

Cyclic Railway Timetabling: a Stochastic Optimization Approach

Leo G. Kroon, Rommert Dekker and Michiel J.C.M. Vromans

ERIM REPORT SERIES <i>RESEARCH IN MANAGEMENT</i>	
ERIM Report Series reference number	ERS-2005-051-LIS
Publication	September 2005
Number of pages	26
Persistent paper URL	
Email address corresponding author	Lkroon@rsm.nl
Address	Erasmus Research Institute of Management (ERIM) RSM Erasmus University / Erasmus School of Economics Erasmus Universiteit Rotterdam P.O.Box 1738 3000 DR Rotterdam, The Netherlands Phone: + 31 10 408 1182 Fax: + 31 10 408 9640 Email: info@erim.eur.nl Internet: www.erim.eur.nl

Bibliographic data and classifications of all the ERIM reports are also available on the ERIM website:
www.erim.eur.nl

REPORT SERIES
RESEARCH IN MANAGEMENT

ABSTRACT AND KEYWORDS	
Abstract	Real-time railway operations are subject to stochastic disturbances. However, a railway timetable is a deterministic plan. Thus a timetable should be designed in such a way that it can absorb the stochastic disturbances as well as possible. To that end, a timetable contains buffer times between trains and supplements in running times and dwell times. This paper first describes a stochastic optimization model that can be used to find an optimal allocation of the running time supplements of a single train on a number of consecutive trips along the same line. The aim of this model is to minimize the average delay of the train. The model is then extended such that it can be used to improve a given cyclic timetable for a number of trains on a common infrastructure. Computational results show that the average delay of the trains can be reduced substantially by applying relatively small modifications to the timetable. In particular, allocating the running time supplements in a different way than what is usual in practice can be useful.
Free Keywords	Railway transportation, Cyclic timetables, Stochastic optimization, Punctuality, Buffer times
Availability	<p>The ERIM Report Series is distributed through the following platforms:</p> <p>Academic Repository at Erasmus University (DEAR), DEAR ERIM Series Portal</p> <p>Social Science Research Network (SSRN), SSRN ERIM Series Webpage</p> <p>Research Papers in Economics (REPEC), REPEC ERIM Series Webpage</p>
Classifications	<p>The electronic versions of the papers in the ERIM report Series contain bibliographic metadata by the following classification systems:</p> <p>Library of Congress Classification, (LCC) LCC Webpage</p> <p>Journal of Economic Literature, (JEL), JEL Webpage</p> <p>ACM Computing Classification System CCS Webpage</p> <p>Inspec Classification scheme (ICS), ICS Webpage</p>

Cyclic Railway Timetabling: a Stochastic Optimization Approach

Leo G. Kroon^{1,2}, Rommert Dekker³, and Michiel J.C.M. Vromans⁴

¹ NS Reizigers, Department of Logistics
3500 HA, Utrecht, The Netherlands

² Rotterdam School of Management

³ Rotterdam School of Economics
Erasmus University Rotterdam
3000 DR, Rotterdam, The Netherlands

⁴ ProRail, Network Planning
3500 GA, Utrecht, The Netherlands

L.Kroon@rsm.nl; R.Dekker@few.eur.nl; Michiel.Vromans@prorail.nl

Abstract. Real-time railway operations are subject to stochastic disturbances. However, a railway timetable is a deterministic plan. Thus a timetable should be designed in such a way that it can absorb the stochastic disturbances as well as possible. To that end, a timetable contains buffer times between trains and supplements in running times and dwell times. This paper first describes a stochastic optimization model that can be used to find an optimal allocation of the running time supplements of a single train on a number of consecutive trips along the same line. The aim of this model is to minimize the average delay of the train. The model is then extended such that it can be used to improve a given cyclic timetable for a number of trains on a common infrastructure. Computational results show that the average delay of the trains can be reduced substantially by applying relatively small modifications to the timetable. In particular, allocating the running time supplements in a different way than what is usual in practice can be useful.

1 Introduction

Punctuality of railway services is a highly important issue, since punctuality is considered as one of its main performance indicators. In the Netherlands, punctuality is defined as the percentage of trains that arrive with a delay of less than 3 minutes at one of the larger railway stations. In several other countries, a 5 minute margin is used. Delays of trains occur since real-time railway operations are subject to stochastic disturbances. However, the underlying railway timetable is a deterministic plan. Therefore, the stochastic disturbances in the operations should be taken into account in the design of a timetable as well as possible. In order to cope with the disturbances in the real-time operations, a timetable contains buffer times between trains and supplements in the running times and in the dwell times of the trains.

Many authors addressed the analysis and the improvement of the punctuality of railway services: several relevant models have been developed to that end. The main examples of these models are the following: (i) simulation models (see Bergmark (1996), König (2001), Middelkoop and Bouwman (2000), and Wahlborg (1996)), (ii) Max-Plus models (see Goverde (1998), De Kort (2000), and Soto Y Koelemeijer et al. (2000)), and (iii) analytical models (see Carey (1999), Higgins and Kozan (1998), and Huisman and Boucherie (2001)). Other relevant literature on stochastic methods for the improvement of railway timetables is Hallowell and Harker (1998), Schwanhäußer (1994), Mühlhans (1990), and Petersen and Taylor (1982). However, a drawback of the existing models is that they are mainly *evaluation* models and that, based on these models, *optimization* of a timetable can only be achieved by trial-and-error. That is, the timetable is modified and then the evaluation model is used afterwards to evaluate the effect of the modification. If necessary, these steps are repeated.

In contrast with the existing models, this paper describes a *stochastic optimization model* (see Birge and Louveaux (1997)) that can be used to modify a given cyclic timetable and, at the same time, to evaluate the modified timetable by operating a number of realizations of the trains in the timetable. These trains are operated as much as possible according to the modified timetable, but subject to external stochastic disturbances. The main criterion that is used to modify the timetable is minimization of the average delay of the trains. Note that other criteria can be handled as well. The structure of the model is such that it is a *symbiosis* of a *timetabling* model and a *simulation* model.

The first model in this paper generates a timetable for a single train that is operated under stochastic external disturbances on a number of consecutive trips along the same line. Here a trip is a movement of a train from one station to the next. The model is used to allocate a fixed total amount of running time supplement to the consecutive trips such that the average delay of the train is minimal. The model is then extended to be applicable in a more complex situation where several trains are operated according to a given cyclic timetable and on a common railway infrastructure. These trains are also operated under stochastic external disturbances. The extended model is used here to improve the timetable with respect to the average delay of the trains by re-allocating buffer times and time supplements. The application of the extended model to a practical case shows that, within the model, the improvement of a given timetable may lead to a substantial reduction of the average delay of the trains.

This paper is structured as follows. Section 2 describes several aspects that are relevant for the allocation of running time supplements. In Section 3 we describe the above mentioned first stochastic optimization model. In Section 4, we prove that, if the train runs over just *two* consecutive trips and if there is a finite probability distribution of the disturbances, then the results of the stochastic optimization model converge to the true optimum if the number of realizations tends to infinity. In Section 5, we present the computational results related to the model of Section 3. Section 6 describes the above mentioned extended stochastic optimization model. Computational results that were obtained by applying

this model to the railway corridor between Haarlem and Maastricht/Heerlen are described in Section 7. The paper is finished with conclusions in Section 8.

2 Running time supplements

2.1 A trade-off

To obtain a high punctuality of the railway services, it is desirable that trains are able to run faster than planned in order to make up for earlier delays. This means that the planned running times should be longer than the technically minimum running times. The difference between the planned running time and the technically minimum running time is the *running time supplement*.

Also other processes (e.g. halting at stations) may obtain time supplements in the planning. However, in this paper we only focus on the allocation of running time supplements. Note that these other process time supplements may be handled in the same way as the running time supplements.

In general, higher running time supplements lead to a better punctuality of the railway services. However, higher running time supplements also lead to higher *planned* running times. This means that the planned travel times of the passengers increase as well. Note that these planned travel times do not only depend on the total *amount* of running time supplements, but also on the *distribution* of the running time supplements among the trips in the timetable. Note further that running time supplements may even have a negative influence on the *realized* travel times. Indeed, each minute of running time supplement in the timetable brings the risk that it is not needed, since there are no disturbances. Furthermore, longer planned running times increase the block occupation times and therewith the track occupation rates. Additionally, longer planned running times require more personnel and rolling stock, hence they are negative for the efficiency of the railway system.

On the other hand, running time supplements add to the predictability of the realized travel times and to the reliability of the railway system as a whole. As a consequence, the running time supplements should be chosen by a trade-off between the above elements.

2.2 Application in practice

In the Netherlands, running time supplements are approximately 7% of the technically minimum running times. This percentage is used nationwide for all types of passenger services. However, due to rounding -because of the integer character of the timetable- and local circumstances, the actual percentages may deviate from this percentage. Furthermore, cargo trains are usually planned 5 kilometer per hour below their maximum speed. Additionally, a running time supplement of 5% of the running times may be used for cargo trains. On top of that, for a cargo train, the planned acceleration and deceleration times are based on a maximum total weight of the train. The difference between this maximum weight and the actual weight of the train acts as an extra running time supplement.

In Switzerland, running time supplements have several components (see Halderman (2003)). First, there is a proportional running time supplement, which equals 7% of the running time for passenger trains and 11% for cargo trains. Secondly, *special operational supplements* are added at highly utilized nodes. Additionally, one minute of supplement is added for each 30 minutes of running time. For trips with high average speeds, the supplements are even higher. In the United Kingdom running times are based on past performance on a railway section (see Rudolph (2003)). Supplements are not explicitly defined here.

Leaflet 451-1 of the UIC (see UIC (2000)) gives recommendations for running time supplements. It recommends a running time supplement to be the sum of a distance dependent supplement and a percentage of the technically minimum running time. The distance dependent supplement is 1.5 minute per 100 kilometer for locomotive-hauled passenger trains and 1 minute per 100 kilometer for multiple unit passenger trains. The running time dependent supplements may vary between 3% for relatively slow trains and 7% for faster trains. For locomotive-hauled trains, this percentage also depends on the weight of the train. For cargo trains, supplements are generally higher. The running time dependent supplement can be replaced by a second distance dependent supplement.

It can be concluded that in practice it is common to allocate the running time supplement on a certain trip to a large extent *in proportion* to the running time on that trip. In this paper, such an allocation is called the *proportional* allocation. However, this paper demonstrates that, from a punctuality point of view, it is better to allocate a somewhat larger part of the total running time supplement to the first trips of the complete route of a train. Indeed, a delay reduction on a certain trip does not only reduce the delay on the respective trip, but also on all subsequent trips. This means that the delay reduction is measured at all subsequent measuring points. Consequently, an early running time supplement is more effective than a late running time supplement. Therefore, one would expect to have a relatively large part of the running time supplements early on. But there is also a downside: if there are no early disturbances, then early supplements are lost. Hence they are useless in that case.

The stochastic optimization model described in this paper can be used to analyze this stochastic trade-off for a cyclic timetable. Also the choice between running time supplements and dwell time supplements can be supported by this model. However, for ease of presentation, we first focus in Section 3 on the allocation of a fixed amount of running time supplement to the consecutive trips of a single train. This allocation is done in such a way that the average arrival delay of the train is minimal. Thereafter, in Section 6, we describe a more complicated situation where several trains are operated on a common infrastructure.

3 A single train on a single line

In this section, we present a stochastic optimization model for allocating a fixed amount of running time supplement S to N consecutive trips of a single train on a single line. On each of the trips, the train is subject to external disturbances,

possibly leading to a delay of the train. This delay is measured at the end of each trip. The objective is to minimize the average delay of the train.

This situation is illustrated in Figure 1. Here the running time supplements on the trips t are denoted by the variables s_t . The external disturbances are denoted by the parameters δ_t , and the resulting delay is represented by the variables D_t . The figure represents the fact that the disturbances δ_t are partially compensated by the running time supplements s_t , so that they can “leave the train” again. The disturbances that cannot be compensated, since the running time supplements have been used completely, accumulate in the delays D_t .

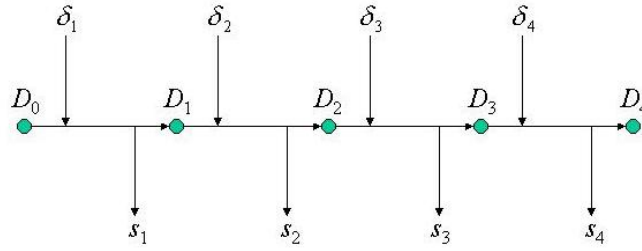


Fig. 1. Relation between disturbances, running time supplements, and delay.

In this paper we assume that all running time supplement allocated to a trip can be used for recovering from a disturbance at the same trip. In other words, the disturbances are assumed to take place at the start of a trip (or at the preceding station), as is shown in Figure 1. This assumption may be relaxed by splitting the trips into a number of smaller trips.

The stochastic optimization model contains a *planning* part for determining the running time supplements and an *evaluation* part for determining the resulting average delay of the train. To that end, at the same time as the running time supplements s_t are determined, R realizations of the train are operated along the N trips subject to externally generated disturbances. Let $\delta_{t,r}$ denote the disturbance incurred on trip t by realization r of the train. Furthermore, the resulting delay of realization r of the train by the end of trip t is denoted by $D_{t,r}$. Then the following relation describes the balance of the external disturbances, the running time supplements, and the delay:

$$D_{t,r} = \max\{ 0, D_{t-1,r} + \delta_{t,r} - s_t \} \text{ for } t = 1, \dots, N; \ r = 1, \dots, R. \quad (1)$$

Note that equation (1) is a mathematical representation of Figure 1. It follows that, if $\delta_{t,r} > s_t$, then the delay of realization r of the train increases over trip t . If $\delta_{t,r} \leq s_t$, then the delay of realization r of the train may decrease over trip t . Note that equation (1) assumes that the train is not influenced by other trains: disturbances are assumed to be autonomous and external. A further

consequence of equation (1) is that the amount of running time supplement $U_{t,r}$ that is actually used by realization r on trip t equals

$$U_{t,r} = \min\{ D_{t-1,r} + \delta_{t,r}, s_t \} \text{ for } t = 1, \dots, N; r = 1, \dots, R. \quad (2)$$

Obviously, equation (1) is not a linear equation. However, it can be linearized easily. Now the complete model can be described as follows:

$$\min D = \sum_{t=0}^N \sum_{r=1}^R w_t D_{t,r} / R \quad (3)$$

subject to

$$D_{t-1,r} + \delta_{t,r} - s_t \leq D_{t,r} \text{ for } t = 1, \dots, N; r = 1, \dots, R \quad (4)$$

$$\sum_{t=1}^N s_t \leq S \quad (5)$$

$$D_{t,r} \geq 0 \text{ for } t = 0, \dots, N; r = 1, \dots, R \quad (6)$$

$$s_t \geq 0 \text{ for } t = 1, \dots, N \quad (7)$$

The objective function (3) indicates that the objective is to minimize the average weighted delay D . For $t = 1, \dots, N$, the weight w_t indicates the weight of the delay at the end of trip t , depending on the number of involved passengers or on the status of the station at the end of trip t (e.g. an ordinary station or a transfer station). Constraints (4) and (6) together give the linearized version of equation (1) relating the delay at the end of trip t to the delay at the end of trip $t-1$. Next, constraint (5) expresses the fact that only a fixed amount of running time supplement is to be distributed among the trips. Finally, constraints (6) and (7) indicate that the variables are to be non-negative.

4 Convergence

In this section we consider the same model as in the previous section, but for the case of a single train that is operated over *two* consecutive trips. We assume that the probability distribution of the disturbances (δ_1, δ_2) has a finite set I of possible values. Each of these values (δ_1^i, δ_2^i) has a probability of occurrence p_i .

For this case we prove that the results of the stochastic programming model converge to the optimal allocation of the running time supplement if the number of realizations tends to infinity. Here we assume that the optimal allocation of the running time supplement is unique. The latter is not essential, but it simplifies the proof somewhat. The assumption holds e.g. if $|I|$ is odd and all disturbances (δ_1^i, δ_2^i) satisfy $\delta_1^i + \delta_2^i > S$, but also in many other situations. Further results on convergence in stochastic optimization can be found in Linderoth et al. (2002).

4.1 Optimal running time supplement

The running time supplement allocated to trip 1 is denoted by s . Then the running time supplement allocated to trip 2 equals $S - s$. Figure 2 shows the

partitioning of the positive (δ_1, δ_2) quadrant for a given value of s into the areas $A_1(s)$, $A_2(s)$, $A_3(s)$, and $A_4(s)$. For example, $A_1(s)$ is the area with relatively small disturbances on both trips. As a consequence, on both trips the delays can be compensated by the running time supplements. Similarly, $A_2(s)$ is the area with relatively small disturbances on the first trip and relatively large disturbances on the second trip. This results in delays on the second trip only.

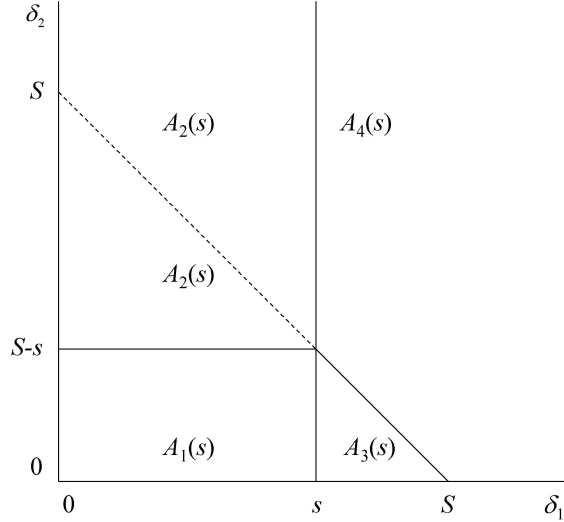


Fig. 2. Partitioning into the areas $A_1(s)$, $A_2(s)$, $A_3(s)$, and $A_4(s)$.

The delay of the train by the end of trip t ($t = 1, 2$) if the disturbances equal (δ_1^i, δ_2^i) is denoted by D_t^i . In that case, the total weighted delay of the train over the two trips is denoted by D^i . As a consequence, for a given value s of the running time supplement on the first trip, the following weighted delays are caused by the disturbances (δ_1^i, δ_2^i) :

- If (δ_1^i, δ_2^i) in $A_1(s)$, then $D_1^i = 0$ and $D_2^i = 0$. Hence $D^i = 0$.
- If (δ_1^i, δ_2^i) in $A_2(s)$, then $D_1^i = 0$ and $D_2^i = \delta_2^i - (S - s)$. Hence $D^i = w_2(\delta_2^i - S + s)$.
- If (δ_1^i, δ_2^i) in $A_3(s)$, then $D_1^i = \delta_1^i - s$ and $D_2^i = 0$. Hence $D^i = w_1(\delta_1^i - s)$.
- If (δ_1^i, δ_2^i) in $A_4(s)$, then $D_1^i = \delta_1^i - s$ and $D_2^i = \delta_1^i + \delta_2^i - S$. Hence $D^i = w_1(\delta_1^i - s) + w_2(\delta_1^i + \delta_2^i - S) = (w_1 + w_2)\delta_1^i + w_2\delta_2^i - w_1s - w_2S$.

From the foregoing it follows that, for a given value s of the running time supplement on the first trip, the average weighted delay $D(s)$ of the train can be expressed as follows:

$$\begin{aligned}
D(s) = & \sum_{i \in A_2(s)} p_i w_2 (\delta_2^i - S + s) + \sum_{i \in A_3(s)} p_i w_1 (\delta_1^i - s) + \\
& \sum_{i \in A_4(s)} p_i ((w_1 + w_2) \delta_1^i + w_2 \delta_2^i - w_1 s - w_2 S). \tag{8}
\end{aligned}$$

The minimization problem to be solved is to find a value s^* for the running time supplement on the first trip such that the average delay $D(s^*)$ is minimal.

It is not difficult to see that the average delay $D(s)$ is a continuous and convex piecewise linear function in s . Furthermore, (8) implies that, if s is not equal to one of the values δ_1^i and $S - s$ is not equal to one of the values δ_2^i , then a slight modification Δs of the running time supplement on the first trip gives the following modification $\Delta D(s)$ of the average delay on the two trips:

$$\begin{aligned}
\Delta D(s) = & \sum_{i \in A_2(s)} p_i w_2 \Delta s - \sum_{i \in A_3(s)} p_i w_1 \Delta s - \sum_{i \in A_4(s)} p_i w_1 \Delta s \\
= & \Delta s \left(\sum_{i \in A_2(s)} p_i w_2 - \sum_{i \in A_3(s) \cup A_4(s)} p_i w_1 \right).
\end{aligned}$$

It follows that the average delay is minimal if the running time supplement s on the first trip is such that, around s , the above expression changes from a negative value (decreasing average delay $D(s)$) to a positive value (increasing average delay $D(s)$). Hence, the optimal running time supplement s^* on the first trip is such that the expression

$$\sum_{i \in A_2(s)} p_i w_2 - \sum_{i \in A_3(s) \cup A_4(s)} p_i w_1$$

is negative for $s = s^* - \Delta s$ and is positive for $s = s^* + \Delta s$ for a sufficiently small value of Δs . It follows that s^* coincides with one of the values $\{\delta_1^i, S - \delta_2^i \mid i \in I\}$. Note that here the assumption is used that there is a unique optimal allocation of the running time supplement.

4.2 Stochastic optimization model

Next, suppose that we have a random sample of R realizations of pairs of disturbances. Let R_i denote the number of occurrences of the pair (δ_1^i, δ_2^i) in this sample. Furthermore, let s denote the proposed value for the running time supplement on the first trip. Then, in the same way as in the previous section, it follows that the average delay $D_R(s)$ can be expressed as follows:

$$\begin{aligned}
D_R(s) = & \sum_{i \in A_2(s)} \frac{R_i}{R} w_2 (\delta_2^i - S + s) + \sum_{i \in A_3(s)} \frac{R_i}{R} w_1 (\delta_1^i - s) + \\
& \sum_{i \in A_4(s)} \frac{R_i}{R} ((w_1 + w_2) \delta_1^i + w_2 \delta_2^i - w_1 s - w_2 S). \tag{9}
\end{aligned}$$

As above, the average delay $D_R(s)$ is a continuous and convex piecewise linear function in s . A similar argument as in the previous section can be used to show that, if the average delay $D_R(s)$ has a unique optimal running time supplement s_R^* , then this optimal value s_R^* is such that the expression

$$\sum_{i \in A_2(s)} \frac{R_i}{R} w_2 - \sum_{i \in A_3(s) \cup A_4(s)} \frac{R_i}{R} w_1$$

is negative for $s = s_R^* - \Delta s$ and is positive for $s = s_R^* + \Delta s$ for a sufficiently small value of Δs . Note that this optimal value s_R^* is the value that is obtained by applying the stochastic optimization model. Figure 3 represents parts of the graphs of the functions $D(s)$ and $D_R(s)$.

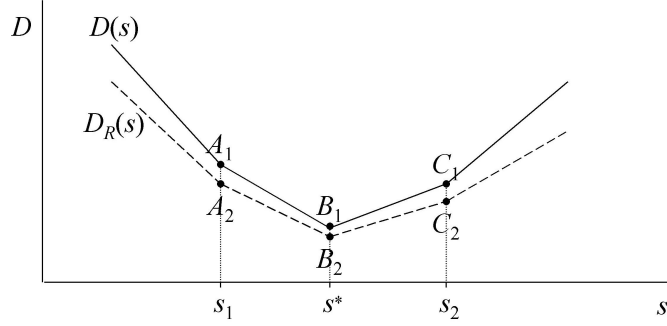


Fig. 3. The convex piecewise linear functions $D(s)$ and $D_R(s)$.

4.3 Proof of Convergence

Theorem 1. *If the minimization problem has a unique optimal solution s^* with $0 < s^* < S$, then $\lim_{R \rightarrow \infty} P(s_R^* = s^*) = 1$.*

Proof. Let $0 \leq s_1 < s^*$ be such that the interval (s_1, s^*) does not contain any value δ_1^i and such that the interval $(S - s^*, S - s_1)$ does not contain any value δ_2^i . Similarly, let $s^* < s_2 \leq S$ be such that the interval (s^*, s_2) does not contain any value δ_1^i and such that the interval $(S - s_2, S - s^*)$ does not contain any value δ_2^i . Next, let Δ_1 and Δ_2 be defined by

$$\Delta_1 := \sum_{i \in A_2(s_1)} p_i w_2 - \sum_{i \in A_3(s_1) \cup A_4(s_1)} p_i w_1,$$

$$\Delta_2 := \sum_{i \in A_2(s_2)} p_i w_2 - \sum_{i \in A_3(s_2) \cup A_4(s_2)} p_i w_1.$$

Since s^* is a unique minimum of the average delay $D(s)$, $\Delta_1 < 0$ and $\Delta_2 > 0$. Note that Δ_1 and Δ_2 are represented in Figure 3 by the differences $D(s^*) - D(s_1)$ and $D(s_2) - D(s^*)$. In other words, the slopes of the solid lines A_1B_1 and B_1C_1 are negative and positive, respectively.

Next we will show that, if R tends to infinity, then the probability that the differences $D_R(s^*) - D_R(s_1)$ and $D_R(s_2) - D_R(s^*)$ are also negative and positive tends to 1. In other words, if R tends to infinity, then the probability that the slopes of the dashed lines A_2B_2 and B_2C_2 in Figure 3 are negative and positive, respectively, tends to 1. A consequence is that, if R tends to infinity, then the probability that $s_R^* = s^*$ tends to 1, as is to be proved.

To that end, first choose $\varepsilon > 0$ and let W be defined by $W := \max\{w_1, w_2\}$. Because of the Law of the Large Numbers, we know that for all $i \in I$ there exists an integer N_i such that for all $R > N_i$ the following holds: $P\left(\left|\frac{R_i}{R} - p_i\right| \geq \frac{-\Delta_1}{W|I|}\right) < \frac{\varepsilon}{2|I|}$. It follows that for all $R > \tilde{R}_1 := \max\{N_i \mid i \in I\}$:

$$\begin{aligned} P(\text{the slope of } A_2B_2 < 0) &= \\ P\left(\sum_{i \in A_2(s_1)} \frac{R_i}{R} w_2 - \sum_{i \in A_3(s_1) \cup A_4(s_1)} \frac{R_i}{R} w_1 < 0\right) &= \\ P\left(\sum_{i \in A_2(s_1)} \left(\frac{R_i}{R} - p_i\right) w_2 - \sum_{i \in A_3(s_1) \cup A_4(s_1)} \left(\frac{R_i}{R} - p_i\right) w_1 < -\Delta_1\right) &\geq \\ P\left(\sum_{i \in I} \left|\frac{R_i}{R} - p_i\right| W < -\Delta_1\right) &\geq P\left(\bigcap_{i \in I} \left\{\left|\frac{R_i}{R} - p_i\right| < \frac{\Delta_1}{W|I|}\right\}\right) = \\ 1 - P\left(\bigcup_{i \in I} \left\{\left|\frac{R_i}{R} - p_i\right| \geq \frac{\Delta_1}{W|I|}\right\}\right) &\geq 1 - \sum_{i \in I} P\left(\left|\frac{R_i}{R} - p_i\right| \geq \frac{\Delta_1}{W|I|}\right) > 1 - \frac{\varepsilon}{2}. \end{aligned}$$

Similarly, there exists an integer \tilde{R}_2 such that for all $R > \tilde{R}_2$

$$P(\text{the slope of } B_2C_2 > 0) = P\left(\sum_{i \in A_2(s_2)} \frac{R_i}{R} w_2 - \sum_{i \in A_3(s_2) \cup A_4(s_2)} \frac{R_i}{R} w_1 > 0\right) > 1 - \frac{\varepsilon}{2}.$$

As a consequence, for all $R > \max\{\tilde{R}_1, \tilde{R}_2\}$ the minimum of $D_R(s)$ is obtained for $s_R^* = s^*$ with probability at least $1 - 2(\frac{\varepsilon}{2}) = 1 - \varepsilon$. \diamond

Theorem 2. *If the minimization problem has a unique optimal solution s^* with $0 < s^* < S$, then for all $\delta > 0$ $\lim_{R \rightarrow \infty} P(|D_R(s_R^*) - D(s^*)| < \delta) = 1$.*

Proof. First, choose $\delta > 0$ and $\varepsilon > 0$, and let the positive number \tilde{R}_0 be such that $P(s_R^* = s^*) > 1 - \frac{\varepsilon}{2}$ for all $R > \tilde{R}_0$. According to the proof of Theorem 1, such a number \tilde{R}_0 exists. Next, we have the following (in)equalities:

$$|D_R(s^*) - D(s^*)| = \left| \sum_{i \in A_2(s^*)} \left(\frac{R_i}{R} - p_i \right) w_2(\delta_2^i - (S - s^*)) + \sum_{i \in A_3(s^*)} \left(\frac{R_i}{R} - p_i \right) w_1(\delta_1^i - s^*) + \sum_{i \in A_4(s^*)} \left(\frac{R_i}{R} - p_i \right) ((w_1 + w_2)\delta_1^i + w_2\delta_2^i - w_1s^* - w_2S) \right| \leq M \times \sum_{i \in I} \left| \frac{R_i}{R} - p_i \right|, \quad (10)$$

where M is an appropriately chosen positive number. Again, because of the Law of the Large Numbers, we know that for all $i \in I$ there exists an integer N_i such that for all integers $R > N_i$ the following holds: $P\left(\left|\frac{R_i}{R} - p_i\right| \geq \frac{\delta}{M|I|}\right) < \frac{\varepsilon}{2|I|}$. Then, it follows that for all integers $R > \max\{N_i \mid i \in I\}$ the following holds:

$$P\left(\sum_{i \in I} \left| \frac{R_i}{R} - p_i \right| < \frac{\delta}{M}\right) \geq P\left(\bigcap_{i \in I} \left\{ \left| \frac{R_i}{R} - p_i \right| < \frac{\delta}{M|I|} \right\}\right) = 1 - P\left(\bigcup_{i \in I} \left\{ \left| \frac{R_i}{R} - p_i \right| \geq \frac{\delta}{M|I|} \right\}\right) \geq 1 - \sum_{i \in I} P\left(\left| \frac{R_i}{R} - p_i \right| \geq \frac{\delta}{M|I|}\right) > 1 - \frac{\varepsilon}{2}. \quad (11)$$

Combining the results in (10) and (11) gives that the following holds for all integers $R > \max\{\{N_i \mid i \in I\} \cup \{\tilde{R}_0\}\}$:

$$\begin{aligned} P(|D_R(s_R^*) - D(s^*)| < \delta) &\geq P((|D_R(s_R^*) - D(s^*)| < \delta) \cap (s_R^* = s^*)) = \\ &P(|D_R(s_R^*) - D(s^*)| < \delta \mid s_R^* = s^*) \times P(s_R^* = s^*) \geq \\ &P\left(\sum_{i \in I} \left| \frac{R_i}{R} - p_i \right| < \frac{\delta}{M}\right) \times P(s_R^* = s^*) \geq \left(1 - \frac{\varepsilon}{2}\right) \left(1 - \frac{\varepsilon}{2}\right) > 1 - \varepsilon. \quad \diamond \end{aligned}$$

Note that the results of Theorems 1 and 2 also hold if the optimal solution s^* is not unique or if s^* equals 0 or S . However, slight modifications of the proofs are required then.

5 Computational results

In this section, we describe the computational results that were obtained by applying the model described in Section 3 to a number of theoretical cases. Computational results for a real-life case are presented in Section 7.

All results in this section are based on equally and exponentially distributed disturbances and on equally weighted delays. However, as was noted earlier already, the latter is certainly *not* essential. It is also possible to use other probability distributions and different weights, including empirical ones.

The results described in this section were obtained by implementing the model in the modeling system OPL Studio and by solving it with the corresponding solver CPLEX 9.0 on an Intel Pentium IV PC with 3.0 GHz processor speed and 512 MB internal memory.

In all cases described in this section, the number of realizations R has been set to 1000. In our experiments, this number turned out to be large enough to generate stable results that were more or less independent of the detailed values of the disturbances. On the other hand, this number of realizations also led to acceptable running times of at most a couple of minutes.

5.1 Optimal allocation of running time supplements

In the experiments described in this section, the running times are subject to exponentially distributed disturbances with an average value of 1 minute. Furthermore, in this section the total amount of running time supplement S equals the total number of trips. That is, the total amount of running time supplement equals the average total disturbances. In Section 5.2, different amounts of running time supplement are applied. In all cases, the objective is to allocate the total amount of running time supplement to the trips in such a way that the average delay is minimal. Note that in the proportional allocation, each trip gets a running time supplement of 1 minute.

The results of the case with 10 trips are shown in Figure 4. The horizontal axis shows the 10 trips, and the vertical axis shows the optimal amounts of running time supplement to be allocated to the trips. The vertical line in the figure represents the weighted average distance of the running time supplements from the starting point. This weighted average distance, WAD , is defined by:

$$WAD = \sum_{t=1}^N \frac{2t-1}{2N} \times s_t. \quad (12)$$

In Figure 4, the value of the WAD equals about 0.425. For the proportional allocation of the running time supplements, the WAD equals exactly 0.5.

Note that, in comparison with the proportional allocation of the running time supplements, the allocation of the running time supplements according to Figure 4 has a negative effect on the average *planned* travel times of the passengers. Indeed, if the numbers of travelers between all O/D-pairs of stations are more or less the same, then minimal average planned travel times of the passengers are obtained either by skipping the running time supplements completely, or by allocating the running time supplements as much as possible to the first or to the last trips along the line. However, it is likely that this allocation of the running time supplements leads to an unreliable timetable.

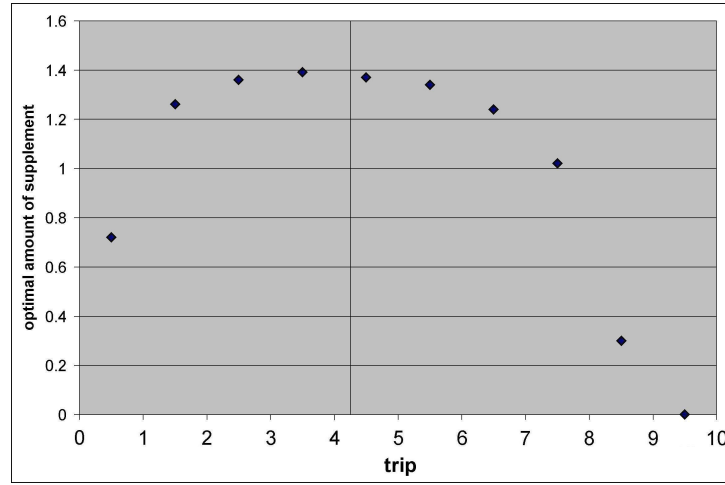


Fig. 4. The optimal allocation of the running time supplements for 10 trips.

In Figure 5, the upward bending line shows the average delay by the end of each trip for the optimal allocation of the running time supplements. The nearly diagonal line in this figure shows the average delay by the end of the trips for the proportional allocation of the running time supplements. Obviously, the optimal allocation performs better on almost all trips. Only on the last trips, the average delay increases quickly for the optimal allocation. This is due to the fact that the supplements have been shifted from the last trips towards the first trips.

Figure 6 presents results that have been obtained by applying the model to 2 to 25 consecutive trips. The average delays of the optimal solutions are compared with the average delays of the proportional allocation of the running time supplements. The figure shows that the average delay decrease is only 1.2% for 2 trips, but the decrease is already 9.5% for 5 trips and 20.1% for 15 trips. Although the shapes of the supplement allocations and the average locations of the supplements are quite similar in all cases, the relative decreases in the average delay are far from equal for the different cases.

5.2 Different amounts of running time supplement

Next, we consider the effect of a different total amount of running time supplement. Figure 7 shows the results of a case with 10 trips, where each trip is subject to exponentially distributed disturbances with an average value of 1 minute. In this case half a minute or two minutes of running time supplement are available per trip. That is, $S = 5$ minutes (dots) or $S = 20$ minutes (diamonds).

If $S = 5$ minutes, the optimal allocation of the running time supplement is even more concentrated on the earlier trips. This is understandable, since early supplements are still more effective and the probability of an excessive early

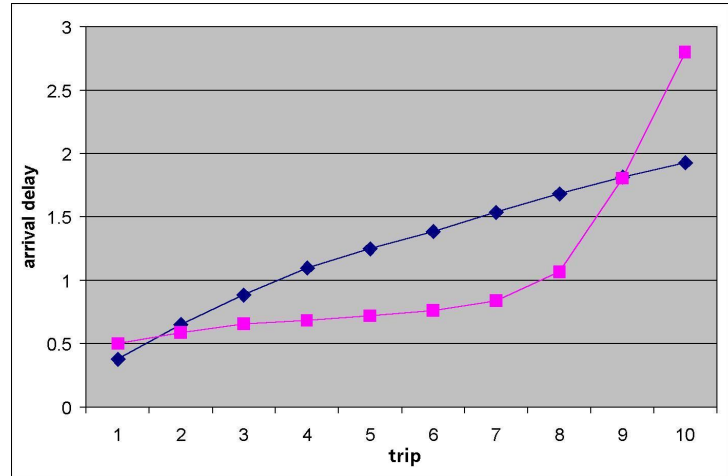


Fig. 5. The average delay by the end of the trips for 10 trips.

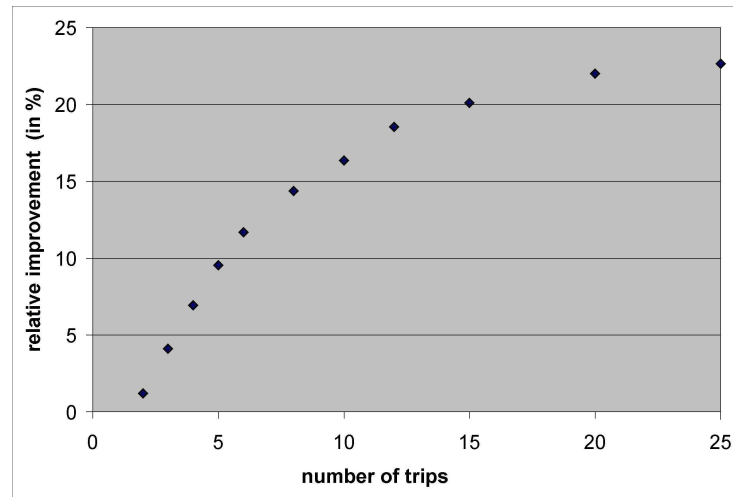


Fig. 6. The decrease in average delay for the optimal supplement allocation in comparison with the average delay for the proportional supplement allocation.

supplement decreases when the total amount of supplement decreases. There is an apparent shift to the left, which is supported by the *WAD*, which is 0.32 in this case. The decrease in the average delay goes up from about 16.3% for the case with $S = 10$ minutes to about 17.8% for the case with $S = 5$ minutes.

In the opposite situation with $S = 20$ minutes, where the total amount of running time supplement is twice as large as the total average disturbances, the allocation of the running time supplement shifts into the opposite direction. Now the *WAD* equals 0.492. The decrease in the average delay goes down to about 2.9% for the case with $S = 20$ minutes. In this case, the difference with the proportional allocation is small.

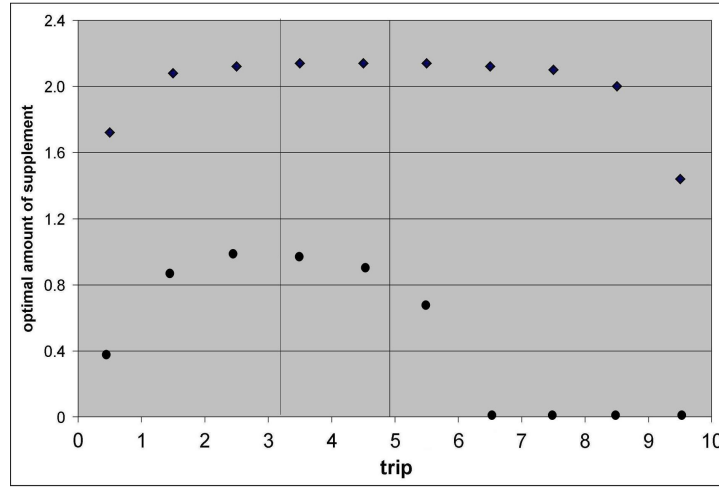


Fig. 7. The optimal allocation for half and double the total amount of supplement.

6 Several trains on a common infrastructure

In this section, we describe an extension of the stochastic optimization model presented in Section 3. Here a given cyclic timetable for a number of trains over a common part of the railway infrastructure is improved with respect to the average weighted delay of the trains. The underlying timetabling model shows some similarity with the well-known Periodic Event Scheduling Model (PESP, see Serafini and Ukovich (1989)). Several researchers have studied the application of PESP for railway timetabling, see e.g. Nachtigall (1996) and Peeters (2003).

The extended model is again a variant of a stochastic optimization model. That is, in order to evaluate and optimize the timetable under construction, R realizations of the processes in each cycle time are operated subject to stochastic disturbances. The R realizations are operated one after another. As a con-

sequence, the delayed trains in realization $r - 1$ may influence the trains in realization r if they use the same parts of the infrastructure. Thus, in contrast with the model described in Section 3, the R realizations are not independent of each other. These interactions between successive realizations are enabled by the cyclic nature of the timetable. They imply that the model is a non-standard variant of a stochastic optimization model (Birge and Louveaux (1997)).

The model aims at improving a given cyclic timetable with respect to the average delay of the trains. The main purpose of the model is to optimally re-allocate buffer times and time supplements to the various process times in the timetable, thereby leaving the orders of the trains on the tracks unchanged. Furthermore, for ease of presentation, we assume that the given timetable does not contain complex cycles, for example caused by the rolling stock circulation or by circular chains of passenger connections. In general, this assumption is valid if the railway network has a tree-like structure and the railway traffic is considered in just one direction. It should be noted that this assumption can be relaxed, but then a more complex model results.

6.1 Timetabling part of the model

We consider a given cyclic timetable with a cycle time T . Such a timetable consists of a number of processes that have to be carried out. For example, trains have to run from one station to another, they have to dwell for a certain period of time in a station, there has to be a certain headway time between two trains on the same infrastructure, etc. For all processes, appropriate process times have to be determined. The begin of a process and the completion of a process are called *events*. Also the corresponding event times are to be determined.

We assume that P processes are to be carried out in each cycle time and that there are E corresponding events. For each process p , the events $b(p)$ and $c(p)$ (with $1 \leq b(p), c(p) \leq E$) denote the begin and completion events of process p . The parameter m_p denotes the technically minimum process time of process p . Furthermore, the variable s_p denotes the planned supplement for the process time of process p . The planned event time of event e is denoted by the variable v_e . We use a linear time axis. This implies that

$$v_{c(p)} - v_{b(p)} = m_p + s_p \text{ for } p = 1, \dots, P.$$

Thus the planned process time of process p equals the difference between the planned completion time and the planned begin time of process p . Most of the constraints to be satisfied in a cyclic timetabling model can be expressed in terms of constraints on the process times (see e.g. Peeters (2003)). In our model, each constraint therefore has the following form:

$$l_p \leq v_{c(p)} - v_{b(p)} \leq u_p \text{ for } p = 1, \dots, P.$$

Here l_p and u_p are appropriate lower and upper bounds. For example, processes such as running along a track, dwelling in a station, and passenger connections between trains can be modeled in this way. Note that for modeling the

headway processes, the assumption that the aim of the model is to improve a given cyclic timetable, thereby leaving the orders of the trains on the tracks unchanged, is essential. Indeed, in a cyclic timetable minimum headway times between trains depend on the orders of the trains on the tracks. If the latter are not known a priori, then additional binary variables are required to model these. This highly complicates the model. However, given the orders of the trains on the tracks, a minimum headway time of at least h minutes between two consecutive trains departing from the same station and entering the same track can be enforced by the following constraint:

$$v_{c(p)} - v_{b(p)} \geq h.$$

Here the events $b(p)$ and $c(p)$ denote the departures of the first and the second train, respectively (that is, the begin and completion events of the corresponding headway process p). Note that the headway times do not have to be bounded from above, since these have only positive effects. This is in contrast with running time supplements, which also have a negative effect, e.g. on the travel times.

In order to allocate a certain amount of time supplement to the process times, Q subsets of processes are selected. Each subset q of processes is connected with a certain amount of time supplement S_q to be allocated to the processes in the set q . Thus the following constraints are to be satisfied:

$$\sum_{p \in q} s_p \leq S_q \text{ for } q = 1, \dots, Q.$$

For example, such a constraint may indicate that a certain total amount of running time supplement is to be allocated to the consecutive running times of a single train. This corresponds with the problem described in Section 3. However, a certain amount of time supplement may also have to be allocated to the trips of a *number* of trains together.

Final relevant constraints specify that, at each part of the infrastructure, the difference between the last and the first planned event time in each cycle time should not exceed the cycle time minus the minimum headway time $T - h$. Moreover, non-negativity constraints have to be imposed on the variables s_p , and if one wants to obtain a timetable that is specified in integer minutes, then integrality constraints have to be imposed on the corresponding event times.

6.2 Evaluation part of the model

In the same way as in the model described in Section 3, the timetable is evaluated during its modification by operating R realizations of the trains in each cycle time subject to predetermined stochastic disturbances. To that end, we use R realizations of the processes and the events in each cycle time.

The stochastic disturbance of realization r of process p is denoted by the parameter $\delta_{p,r}$ for $p = 1, \dots, P$ and $r = 1, \dots, R$. Furthermore, the realized event time of realization r of event e is denoted by the variable $\tilde{v}_{e,r}$ for $e = 1, \dots, E$ and $r = 1, \dots, R$.

Mainly the delays of the events corresponding to arrivals of trains are measured, but also other delays can be taken into account. The measured events are called the *relevant* events. The set of relevant events is denoted by \tilde{E} . The delay of realization r of relevant event e is denoted by the variable $D_{e,r}$. The average weighted delay over all processes is denoted by D .

The variables $D_{e,r}$ and D are assumed to be non-negative. The objective is to minimize the average weighted delay of the trains. Thus the objective is to

$$\text{minimize } D = \sum_{r=1}^R \sum_{e \in \tilde{E}} w_e D_{e,r} / R. \quad (13)$$

Here the weights w_e indicate the weights of the different delays. The constraints linking the event times of the processes to the technically minimum process times and the disturbances are the following:

$$\tilde{v}_{c(p),r} - \tilde{v}_{b(p),r} \geq m_p + \delta_{p,r} \text{ for } p = 1, \dots, P; \ r = 1, \dots, R. \quad (14)$$

Furthermore, processes do not begin too early, and a delay occurs if a relevant process ends too late. This results in the following constraints:

$$v_{b(p)} + rT \leq \tilde{v}_{b(p),r} \text{ for } p = 1, \dots, P; \ r = 1, \dots, R, \quad (15)$$

$$\tilde{v}_{e,r} - v_e - rT \leq D_{e,r} \text{ for } e \in \tilde{E}; \ r = 1, \dots, R. \quad (16)$$

Note that here we use the cyclic character of the timetable, since the planned event time of realization r of event e equals $v_e + rT$.

As was mentioned earlier, the R realizations of the cycle times are operated one after another. As a consequence, the delayed trains in realization $r - 1$ may influence the trains in realization r if they use the same parts of the infrastructure. To that end, let e_1 be the *first* planned event in a cycle time on a certain part of the infrastructure, and let e_2 be the *last* planned event in a cycle time on the *same* infrastructure. Then the following constraint is to be satisfied:

$$\tilde{v}_{e_1,r} - \tilde{v}_{e_2,r-1} \geq h. \quad (17)$$

In other words, realization r of event e_1 cannot be carried out earlier than the headway time h after realization $r - 1$ of event e_2 has taken place. Also other interactions between successive realizations can be modeled. For example, realization $r - 1$ of a train that is one of the last trains in each cycle time may have a passenger connection with realization r of a train that is one of the first trains in each cycle time. As was noted earlier, these interactions imply that the model is a rather non-standard stochastic optimization model.

7 Computational results

In this section, we describe the computational results that were obtained by applying the model described in Section 6 to a real-life case of NS Reizigers, the

main Dutch operator of passenger trains. This analysis was carried out for study purposes. The results have not yet been implemented in the real timetable.

For solving the model to optimality, we again used the modeling system OPL Studio and CPLEX 9.0 on an Intel Pentium IV PC with 3.0 GHz processor speed and 512 MB internal memory. For this case we always used 500 realizations per run, because this gave sufficiently stable results and the computation times remained at an acceptable level of about one hour on the indicated hardware.

7.1 Case: Haarlem–Maastricht/Heerlen

The model was applied to improve the 2004 timetable of NS Reizigers on the corridor from Haarlem (Hlm) in the western part of the Netherlands to Maastricht (Mt) and Heerlen (Hrl) in the southern part. Throughout this section, this original timetable is called the *reference* timetable. All passenger trains on this corridor were included in the model. Cargo trains were left out, both from the reference timetable and from the improved one. As a consequence, the results for the two timetables are still comparable.

The lines on the studied corridor are shown in Figure 8. In this figure, the dotted lines indicate the stoptrains, the other lines indicate the intercity lines. The main stations, represented by boxes, are the stations where the delays are measured. All trains dwell at these stations. The black dots indicate the other dwell stations of the trains.

The main lines on the studied corridor are the intercity lines 800 (Haarlem–Maastricht) and 900 (Haarlem–Heerlen), which are both operated once per hour. These lines have the corridor from Haarlem to Sittard (Std) in common. On this corridor, there is nearly a 30 minute cycle time, since almost all lines are operated there twice per hour with an exact 30 minute cycle time. The order of the trains on the different trips follows from the reference timetable. The overtakings of the stoptrains in Abcoude (Ac), Geldermalsen (Gdm) and 's-Hertogenbosch (Ht) remain unchanged.

Turn-around constraints at the line endpoints have not been taken into account in the model. This means that the southbound and the northbound trains are almost independent of each other. Therefore, two nearly independent problems are created: the southbound problem and the northbound problem. Only the southbound problem is described here.

The planned running times in the reference timetable include on average 7.92% of running time supplement on each trip. The only exceptions can be found on the line 3500, which bears additional supplements of 1 minute between Duivendrecht (Dvd) and Abcoude, and 4 additional minutes between Abcoude and Breukelen (Bkl). Only the trains that are overtaken have dwell time supplements. The other planned dwell times are equal to the minimum dwell times.

For the lines that are not fully covered by the corridor, the departure and arrival times at the stations where these lines enter or leave the model have been fixed to the event times in the reference timetable. In the realizations, these trains are assumed to enter the model at these stations without a delay.

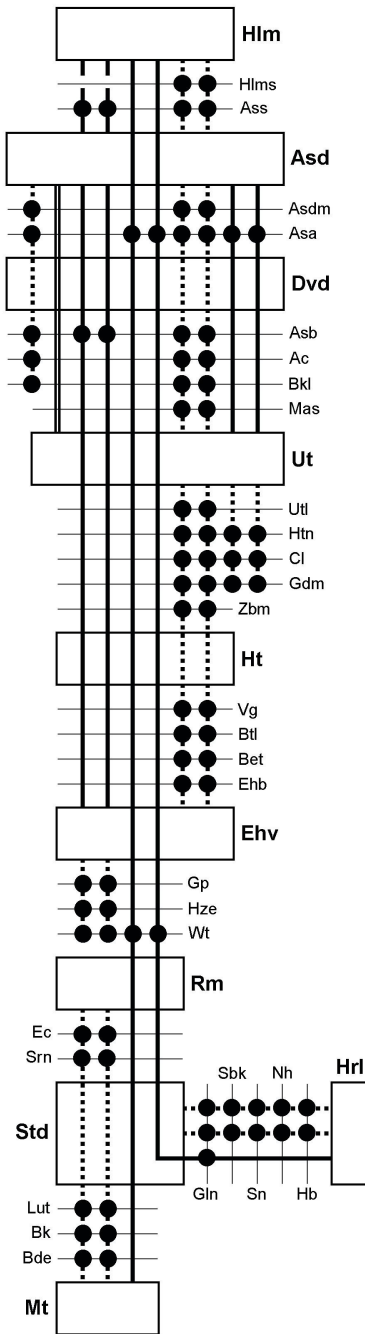


Fig. 8. The lines on the corridor Haarlem-Maastricht/Heerlen.

In this case study, we mainly used exponential distributions for generating the disturbances to the process times. However, as was noted earlier, also other probability distributions could have been used. Anyway, in the sensitivity analysis, other probability distributions were used as well.

7.2 Results

First, the reference timetable was *evaluated*. That is, after fixing all event times according to this timetable, the stochastic optimization model was applied to operate 500 realizations of the trains according to this timetable but under stochastic disturbances.

Each trip between two measuring stations was disturbed with an average disturbance of 7.92% of the minimum running time. This average is the same as the planned running time supplement in the reference timetable. This percentage was also chosen in such a way that the evaluation led to a punctuality of 83.7%, which is comparable to the punctuality observed in practice. The evaluation led to an average delay of 1.38 minutes.

Next, the timetable was *optimized* by unfixing most of the event times and by operating 500 realizations of the trains under the *same* stochastic disturbances as in the evaluation of the reference timetable. The objective was to modify the timetable by re-allocating the running time supplements and the buffer times in such a way that the average delay at the ten measuring stations was minimal. Other relevant measures, such as the punctuality, were determined afterwards.

The optimization led to a model with 160,000+ variables and 300,000+ constraints. Because of the size of the model, the event times were allowed to be fractional, so that the model could be solved by Linear Programming. By the optimization, the average delay was reduced to 0.947 minutes, which is 31.4% less than the average delay in the reference timetable. The 3-minute punctuality increased from 83.7% for the reference timetable to 89.5%: this is a reduction of the number of late trains by 35.2%.

The optimal running time supplements for the lines 800 and 900 obtained by the optimization are shown in the last column of Table 1. The third and fourth column (“per line”) show the results that were obtained by applying the single line model described in Section 3 to the lines 800 and 900 separately.

As in the reference timetable, an exact 30-minute cycle time was enforced for most lines on the corridor from Haarlem to Sittard, leading to identical supplements there for the lines 800 and 900 up to Sittard. Because of the longer running time between Sittard and Heerlen as compared to the running time between Sittard and Maastricht, there was 0.22 minute more running time supplement available for the line 900 than for the line 800. The latter could only be allocated to the trip Sittard-Heerlen, due to the 30-minute cycle time between Haarlem and Sittard.

For the rest of the corridor, the supplement allocation is very similar to the one found by the single line model described in Section 3. The only remark that can be made in this respect is that slightly larger supplements are found for

the most busy parts of the infrastructure between Amsterdam and Utrecht, and smaller supplements for the somewhat quieter parts south of 's-Hertogenbosch.

Table 1. The running time supplements in minutes for the lines 800 and 900 obtained by single line optimization and by corridor optimization.

trip	avg. disturbance	running time supplements		
		per line 800	per line 900	corridor 800 & 900
Hlm-Asd	1.03	0.85	0.87	0.90
Asd-Dvd	0.81	1.01	1.02	1.16
Dvd-Ut	1.25	1.43	1.44	1.96
Ut-Ht	2.05	2.63	2.67	2.51
Ht-Ehv	1.32	1.71	1.72	1.55
Ehv-Rm	2.27	2.57	2.64	2.18
Rm-Std	1.10	0.72	0.78	0.67
Std-Mt	1.10	0.00	-	0.00
Std-Hrl	1.32	-	0.00	0.22

7.3 Sensitivity Analysis

The timetable found by the stochastic optimization model is only optimal with respect to the *applied* disturbances. Therefore we carried out a sensitivity analysis in order to investigate the behavior of this timetable under other disturbances from the same distribution and under disturbances from other distributions. For this analysis, again only the southbound timetable was considered. In the following description, the *preferred* timetable is the timetable which is optimal with respect to exponentially disturbed running times with an average disturbance of 7.92% of the respective minimum running time.

First, we analyzed the consequences of other sets of random disturbances from the same disturbance distribution. The timetable was not optimized again for these other sets of disturbances, but both for the reference timetable and for the preferred timetable, the delay propagation resulting from these disturbances was evaluated. Ten random sets of disturbances from the same distribution were used, leading to ten evaluations of both timetables.

This led to the results shown in Table 2. The range of the average delay and the unpunctuality has a width of at most 10%. This is relatively small in comparison with the differences between the reference and the preferred timetable.

Next, we evaluated the preferred timetable, but now for sets of disturbances from other probability distributions. All distributions described here were multiplied by 0.0792 times the technically minimum running time. In this way, the original disturbance distribution could be described as an exponential distribution with an average value of 1 minute. First, the timetable was evaluated for the

Table 2. The influence of randomness on the punctuality measures.

measure	timetable	avg.	min.	max.	range	σ
average delay	reference	1.34	1.30	1.38	6.0%	0.03
average delay	preferred	0.92	0.89	0.95	5.8%	0.02
unpunctuality	reference	15.6%	15.0%	16.3%	8.0%	0.5%
unpunctuality	preferred	10.0%	9.5%	10.5%	10.1%	0.4%

situation where a large part of the running times was not disturbed, and the rest was disturbed again by exponentially distributed disturbances. Furthermore, a uniform distribution and a triangular distribution were applied.

The results of this sensitivity analysis are summarized in Table 3. This table shows that the preferred timetable outperformed the reference timetable for all disturbance distributions that were applied. The worst results were obtained for the relatively large disturbances that occur with a low probability (80% 0 and 20% exp. 6). This was to be expected, since the running time supplements are intended for handling small disturbances only.

Table 3. Punctuality gain for different disturbance distributions. The parameter for the exponential distributions is the average (not the reciprocal of the average).

disturbance distribution	reference timetable		preferred timetable		improvement	
	avg. delay	punct.	avg. delay	punct.	avg. delay	punct.
exp. 1	1.38	83.7%	0.95	89.5%	31.4%	35.2%
50% 0 and 50% exp. 1.5	1.43	83.6%	1.05	87.9%	26.4%	26.2%
80% 0 and 20% exp. 6	1.50	89.9%	1.24	91.4%	17.0%	14.9%
uniform (0,2.5)	1.66	79.8%	1.01	91.1%	39.1%	55.7%
triangular (0,0,4)	2.02	75.4%	1.33	85.3%	34.4%	40.2%

Finally, we compared the preferred timetable with the timetables that were obtained by *optimizing* the timetable under different disturbance distributions. Again, the optimization was carried out with respect to the average delay, and the punctuality was determined afterwards. The obtained results are shown in Table 4. In this table, the column “optimal timetable” represents the optimal timetable for our samples of the corresponding disturbance distribution. Hence, each row in this column corresponds to a different optimal timetable. Table 4 shows that the preferred timetable is close to the optimal timetable for each of the applied distributions.

It can be concluded that the preferred timetable has a relatively high quality under a range of disturbance distributions. For further details of the experiments

Table 4. Optimality gap of the preferred timetable.

disturbance distribution	preferred timetable		optimal timetable		improvement	
	avg. delay	punct.	avg. delay	punct.	avg. delay	punct.
exp. 1	0.95	89.5%	0.95	89.5%	0.0%	0.0%
50% 0 and 50% exp. 1.5	1.05	87.9%	1.04	88.0%	1.1%	1.1%
80% 0 and 20% exp. 6	1.24	91.4%	1.17	92.2%	6.1%	10.0%
uniform (0,2.5)	1.01	91.1%	1.00	91.2%	1.6%	1.5%
triangular (0,0,4)	1.33	85.3%	1.31	85.3%	1.2%	0.2%

that were described here and for a description of further experiments that have been carried out we refer to Vromans (2005).

8 Final remarks and further research

In this paper, we showed that stochastic optimization is a useful approach for improving cyclic railway timetables. We first described a simple model for allocating a fixed amount of running time supplement to the consecutive trips of a single train. Thereafter, we extended the model such that it can be applied for optimally allocating time supplements and buffer times if several trains are operated according to a given cyclic timetable on a common infrastructure. This stochastic optimization model is a *symbiosis* of a *timetabling* model and a *simulation* model. Indeed, during the modification of the timetable, a number of realizations of the timetable under construction is operated under stochastic disturbances. The time supplements and buffer times are selected by the model such that the resulting total average delay is minimal.

The results obtained by the first model indicate that a proportional distribution of the running time supplements does not lead to a minimum average delay. A relatively large part of the running time supplements has to be shifted towards the earlier trips of a train. The motivation is that delay reductions early on are counted on all subsequent trips. A consequence is that running time supplements on the last trips are relatively low, since these decrease the delay on the last trips only. The difference between the optimal allocation and the proportional allocation of the running time supplements is largest when the average amount of running time supplement per trip is relatively small in comparison with the average size of the disturbances.

The results obtained by the second model indicate that a significant reduction of the average delay can be achieved by re-allocating time supplements and buffer times in a given timetable. We focused on the timetable on the corridor between Haarlem and Maastricht/Heerlen in the Netherlands. On this corridor, the average delay could be reduced within the model by about 30%. Further experiments, for example including both directions of the studied corridor and integer departure times, will be carried out to further validate the model.

So far, we mainly experimented with disturbances from exponential probability distributions. However, the latter is not at all essential. Any probability distributions, including empirical ones, can be used. Moreover, each process can be subject to disturbances from its own probability distribution. Obviously, the applied probability distributions should be such that the disturbances are as much as possible realistic in practice. This may require a lot of field research in order to determine the appropriate probability distributions.

As was noted earlier, the allocation of the time supplements that leads to a minimal average delay of the trains need not lead to minimal average *planned* or *realized* travel times of the passengers. Indeed, if the objective is to minimize these, then in principle it is optimal to skip all time supplements: each minute of time supplement brings the risk that it is redundant in certain realizations, since there are no disturbances. Moreover, if the objective is to minimize the average travel times of the passengers, given the allocation of a certain *fixed* amount of running time supplements, then these running time supplements should mainly be allocated to the first or the last trips of the trains: these are usually the parts of a line with the lowest numbers of passengers. However, it is likely that a timetable with such an allocation of the running time supplements or without any time supplements is quite unreliable. Therefore, a trade-off has to be made between the improved reliability that is caused by the time supplements and several other criteria, such as the travel times of the passengers and the efficiency of rolling stock and crew. This is a subject for further research.

In our further research, we will also experiment with relaxations of the assumptions described in Section 6. That is, we will first assume that the given timetable may contain cycles that are caused by the rolling stock circulation or by chains of passenger connections. Under certain conditions, this seems to be a relatively easy extension. Next, we will assume that the orders of the trains on the tracks have not been fixed a priori. Since this extension requires many additional binary variables, it strongly reduces the computability of the timetable. However, the gain will be that the model will not only be able to *improve* a given timetable, but even to generate *from scratch* a timetable that is optimal with respect to the average delay of the trains.

References

- Bergmark, R.: Railroad capacity and traffic analysis using SIMON. In: J. Allan, C.A. Brebbia, R.J. Hill, and G. Sciutto (eds.), *Computers in Railways V*, (1996) 183-191. WIT Press, Ashurst, UK.
- Birge, J.R. and F. Louveaux: *Introduction to Stochastic Programming*, (1997). Springer, New York.
- Carey, M.: Ex ante heuristic measures of schedule reliability, *Transportation Research B*, **33(7)** (1999) 473-494.
- Goverde, R.M.P.: The Max-Plus Algebra approach to railway timetable design. In: B. Mellitt, R.J. Hill, J. Allan, G. Sciutto, and C.A. Brebbia, (eds.), *Computers in Railways VI*, (1998) 339-350. WIT Press, Ashurst, 1998
- Haldeman, L.: *Automatische Analyse von IST-Fahrplänen*, Master's thesis, ETH Zürich (2003), Switzerland, (in German).

- Hallowell, S.F. and P.T. Harker: Predicting on-time performance in scheduled rail-road operations: methodology and application to train scheduling, *Transportation Research A*, **32(6)** (1998) 279–295.
- Higgins, A. and E. Kozan: Modelling train delays in urban networks, *Transportation Science*, **32(4)** (1998) 346–357.
- Huisman, T. and R.J. Boucherie: Running times on railway sections with heterogeneous train traffic. *Transportation Research Part B*, **35** (2001) 271–292.
- Jochim, H.E.: Verspätung auf Stammstrecken: wie man ihrer nicht Herr wird. *Der Nahverkehr*, **3** (2004), (in German).
- König, H.: VirtuOS: Simulieren von Bahnbetrieb. *Betrieb und Verkehr*, **50(1-2)** (2001) 44–47, (in German).
- Kort, A.F. de: Advanced railway planning using Max-Plus algebra. In: J. Allan, C.A. Brebbia, R.J. Hill, and G. Sciutto (eds.), *Computers in Railways VII*, (2000) 257–266. WIT Press, Ashurst, UK.
- Linderoth, J.T., A. Shapiro, and S.J. Wright: The Empirical Behavior of Sampling Methods for Stochastic Programming. Optimization Technical Report 02-01. University of Wisconsin-Madison, (2002).
- Middelkoop, D. and M. Bouwman: Train network simulator for support of network wide planning of infrastructure and timetables. In: J. Allan, C.A. Brebbia, R.J. Hill, and G. Sciutto (eds.), *Computers in Railways VII*, (2000) 267–276. WIT Press, UK.
- Mühlhans, E.: Berechnung der Verspätungsentwicklung bei Zugfahrten. *Eisenbahntechnische Rundschau*, **39** (1990) 465–468.
- Nachtigall, K.: Periodic network optimization with different arc frequencies. *Discrete Applied Mathematics*, **69** (1996) 1–17.
- Peeters, L.W.P.: Cyclic railway timetable optimization. Ph.D. thesis, Erasmus University Rotterdam (2003), Rotterdam School of Management.
- Petersen E.R. and A.J. Taylor: A structured model for rail line simulation and optimization. *Transportation Science*, **18** (1982) 192–206.
- Rudolph, R.: Allowances and margins in railway scheduling. *Proceedings of WCRR 2003*, (2003) 230–238, Edinburgh, UK.
- Schwanhäußer, W.: The status of German railway operations management in research and practice. *Transportation Research A*, **28(A)** (1994) 495–500.
- Serafini, P. and W. Ukovich: A mathematical model for Periodic Event Scheduling Problems. *SIAM Journal on Discrete Mathematics*, **2** (1989) 550–581.
- Soto y Koelemeijer, G., A.R. Iounousov, R.M.P. Goverde, and R.J. van Egmond: PETER, a performance evaluator for railway timetables. In: J. Allan, C.A. Brebbia, R.J. Hill, and G. Sciutto (eds.), *Computers in Railways VII*, (2000) 405–414. WIT Press, Ashurst, UK.
- U.I.C.: Timetable recovery margins to guarantee timekeeping - Recovery margins, Leaflet 451-1, (2000) U.I.C., Paris, France.
- Vromans, M.J.C.M.: Reliability of railway services. Ph.D. thesis, Erasmus University Rotterdam, Rotterdam School of Management, (2005).
- Wahlborg, M.: Simulation models: important aids for Banverket's planning process. In: J. Allan, C.A. Brebbia, R.J. Hill, and G. Sciutto (eds.), *Computers in Railways V*, (1996) 175–181. WIT Press, Ashurst, UK.

Publications in the Report Series Research* in Management

ERIM Research Program: "Business Processes, Logistics and Information Systems"

2005

On The Design Of Artificial Stock Markets

Katalin Boer, Arie De Bruin And Uzay Kaymak

ERS-2005-001-LIS

<http://hdl.handle.net/1765/1882>

Knowledge sharing in an Emerging Network of Practice: The Role of a Knowledge Portal

Peter van Baalen, Jacqueline Bloemhof-Ruwaard, Eric van Heck

ERS-2005-003-LIS

<http://hdl.handle.net/1765/1906>

A note on the paper Fractional Programming with convex quadratic forms and functions by H.P.Benson

J.B.G.Frenk

ERS-2005-004-LIS

<http://hdl.handle.net/1765/1928>

A note on the dual of an unconstrained (generalized) geometric programming problem

J.B.G.Frenk and G.J.Still

ERS-2005-006-LIS

<http://hdl.handle.net/1765/1927>

Privacy Metrics And Boundaries

L-F Pau

ERS-2005-013-LIS

<http://hdl.handle.net/1765/1935>

Privacy Management Contracts And Economics, Using Service Level Agreements (Sla)

L-F Pau

ERS-2005-014-LIS

<http://hdl.handle.net/1765/1938>

A Modular Agent-Based Environment for Studying Stock Markets

Katalin Boer, Uzay Kaymak and Arie de Bruin

ERS-2005-017-LIS

<http://hdl.handle.net/1765/1929>

Lagrangian duality, cone convexlike functions

J.B.G. Frenk and G. Kassay

ERS-2005-019-LIS

<http://hdl.handle.net/1765/1931>

Operations Research in Passenger Railway Transportation

Dennis Huisman, Leo G. Kroon, Ramon M. Lentink and Michiel J.C.M. Vromans

ERS-2005-023-LIS

<http://hdl.handle.net/1765/2012>

Agent Technology Supports Inter-Organizational Planning in the Port
Hans Moonen, Bastiaan van de Rakt, Ian Miller, Jo van Nunen and Jos van Hillegersberg
ERS-2005-027-LIS
<http://hdl.handle.net/1765/6636>

Faculty Retention factors at European Business Schools
Lars Moratis, Peter van Baalen, Linda Teunter and Paul Verhaegen
ERS-2005-028-LIS
<http://hdl.handle.net/1765/6559>

Determining Number of Zones in a Pick-and-pack Orderpicking System
Tho Le-Duc and Rene de Koster
ERS-2005-029-LIS
<http://hdl.handle.net/1765/6555>

Integration of Environmental Management and SCM
Jacqueline Bloemhof and Jo van Nunen
ERS-2005-030-LIS
<http://hdl.handle.net/1765/6556>

On Noncooperative Games and Minimax Theory
J.B.G. Frenk and G.Kassay
ERS-2005-036-LIS
<http://hdl.handle.net/1765/6558>

Optimal Storage Rack Design for a 3-dimensional Compact AS/RS
Tho Le-Duc and René B.M. de Koster
ERS-2005-041-LIS
<http://hdl.handle.net/1765/6730>

Strategies for Dealing with Drift During Implementation of ERP Systems
P.C. van Fenema and P.J. van Baalen
ERS-2005-043-LIS
<http://hdl.handle.net/1765/6769>

Modeling Industrial Lot Sizing Problems: A Review
Raf Jans and Zeger Degraeve
ERS-2005-049-LIS
<http://hdl.handle.net/1765/6912>

Cyclic Railway Timetabling: a Stochastic Optimization Approach
Leo G. Kroon, Rommert Dekker and Michiel J.C.M. Vromans
ERS-2005-051-LIS

* A complete overview of the ERIM Report Series Research in Management:
<https://ep.eur.nl/handle/1765/1>

ERIM Research Programs:
LIS Business Processes, Logistics and Information Systems
ORG Organizing for Performance
MKT Marketing
F&A Finance and Accounting
STR Strategy and Entrepreneurship

Publications in the Report Series Research* in Management

ERIM Research Program: "Business Processes, Logistics and Information Systems"

2005

On The Design Of Artificial Stock Markets

Katalin Boer, Arie De Bruin And Uzay Kaymak

ERS-2005-001-LIS

<http://hdl.handle.net/1765/1882>

Knowledge sharing in an Emerging Network of Practice: The Role of a Knowledge Portal

Peter van Baalen, Jacqueline Bloemhof-Ruwaard, Eric van Heck

ERS-2005-003-LIS

<http://hdl.handle.net/1765/1906>

A note on the paper Fractional Programming with convex quadratic forms and functions by H.P.Benson

J.B.G.Frenk

ERS-2005-004-LIS

<http://hdl.handle.net/1765/1928>

A note on the dual of an unconstrained (generalized) geometric programming problem

J.B.G.Frenk and G.J.Still

ERS-2005-006-LIS

<http://hdl.handle.net/1765/1927>

Privacy Metrics And Boundaries

L-F Pau

ERS-2005-013-LIS

<http://hdl.handle.net/1765/1935>

Privacy Management Contracts And Economics, Using Service Level Agreements (Sla)

L-F Pau

ERS-2005-014-LIS

<http://hdl.handle.net/1765/1938>

A Modular Agent-Based Environment for Studying Stock Markets

Katalin Boer, Uzay Kaymak and Arie de Bruin

ERS-2005-017-LIS

<http://hdl.handle.net/1765/1929>

Lagrangian duality, cone convexlike functions

J.B.G. Frenk and G. Kassay

ERS-2005-019-LIS

<http://hdl.handle.net/1765/1931>

Operations Research in Passenger Railway Transportation

Dennis Huisman, Leo G. Kroon, Ramon M. Lentink and Michiel J.C.M. Vromans

ERS-2005-023-LIS

<http://hdl.handle.net/1765/2012>

Agent Technology Supports Inter-Organizational Planning in the Port
Hans Moonen, Bastiaan van de Rakt, Ian Miller, Jo van Nunen and Jos van Hillegersberg
ERS-2005-027-LIS
<http://hdl.handle.net/1765/6636>

Faculty Retention factors at European Business Schools
Lars Moratis, Peter van Baalen, Linda Teunter and Paul Verhaegen
ERS-2005-028-LIS
<http://hdl.handle.net/1765/6559>

Determining Number of Zones in a Pick-and-pack Orderpicking System
Tho Le-Duc and Rene de Koster
ERS-2005-029-LIS
<http://hdl.handle.net/1765/6555>

Integration of Environmental Management and SCM
Jacqueline Bloemhof and Jo van Nunen
ERS-2005-030-LIS
<http://hdl.handle.net/1765/6556>

On Noncooperative Games and Minimax Theory
J.B.G. Frenk and G.Kassay
ERS-2005-036-LIS
<http://hdl.handle.net/1765/6558>

Optimal Storage Rack Design for a 3-dimensional Compact AS/RS
Tho Le-Duc and René B.M. de Koster
ERS-2005-041-LIS
<http://hdl.handle.net/1765/6730>

Strategies for Dealing with Drift During Implementation of ERP Systems
P.C. van Fenema and P.J. van Baalen
ERS-2005-043-LIS
<http://hdl.handle.net/1765/6769>

Modeling Industrial Lot Sizing Problems: A Review
Raf Jans and Zeger Degraeve
ERS-2005-049-LIS
<http://hdl.handle.net/1765/6912>

Cyclic Railway Timetabling: a Stochastic Optimization Approach
Leo G. Kroon, Rommert Dekker and Michiel J.C.M. Vromans
ERS-2005-051-LIS

* A complete overview of the ERIM Report Series Research in Management:
<https://ep.eur.nl/handle/1765/1>

ERIM Research Programs:
LIS Business Processes, Logistics and Information Systems
ORG Organizing for Performance
MKT Marketing
F&A Finance and Accounting
STR Strategy and Entrepreneurship