and Huisman use train series that are equally spaced to retain the timetable regularity but the scheduling model itself is non-cyclic. Van Aken et al. (2017b) reduce the complexity of the cyclic timetable adjustment problem by various aggregation techniques. Since relatively small time adjustments are studied (max 10 min), many period-deciding variables and headway constraints can be avoided which enables faster solving and larger problem instances.

2.3 Combined scheduling of maintenance and trains

Here we list references that schedule *both* trains and maintenance in the same model. None of these treat cyclic schedules. First, there is a group of papers that introduce a small number of work possessions into an existing train timetable, by allowing different types of adjustments to the trains. Ruffing (1993) is an early paper and one of few that consider operational restrictions (reduced speed) for trains passing a work site. Albrecht et al. (2013) address the real-time operational control case for a single-track line, treat maintenance as pseudo trains and allow train times to be adjusted but not cancelled. Forsgren et al. (2013) treat the tactical timetable revision planning case with a model that can handle a network with both single and multi-track lines, can shift the work start time, allow trains to be rerouted or cancelled and consider different running times depending on train stops. Luan et al. (2017) and D'Ariano et al. (2019) also address the timetable revision planning case for mixed networks, the former treating maintenance activities as pseudo trains, while the latter use separate maintenance tasks to be scheduled consecutively whenever possible and also apply a stochastic approach in order to find robust scheduling solutions.

Lidén and Joborn (2017) treat a long-term tactical planning case where a traffic plan does not yet exist and many work slots shall be coordinated with the train traffic. An exact model is used but the network capacity is controlled with an aggregated approach that does not guarantee a conflict-free timetable. A mathematically stronger model is presented in Lidén et al. (2018), which is extended with a maintenance crew resource assignment model such that resource-efficient planning is achieved, while respecting work and rest time regulations. Both these papers use synthetic test instances to show that multi-day problems can be solved optimally with reasonable solution times in a commercial mixed integer solver. As an example, a 24 h single-track line instance with 18 links and 80 trains can be solved optimally without resource considerations, but does, however, not reach acceptable optimality gaps within a time limit of 1 h when including resources. This size is comparable with the case study for Trafikverket and since we will also introduce cyclic handling,

we expect that solution times of several hours are needed in order to find good quality solutions.

2.4 Findings

Section 2.1 shows that maintenance planning rarely treats a cyclic problem. The one example found (van Zante-de Fokkert et al. 2007) uses a periodicity of four weeks, which is too long for our purposes, although the ratio of maintenance time to scheduling period is similar. Also, we want the window patterns to be constructed automatically rather than given as predefined input variants. As discussed in Sect. 2.2, cyclic train scheduling has been extensively studied, but the periodicity for these problems are usually shorter (1 h) than what we want (1 day). Only one example has been found (Van Aken et al. 2017a, b) that considers maintenance and treats a cyclic problem, but that model is not suitable in our case since the maintenance closures are longer than the scheduling period. The results in Caprara et al. (2007) regarding the added complexity when introducing period-deciding variables has motivated a modeling approach that avoids such variables—as described in Sect. 4. Finally, from previous computational results (Lidén et al. 2018), we got an indication of the solution times that can be expected in our case study.

3 Problem description

The planning problem we consider applies to organisations that are responsible for coordinating railway traffic and network maintenance, such as infrastructure managers, transport administrators or railway companies that own and operate the infrastructure network. The planning horizons of train services and maintenance tasks can differ substantially, which—depending on the planning procedure—may favor early applicants and leave costly or even insufficient track access possibilities for other actors. This situation has been observed in Sweden where the increase in rail traffic together with the previous planning regime has forced maintenance to be performed at odd times and/or in shorter time slots which lead to inefficiency and cost increases for the maintenance contractors, potentially even reduced track quality, resulting in higher governmental spending.

To increase the possibility of suitable work possessions, a new planning regime is being introduced, where Trafikverket (the Swedish Transport Administration) proposes regular, 2–6 h train-free maintenance windows before the timetable is constructed. The maintenance windows are given as a prerequisite for: (a) the procurement of multi-year maintenance contracts, and (b) the yearly timetable process, which give stable quotation and planning conditions for the contractors. The overall aim is to increase efficiency, reduce the cost and planning burden as well as to improve robustness and punctuality. However, since maintenance windows will reduce the train scheduling possibilities, the window patterns should be designed such that maintenance activities and train operations are coordinated in a well-balanced manner, which is non-trivial.

We address the coordination of maintenance windows and train traffic as an optimization problem for a railway network. The purpose of the optimization model is to find a pattern of maintenance windows that allows a desired set of train services to be run, and that minimizes the total cost for train operations and maintenance. In this study, the train operating cost is measured by total running time and deviation from preferred departure, while the maintenance cost consists of direct work time, indirect setup/overhead time and crew usage costs. Scheduling of trains must respect given minimum travel and dwelling durations as well as line capacity limitations imposed by other trains and the maintenance scheduling. The maintenance window schedule shall fulfill given work volumes, window requirements, as well as maintenance resource considerations.

A macroscopic infrastructure model is used, which allows for networks of arbitrary size and granularity. Nodes are placed where train services start, end or may change route, as well as between different maintenance areas. The nodes are connected by links which correspond to single-, double- or multi-track lines that may contain intermediate stations (meet/pass loops). Traffic capacity restrictions are modeled as limitations on the number of trains that can be scheduled over each link per time period, both in each direction and as a sum of both directions. The traffic capacity is reduced when maintenance windows are scheduled on the links. Single-track links will be completely blocked, while double-track links will be reduced to single-track capacity, by a maintenance window. This modeling approach gives a basic traffic capacity control, but not the full meet-pass planning necessary for a conflict-free timetable.

In summary, the problem properties are as follows: First of all, both maintenance windows and train services are to be scheduled over a railway network for a period of one or more days. For the case study, we will consider a railway line and a one-day schedule which is cyclically repeated with a 24 h period in order to get a production plan for a standard work day. The tasks are typically from one to several hours long, where the train services have continuous (real-valued) start/end times and durations over the links, and the maintenance windows use discrete (1 h) event times, but may have a varying length. A moderately large scheduling flexibility (± 2 h) will be used for the train services, while maintenance windows can be freely scheduled in the planning period. Maintenance crew resources will be considered in some experiments, but only their spatial availability is given beforehand. The temporal crew scheduling and location sequencing shall respect a limitation of available crew resources, as well as work time regulations regarding maximum work hours per working day and minimum rest time between working days. Thus, the crew schedules that result from the resource assignments may start any time on the day, but each working shift will not be longer than the work day limitation (typically 8–10 h), and the non-working time between two shifts will be at least as long as the minimum required rest time (typically 10–12 h). These constraints will be enforced cyclically. A delimitation is that different crew types and varying travel time between work locations are not considered.

4 Mathematical model

In this section we describe the optimization model that has been used. It is essentially the same model as presented in Lidén et al. (2018), but with minor adjustments in the objective function and some crucial changes in the constraints, which enable a cyclic schedule and varying maintenance window size. We label the model CISM_R as an abbreviation for the Cyclic Integrated train Service and railway Maintenance planning problem with Resource considerations. In Sect. 4.1 we give a simplified formulation of CISM_R, so as to show the overall structure, while Sect. 4.2 describes the model adjustments for cyclic schedules and varying maintenance window size. The complete model with all details is given in Appendix A. Finally, Sect. 4.3 discusses mathematical and practical properties of the model.

4.1 Model structure

The railway network is modeled by a link set L . The scheduling problem has a planning horizon of length H , divided into a sequence $T = \{0, \dots, H-1\}$ of unit size time periods $t \in T$, each covering real-valued event times between t and $t+1$. Start index 0 is used here since we will treat all time indices as modulo H to get cyclic constraints.

For the train traffic we have a set S of train services. Each train service $s \in S$ has a set R_s of possible routes, defined as a sequence of links, which gives the set L_s of all possible links that train service s can traverse. The scheduling of trains shall be done by selecting one route $r \in R_s$ and deciding entry and exit times for each link in that route, with a preferred departure time τ_s from the origin. (Although the optimization model can select among different train routes (including cancellations) for each train service, the case study will not explore these possibilities and will only use one route per train service. We still describe the general model, so as to be consistent with previous publications.)

For the maintenance, a subset $L^M \subseteq L$ of the links shall have maintenance windows. Each link $l \in L^M$ shall have a volume of at least V_l time periods covered by windows and the scheduling shall be done according to a set W_l of window options. Each option $o \in W_l$ is defined by a tuple $o = (\theta_o^{\min}, \theta_o^{\max})$ that gives the shortest θ_o^{\min} and longest θ_o^{\max} allowed window length. As an example, if the maintenance volume V_l is 4 h on link l and the window options $W_l = \{(4, 4), (1, 3)\}$, then the first option ($\theta_1^{\min} = \theta_1^{\max} = 4$) means that one 4 h long window will be scheduled, while the other option ($\theta_2^{\min} = 1, \theta_2^{\max} = 3$) will give one of the following window combinations: 1+3, 2+2, 1+1+2 or 1+1+1+1 h.

For the resource considerations, we have a set K of maintenance crews and subsets $K_l \subseteq K$ stating which crews can maintain link l . The crews are partitioned into crew bases, such that all crews belonging to a base can maintain the same set of links. The resource constraints to handle are: (a) a limited number of maintenance crews, with (b) work time limitations enforcing a maximum length of the work day and a minimum rest time between the work days.

The model has three groups of variables, which concern the scheduling of train services, maintenance windows and maintenance crews, respectively. The main variables for scheduling train services are:

z_{sr}	Route choice: whether train service s uses route r or not
e_{sl}^+, e_{sl}^-	Event time: entry(+)/exit(−) time for train service s on link l
e_s^O, e_s^D	Event time at the origin (O) and destination (D) for train service s
x_{slt}^+, x_{slt}^-	Link entry/exit: whether train service s enters/exits link l in time period t or not
u_{slt}	Link usage: whether train service s uses link l in time period t or not
n_{lt}^h	Number of train services traversing link l in direction h during time period t

The variables for scheduling maintenance windows are:

w_{lo}	Maintenance window option choice: whether link l is maintained with window option o or not
y_{lt}	Maintenance work: whether link l is maintained in time period t or not
v_{lot}	Work start: whether maintenance on link l according to window option o is started in time period t or not

The main variables for scheduling maintenance crews are:

q_k	Crew usage: whether crew k is used/assigned to any work or not
\check{y}_{kt}	Crew availability: whether crew k is available for maintenance work in time period t or not
\check{y}_{kt}	Start of work day: whether time period t is the first in a working day for crew k or not
d_{lkt}	Crew assignment: whether maintenance on link l in time period t is done by crew k or not
q_{kl}^L	Link assignments: whether crew k is assigned to work on link l or not

The complete formulation of CISMIR can now be summarized as follows:

$$\text{minimize } c(\mathbf{z}, \mathbf{e}, \mathbf{y}, \mathbf{v}, \mathbf{q}, \check{\mathbf{y}}) \quad (1)$$

$$\text{subject to } \mathbf{A}(\mathbf{z}, \mathbf{e}, \mathbf{x}, \mathbf{u})^{\text{route}} \quad (2)$$

$$\mathbf{A}(\mathbf{z}, \mathbf{e})^{\text{trains}} \quad (3)$$

$$\mathbf{A}(\mathbf{w}, \mathbf{y}, \mathbf{v})^{\text{maintenance}} \quad (4)$$

$$\mathbf{A}(\mathbf{u}, \mathbf{n}, \mathbf{y})^{\text{capacity}} \quad (5)$$

$$\mathbf{A}(\mathbf{y}, \mathbf{d}, \check{\mathbf{y}})^{resource} \quad (6)$$

$$\begin{aligned} &\mathbf{A}(\mathbf{q}, \check{\mathbf{y}}, \check{\mathbf{v}})^{work\ day} \\ &\mathbf{z}, \mathbf{x}, \mathbf{u}, \mathbf{w}, \mathbf{y}, \mathbf{v}, \mathbf{q}, \check{\mathbf{y}}, \check{\mathbf{v}}, \mathbf{d} \text{ binary} \\ &\mathbf{e}, \mathbf{n} \text{ non-negative} \end{aligned} \quad (7)$$

where $c(\cdot)$ is a linear objective function and $\mathbf{A}(\cdot)$ are linear constraint functions over the indicated variables. The event variables \mathbf{e} are continuous (real-valued) and give the detailed scheduling times. The counting variables \mathbf{n} are also continuous but will take integer values since they are summations over the binary usage variables \mathbf{u} .

The objective (1) is a linear combination of the train, maintenance and crew scheduling variables. In this paper, we use the objective function

$$c(\cdot) = c_1(\mathbf{e}) + c_2(\mathbf{y}, \mathbf{v}) + c_3(\mathbf{q}, \check{\mathbf{y}})$$

where

$$\begin{aligned} c_1(\mathbf{e}) &= \sum_{s \in S} \left[\sigma_s^{time} (e_s^D - e_s^O) + \sigma_s^{dev} |e_s^O - \tau_s| \right], \\ c_2(\mathbf{y}, \mathbf{v}) &= \sum_{l \in L} \left[\sum_{t \in T} \lambda_{lt}^{time} y_{lt} + \sum_{o \in W_l} \sum_{t \in T} \lambda_{lot}^{start} v_{lot} \right], \text{ and} \\ c_3(\mathbf{q}, \check{\mathbf{y}}) &= \lambda^{use} \sum_{k \in K} q_k + \sum_{k \in K} \sum_{t \in T} \lambda_t^{avail} \check{y}_{kt} + \lambda^{link} \sum_{k \in K} \sum_{l \in L_k^K} q_{kl}^L \end{aligned}$$

measure the train cost (c_1), window cost (c_2) and crew cost (c_3), respectively. The train cost uses cost factors σ_s^{time} for total running time between origin and destination, and σ_s^{dev} for deviation from the preferred departure time. The window cost uses cost factors λ_{lt}^{time} for work time, and λ_{lot}^{start} for setup/overhead time. The crew cost uses cost factors λ^{use} for the number of crews used, λ_t^{avail} for crew availability time, and λ^{link} for the number of links scheduled per crew. The values for these cost factors are given in Table 2 and Sect. 5.4.

The constraints enforce: (2) correct (feasible) bounds on the train events and linking of entry/exit and usage variables according to the selected route, (3) sufficient travel durations and dwell times along the chosen route, (4) sufficient maintenance windows scheduled according to the chosen option, (5) that the available network capacity is respected, (6) assignment of crew resources, and (7) that work time regulations are respected (max work day length and min rest time between work days).

The full details of the model is presented in Appendix A.

4.2 Adjustments for cyclic schedules and variable window size

We now describe the model adjustments (as compared to Lidén et al. 2018) that enable a cyclic schedule and variable window sizes. For the train scheduling, the event times must be within a band given by the preferred departure time τ_s and a

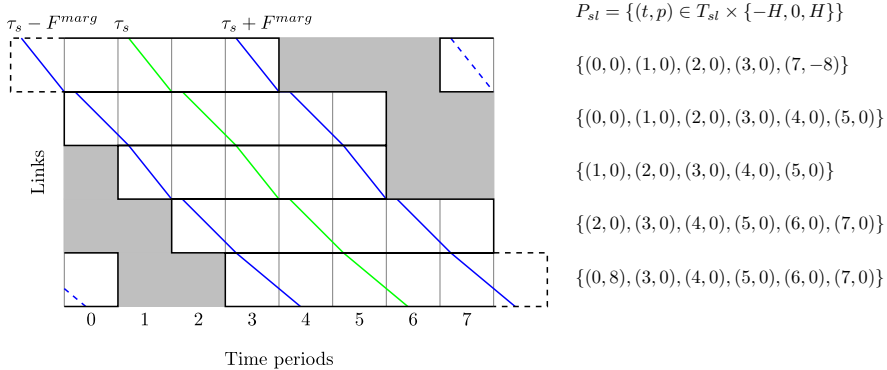


Fig. 1 Illustration of T_{sl} and P_{sl} for a case with 5 links, $H = 8$ and $F^{marg} = 2$. Tilted lines indicate the earliest (blue), preferred (green) and latest (blue) possible train schedule, while the white boxes indicate the affected time periods. Dashed lines show the wrap around for the train scheduling and time periods, respectively

scheduling margin $\pm F^{marg}$. Each train scheduling band corresponds to a set T_{sl} of allowed time periods for train service s on link l , which may wrap around the start or end of the scheduling horizon. Event times are kept consistent with the time periods T_{sl} by compensation factors p defined by $P_{sl} = \{(t, p) \in T_{sl} \times \{-H, 0, H\}\}$. This formulation allows cyclic schedules without any additional variables, under the restriction that all train scheduling bands fit within the range $[-H, H]$, which limits the selection of F^{marg} . Figure 1 illustrates this for a train service, where there is a wrap around for the first and last link. The tilted lines show the earliest possible (blue), preferred (green), and latest possible (blue) train schedule, while the white boxes show the corresponding time periods where the link usage will be counted. Greyed out boxes indicate forbidden time periods for this train service. As an example, the topmost link has compensation factor $p = 0$ for time periods $0, \dots, 3$ while time period 7 has $p = -8$, which means that trains with event times $e < 0$ will affect the capacity count for period 7.

The compensation factors P_{sl} are used in the bound constraints for the event variables as follows:

$$\sum_{(t,p) \in P_{sl}} (t+p)x_{slt}^+ \leq e_{sl}^+ \leq \sum_{(t,p) \in P_{sl}} (t+1+p-\xi)x_{slt}^+ \quad \forall s \in S, l \in L_s$$

$$\sum_{(t,p) \in P_{sl}} (t+p+\xi)x_{slt}^- \leq e_{sl}^- \leq \sum_{(t,p) \in P_{sl}} (t+1+p)x_{slt}^- \quad \forall s \in S, l \in L_s$$

where ξ is a small positive number, which ensures that $x_{slt}^+ = 1 \iff t \leq e_{sl}^+ < t+1$ and $x_{slt}^- = 1 \iff t < e_{sl}^- \leq t+1$. (The value of ξ has been set to 0.001 which gives a time accuracy of a couple of seconds. It could also be set to a value that corresponds to the signalling margin before and after a train.)

For the window and crew scheduling, all sequence constraints over time period variables will be applied for all time periods, and period indices are treated as

modulo H . Thus, the cyclic model gets slightly more constraints and also more variables participating in the constraints that wrap around the schedule horizon, as compared to a non-cyclic model. Finally, the shortest and longest allowed window lengths (θ_o^{\min} and θ_o^{\max}) are used in the window length and separation constraints as follows:

$$\sum_{o \in W_l} \left[\sum_{t'=t+1-\theta_o^{\min}}^t v_{lot'} \right] \leq y_{lt} \quad \forall l \in L, t \in T$$

$$\sum_{t'=t+1-\theta_o^{\max}}^t v_{lot'} + 1 \geq y_{lt} + w_{lo} \quad \forall l \in L, o \in W_l, t \in T$$

$$\sum_{t'=t+1}^{t+MS_o} v_{lot'} \geq w_{lo} - y_{lt} \quad \forall l \in L, o \in W_l, t \in T$$

where MS_o is the max separation between two windows for window option o .

4.3 Mathematical and practical properties

Our modeling approach combines a detailed representation of each train schedule (using real valued event variables) with an aggregated representation of the capacity usage—both temporally and spatially. Some of the advantages of this approach are that larger problem instances can be handled, and that integrated train and maintenance scheduling is enabled. Also, as seen above, the modeling adjustments are straightforward for cyclic planning. The obvious drawback is that the coarse capacity representation does not give conflict-free schedules and that train running times and adjustments may be erroneously estimated. As a consequence, the minimal travel durations should include suitable margins which will be necessary for train separation, meetings, etc, in the final timetable. In the case study we will use data obtained from real timetables to establish reasonable travel durations over the links (see Appendix B).

Figure 2 shows a non-cyclic example where a set of train services are homogeneously spaced in the left graph, along with the resulting capacity usage for each link and period. In the right graph, a 2 h maintenance window has been introduced (the yellow rectangle) and, as a consequence, several trains need to be rescheduled (marked with blue color). Trains C and D are moved earlier, while E is moved later. Train C will overlap train D in the model, but we show them slightly separated for sake of clarity. If the capacity restriction is 3 trains/h, then train A will also be moved earlier (overlapping its preceding train) due to trains B, C and D consuming all capacity in the vicinity of the maintenance window. As a chain reaction, the star-marked train will also be moved slightly earlier. The capacity-restricted link-period combinations that cause these schedule adjustments are marked with light grey in the figure. In this example, the scheduling adjustments are underestimated for trains B and C, while the adjustment of A and the star-marked train is overestimated. This is a classic discretization error, that will arise in most time-space

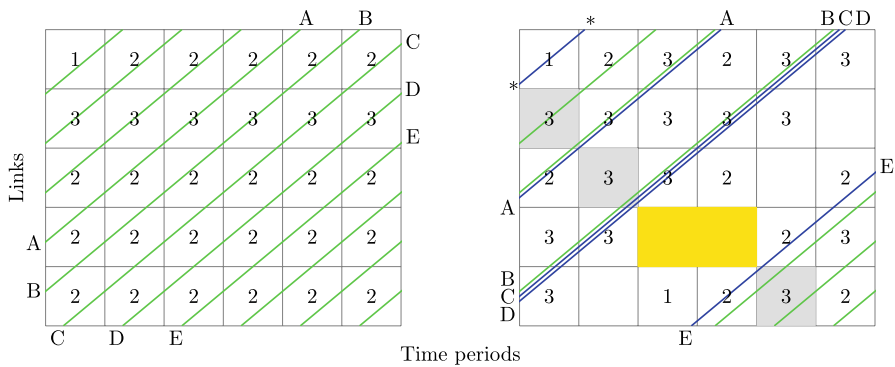


Fig. 2 Train adjustments due to capacity restrictions (≤ 3)—non-cyclic case. Numbers indicate capacity usage, letters indicate trains. Green lines indicate trains with their original schedule and blue lines the rescheduled trains. The yellow box marks the introduced maintenance window while the light grey boxes are the link-period combinations which cause the train adjustments

network representations. Boland et al. (2019) show that the discretization error can be 20% for service network design problems. Our problem is slightly different, but clearly suffers from a similar deficiency. In cases with homogenous or very similar train services it is possible to introduce separation constraints (see Lidén and Joborn, 2017) that achieve more realistic plans, but for cases with mixed traffic and varying train costs it is more cumbersome to resolve this issue. Hence, it is necessary to verify whether feasible timetables are possible to achieve, either by using manual reviews or some technical method such as simulation or a timetabling tool.

The adjustments for achieving a cyclic schedule also have some important implications. It is a great advantage that no extra variables are needed for the cyclic train scheduling.¹ However, the train service scheduling restrictions limit the usage of this model to problems where the total running time of each train is less than, or in the same order of magnitude as the planning period. Thus, it will not fit problems with very short (e.g. hourly) periodic schedules and long train running times.

The cyclic sequence constraints for the maintenance and resource handling has an important mathematical drawback. In the non-cyclic case, these constraints define the convex hull and are, therefore, naturally integral as shown by Queyranne and Wolsey (2017), while in the cyclic case this is not the case—despite that slightly more constraints are used and more variables participate in the constraints that wrap around the schedule boundaries. This is further analyzed in Kalinowski et al. (2018). Thus, an increase in solution time might be expected, since more branching will be necessary in the solver.

In some situations it is possible to relax the problem by only applying cyclic handling for the train scheduling and let the maintenance scheduling remain non-cyclic.

¹ Note that it is the complete schedule that is cyclic. Groups of train services with equal spacing, i.e. having the same schedule repeated with a time shift (e.g. 30 min), where the interval is shorter than the schedule horizon are currently not handled—but could also be included in the model.

This can be done when the cost structure for the maintenance work will push the maintenance windows together. In our case, scheduling maintenance during day time is encouraged since evening and night time is substantially more expensive. Given that the number of expensive time periods are close to the required rest time between work days, the chances of finding cyclically feasible schedules are fairly good even without applying cyclic sequence constraints. We have experimented with this setting and for some instances there is a slight performance gain with this approach—but the solutions are often cyclically infeasible until the optimality gap has been sufficiently reduced. Our standard approach is, therefore, to use cyclic sequence constraints. The small performance differences indicate that the problem complexity is dominated by the train scheduling part and since our focus is on the practical case study, we have not performed any computational experiments on cyclic versus non-cyclic problem instances.

5 Case study

In this case study, the experimental design has been to produce optimized plans for different window options and resource considerations, study the cost sensitivity, evaluate and compare the obtained plans to current timetables, and estimate the benefits of doing integrated rather than sequential planning. The solutions have been reviewed by a reference group from Trafikverket in order to judge how well they could work in practice.

The remainder of this section is organized as follows: The input data and cost settings are described in Sect. 5.1. Basic scheduling results are presented in Sect. 5.2, followed by a cost-sensitivity study in Sect. 5.3 and plans which consider the crew resources in Sect. 5.4. Then an assessment of the produced plans is given in Sect. 5.5 by comparing with the actual outcome of the capacity planning at Trafikverket. In Sect. 5.6 we estimate how large the benefit is when doing an integrated planning of traffic and maintenance. Finally, a summary of the results is given in Sect. 5.7.

5.1 Input data and cost settings

The area of study is the major freight transportation line through the northern part of Sweden as shown in Fig. 3, between Storvik and Boden, where Trafikverket has had difficulty in scheduling maintenance windows as wanted.

The total length is 913 km and most of the line is single-track (780 km). The traffic consists of 82 trains, and is a mixture of long- and medium-distance freight trains, intercity and regional passenger trains. In Appendix B, more details are given regarding the network and the wanted train traffic, together with capacity settings and train schedules obtained when no maintenance windows are planned. Appendix B also contains information about problem sizes and performance statistics for all conducted tests.

In this study, we let the model schedule maintenance windows only on the single-track links. Since safety regulations in Sweden do allow for maintenance to be

Fig. 3 Map of the study area**Table 1** Yearly maintenance volume, Ockelbo–Ljusdal (167 km with 21 stations)

	Volume (h)	Volume per km	Volume per station stretch
Half hour tasks	$0.5 \times 696 = 348$	2.1	17.4
1 h tasks	$1 \times 1376 = 1376$	8.2	68.8
2 h tasks	$2 \times 843 = 1686$	10.1	84.3
Sum	3410	20.4	170.5

The two last columns (obtained by dividing the volume figures with 167 km and 21 stations, respectively) give two different estimates of the maintenance volume when applied to other links

done on one track while trains are run on neighboring tracks (often at a reduced speed), the double-track links are relatively easy to find windows for. Hence we do not schedule any windows on the double track-links, but assume that they can be scheduled after the single-track patterns have been constructed.

The yearly maintenance volumes have been estimated by using budget data from Trafikverket for the stretch Ockelbo to Ljusdal (see Lidén and Joborn, 2016) concerning 30-min, 1- and 2-h tasks as shown in Table 1. We assume that these figures are representative for the whole line and scale the values accordingly to get the yearly volumes for each link. Furthermore, we assume that this work shall be performed during the weekdays of 48 normal working weeks, i.e. 240 working days. Under these assumptions, the necessary work time per working day (for one

Table 2 Cost factors in the objective function

Description	Value
Train running per hour (σ_s^{time})	1
Deviation cost per hour (σ_s^{dev})	
Regional passenger trains	1
Intercity passenger trains, prioritized freight trains	0.5
Regular freight trains	0.1
Work cost per hour and team (λ_{li}^{time})	0.5
Setup cost per maintenance window (λ_{lot}^{start})	0.5
Compensation for late work hours	
Evening, 18–22	+ 25%
Night, 22–06	+ 60%

maintenance team) ranges between 2 and 6 h for the different links, of which half is required for the 2 h tasks. Other work distributions can be considered, such as moving some work to weekends and concentrated maintenance weeks, but here we focus on finding solutions where the bulk of work is performed on daily maintenance windows.

For the cost settings, we use a monetary scale where 1 equals the hourly cost for running a train. All the cost factors, as given in Table 2, have been decided in cooperation with Trafikverket. According to the guidelines for cost-benefit appraisals used in Sweden ASEK (2018, chapter 13 and 14), the time-dependent train running cost varies between 1800 (intercity passenger trains) and 4700 (overnight sleeper trains) SEK/h, with most freight trains at 2700 SEK/h. We make no distinction between train types and use 2700 SEK/h as our reference value. For the maintenance windows, we use 0.5 as our standard hourly cost rate which corresponds with current contractual levels in Sweden. Compensation factors for evening and night work along with the setup time of 1h for each maintenance window have been obtained from maintenance contractors in Sweden.

5.2 Basic scheduling results

Our first option is to schedule one 2 h long maintenance window on all links, which is sufficient for all 2 h tasks, while assuming that shorter tasks can be fitted into slots where no trains are scheduled. (According to our experiments, this is indeed the case, since such open slots will appear in the shadow of windows scheduled at other links.) We label this option L (for low window volume) and note that it only requires one maintenance team per link and day.

If all tasks are to be done within scheduled maintenance windows, we can either stay with 2 h long windows and use 1–3 maintenance teams in parallel, or we can schedule 2–4 h long maintenance windows with only one team per window (again assuming that some open slots will become available for the shorter tasks). In our experiments, the alternative with one 2 h window and parallel teams have been modeled by changing the cost factors for the affected links, but the resulting schedules

are very similar to the L-option and we do not discuss them further here. The alternative with windows of length 2–4 h is labelled H (for high window volume).

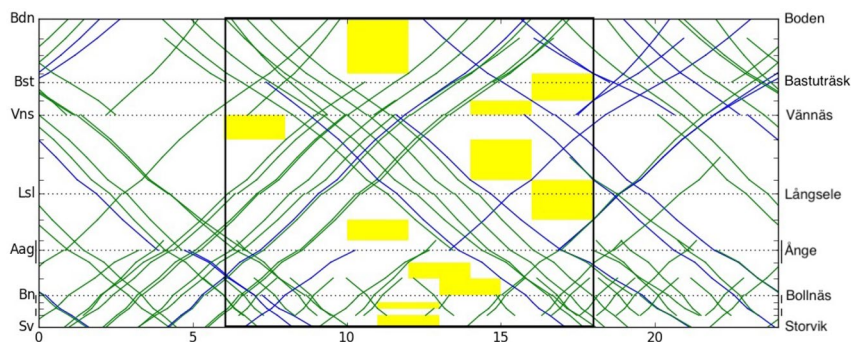
The minimum train running cost (while respecting the capacity limits) is 448.4 (see Appendix B). For case L (with one 2h window per link) the minimum window cost is $16 \times (2 + 1) \times 0.5 = 24$ (16 links with 2h windows and 1h setup at cost 0.5), while case H (with 2–4h windows) requires 49 h of window time, which gives a minimum window cost of $(49 + 16) \times 0.5 = 32.5$. In the latter case we allow solutions with split windows, such as 1 + 3 and 2 + 2, but with extra cost for additional setup times. Thus, it is preferred to schedule one single window per link. For case L we only allow one single window per link.

We note that the total train cost is an order of magnitude larger than the window costs. However, this does not necessarily mean that the trains will have precedence over the maintenance windows in the scheduling problem. From Table 2, we see that the cost increase for placing windows on late working hours or splitting windows (which require more setup) is larger than adjusting the departure times for regular freight trains. Hence, the freight trains are more flexible and, as a consequence, we get solutions where: (1) primarily the freight trains get adjusted departure times, (2) most maintenance windows are placed on regular working hours, and (3) few trains get prolonged running times. At first sight this might seem non-intuitive, but following the global optimality view the model takes and shows the importance of how the cost structure is chosen.

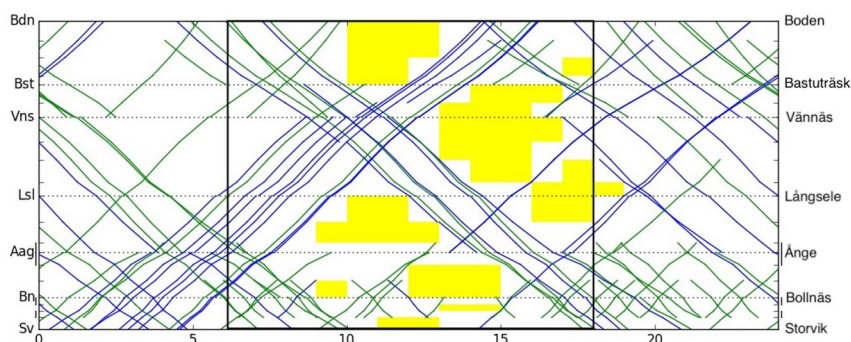
The obtained solutions are plotted in Fig. 4a, b, with time on the horizontal axis and distance on the vertical axis. The normal working hours between 6 and 18 have been marked, and we see that case L has all windows scheduled within this time frame, while case H has only one window outside it. The reference group from Trafikverket assesses the train scheduling as reasonable for case L, while case H looks too cramped. It is doubtful if a feasible timetable can be found (with reasonable time supplements), since both a north- and south-bound group of trains need to be fitted into 2h wide corridors. Case H has, however, been stopped after 8 h of computation, when an optimality gap of 0.78% had been achieved; so there might be better solutions.

In Table 3 the cost results are summarized. The upper half tabulates the train costs along with runtime additions (row 2), departure deviations (row 3) and departure statistics (row 4). The lower half tabulates the maintenance window costs along with cost additions due to late work hours (row 6) and extra setup/overhead time (row 7). These results are directly taken from the different parts of the objective function (see Sect. 4.1). The cost additions are given as “volume \Rightarrow cost”. As an example, case L has a runtime addition of 0.2 h which gives a cost addition of 0.2, and a total amount of departure deviations of 13.9 h which gives a cost addition of 2.3. The deviation statistics are given as “#trains: mean \pm stdev [max]”, measured in minutes. (The mean and standard deviation results have been calculated for the adjusted trains only.) Even without any maintenance windows (first column), 10 trains get their departure adjusted with an average of 17 min, standard deviation of 16 min and max value of 46 min.

Firstly, we observe that the window costs are at their minimum for case L and almost at minimum for case H (windows have been split at four links). The train



(a) Basic solution, case L. (Solution found after 40 min, proven optimal with gap 0.1% after 2 hours.)



(b) Basic solution, case H. (Solution found after 5.5 hours, stopped after 8 hours with optimality gap 0.78%. The train schedule is cramped.)

Fig. 4 Basic solutions

Table 3 Cost comparison, basic solutions

	No windows	Case L	Case H
<i>Train cost</i>	448.4	450.5 (+2.1)	454.2 (+5.8)
Run time addition	0	0.2 h \Rightarrow 0.2	0.3 h \Rightarrow 0.3
Departure deviation	2.9 h \Rightarrow 0.4	13.9 h \Rightarrow 2.3	23.4 \Rightarrow 5.8
Deviation statistics	10: 17 \pm 16 [46]	24: 34 \pm 31 [102]	31: 46 \pm 32 [112]
<i>Window cost</i>		24	34.6 (+2.1)
Late work hour cost		0	1 h \Rightarrow 0.1
Additional setup cost		0	4 times \Rightarrow 2

Cost additions are given as volume \Rightarrow cost. Deviation statistics are given as #trains: mean \pm stdev [max] minutes

costs have a slight increase, mostly due to an increase of departure deviations, but since they mostly affect the freight trains, the cost effect is small. The relative cost increase (compared to the case with no windows) is $2.1/448.4 = 0.5\%$ and $5.8/448.4 = 1.3\%$ for case L and H, respectively. For case L, 24 trains get their

Table 4 Cost sensitivity, train cost scale factor: case L

Train cost scale factor, sf	$1 \leq sf \leq 1.3$	$1.4 \leq sf \leq 2.0$
<i>Train cost</i>	450.5 (+2.1)	449.7 (+1.3)
Run time addition	0.2 h \Rightarrow 0.2	0.2 h \Rightarrow 0.2
Departure deviation	13.9 h \Rightarrow 2.3	8.7 h \Rightarrow 1.5
Deviation statistics	24: 34 \pm 31 [102]	21: 25 \pm 25 [102]
<i>Window cost</i>	24	24.9 (+0.9)
Late work hour cost	0	0.9
Additional setup cost	0	0

departure adjusted with 34 ± 31 [102] (mean \pm stdev [max]) minutes and for case H, 31 trains are adjusted with 46 ± 32 [112] min.

From these results we conclude that: (1) it is possible to schedule 2-h windows at all single-track links by adjusting the trains moderately and with a 0.5% cost increase for the trains, (2) windows longer than 2 h increase the train adjustment costs and give a cramped schedule for which it is unlikely that a feasible timetable can be found.

5.3 Cost sensitivity

We now analyze how sensitive the solutions are to changes in the cost settings. This is of general interest, since the uncertainties in cost data, train running times and maintenance prices can be substantial. First, we note that for case L the window costs are already minimal, and it will, therefore, have no effect to increase the maintenance costs factors for this case. For case H an increase in maintenance cost factors will reduce the maintenance costs and increase the train costs—which has been confirmed experimentally. However, the train costs are already high and, therefore, such solutions are of little interest, so we do not present them here. We only note that even if the maintenance cost factors are doubled, there are still split windows in the solution and the train costs are increasing a bit more than what is gained for the window costs.

Instead, we study the effect of increasing the train cost factors (which is the same as reducing the maintenance cost factors). We perform a series of experiments where all train cost factors are increased with a common factor and note when the structure of the window solution changes. The result is presented in Table 4, which shows that the solution stays the same, even with a 30% increase in train cost factors. At 40% some windows are scheduled at the early morning and late evening, but that solution remains even if the cost factors are doubled. Hence, we conclude that the scheduling solutions are stable to train cost differences even as large as 30% for this case. However, it can be anticipated that the details of the train schedule may change considerably if the cost factors for different train types are modified, or if the number of trains is changed. Such a detailed analysis has not been done yet.

Table 5 Maintenance bases

Station stretches	# Maintenance teams
Boden–Älvsbyn–Storblåliden	2
Storblåliden–Jörn–Bastuträsk–Hällnäs	3
Hällnäs–Vännäs–Trehörningssjö	2
Trehörningssjö–Mellansel–Aspeå–Långsele	3
Långsele–Roback–Bräcke	2
Ramsjö–Ljusdal–Bollnäs	2
Kilafors–Holmsveden, Ockelbo–Storvik	2

5.4 Plans considering maintenance resources

All the solutions obtained so far have several concurrent maintenance windows, which require more crew resources for the maintenance contractors. As an example, the basic solution for case L shown in Fig. 4a requires 14 maintenance teams. Therefore, we now consider limitations and costs for maintenance resources and try to find solutions that are more resource-efficient for the contractors. The network has been divided into seven crew bases as described in Table 5. The number of maintenance teams is equal to the number of links (so that a feasible solution always exists), but each team that is utilized, incurs a cost of 1, which is equivalent to a 25% overhead cost for an 8 h working day (at work time cost 0.5/h). The crew cost counts from the start of the first window to the end of the last, even if the team is not assigned to maintenance windows during their whole work day. Half of the setup times incur a cost in order to avoid double counting preparations that can be done with an unassigned time that are within the working day limit. The maximum working hours are set to 10 h and the minimum rest time between working days is set to 12 h. Apart from this, the same maintenance cost factors are used as previously, and the same compensation factors are applied for scheduling crew resources during early mornings, late evenings and night time.

Since the resource considerations tend to spread out the maintenance windows, we anticipate that the train scheduling will need larger adjustments. To this end, we extend the train scheduling window, but not too much, since the problem size (and solution time) grows quickly with such an increase. After some calibration experiments we have chosen $(\lfloor \tau_s - 3 \rfloor, \lceil \tau_s + 3 \rceil)$, i.e. ± 3 h, for the train scheduling window in these computations.

The problem complexity and solution times increase considerably when the resource scheduling is included. In our experiments we have, therefore, used a relative optimality gap of 0.5%, but still the computations need to run for several days before solutions with reasonable gaps are obtained. Even then the wanted optimality gap is usually not reached. Providing a reasonably good initial solution greatly helps, and we have either used the basic solution obtained from case L or a good resource solution from previous runs. In some cases it helps to use the semi-cyclic approach, where cyclic constraints are only applied to the train scheduling.

The resource considerations will further emphasize the maintenance part and we expect to see even larger train cost increases (and possibly unreasonable train schedules). To counter this effect a limitation on the maximum train cost has been included, which enables us to find solutions that give different balances between train and maintenance cost. In effect, this is the same as treating the problem as a bi-objective optimization and finding the Pareto-optimal solutions. In order to explore the Pareto front, other efficient multi-objective optimization methods should be considered, like the ϵ -constraint method, the hybrid method or the elastic constraint method (Ehrgott 2005).

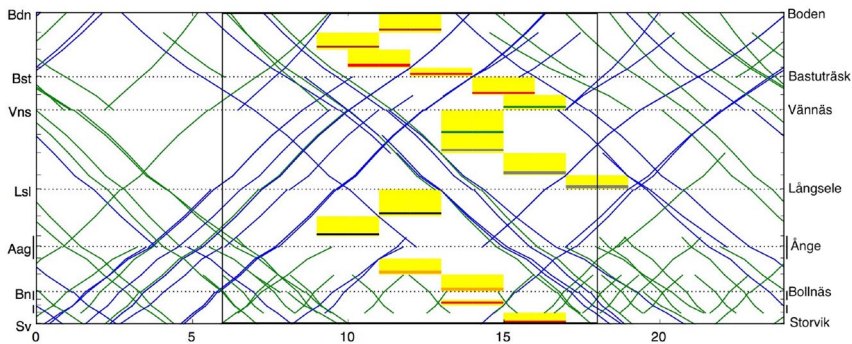
Figure 5a–c show the solutions found with no train cost limit, train cost below 452 and 451, respectively. The same solutions are compared in Table 6 where we also include the basic solutions for case L as comparison. The expected solution structure is clearly seen: With no train cost limit, the model is able to find a solution with only 7 maintenance teams, but here groups of three trains have been scheduled “on top of each other” and then spaced with 1 h to comply with the capacity requirements. This is a cramped solution and even though it might be possible to create a feasible timetable, the train costs are clearly underestimated. The two solutions where the train cost has been limited are better, but need more maintenance teams.

From these results, we conclude that: (1) resource considerations give much better solutions for the contractors, (2) the schedule changes for the trains must be limited in order to avoid cramped and unreasonable plans, and (3) solution times increase substantially when including resource considerations. The latter indicate that heuristic approaches should be investigated.

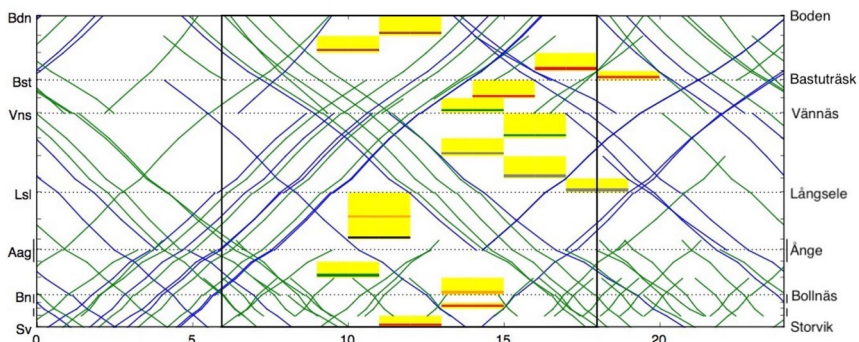
5.5 Comparison with actual plans

In this section, we study the window solution constructed manually by Trafikverket with the ones produced by the optimization model. The window plan from Trafikverket was made for the 2018 timetable period, using the actual train path requests for that year, and is shown in Fig. 6a. Several of the windows only apply for a limited number of operating days, as indicated in the figure. The traffic shown in Fig. 6a is the one used in our case study (based on data from 2017) without any adjustment for the windows. Hence, the figure is not a valid plan, but rather it shows how well the actual window plan matches the traffic patterns that we have been using. This highlights the problem of using historic traffic data for designing window patterns before the actual train path requests are known. Nevertheless, the windows match the traffic relatively well, but we also see some possible open slots in the afternoon that have not been utilized by Trafikverket—which of course can be due to an increase in actual traffic need or other unknown factors. (The lack of windows between Långsele (Lsl) and Ånge (Aag) is because maintenance windows will be introduced later for that part of the network.)

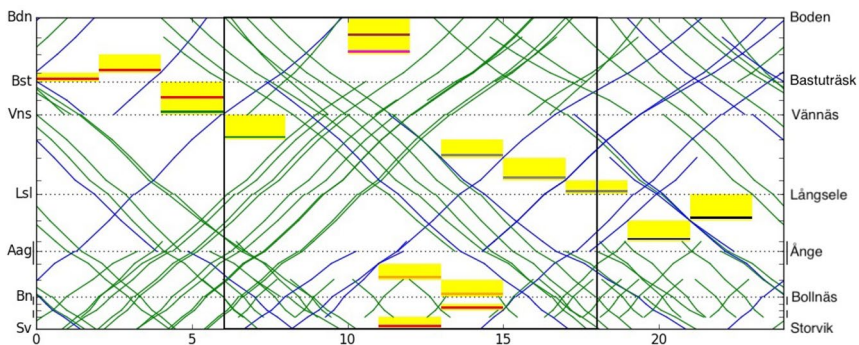
In Fig. 6b we compare the windows planned by Trafikverket with the optimization solution, which considers resources and limits the total train cost to 451. The manual solution has less window time (on a weekly basis) and more night windows, which gives a higher window cost than the optimized solution (also after



(a) No train cost limit. (Solution found after 40 hours, stopped after 140 hours with optimality gap 0.62%. The train schedule is cramped.)



(b) Train cost ≤ 452 . (Solution found after 168 hours, stopped after 200 hours with optimality gap 0.62%. The train schedule is partly cramped.)



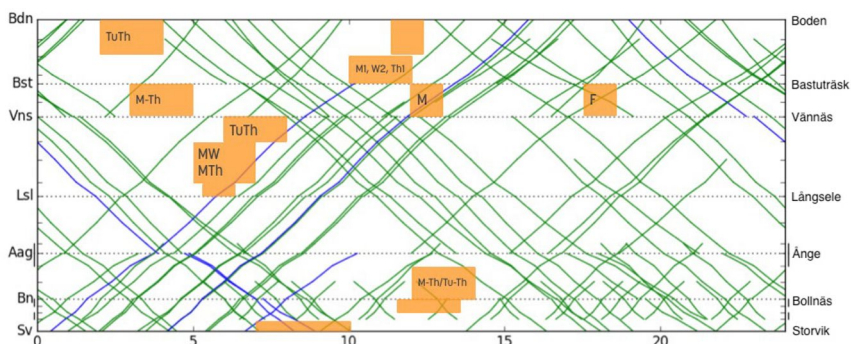
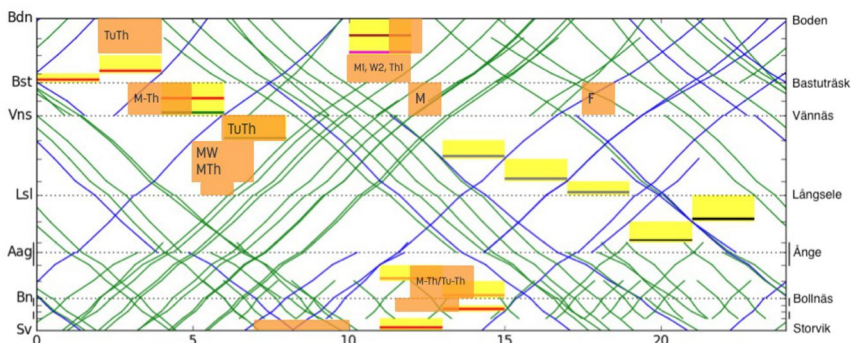
(c) Train cost ≤ 451 . (Solution found after 24 hours, stopped after 164 hours with optimality gap 0.98%.)

Fig. 5 Solutions with resource considerations. Colored bars on the maintenance windows indicate the maintenance teams

being compensated for the volume difference). Although the optimization model has scheduled a larger volume of windows, there are several similarities between the plans. This comparison shows that there is a gain in using the integrated

Table 6 Cost comparison, solutions with resource considerations

	No train cost limit	Train cost ≤ 452	Train cost ≤ 451	Case L
<i>Train cost</i>	454.0 (+5.6)	451.9 (+3.5)	451.0 (+2.6)	450.5 (+2.1)
Run time addition	0.5 h \Rightarrow 0.5	0	0.2 h \Rightarrow 0.2	0.2 h \Rightarrow 0.2
Departure deviation	26.5 h \Rightarrow 5.5	21.5 h \Rightarrow 3.8	15.6 h \Rightarrow 2.8	13.9 h \Rightarrow 2.3
Deviation statistics	36: 44 \pm 33 [140]	29: 44 \pm 38 [158]	26: 36 \pm 32 [124]	24: 34 \pm 31 [102]
<i>Window cost</i>	28.7 (+0.1)	31.0 (+2.4)	33.5 (+4.9)	35.6 (+7.6)
Late work hour cost	1 h \Rightarrow 0.1	3 h \Rightarrow 0.4	13 h \Rightarrow 3.9	0
Crew cost	7 teams \Rightarrow 8.6	9 teams \Rightarrow 10.6	8 teams \Rightarrow 9.6	14 teams \Rightarrow 15.6

**(a)** Window plan for 2018, constructed manually by Trafikverket compared with the traffic input in the case study (based on data from 2017)**(b)** Manually constructed windows for 2018 compared with optimization solution (resource consideration with train cost ≤ 451)**Fig. 6** Actual maintenance window plans

optimization approach (more windows can be scheduled at a lower cost) and that it seems plausible that the obtained solutions can be used as a starting point and guidance when constructing the real window patterns. A quantitative cost

Table 7 Cost comparison, integrated versus sequential planning

	Without resources		With resource considerations	
	Integrated	Trains first	Integrated	Trains first
<i>Train cost</i>	450.5	450.6 (+0.1)	451.0	450.3 (−0.7)
Run time addition	0.2 h \Rightarrow 0.2	1.9 h \Rightarrow 1.9	0.2 h \Rightarrow 0.2	1.5 h \Rightarrow 1.5
Departure deviation	13.9 h \Rightarrow 2.3	1.9 h \Rightarrow 0.7	15.6 h \Rightarrow 2.8	1.8 h \Rightarrow 0.8
Deviation statistics	24: 34 \pm 31 [102]	11: 10 \pm 5 [19]	26: 36 \pm 32 [124]	11: 10 \pm 5 [19]
<i>Window cost</i>	24	26.9 (+2.9)	33.5	40.6 (+7.1)
Extra work cost	0	2.9	3.9	5.0
Crew cost			8 teams \Rightarrow 9.6	14 teams \Rightarrow 15.6

comparison for the train services has not been done, since the plans are based on different traffic conditions.

5.6 Integrated versus sequential planning

We now investigate the effect of using an existing timetable and making very small adjustments to the trains in order to make room for the maintenance windows. This corresponds to the process used by Trafikverket or a sequential planning where the train traffic is constructed first and then minimally adjusted for the necessary maintenance. Such a process will increase the cost for maintenance and we want to investigate how large that effect can be.

In this experiment, we change the cost factors so that all trains have the same deviation cost per hour and equal to the train running cost ($\sigma_s^{dev} = \sigma_s^{time}$). Furthermore, we scale the train costs by factors 2 and 3 for the basic case and when considering resources, respectively, which makes it very costly to change the train scheduling. We also reduce the train scheduling window to be ± 1 h, since small train adjustments will be used. With these settings, more windows are overlapping and placed at costly working hours in the obtained solutions. For the resource solution 14 teams are needed instead of 8. On the other hand the adjustments to the trains are reduced substantially.

In Table 7 the cost effects are shown. For comparison reasons we use the original cost factors, although it should be noted that the rescaling does not guarantee that the train costs are reduced—only the actual scheduling changes—which explains why there is a small train cost addition for the sequential planning case without resources.

For the trains we see that the cost difference is rather small despite the relatively large reduction in adjustments. This is due to an increase in train running times and that departure adjustments are placed also on high priority trains. By contrast, the window costs increase substantially. Without resource considerations the relative maintenance cost increase is $2.9 / 24 = 12\%$, while plans that consider resources get an increase of $7.1 / 33.5 = 21\%$. Thus, we conclude that sequential planning may drive up maintenance costs with 12–21%, with little or no gain in train costs (but of

course less changes to the timetable). Conversely, an integrated planning may give savings for the maintenance costs in the order of 11–17% as compared to sequential planning—without incurring any large cost increases for the train traffic. Note that these figures have been obtained by comparing optimized production plans.

5.7 Summary of results

The primary result is that optimal cyclic one-day plans can be produced for trains, maintenance windows and maintenance resources. For the studied scenario, 2 h long windows can be achieved with small adjustments of the trains (amounting to an increase in train running costs of 0.5%). Longer windows will increase the train costs and give a cramped plan for which it is unlikely that a feasible timetable can be found. The results are stable for cost changes, and the train costs must increase by more than 30% in order to change the structure of the window solutions. By considering the maintenance resources, it is possible to produce window schedules that use less maintenance crews and are much more efficient for the contractors. However, the cost increase for the trains then needs to be constrained in order to get reasonable traffic plans and avoid cramped schedules. A comparison with manually constructed plans from Trafikverket indicates that the optimization model can schedule larger window volumes at a lower cost and that reasonable solution structures are obtained, which can be used as guidance for constructing the real window patterns. Finally, we estimate that using an integrated planning approach (where maintenance windows and trains are jointly planned) instead of a sequential approach (where an existing or new train timetable has precedence over the maintenance windows), will give maintenance cost savings of 11–17%, without incurring any large cost increases for the train traffic.

In our discussions with Trafikverket, the resource solution shown in Fig. 5c appears to be the most promising. It also has a nice structure where it is easy to achieve longer train-free windows (4 or 6 h) by cancelling one or two trains only. Thus, much less train adjustments are needed when longer work tasks are required or when concentrated maintenance weeks are to be scheduled, which is highly appreciated also by the train operators. Our reference persons confirm that the solution structure looks reasonable, but a more detailed assessment is needed—also on other traffic and network situations.

Although we have included some margins in the travel durations to account for daily variations and short stops due to meetings and overtakings, a feasible timetable will need further modifications and additional train stops. Based on existing timetables (see Appendix B) and possible discretization errors in the capacity model (see Sect. 4.3) we estimate that an 8–20% increase in total train running time is necessary in order to obtain a conflict-free timetable from the obtained solutions. Since it is unknown which trains will be affected by these modifications, some of the maintenance windows might also need to be adjusted in the final timetable.

With these experiments we have applied and partly validated the optimization model on a demanding planning problem. Furthermore, several interesting results

have been obtained which we believe are of general interest and apply in similar planning situations—which was the purpose of this case study.

6 Conclusions and future directions

We have addressed the integrated scheduling of rail traffic and network maintenance with a mixed integer optimization model suitable for long-term tactical capacity planning. A cyclic scheduling model, without any period-deciding variables, has been used on a real-life case study concerning an important railway line in Sweden. The results show that it is indeed possible to produce globally optimal schedules for the coordination of trains and maintenance, also including crew resource considerations. However, the plans have an aggregated capacity usage model which means that the detailed meet-pass planning is not included and, therefore, fully conflict-free timetables are not obtained. Nevertheless, the results show that substantial maintenance cost savings can be achieved with only small cost increases for the train traffic. This is a strong argument for developing and using integrated scheduling solutions. Furthermore, timetable construction can use these schedules for finding good initial solutions or for reducing the search space.

An important conclusion from this study is the fundamental impact that the cost structure has on the obtained solutions. It has become obvious that the marginal cost changes when adjusting trains and maintenance windows is more important than the absolute cost levels. Despite that the total train costs have been an order of magnitude larger than the total maintenance cost, the latter have still been scheduled close to their minimum cost—since the total adjustment of trains has been cheaper than moving windows to lowly utilized but costly evening and night hours. A more detailed cost model for the train scheduling, which considers the cost differences for train driving that is done during daytime, evening and night time, would change this balance and make the cost comparison between train services and maintenance windows more fair. Furthermore, it should be noted that we have not considered connection times between turning trains or the rolling stock circulations.

The discretization problem of our aggregated capacity usage model has been discussed, which gives a tendency to group train services “on top of each other” in the solutions. In future research a closer analysis of these aspects would be useful, with the purpose of suggesting how to appropriately set the capacity limits—possibly also studying alternate or adjusted model approaches. Furthermore, long solution times have been observed for several instances, particularly when including resource considerations. Although several hours or days of computation could be acceptable when working with long-term planning problems, the usability would be improved with heuristic methods that give high-quality solutions quicker and enable interactive experiments with end users and capacity planners.

A long single-track line has been extensively studied, but other cases, such as double- or multi-track lines, networks etc., should also be studied under varying traffic patterns. Such work is planned together with Trafikverket, with the intention of paving the way for the possibility of using optimization-based methods as scheduling and decision support in the regular capacity planning process.

Acknowledgements Open access funding provided by Swedish National Road and Transport Research Institute (VTI). This work is performed as part of the research project “Efficient planning of railway infrastructure maintenance”, funded by Trafikverket with the grant TRV 2013/55886 and conducted within the national research program “Capacity in the Railway Traffic System”. Contributions, support and comments from several contact and reference persons are thankfully noted, especially Lars Brunsson, Ulrica Söderström, Kristina Eriksson, Jan Sköld and Cliff Eriksson at Trafikverket and Martin Joborn and Jan Lundgren at Linköping University. We also thank the anonymous reviewers for constructive and helpful comments which greatly helped to improve the quality of the paper. All remaining errors are the responsibility of the author.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendices

A. Detailed model description

In this appendix we give the complete and detailed description of the mathematical model. The structure of the model is illustrated in Fig. 7 as a constraint (or co-occurrence) graph, with vertices for the variables and edges connecting variables that occur in the same constraint. The purpose of the illustration is to show how the different (groups of) variables are connected to each other. The constraints correspond to cliques in the graph as indicated by the equation references in the figure.

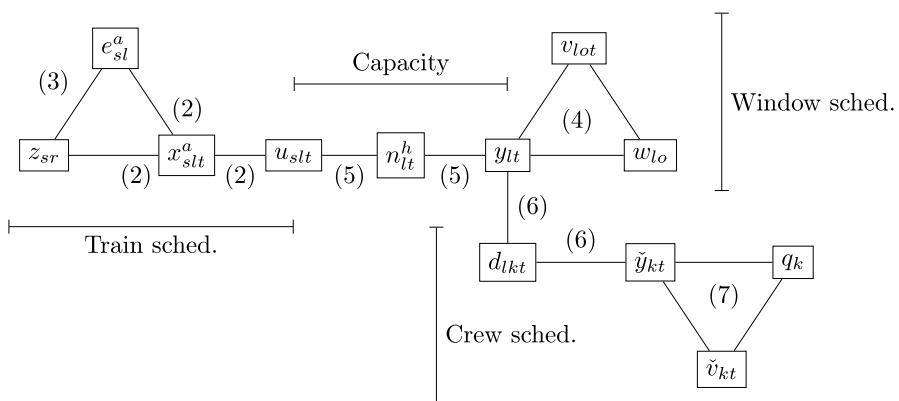


Fig. 7 Variable and constraint graph. The edge labels refer to the constraint numbering

The figure shows that the central variables are u_{slt} (link usage), n_{lt}^h (train counts), y_{lt} (maintenance work) and d_{lkt} (crew assignment).

Sets and parameters

The scheduling shall be done within a planning period defined by

- H the length of the planning horizon
 T a sequence $\{0, \dots, H-1\}$ of unit size time periods t , covering real-valued event times between t and $t+1$.

Railway network

The network is defined by

- N A set of network nodes
 L A set of network links $l = (i, j) : i, j \in N \wedge i \neq j \wedge i < j$
 $C_l^{Nom,d}$ Nominal track capacity (number of trains per time unit) *in each direction* for link $l \in L$
 $C_l^{Red,d}$ Reduced track capacity *in each direction* for link $l \in L$, when maintenance is carried out
 C_l^{Nom} Nominal link capacity (sum for *both directions*) for link $l \in L$
 C_l^{Red} Reduced link capacity (sum for *both directions*) for link $l \in L$

Train services

The scheduling of train services is defined by

- R A set of routes r , each defined as an ordered set of link-direction pairs $(l, h) \in L \times \{0, 1\}$, where $h = 1$ if the route traverses link l in the i, j (forward) direction, and 0 otherwise. One route alternative may be a cancellation—consisting of no links.
 L_r, N_r The links and nodes in route $r \in R$, given by $L_r = \{l \in L : (l, h) \in r\}$ and $N_r = \{i, j \in N : (i, j) \in L_r\}$
 S A set of train services s to be scheduled
 $R_s \subset R$ The set of possible routes r for train service $s \in S$
 L_s, N_s All links and nodes that can be visited by train service s , given by $L_s = \bigcup_{r \in R_s} L_r$; $N_s = \bigcup_{r \in R_s} N_r$
 N_s^O, N_s^D The origin and destination node for train service s
 T_{sl} The allowed time periods for train service s on link l
 P_{sl} Time period compensation factors, defined by $P_{sl} = \{(t, p) \in T_{sl} \times \{-H, 0, H\}\}$
 τ_s The preferred departure time for train service s
 π_{srl} Minimum duration over link $l \in L_r$ when train service s uses route $r \in R_s$

ϵ_{sn}	Minimum duration (dwell time) at node $n \in N_s$ for train service s
σ_s^{time}	Cost per travel time for train service s
σ_s^{dev}	Cost per time unit for train service s , when deviating from the preferred departure time τ_s
σ_{sr}^{route}	Cost for using route r for train service s

Maintenance windows

The scheduling of maintenance windows is defined by

W	A set of maintenance window options
$W_l \subseteq W$	The set of possible maintenance window options for link $l \in L$
V_l	The required maintenance volume (total time) for link $l \in L$
$\theta_o^{min}, \theta_o^{max}$	Minimum and maximum number of time periods for maintenance window option $o \in W$
MS_o	Maximum separation between two maintenance windows for option o
λ_{lt}^{time}	Cost per time unit when performing maintenance work on link l in time period t
λ_{lot}^{start}	Cost for starting maintenance work on link l with window option o in time period t

Crew resources

The availability and assignment of crew resources is defined by

K	A set of maintenance crews
B	The number of crew bases
$K_l \subseteq K$	The maintenance crews that can maintain link $l \in L$
L_k^K	The set of links that can be maintained by crew k , given by $L_k^K := \{l \in L : k \in K_l\}$
$K_b^{base} \subseteq K$	The ordered set of crews belonging to crew base b , with the requirement that no crew can belong to more than one crew base, i.e., $K_b^{base} \cap K_{b+i}^{base} = \emptyset, \forall b = 1, \dots, B-1; i = 1, \dots, B-b$
Ω, Ψ	Maximum length for, and minimum rest after, a crew work day (number of periods)
D_k^{min}	The shortest possible work day length for crew k
Δ_k^{max}	Largest possible separation crew k can have between two work days
λ^{use}	Cost for using a crew
λ^{avail}	Cost per time unit for having a crew available for work
λ^{link}	Cost for assigning one crew to do maintenance on one link

Variables and bounds

The variables and bounds are defined as follows:

z_{sr}	Route choice: = 1 if train service s uses route r , 0 otherwise	$\in \{0, 1\} \forall s \in S, r \in R_s$
e_{sl}^+, e_{sl}^-	Event time: entry(+)/exit(-) time for service s on link l , = 0 if not using link l	$\in [0, H] \forall s \in S, l \in L_s$
e_s^O, e_s^D	Departure/arrival time at the origin/destination for service s	$\in [0, H] \forall s \in S$
f_s	Departure time deviation for train service s , i.e., how much e_s^O deviates from τ_s	$\in [0, H] \forall s \in S$
x_{slt}^+, x_{slt}^-	Link entry/exit: = 1 if train service s enters/exits link l in time period t , 0 otherwise	$\in \{0, 1\} \forall s \in S, l \in L_s, t \in T_s$
u_{slt}	Link usage: = 1 if train service s uses link l in time period t , 0 otherwise	$\in \{0, 1\} \forall s \in S, l \in L_s, t \in T_s$
n_{lt}^h	Number of train services traversing link l in direction h during time period t	$\in [0, S] \forall l \in L, t \in T, h \in \{0, 1\}$
w_{lo}	Maintenance window option choice: = 1 if link l is maintained with window option o , 0 otherwise	$\in \{0, 1\} \forall l \in L, o \in W_l$
y_{lt}	Maintenance work: = 1 if link l is maintained in time period t , 0 otherwise	$\in \{0, 1\} \forall l \in L, t \in T$
v_{lot}	Work start: = 1 if maintenance on link l according to window option o is started in time period t , 0 otherwise	$\in \{0, 1\} \forall l \in L, o \in W_l, t \in T$
q_k	Crew usage: = 1 if crew k is used/assigned to any work, 0 otherwise	$\in \{0, 1\} \forall k \in K$
\check{y}_{kt}	Crew availability: = 1 if crew k is available for maintenance work in time period t , 0 otherwise	$\in \{0, 1\} \forall k \in K, t \in T$
\check{v}_{kt}	Start of work day: = 1 if time period t is the first in a working day for crew k , 0 otherwise	$\in \{0, 1\} \forall k \in K, t \in T$
d_{lkt}	Crew assignment: = 1 if maintenance on link l in time period t is done by crew k , 0 otherwise	$\in \{0, 1\} \forall k \in K, l \in L_k^K, t \in T$
q_{kl}^L	Crew link usage: = 1 if crew k is used/assigned to any work on link l , 0 otherwise	$\in \{0, 1\} \forall k \in K, l \in L_k^K$

Objective function

The objective function, to be minimized, is formulated as follows:

$$\begin{aligned}
 & \sum_{s \in S} \left[\sigma_s^{time} (e_s^D - e_s^O) + \sigma_s^{dev} f_s + \sum_{r \in R_s} \sigma_{sr}^{route} z_{sr} \right] \\
 & + \sum_{l \in L} \left[\sum_{t \in T} \lambda_{lt}^{time} y_{lt} + \sum_{o \in W_l} \sum_{t \in T} \lambda_{lot}^{start} v_{lot} \right] \\
 & + \lambda^{use} \sum_{k \in K} q_k + \lambda^{avail} \sum_{k \in K} \sum_{t \in T} \check{y}_{kt} + \lambda^{link} \sum_{k \in K} \sum_{l \in L_k^K} q_{kl}^L
 \end{aligned} \tag{1}$$

The function consists of three traffic costs (total running time, deviation from the preferred departure time and route choice), two maintenance window costs (work time and setup/overhead time) and three crew costs (number of crews used, crew availability time and number of links scheduled per crew). In the case study the route cost is always zero.

Constraints

The constraints are grouped into categories as described in Sect. 4.1. Time period indices are treated as modulo H .

Train routing, link usage and event bounds

$$\sum_{r \in R_s} z_{sr} = 1 \quad \forall s \in S \quad (2.1)$$

$$\sum_{t \in T_s} x_{slt}^a \geq \sum_{r \in R_l} z_{sr} \quad \begin{array}{l} \forall s \in S, l \in L_s, a \in \{+, -\} \\ R_l = \{r \in R_s : l \in L_r\} \end{array} \quad (2.2)$$

$$u_{slt} = \sum_{t' \in T_s : t' \leq t} x_{slt'}^+ - \sum_{t' \in T_s : t' \leq t-1} x_{slt'}^- \quad \forall s \in S, l \in L_s, t \in T_s \quad (2.3)$$

$$\sum_{(t,p) \in P_{sl}} (t+p)x_{slt}^+ \leq e_{sl}^+ \leq \sum_{(t,p) \in P_{sl}} (t+1+p-\xi)x_{slt}^+ \quad \forall s \in S, l \in L_s \quad (2.4)$$

$$\sum_{(t,p) \in P_{sl}} (t+p+\xi)x_{slt}^- \leq e_{sl}^- \leq \sum_{(t,p) \in P_{sl}} (t+1+p)x_{slt}^- \quad \forall s \in S, l \in L_s \quad (2.5)$$

This group of constraints ensures correctness of: (2.1) route choice, (2.2) linking of entry/exit variables to route choice, (2.3) calculation of usage values, and (2.4, 2.5) bounds for the event variables. Here ξ is a small positive number, which ensures that $x_{slt}^+ = 1 \iff t-1 \leq e_{sl}^+ < t$ and $x_{slt}^- = 1 \iff t-1 < e_{sl}^- \leq t$.

Train scheduling and durations

$$e_s^O = \sum_{l \in OUT_{s,N_s^O}} e_{sl}^+ \quad \forall s \in S \quad (3.1)$$

$$e_s^D = \sum_{l \in IN_{s,N_s^D}} e_{sl}^- \quad \forall s \in S \quad (3.2)$$

$$f_s \geq e_s^O - \tau_s \quad \forall s \in S \quad (3.3)$$

$$f_s \geq \tau_s - e_s^O \quad \forall s \in S \quad (3.4)$$

$$e_{sl}^- - e_{sl}^+ \geq \pi_{sr} z_{sr} \quad \forall s \in S, r \in R_s, l \in L_r \quad (3.5)$$

$$\epsilon_{sn} \sum_{r \in R_s : n \in N_r} z_{sr} \leq \sum_{l \in OUT_{sn}} e_{sl}^+ - \sum_{l \in IN_{sn}} e_{sl}^- \quad \forall s \in S, n \in N_s \setminus \{N_s^O, N_s^D\} \quad (3.6)$$

This group of constraints handles: (3.1, 3.2) calculation of departure and arrival times, (3.3, 3.4) deviation from preferred departure time, (3.5) sufficient travel durations, and (3.6) dwell times along the chosen route. Here IN_{sn} and OUT_{sn} are functions returning all possible incoming and outgoing links for train service s at node n .

Maintenance window scheduling

$$\sum_{o \in W_l} w_{lo} = 1 \quad \forall l \in L \quad (4.1)$$

$$\sum_{t \in T} y_{lt} \geq V_l \quad \forall l \in L \quad (4.2)$$

$$\sum_{o \in W_l} v_{lot} \geq y_{lt} - y_{l,t-1} \quad \forall l \in L, t \in T \quad (4.3)$$

$$v_{lot} \leq w_{lo} \quad \forall l \in L, o \in W_l, t \in T \quad (4.4)$$

$$\sum_{o \in W_l} v_{lot} \leq y_{lt} \quad \forall l \in L, t \in T \quad (4.5)$$

$$\sum_{o \in W_l} v_{lot} \leq 1 - y_{l,t-1} \quad \forall l \in L, t \in T \quad (4.6)$$

$$\sum_{o \in W_l} \left[\sum_{t'=t+1-\theta_o^{\min}}^t v_{lot'} \right] \leq y_{lt} \quad \forall l \in L, t \in T \quad (4.7)$$

$$\sum_{t'=t+1-\theta_o^{\max}}^t v_{lot'} + 1 \geq y_{lt} + w_{lo} \quad \forall l \in L, o \in W_l, t \in T \quad (4.8)$$

$$\sum_{t'=t+1}^{t+MS_o} v_{lot'} \geq w_{lo} - y_{lt} \quad \forall l \in L, o \in W_l, t \in T \quad (4.9)$$

This constraint group concerns: (4.1) choice of maintenance window option; (4.2) work volume; (4.3–4.6) consistency of work start variables v_{lot} , window choice w_{lo} and work variables y_{lt} ; and (4.7–4.9) length of maintenance windows and the separation between them.

Train counting and link capacity

$$n_{lt}^h = \sum_{s \in S_{lh}^{dir} : t \in T_s} u_{slt} \quad \forall l \in L, t \in T, h \in \{0, 1\} \quad (5.1)$$

$$S_{lh}^{dir} = \{s \in S : (l, h) \in r \ \forall r \in R_s\}$$

$$n_{lt}^h \leq (1 - y_{lt})C_l^{Nom,d} + y_{lt}C_l^{Red,d} \quad \forall l \in L, t \in T, h \in \{0, 1\} \quad (5.2)$$

$$n_{lt}^0 + n_{lt}^1 \leq (1 - y_{lt})C_l^{Nom} + y_{lt}C_l^{Red} \quad \forall l \in L, t \in T \quad (5.3)$$

These constraints control: (5.1) train counting for each link and direction, and (5.2, 5.3) capacity usage under normal and maintenance conditions.

Crew resource assignment

$$\sum_{k \in K_l} d_{lkt} = y_{lt} \quad \forall l \in L, t \in T \quad (6.1)$$

$$\sum_{l \in L_k^K} d_{lkt} \leq \check{y}_{kt} \quad \forall k \in K, t \in T \quad (6.2)$$

$$q_k \leq q_{k-1} \quad \forall b = 1, \dots, B; k \in K_b^{base} \setminus \text{first}(K_b^{base}) \quad (6.3)$$

$$d_{lkt} \leq q_{kl}^L \quad \forall k \in K, l \in L_k^K, t \in T \quad (6.4)$$

These constraints control: (6.1, 6.2) assignment of crew resources to each maintenance window period, (6.3) ordering of crew resource assignments, and (6.4) assignment of crews to links.

Work day regulations

$$\check{y}_{kt} \geq \check{y}_{kt} - \check{y}_{k,t-1} \quad \forall k \in K, t \in T \quad (7.1)$$

$$\check{y}_{kt} \leq \check{y}_{kt} \quad \forall k \in K, t \in T \quad (7.2)$$

$$\sum_{t'=t+1-D_k^{\min}}^t \check{y}_{kt'} \leq \check{y}_{kt} \quad \forall k \in K, t \in T \quad (7.3)$$

$$\sum_{t'=t+1-\Omega}^t \check{y}_{kt'} \geq \check{y}_{kt} \quad \forall k \in K, t \in T \quad (7.4)$$

$$\sum_{t'=t+1}^{t+\Psi} \check{v}_{kt'} \leq q_k - \check{y}_{kt} \quad \forall k \in K, t \in T \quad (7.5)$$

$$\sum_{t'=t+1}^{t+\Delta_k^{\max}} \check{v}_{kt'} \geq q_k - \check{y}_{kt} \quad \forall k \in K, t \in T \quad (7.6)$$

These constraints control: (7.1, 7.2) coupling between crew availability and start of work day, (7.3)/(7.4) minimum/maximum work day length, and (7.5)/(7.6) minimum/maximum rest time between work days.

Variable types

$\mathbf{z}, \mathbf{x}, \mathbf{w}, \mathbf{v}, \mathbf{y}$ binary

$\mathbf{e}, \mathbf{f}, \mathbf{n}$ non-negative

$\mathbf{q}, \check{\mathbf{y}}, \check{\mathbf{v}}, \mathbf{d}, \mathbf{q}^L$ binary

B. Instance data and performance statistics

The studied railway network consists of a sequence of lines, mostly single-track, with a total length of 913 km between Storvik and Boden. For this case study, the network has been divided into 20 links (16 single-track) corresponding to the entry/exit stations for the traffic, single/double-track changes and some minor stations that give a reasonable partitioning into maintenance areas. The link lengths range between 16 and 69 km (mean value 47) and the single-track links having between 3 and 10 station-to-station sections (mean value 6); all in all there are 110 stations and meet-pass loops along the complete network stretch.

There are 11 stations where traffic originates/enters or terminates/exits, and the traffic consists of 82 trains in 15 relations. The longest train run requires more than 13 h and the shortest 36 min. The traffic data has been collected from the published timetable for a standard weekday (Thursday) in March 2017. Some short-distance freight trains that are regarded as easy to schedule have been removed but otherwise this is the complete traffic running over the studied line.

The minimum duration times for the trains over each link include margins for daily variations and short additional stops due to meetings and overtakings—as scheduled in the published timetable. Dwell times for critical stops due to passenger and crew exchanges etc have been included, but technical stops longer than 30 min have been removed since the placement of them cannot be done until the train schedule is decided. About half of the trains have such long technical stops (mean value 50 min) which have been removed, and the total sum of all these adjustments amounts to 37 h. After these adjustments the total train running time is 448 h (or an average of 5.5 h per train). Based on these figures, we expect a feasible

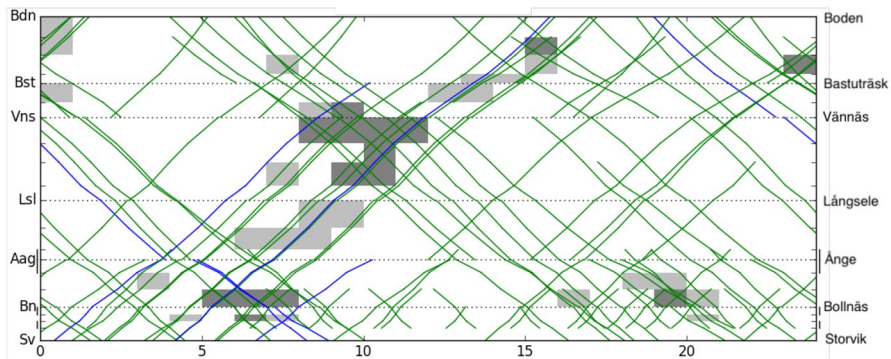


Fig. 8 Scheduling trains only. Active capacity limits marked with grey color (light grey for one direction, dark grey for sum of trains)

and conflict-free timetable to have a train running time that is at least $37/448 = 8\%$ higher than obtained from our optimization model.

All trains have a preferred departure time τ_s (taken from the timetable), and we adopt a normal scheduling window of $(\lfloor \tau_s - 2 \rfloor, \lceil \tau_s + 2 \rceil)$, i.e. ± 2 h (adjusted for some of the test cases). Although it is possible to choose between alternate routes for some trains (e.g. running over the coastal lines), these rerouting possibilities have not been included in the study. Also, we have not included the possibility to cancel any trains.

We impose limitations on the traffic capacity, expressed as the number of train movements per link and hour, both for the number of trains in each direction and for the total number of trains. These limits have been set to, respectively, 4 and 6 trains per link and hour for the single-track links, which means that at most 4 trains can pass over a link each hour in one direction and at most 6 trains in total. This is roughly equivalent to a train separation of 15 min for one-way traffic and 20 min for two-way traffic, which is a slightly conservative setting, given that the longest station-to-station distance is about 10 km, which takes about 7.5 min to pass at 80 km/h.

As a baseline, we first produce a schedule for only the trains. To respect the capacity limits, the optimization model adjusts the departures for about 10 trains with up to 45 min each. The total departure adjustment is a little less than 3 h, which, with the chosen cost factors (see Sect. 5.1), amounts to a cost addition of 0.4—a marginal value compared to the minimum train running cost of 448 (given by the minimum train running time of 448 h and train running cost 1/h). Figure 8 shows the obtained train schedule, where grey boxes indicate time periods where the capacity limits are active (dark grey for the sum of trains and light grey for one direction). Trains with a departure adjustment of more than 15 min are depicted with blue color.

All experiments have been run with Gurobi 7.5.2, limited to use at most 4 threads, on a network server with 12 CPU's, 24 Gb memory and Ubuntu 16.04.3.

Table 8 Problem statistics

Instance	Size	Presolve size	Non zeroes	IP time	Sol time	Gap (%)	Figures	Tables
No windows	13×16	9×12	48	1 min	1 min	0.01	8	3
Case L	16×17	12×13	64	0.6 h	2 h	0.1	4a	3
Case H	16×17	12×13	65	5.5 h	8 h	0.78	4b	3
Res, no train cost limit	21×23	17×20	116	40 h	140 h	0.62	5a	6
Res, train cost ≤ 452	20×23	16×20	99	168 h	200 h	0.62	5b	6
Res, train cost ≤ 451	22×28	18×24	151	24 h	164 h	0.98	5c	6
Trains first, no res	15×12	9×8	32	1 min	1 min	0.1		7
Trains first, with res	18×14	11×9	55	11 min	40 min	0.05		7

Size and presolve size are given as $nrows \times ncols$ (10^3) and non-zeroes are given as (10^3). “IP time” is the time when the integer solution was found, while “Sol time” gives the time when optimality was proven or the run was stopped

Apart from setting the relative MIP gap threshold, only default options have been used. Performance statistics for all the experiments are given in Table 8, which lists problem sizes, solution times and optimality gaps—along with references to solution figures and result tables.

References

- Albrecht AR, Panton DM, Lee DH (2013) Rescheduling rail networks with maintenance disruptions using Problem Space Search. *Comput Oper Res* 40(3):703–712. <https://doi.org/10.1016/j.cor.2010.09.001>
- Arenas D, Pellegrini P, Hanafi S, Rodriguez J (2018) Timetable rearrangement to cope with railway maintenance activities. *Comput Oper Res* 95:123–138. <https://doi.org/10.1016/j.cor.2018.02.018>
- ASEK (2018) Appraisal guidelines. Swedish Transport Administration. <http://www.trafikverket.se/ASEK>
- Boland N, Hewitt M, Marshall L, Savelsbergh M (2019) The price of discretizing time: a study in service network design. *EURO J Transp Logist* 8:195–216. <https://doi.org/10.1007/s13676-018-0119-x>
- Boland N, Kalinowski T, Waterer H, Zheng L (2013) Mixed integer programming based maintenance scheduling for the Hunter Valley coal chain. *J Sched* 16(6):649–659. <https://doi.org/10.1007/s10951-012-0284-y>
- Borraz-Sánchez C, Klabjan D (2012) Strategic gang scheduling for railroad maintenance. CCITT, Center for the Commercialization of Innovative Transportation Technology, Northwestern University. http://ntl.bts.gov/lib/46000/46100/46128/CCITT_Final_Report_Y201.pdf
- Brucker P, Heitmann S, Knust S, (2005) Scheduling railway traffic at a construction site. In: Günther HO, Kim KH (eds) *Container terminals and automated transport systems*. Springer, Berlin Heidelberg, pp 345–356. https://doi.org/10.1007/3-540-26686-0_15
- Budai G, Huisman D, Dekker R (2006) Scheduling preventive railway maintenance activities. *J Oper Res Soc* 57(9):1035–1044. <https://doi.org/10.1057/palgrave.jors.2602085>
- Cacchiani V, Caprara A, Galli L, Kroon L, Maróti G, Toth P (2012) Railway rolling stock planning: robustness against large disruptions. *Transp Sci* 46(2):217–232. <https://doi.org/10.1287/trsc.1110.0388>
- Cacchiani V, Huisman D, Kidd M, Kroon L, Toth P, Veelenturf L, Wagenaar J (2014) An overview of recovery models and algorithms for real-time railway rescheduling. *Transp Res Part B Methodol* 63:15–37. <https://doi.org/10.1016/j.trb.2014.01.009>

- Cacchiani V, Toth P (2012) Nominal and robust train timetabling problems. *Eur J Oper Res* 219(3):727–737. <https://doi.org/10.1016/j.ejor.2011.11.003>
- Cadarso L, Marín Á (2011) Robust rolling stock in rapid transit networks. *Comput Oper Res* 38(8):1131–1142. <https://doi.org/10.1016/j.cor.2010.10.029>
- Camci F (2015) Maintenance scheduling of geographically distributed assets with prognostics information. *Eur J Oper Res* 245(2):506–516. <https://doi.org/10.1016/j.ejor.2015.03.023>
- Caprara A, Kroon L, Monaci M, Peeters M, Toth P (2007) Passenger railway optimization (ch 3). In: *Handbooks in operations research and management science*, vol 14. Elsevier, pp 129–187. [https://doi.org/10.1016/S0927-0507\(06\)14003-7](https://doi.org/10.1016/S0927-0507(06)14003-7)
- Caprara A, Kroon L, Toth P (2011) Optimization problems in passenger railway systems. *Wiley Encycl Oper Res Manag Sci*. <https://doi.org/10.1002/9780470400531.eorms0647>
- Caprara A, Monaci M, Toth P, Guida PL (2006) A Lagrangian heuristic algorithm for a real-world train timetabling problem. *Discrete Appl Math* 154(5):738–753. <https://doi.org/10.1016/j.dam.2005.05.026>
- Corman F, Meng L (2014) A review of online dynamic models and algorithms for railway traffic management. *IEEE Trans Intell Transp Syst* 16(3):1274–1284. <https://doi.org/10.1109/TITS.2014.2358392>
- D'Ariano A, Meng L, Centulio G, Corman F (2019) Integrated stochastic optimization approaches for tactical scheduling of trains and railway infrastructure maintenance. *Comput Ind Eng* 127:1315–1335. <https://doi.org/10.1016/j.cie.2017.12.010>
- den Hertog D, van Zante-de Fokkert JI, Sjamaar SA, Beusmans R (2005) Optimal working zone division for safe track maintenance in The Netherlands. *Accid Anal Prev* 37(5):890–893. <https://doi.org/10.1016/j.aap.2005.04.006>
- Ehrgott M (2005) *Multicriteria optimization*. Springer, Berlin. <https://doi.org/10.1007/3-540-27659-9>
- EIM-EFRTC-CER Working Group, (2012) Market strategies for track maintenance and renewal. Tech. Rep. 2353 7473-11, CER-Community of European Railway and Infrastructure Companies. <http://www.cer.be/publications/brochures-studies-and-reports/report-eim-efrtc-cer-working-group-market-strategies>
- Forsgren M, Aronsson M, Gestrelus S (2013) Maintaining tracks and traffic flow at the same time. *J Rail Transp Plan Manag* 3(3):111–123. <https://doi.org/10.1016/j.jrtpm.2013.11.001>
- Giacco GL, D'Ariano A, Pacciarelli D (2014) Rolling stock rostering optimization under maintenance constraints. *J Intell Transp Syst* 18(1):95–105. <https://doi.org/10.1080/15472450.2013.801712>
- Gustavsson E (2014) Scheduling tamping operations on railway tracks using mixed integer linear programming. *EURO J Transp Logist* 4(1):97–112. <https://doi.org/10.1007/s13676-014-0067-z>
- Higgins A (1998) Scheduling of railway track maintenance activities and crews. *J Oper Res Soc* 49(10):1026–1033. <https://doi.org/10.2307/3010526>
- Kalinowski T, Lidén T, Waterer H (2018) Tight MIP formulations for bounded up/down times with cyclic boundary conditions. *ArXiv e-prints*. <https://ui.adsabs.harvard.edu/abs/2018arXiv180901773K>
- Lai Y-C, Fan D-C, Huang K-L (2015) Optimizing rolling stock assignment and maintenance plan for passenger railway operations. *Comput Ind Eng* 85:284–295. <https://doi.org/10.1016/j.cie.2015.03.016>
- Lidén T (2015) Railway infrastructure maintenance—a survey of planning problems and conducted research. *Transp Res Proc* 10:574–583. <https://doi.org/10.1016/j.trpro.2015.09.011>
- Lidén T (2016) Towards concurrent planning of railway maintenance and train services. Licentiate thesis, Linköping University. <https://doi.org/10.3384/lic.diva-128780>
- Lidén T, Joborn M (2016) Dimensioning windows for railway infrastructure maintenance: cost efficiency versus traffic impact. *J Rail Transp Plan Manag* 6(1):32–47. <https://doi.org/10.1016/j.jrtpm.2016.03.002>
- Lidén T, Joborn M (2017) An optimization model for integrated planning of railway traffic and network maintenance. *Transp Res Part C Emerg Technol* 74:327–347. <https://doi.org/10.1016/j.trc.2016.11.016>
- Lidén T, Kalinowski T, Waterer H (2018) Resource considerations for integrated planning of railway traffic and maintenance windows. *J Rail Transp Plan Manag* 8(1):1–15. <https://doi.org/10.1016/j.jrtpm.2018.02.001>
- Louwerse I, Huisman D (2014) Adjusting a railway timetable in case of partial or complete blockades. *Eur J Oper Res* 235(3):583–593. <https://doi.org/10.1016/j.ejor.2013.12.020>
- Luan X, Miao J, Meng L, Corman F, Lodewijks G (2017) Integrated optimization on train scheduling and preventive maintenance time slots planning. *Transp Res Part C Emerg Technol* 80:329–359. <https://doi.org/10.1016/j.trc.2017.04.010>

- Peng F, Ouyang Y (2012) Track maintenance production team scheduling in railroad networks. *Transp Res Part B Methodol* 46(10):1474–1488. <https://doi.org/10.1016/j.trb.2012.07.004>
- Queyranne M, Wolsey LA (2017) Tight MIP formulations for bounded up/down times and interval-dependent start-ups. *Math Program A* 164(1–2):129–155. <https://doi.org/10.1007/s10107-016-1079-2>
- Ruffing JA (1993) An analysis of the scheduling of work windows for railroad track maintenance gangs. In: *Proceedings of the transportation research forum*, vol 35, pp 307–314. <http://trid.trb.org/view.aspx?id=551772>
- Savelsbergh M, Waterer H, Dall M, Moffiet C (2015) Possession assessment and capacity evaluation of the Central Queensland Coal Network. *EURO J Transp Logist* 4(1):139–173. <https://doi.org/10.1007/s13676-014-0066-0>
- Serafini P, Ukovich W (1989) A mathematical model for periodic scheduling problems. *SIAM J Discrete Math* 2(4):32. <https://doi.org/10.1137/0402049>
- Su Z, Jamshidi A, Núñez A, Baldi S, De Schutter B (2017) Multi-level condition-based maintenance planning for railway infrastructures—a scenario-based chance-constrained approach. *Transp Res Part C Emerg Technol* 84:92–123. <https://doi.org/10.1016/j.trc.2017.08.018>
- Vale C, Ribeiro IM (2014) Railway condition-based maintenance model with stochastic deterioration. *J Civ Eng Manag* 20(5):686–692. <https://doi.org/10.3846/13923730.2013.802711>
- Vale C, Ribeiro IM, Calçada R (2012) Integer programming to optimize tamping in railway tracks as preventive maintenance. *J Transp Eng* 138(1):123–131. [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000296](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000296)
- Van Aken S, Bešinović N, Goverde RM (2017a) Designing alternative railway timetables under infrastructure maintenance possessions. *Transp Res Part B Methodol* 98:224–238. <https://doi.org/10.1016/j.trb.2016.12.019>
- Van Aken S, Bešinović N, Goverde RM (2017b) Solving large-scale train timetable adjustment problems under infrastructure maintenance possessions. *J Rail Transp Plan Manag* 7(3):141–156. <https://doi.org/10.1016/j.jrtpm.2017.06.003>
- van Zante-de Fokkert JJ, den Hertog D, van den Berg FJ, Verhoeven JHM (2007) The Netherlands schedules track maintenance to improve track workers' safety. *Interfaces* 37(2):133–142. <https://doi.org/10.1287/inte.1060.0246>
- Vansteenwegen P, Dewilde T, Burggraeve S, Cattrysse D (2015) An iterative approach for reducing the impact of infrastructure maintenance on the performance of railway systems. *Eur J Oper Res* 252(1):39–53. <https://doi.org/10.1016/j.ejor.2015.12.037>
- Veelenturf LP, Kidd MP, Cacchiani V, Kroon LG, Toth P (2016) A railway timetable rescheduling approach for handling large-scale disruptions. *Transp Sci* 50(3):841–862. <https://doi.org/10.1287/trsc.2015.0618>
- Wen M, Li R, Salling KB (2016) Optimization of preventive condition-based tamping for railway tracks. *Eur J Oper Res* 252(2):455–465. <https://doi.org/10.1016/j.ejor.2016.01.024>
- Zhu Y, Goverde RMP, Quaglietta E (2018) Railway timetable rescheduling for multiple simultaneous disruptions. In: *CASPT 2018: 14th conference on advanced systems in public transport and Transit-Data 2018*. <http://resolver.tudelft.nl/uuid:1664b528-4e12-4745-bf87-a798a7e878ef>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.