wctr'04
4-8 july 2004 / istanbul-turkey

RELIABILITY AND HETEROGENEITY OF RAILWAY SERVICES

Michiel Vromans[a], Rommert Dekker[b], Leo Kroon[*a,c]
[a]Erasmus University Rotterdam
School of Management, Department of Decision and Information Management
P.O. Box 1738, 3000 DR Rotterdam, the Netherlands
[b]Erasmus University Rotterdam
School of Economics, Department of Econometrics and Operations Research
P.O. Box 1738, 3000 DR Rotterdam, the Netherlands
[c]Netherlands Railways
Logistics Department, HGB IV, room 9.30a
P.O. Box 2025, 3500 HA Utrecht, the Netherlands

mvromans@fbk.eur.nl, rdekker@few.eur.nl, lkroon@fbk.eur.nl

Abstract

Reliability is one of the key factors in transportation, both for passengers and for cargo. This paper examines reliability in public railway systems. Reliability of railway services is a complex matter, since there are many causes for disruptions and at least as many causes for delays to spread around in space and time.

One way to increase the reliability is to reduce the propagation of delays due to the interdependencies between trains. In this paper we attempt to decrease these interdependencies by reducing the running time differences per track section, i.e. by creating more homogeneous timetables, as opposed to the present day heterogeneous ones.

Because of the complexity of railway systems, we use network wide simulation for the analysis of the alternative timetables. We report on both theoretical and practical cases. Besides a comparison of different timetables, also general timetabling principles are deduced.

Keywords: Railways; Transportation; Reliability; Heterogeneity; Simulation
Topic Area: E2 Performance Measurement

1. Introduction

Railway infrastructure capacity is limited and has therefore to be used carefully. Due to the increased utilization of the railway infrastructure in the Netherlands over the past years, the railway system has become quite vulnerable to disruptions. This has resulted in a low punctuality and in many customer complaints. Also in many other countries the reliability of the railway traffic is an important topic.

In the Netherlands, the railway system is characterized by many interconnected relations. Passenger transfers, rolling stock circulations and crew schedules all play their role in the relations between trains. However, the shared use of the same infrastructure by different railway services, with different origins and destinations, different speeds, and different dwell patterns, is probably the main reason for the propagation of delays

---

[*] Corresponding Author

throughout the network. This makes such a national railway system much more vulnerable to disruptions than metro systems.

Reliability is, together with door-to-door travel time, one of the predominant performance measures in railway traffic. The predictability of the arrival times is a big factor in deciding to use rail or road transport, both for passengers and for cargo.

In this paper we investigate the effect of the heterogeneity of the timetable on the reliability of the railway system. This is a subject on which a lot of practical intuition exists, but hardly any scientific literature can be found. Railway traffic is considered to be heterogeneous when trains have large running time differences on the same track sections. When running times per track section are more or less equal for all trains, then the timetable is called homogeneous. The research goal of this paper implies that we are mainly interested in relatively small initial (primary) disturbances, because no reasonable timetable is robust enough to handle large disturbances or disruptions without severe on-line adjustments of the railway traffic.

Many characteristics and details of a railway timetable have their influence on the reliability. Therefore, to understand the relations properly, one first needs to know how a timetable is constructed (Figure 1). Starting from the market demand, line planning is the first phase in railway planning, where train connections are determined: starting and ending stations of lines are chosen, including the routes, and the stations at which has to be stopped for alighting and boarding. The chosen lines determine the service differentiation – intercity trains, local trains, cargo, etc.– and consequently a large part of the heterogeneity of the timetable. The line planning step is followed by the timetabling step, where departure and arrival times are chosen. There are often several iterations between these two steps, because a preferred line plan does not imply a feasible timetable. Feedback loops may have to be executed.

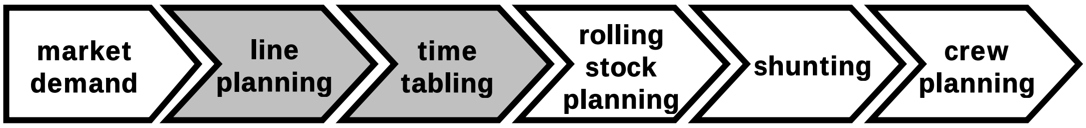

Figure 1: Sequence of interdependent railway planning phases

When the timetable is finished, the rolling stock circulation is planned. This step also includes shunting and scheduling repositioning trips. Both for the regular trains and for the shunting work, train drivers have to be scheduled. For passenger trains, also conductors are needed. Although the steps presented in the flow-diagram above depend on each other in the presented order, coordination between these steps is not easy. This means that not all the forward consequences of a certain planning step are taken into account immediately. For example, in the timetabling phase, one cannot always foresee the resulting impact on the rolling stock circulations nor the impact on the crew schedules. Again, feedback loops may be necessary.

The line plan and the timetable determine the heterogeneity of the railway system to a large extent. Since we investigate the influence of the heterogeneity on the reliability and the punctuality, we focus on the line planning and the timetabling steps.

In this paper we develop line plans which mainly differ from each other in heterogeneity. While keeping the numbers of stops per hour equal for all stations, we develop more heterogeneous and more homogeneous timetables for a practical case. A theoretical case is described by Vromans, Dekker and Kroon (Vromans et al., 2003).

To quantify the differences between these timetables, two new heterogeneity measures are proposed. These new measures follow the ideas described by Carey (Carey, 1999), but they also take into account several other characteristics of the timetable. These two new heterogeneity measures are not only used for the evaluation of the timetables, but also for the prediction of the reliability.

In this paper, detailed simulation of the railway systems is used for the comparison of the heterogeneous and the homogeneous timetables. A wide range of disturbance distributions and disturbance levels is used for this evaluation.

Because cargo trains cover less than ten percent of the Dutch railway traffic, we concentrate on passenger trains in this paper. Furthermore, we assume, as is common in several European countries, that the timetables are cyclic. This means that passenger services are repeated every cycle time, typically every hour.

When researching heterogeneity, double track sections are more relevant than single track sections or sections with four parallel tracks. Indeed, the timetable for single-track lines is mostly dictated by the distances between passing points. In the case of four parallel tracks, trains with different speeds are usually already separated, and each track has its own speed: one track for slow traffic, and one for fast traffic. The interesting part is where all trains for one direction run on one track: double track lines. Notice that many of Europe’s main lines are, at least, double track lines indeed.

This paper starts with an introduction on railway reliability, including a literature overview. This is followed in section 3 by a discussion on heterogeneity and its influence on reliability. Also the new heterogeneity measures are introduced here. Sections 4 and 5 present a theoretical and a practical case, respectively. The principle of homogenization is further discussed in section 6. Conclusions follow in section 7.

2. Punctuality and reliability

When investigating railway reliability, it is important to make a distinction between primary and secondary delays. Primary delays are initial delays caused on a train from the outside and not by other trains. These delays are caused by malfunctioning rolling stock, malfunctioning infrastructure, bad weather conditions, excessive alighting and boarding times of passengers, accidents at road-railroad crossings, and so on. Secondary delays are those delays of trains that are caused by earlier delays of other trains. They are also referred to as knock-on delays. Secondary delays appear because of the shared use of the same infrastructure, rolling stock connections, transfers in crew schedules, passenger transfers, dispatching actions, and so on. In our study we consider the primary delays as given, and we aim at developing timetables which both absorb primary delays fast and cause as few secondary delays as possible.

The measures which are chosen for evaluating the reliability are the average delays and the observed punctuality. Punctuality is probably the most widely used reliability measure in practice (Schaafsma, 2001), both in the Netherlands and abroad. This measure calculates the percentage of trains arriving within a certain number of minutes from the scheduled arrival time. In practice in the Netherlands, a three-minute margin is used. However, in most other countries –as well as for international comparisons– a five minute margin is more common.

Besides the arrival punctuality, also the departure or the start-up punctuality can be calculated. Furthermore, it is important at which stations the punctuality is measured.

In simulation research, it is quite easy to compare punctualities on different punctuality margins and on different sets of stations, train types or lines. Other possible measures for reliability, which are not considered in this paper, are the percentage of realized passenger transfers and the average delays of the passengers.

Literature overview

Over the last fifteen years, a wide range of researchers has studied railway timetabling and punctuality issues. Literature reports on different types of timetable evaluation models. In the following, we first describe the analytical delay models, starting with max-plus algebra. Thereafter, stochastic models are described, and finally the focus is on railway simulation.

Max-plus algebra is an analytical approach for evaluating the robustness of a timetable. Some relevant key characteristics, such as the minimum cycle time, can be calculated with max-plus algebra (Subiono, 2000; Goverde, Soto y Koelemeijer, 2000; Van Egmond, 2000; De Kort, 2000). PETER, an analysis tool based on max-plus algebra, is a performance evaluator for timetables (Soto y Koelemeijer et al., 2000; Goverde, Odijk, 2002). Whereas max-plus algebra cannot handle stochastic elements, Hansen (2000) uses both queuing theory and max-plus algebra to study the capacity and stability of railway systems, but only in stations.

Huisman, Boucherie and Van Dijk developed a stochastic analytical waiting line model for analyzing delays at a double track section (Huisman, Boucherie, 2001; Huisman et al., 2002). Their models are based on train frequencies and running times only, not on detailed timetables with arrival and departure times.

Higgins, Kozan and Ferreira come up with a model to quantify the risk of delays on a single track line (Higgins et al., 1995). Higgins and Kozan also developed an analytical model to quantify the expected delays of individual passenger trains in an urban rail network (Higgins, Kozan, 1998).

Carey and Kwieciński (1995) mainly focus on recovery times in their stochastic analysis. Carey also uses heuristic measures for timetable reliability (Carey, 1999) and includes behavioral response (Carey, 1998). These approaches are rather simplified, and they lack verification with reality.

Other researchers use simulation as a tool to analyze the influence of delays on the train circulation, given some traffic scenario. SIMON is a Swedish software tool using simulation of the whole network (Wahlborg, 1996; Bergmark, 1996). Amongst others, VirtuOS (König, 2001) and SABINE (Fischer et al., 1995) are used in Germany, and Open Track (Hürlimann, 2001) is a railway simulation program developed at ETH Zürich. UX-SIMU is used for simulation of railway traffic in Denmark (Kaas, 2000).

This literature is mainly focused on the simulation software itself and sometimes on a simple comparison of multiple timetables. More thorough research of the impact of timetabling principles on the corresponding reliability is hardly found.

However, Middelkoop and Bouwman describe the use of SIMONE (Middelkoop, Bouwman, 2000, 2001) for the evaluation of traffic scenarios in the Netherlands. SIMONE is capable of simulating the entire Dutch railway network. In this paper we also use SIMONE to execute theoretical analyses on the basis of simulation.

3. Homogeneity, heterogeneity, and headways

Railway traffic is considered to be homogeneous if all trains have similar characteristics, especially the *same average speed* per track segment, resulting from the running times and the stopping times. Appropriate examples of homogeneous railway traffic are metro systems where all trains have the same running times per track and where all trains stop at all stations. However, for national railway networks, railway traffic cannot be fully homogeneous. Usually cargo trains and passenger trains share the same infrastructure. But probably a more important factor is the large differentiation in passenger services, ranging from short distance trains –which dwell at nearly all stations underway–, via intercity trains, to international high speed connections –with high speeds, only stopping at a few large stations–, partly sharing the same infrastructure. If there are large differences in the timetable characteristics of the trains on the same track, then the railway traffic is called heterogeneous.

Possibilities for homogenization

Homogenization of a railway system means that differences in running times per track section of different trains along a railway line are decreased. There are several alternative options for homogenization:

- Slowing down express trains:

Decreasing the speed of an express train means longer running times for these, usually considered more prestigious, services. On the other hand, besides the homogenization effect researched in this paper, the extra running time supplement created in this way will increase reliability.

- Speeding up local or stopping services:

Decreasing running times can only be achieved by decreasing running time supplements or by using faster rolling stock. The first option can be very hurtful for reliability and is very restricted in size, the second option is probably very costly.

- Overtaking:

When slower services are overtaken by faster services, running time differences should only be regarded from or up to this overtaking station. An important prerequisite is the presence of an overtaking track: a second track for the same direction is needed. One of the disadvantages is the interdependency between both trains at the overtaking station. It also leads to a time loss of the stopping service.

- Shorter lines for the stopping services:

By decreasing the length of stopping services, stopping and express services share the same infrastructure for shorter distances and the difference in number of stops decreases. Unfortunately this leads to more passenger transfers. In theory, shorter stopping services is almost the same as overtaking, if the shorter services have the overtaking stations as start and ending points. The difference is found in the rolling stock circulations. At first glance shorter services leads to less dependencies in the network, but turning around at line-endpoints can also lead to additional delays and additional conflicting routes.

- Equalizing the number of stops:

Adding some stops to the express services leads to smaller differences between services. The small stations which are now serviced by the express trains, can be skipped by the stopping services. Repeating this until the number of stops of the stopping trains is equal to that of the express services, a maximal homogenization can be reached. One cannot speak about stopping trains and express trains anymore!

A very harsh way to equalize the number of stops is to close down some minor stations. These could be serviced by busses.

For this paper the last option is chosen: we equalize the number of stops to create more homogeneous timetables. Using different amounts of running time supplements or different kinds of rolling stock makes comparisons between heterogeneous and homogeneous timetables unfair. This discards the options of slowing down express trains or speeding up local trains. Additional overtaking or shortening local services is often impossible without large investments for new sidings, where trains can be overtaken orturn around. The final option seems the easiest to implement on the short term and is, maybe together with additional overtakings, the most promising on the long term.

Hypothesis

When heterogeneous services share the same infrastructure over large distances, timetabling becomes very complicated. Heterogeneity usually leads to many small headway times, which may increase delay propagation in the operations. Therefore, we propose the following hypothesis, which is studied in the remainder of this paper

Hypothesis: *the heterogeneity resulting from the line plan and the timetable has a negative influence on the punctuality and the reliability of a railway system.*

In order to support this hypothesis, we first develop two heterogeneity measures. Then simulation of both theoretical and practical cases is used to show the importance of homogeneity of a timetable. Besides the fact that we use the heterogeneity measures in this paper for a theoretical comparison of different timetables, it is also intended to be useful for the development of timetables for real world operations.

Heterogeneity measures

Given the train frequency of a line, the average headway at a location along that line is simply equal to the cycle time divided by the frequency. More useful headway measures are described by Carey (Carey, 1999). He shows that equalizing scheduled headways for one station has a positive influence on the punctuality when train delay distributions are equal for all trains and sloping downward. The measures he describes are based on this principle. These measures include:

- the percentage of headways smaller than a certain size;
- the percentiles of the headway distribution;
- range, standard deviation, variance, or mean absolute deviation of the headways.

The further description of these measures in the mentioned paper implies that the headways are measured at one single location only.

An important disadvantage of measuring headways at only one single location is that this does not tell anything about the behavior of the trains on the surrounding track sections. Therefore we consider the smallest headways between two consecutive trains on a certain track section instead of at one single location. When the trains on a certain track section are completely homogeneous, then the sum of the smallest headways on this track section is equal to the cycle time. On the contrary, when traffic on a certain track is highly heterogeneous, then the short distance trains depart just after the long distance trains at the start of the track section, and the long distance trains arrive just after the short distance trains at the end of the track section. This leads to a small total sum of smallest headways.

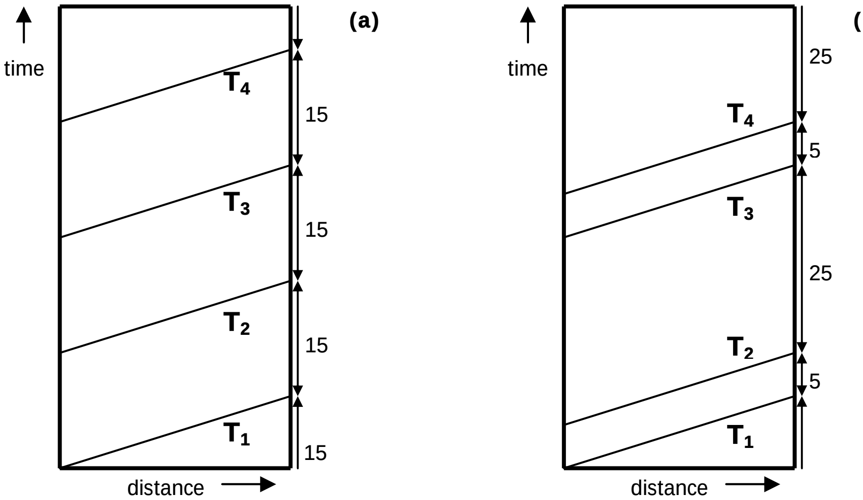

Figure 2: time-distance diagram for two homogeneous situations with a different headway distribution.

The disadvantage of just taking the sum of the smallest headways in a linear way is that it does not take into account how the trains are spread over the cycle time. With a cycle time of sixty minutes and four homogeneous trains, one will always have a total sum of (smallest) headways of 60', whether these trains are nicely spread (four 15'-intervals; Figure 2(a)) or not (e.g. 5', 25', 5' and 25'-intervals; Figure 2(b)). However, taking the sum of the reciprocals gives a clear distinction between these situations. In particular, the examples in Figure 2 lead to $\frac{1}{15}+\frac{1}{15}+\frac{1}{15}+\frac{1}{15}=0.27$, and $\frac{1}{5}+\frac{1}{25}+\frac{1}{5}+\frac{1}{25}=0.48$, respectively.

This leads to our first heterogeneity measure, based on both the heterogeneity and the spread of the trains over the hour. This measure is applicable to railway tracks between two neighboring railway nodes. The Sum of Shortest Headway Reciprocals (SSHR) is defined as follows:

$$SSHR=\sum_{i=1}^{n}\frac{1}{h_i^-} \tag{1}$$

with h_i^- the smallest scheduled headway between train i and $i+1$ on the track section, and train n is followed by train 1, due to cyclicity. It is not difficult to see that, given the number of trains on a certain track section together with their running times, for minimizing the SSHR one should minimize the running time differences between subsequent trains and one should equalize the minimum headways between subsequent trains. If the order of the trains has been pre-specified already, then one should equalize the minimum headways between subsequent trains.

As stated earlier, the SSHR is not only capable of representing the distribution of the trains over the hour on a track, but it is also capable of including the heterogeneity of these trains on this track. The homogeneous situation in Figure 2(a) gives an SSHR of 0.27. The slightly heterogeneous situation in Figure 3(a) leads to an SSHR of $\frac{1}{9}+\frac{1}{9}+\frac{1}{9}+\frac{1}{9}=0.44$. Figure 3(b) represents a very heterogeneous situation with an SSHR of $\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}=2$.

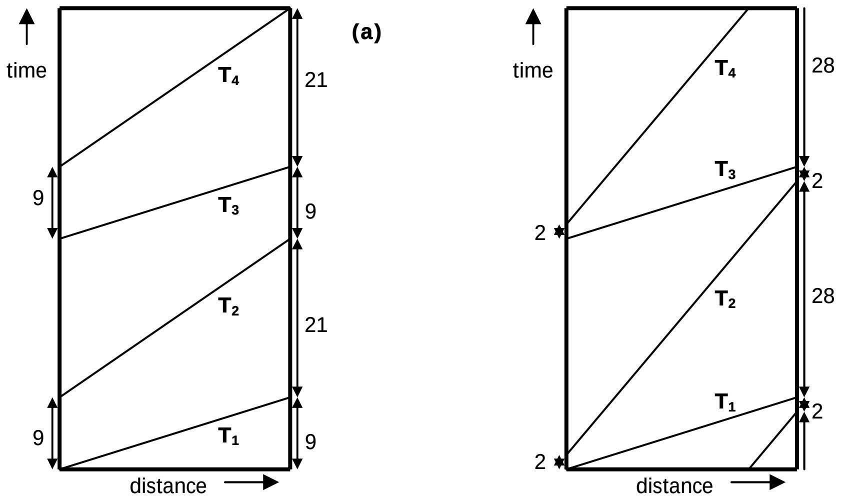

Figure 3: Two time-distance diagrams with slightly heterogeneous (a) and very heterogeneous (b) railway traffic.

A disadvantage of the SSHR is that headways at departure are penalized as heavily as headways at arrival. However, headways at arrival seem to be more important than headways at departure. The first reason is that delays at arrival are, on average, larger than delays at departure. Secondly, faster long distance trains can be caught behind short distance trains towards the end of a railway section. Therefore we propose a second measure, which only depends on the arrival headways between every pair of subsequent trains, the Sum of Arrival Headway Reciprocals (SAHR):

$$SAHR = \sum_{i=1}^{n} \frac{1}{h_i^A} \qquad (2)$$

with h_i^A the headway at arrival between train i and $i+1$.

In homogeneous cases, the SAHR is equal to the SSHR, so the SAHR is 0.27 in Figure 2(a) and 0.48 in Figure 2(b). In heterogeneous cases, the SAHR is always less than the SSHR. The timetables represented in Figures 3(a) and 3(b) have an SAHR of $\frac{1}{9}+\frac{1}{21}+\frac{1}{9}+\frac{1}{21}=0.32$ and $\frac{1}{2}+\frac{1}{28}+\frac{1}{2}+\frac{1}{28}=1.07$, respectively.

Unfortunately, the SAHR does not take into account the track section anymore and it is therefore in fact a single location measure. Still, the arrival distribution can only be *nice* if the timetable is not too heterogeneous. This means that heterogeneity is implicitly taken into account in the SAHR. However, an improved measure may be attained by taking the weighted average of the two measures above.

The two measures developed above are not absolute measures. They are mainly meant to be able to compare different timetables for the same track or as an indication how to produce a reliable timetable for a certain track.

Experiments

In this paper we study the effect of homogenization on the reliability and the punctuality of a railway system. For obtaining more homogeneous timetables, we have chosen the option of shifting stops from the short distance services to the long distance services until the numbers of stops of the trains per track section are as much equal as possible. This leads to more equal running times per track section.

The cyclic timetables were developed with the automatic timetabling tool DONS (Hooghiemstra, 1994). Any real-life or artificial railway network can be defined with this tool. On this infrastructure, train lines can be defined, including their intermediate stops, their rolling stock types, transfer connections and other characteristics. DONS will then provide a feasible timetable, or, if this is not possible, tell which constraints following from the input make a timetable infeasible.

For the comparison of the timetables, simulation of railway traffic has been used. The simulations reported on in this paper are performed with SIMONE. This simulation model is in use both by ProRail, the Dutch railway infrastructure manager, and by Netherlands Railways, the predominant Dutch operator of passenger trains. It is used both for timetable comparisons and for scientific research. A more detailed description of SIMONE is given by Middelkoop and Bouwman (2000, 2001).

4. A practical case

For the evaluation of the reliability a practical case has been worked out. We compared a real-life heterogeneous timetable with a more homogeneous timetable for a busy line in the Netherlands. Some details of the real timetable have been adjusted slightly for the simulations (Nederlandse Spoorwegen, 2002).

The case which is elaborated here consists of the lines from The Hague Central (Gvc) and Rotterdam Central (Rtd) to Utrecht Central (Ut), which merge at Moordrecht Junction (Mda). These lines are represented by the bold lines in Figure 4. This part of the network has double track everywhere, except for the section between Moordrecht Junction and Gouda Goverwelle (Gdg), which has four parallel tracks. Moordrecht Junction is a non-level crossing. The distance between The Hague and Utrecht is 61 km and the distance between Rotterdam and Utrecht is 55.8 km.

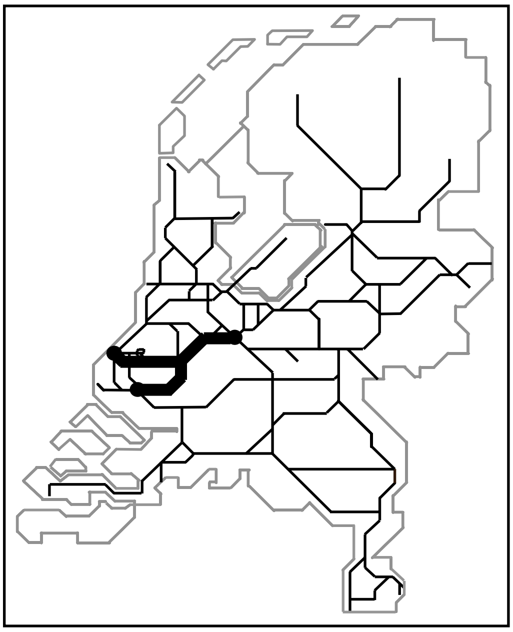

Figure 4: the railway network served by Netherlands Railways. The bold lines represent the tracks between The Hague and Utrecht and between Rotterdam and Utrecht that are considered in the case.

The lines that are operated on these tracks are shown in Figure 5. All lines have a cycle time of 30 minutes, which leads to, for example, eight trains per hour between Rotterdam Central and Moordrecht Junction, and twelve trains per hour between Woerden (Wd) and Utrecht Central.

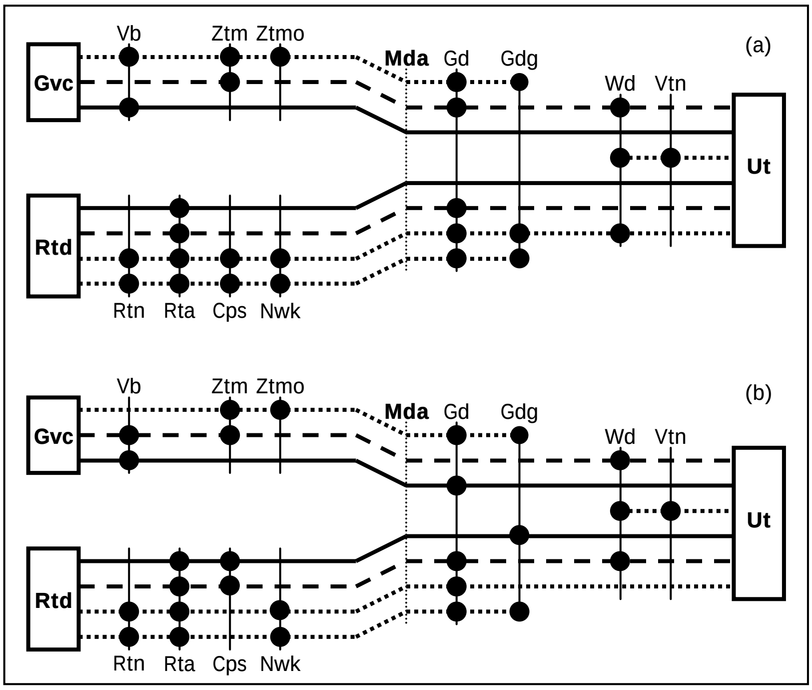

Figure 5: Train lines and dwelling patterns of the heterogeneous situation (a) and the homogeneous situation (b) of the practical case. Besides the junction Mda, all abbreviations indicate a station.

Heterogeneous situation

With some adjustments, the 2003 rush hour timetable has been taken for the heterogeneous situation (Nederlandse Spoorwegen, 2002). Cargo trains are skipped, resulting in a *three-train-system*: long distance trains, interregional trains and short distance trains, represented in Figure 5 by solid (——), dashed (– – –), and dotted lines (·········), respectively.

Every 30 minutes there is one short distance train from Gvc to Gdg, there is one interregional train from Gvc to Ut, and there is one long distance train from Gvc to Ut. Starting from Rtd, there is one short distance train running to Gdg, a second short distance train running to Ut (not dwelling at Vleuten (Vtn)), one interregional train to Ut, and one long distance train to Ut. Additionally there is a short distance train from Wd to Ut. This adds up to 16 trains per hour per direction, with a Basic One-hour Timetable given by Figure 6.

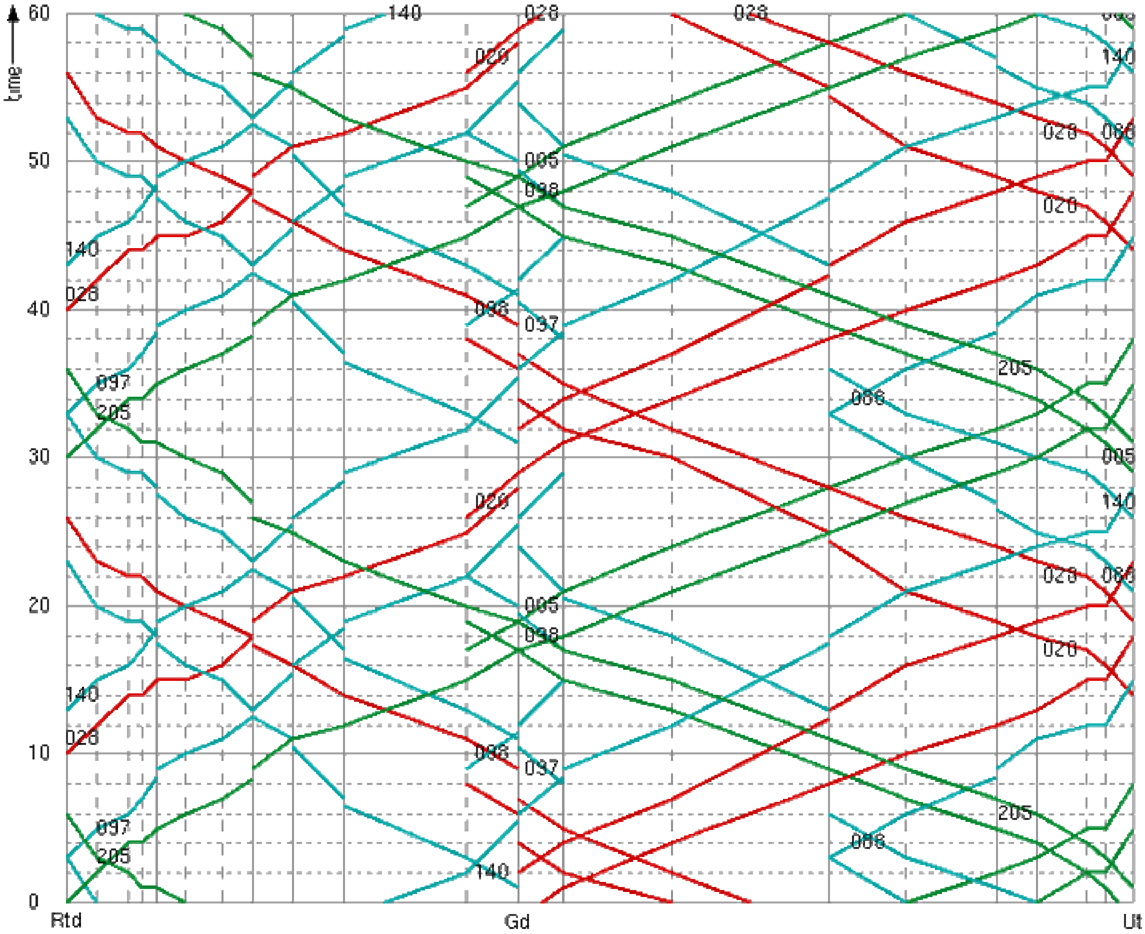

Figure 6: Basic One-hour Timetable of the heterogeneous situation for the Rotterdam branch. The solid vertical lines represent stations, the broken lines other timetabling points. The numbers in the graph are the line numbers. Note that the sections around Gd have four tracks.

Unlike the theoretical case, the lines in this case are run by different rolling stock types. These rolling stock types have their own specific characteristics concerning acceleration and top speed. These are, according to the real-life situation, matched with the service provided.

Homogeneous situation

As in the theoretical case, the heterogeneous situation is homogenized by decreasing the number of stops of the short distance services and by compensating those by additional stops of the faster services. In the end, the total number of stops per station is equal in both situations. The final dwell pattern is shown in Figure 5(b), and the accompanying time-distance diagram in Figure 7.

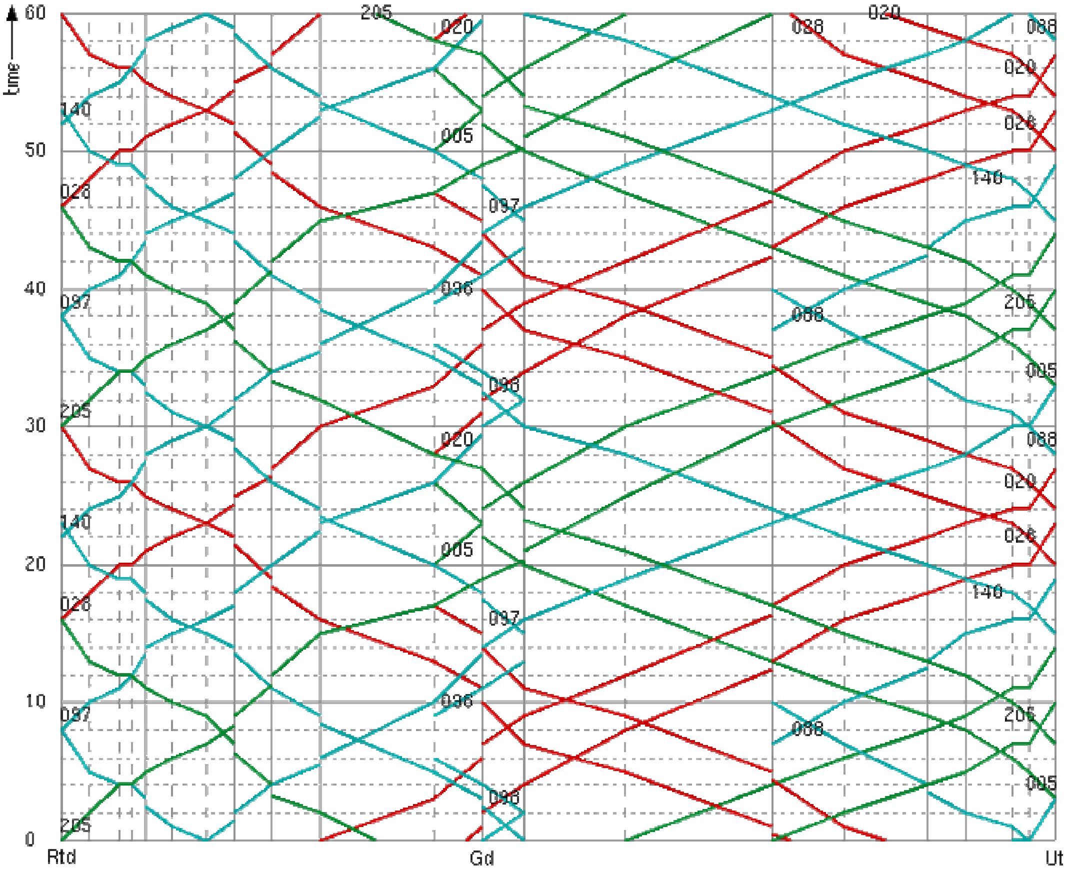

Figure 7: Basic One-hour Timetable of the homogenized situation for the Rotterdam branch.

Because of the homogenization, there is no clear distinction anymore between slower and faster services. Therefore, the necessity for different types of rolling stock has gone. However, for a fair comparison, the same rolling stock has been used for both situations. The SSHR and the SAHR for the practical case are given in table 1.

Table 1: Tables of SSHR and SAHR for the different tracks between The Hague, Rotterdam and Utrecht

SSHR from	 to	Hetero- geneous	Homo- geneous
Gvc	Gd	0.76	0.67
Rtd	Gd	3.18	1.47
Gd	Ut	3.47	2.97
Ut	Gdg	5.52	3.02
Gd	Gvc	1.06	0.69
Gd	Rtd	2.83	1.19

SAHR from	 to	Hetero- geneous	Homo- geneous
Gvc	Gd	0.64	0.62
Rtd	Gd	1.36	1.17
Gd	Ut	2.70	2.52
Ut	Gdg	2.55	1.77
Gd	Gvc	0.67	0.61
Gd	Rtd	1.74	1.07

Experimental Design

Again, the simulation experiments consist of fifty runs of 1320 minutes, including 120 minutes of warm-up time. The primary delay distributions, all exponential again, are given in table 2.

Table 2: the experiments and their primary disturbances

Experiment	DWELL TIME DISTURBANCE [all stations]		ABSOLUTE RUNNING TIME DISTURBANCE		RELATIVE RUNNING TIME DISTURBANCE		TOTAL PRIMARY DISTURBANCES (in minutes per hour)
	probability for disturbance	average disturbance (in minutes)	probability for disturbance	average disturbance (in minutes)	probability for disturbance	average disturbance (in % of running time)	
1	100%	0.6					95.9
2	100%	0.8					128.0
3	10%	1	5%	1			38.7
4	20%	1	10%	1			78.2
5	50%	0.5	20%	0.5			86.2
6	75%	0.5	30%	0.5			129.2
7					30%	40%	112.8
8					60%	20%	114.0
9	20%	1			30%	20%	89.3
10	50%	0.5			30%	20%	96.9

The first two experiments have dwell time disturbances at all stations. The next four experiments have a combination of dwell time disturbances and absolute running time disturbances. Absolute running time disturbances are independent of the running time and have the averages given in table 2. The disturbances of experiments 5 and 6 have a larger probability of occurring, but are smaller than those of experiments 3 and 4.

Table 3: the results of the practical case. The average delays have standard deviations between 0.3% and 2.4% in the heterogeneous case, and between 0.3% and 3.3% in the homogeneous case.

Experiment	AVERAGE ARRIVAL DELAY (per train measurement in minutes)			3-MINUTE DISPUNCTUALITY (% of trains delayed)			TOTAL INCURRED SECONDARY DELAYS (in minutes per hour)		
	Heterogeneous	Homogeneous	*improvement (relative in %)*	Heterogeneous	Homogeneous	*improvement (relative in %)*	Heterogeneous	Homogeneous	*improvement (relative in %)*
1	1.08	0.72	*33.4*	8.8	3.1	*64.6*	30.0	3.2	*89.3*
2	1.85	1.22	*34.0*	21.1	10.4	*50.8*	50.3	8.4	*83.4*
3	0.60	0.40	*33.7*	5.0	3.0	*39.9*	20.9	4.2	*80.2*
4	1.40	0.92	*34.5*	15.0	9.0	*39.9*	41.3	10.8	*74.0*
5	1.13	0.75	*33.6*	8.7	3.9	*55.4*	28.8	3.9	*86.6*
6	2.02	1.35	*33.2*	23.1	12.0	*48.2*	47.5	8.7	*81.7*
7	2.21	1.55	*29.9*	25.9	17.6	*32.1*	60.2	22.9	*62.0*
8	1.71	1.22	*28.6*	18.3	10.8	*41.1*	38.7	10.4	*73.1*
9	1.30	0.89	*31.4*	12.4	7.5	*39.6*	35.0	8.7	*75.2*
10	1.28	0.88	*31.0*	11.4	5.9	*48.0*	32.1	6.3	*80.4*

Experiments 7 and 8 have relative running time disturbances, which depend on the scheduled running times of the trains on the track. The average delay equals a certain percentage of the running time. Although the total primary delays are equal, experiment 8 has more but smaller disturbances than experiment 7. Experiments 9 and 10 have both dwell time disturbances and relative running time disturbances.

Table 3 shows the results of the simulations. The homogeneous situation leads to a significantly better reliability than the heterogeneous situation.

The approximately 30% decrease in average arrival delay is quite substantial, but the most surprising result is the consistency of this size for the different delay distributions. The average delay per train decreases with 28% to 35% for all experiments.

Obviously, because the primary delays are equal for the heterogeneous and the homogeneous case, the difference is made by the secondary delays. Delay propagation drops by 62% to 90%.

Another point worth noticing is the fact that a few large primary delays have a larger negative impact on the reliability than many small primary delays with the same total size of the initial delays. See for example experiments 4 and 5. One explanation is the fact that the timetable can absorb one small delay here and another small delay somewhere else. However, it may be impossible to absorb one large delay.

A final remark should be made on the dwell time disturbances. These disturbances are randomly generated at all stops, which means that local stop trains are more vulnerable for this kind of disturbance. Implicitly, average realized running time differences are increased, which means that this disturbance has an additional heterogenization effect.

SSHR, SAHR and the results

Furthermore, Figures 8 and 9 show the improvements in the SSHR and in the SAHR in comparison with the delay reductions. For this comparison, the network has been divided into six sections: Gvc-Gd, Rtd-Gd and Gd-Ut and vice versa. The average delay for each section, as given in the Figures, is the average delay of all trains at the endpoint of the section minus the average delay of these trains at the start of the section. Thus it is the *increase* in average delay on the corresponding section.

The upper right of each line-segment in Figure 8, represented by a circle, is the SSHR and the average delay in the homogeneous situation. The other endpoint, represented by a square, is the result of the homogenization. Figure 9 shows the same comparison for the SAHR. The graphs represent the results of experiment 7, but the graphs for the other experiments are quite similar.

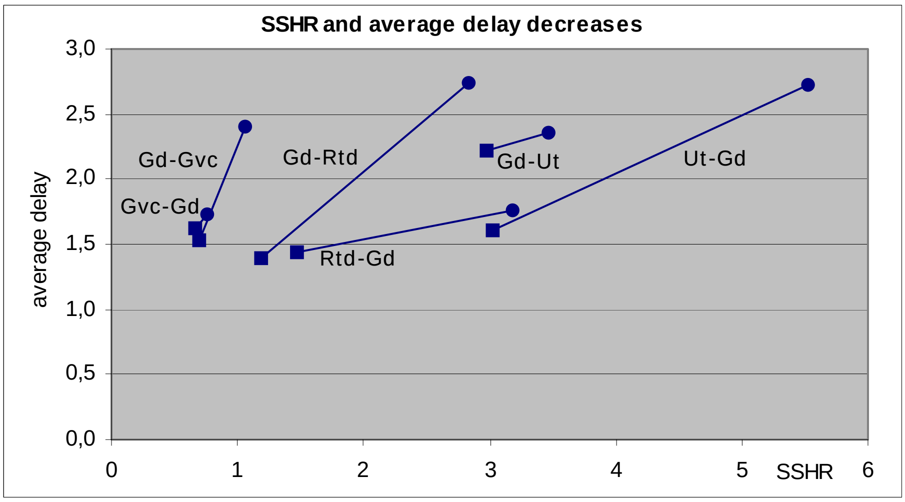

Figure 8: The relation between the decrease in SSHR and the reduction of the average delay.

Figures 8 and 9 show that homogenization of the timetable leads to a reduction both in the SSHR and in the SAHR. It also shows that it leads to a reduction in the average delays on all track sections.

The decrements in the SSHR are quite different for the different track sections: Gvc-Gd only shows a small difference, while Ut-Gd shows a large decrease. The same is true for the SAHR.

The relative reductions in the SSHR are not equal, or almost equal, to those of the SAHR. This means that the SSHR and the SAHR are two rather distinct measures. See for example the track section Rtd-Gd.

In general, a larger decrease in the SSHR leads to a larger reduction in the delays. Still, the line-segments in Figure 8 for the track sections Rtd-Gd and Gd-Ut are rather flat. This means that the delays do not always decrease as much as the SSHR suggests. Therefore the SSHR can be used as an indication in what direction the reliability goes, but it is not an absolute measure.

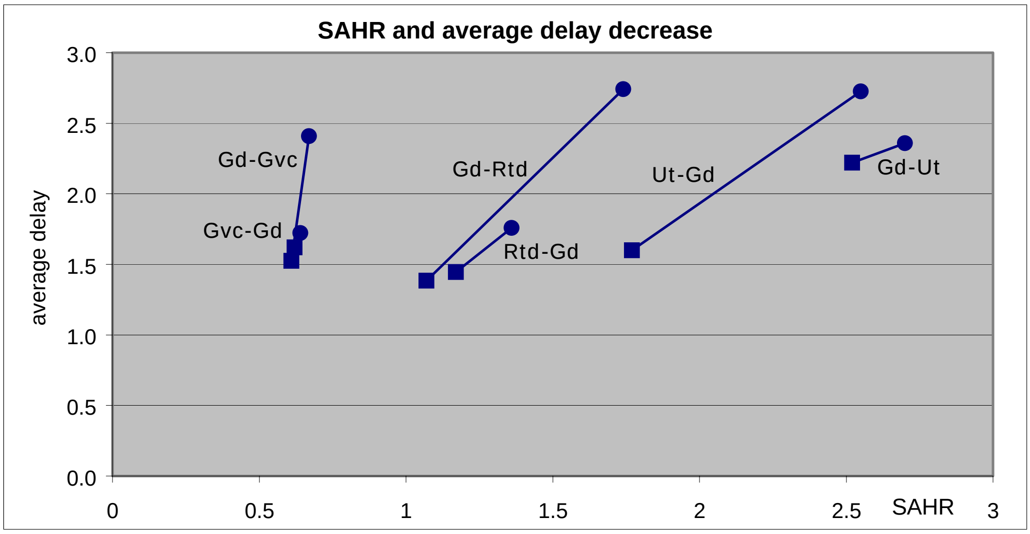

Figure 9: The relation between the decrease in SAHR and the reduction of the average delay.

For the starting sections, the departure reliability from the first stations is very high. This means that small departure headways hardly have a negative influence on the reliability. The reduction of the SSHR on Rtd-Gd is mainly based on a fairer distribution of the departures; arrivals are hardly affected by the homogenization. This may explain the relatively small reliability improvement of this section.

Also, a larger decrease in the SAHR leads to a larger reduction in delays. Still, sections with a relatively large delay reduction can be observed (e.g. Gd-Gvc), as well as sections with a relatively small delay reduction (e.g. Gd-Ut).

Due to the large reduction of the delays on the section Ut-Gd, predicted by the SAHR, there is also a large reduction in the departure delays for the section Gd-Gvc. A better starting reliability implies fewer secondary delays, which explains the large delay reduction on Gd-Gvc, and, to a lesser extent, on Gd-Rtd.

Both Figure 8 and Figure 9 have a minimal level of average delay for all sections. This is due to the amount of disturbances, which exceeds the amount of running time supplements. The running time supplements equal 7% of the running times, whereas the disturbances average 12% of the running times. This means that running times are, on average, 5% extended. Whereas all considered sections of around 20 minutes, this amounts to approximately 1 minute as the lower bound for average delays.

5. Discussion

The measures: SSHR and SAHR

The measures SSHR and SAHR are both able to predict reliability changes. The SSHR is applicable to track sections between stations, whereas the SAHR can be used for a station, or for all arrivals from a certain track at that station.

Although a large decrease in the SSHR or in the SAHR leads to a large decrease in delays in most cases, it is hard to predict the exact size of the delay reduction. Using a weighted average of the two measures may be advantageous, because it takes heterogeneity into account, and it weighs the arrival headways more heavily than the departure delays.

Equalizing headways

Minimizing the SSHR or the SAHR implies equalizing the headways. Although a reduction in these measures indicates a reduction in delays, minimizing the sum of SSHRs or SAHRs over the network is not necessarily optimal. This can for example be seen from the SSHR in the practical case, where a large reduction on one section (Rtd-Gd) has much less influence on the reliability than a small reduction of the SSHR on another track (Gd-Gvc).

Utilize and Build

“Benutten en Bouwen” (Utilize and Build) is the vision and the intended direction of the combined Dutch railway branch for the future, up to 2020. Experts from the ministry of Traffic and Waterworks, the railway infrastructure manager ProRail, the passenger operator Netherlands Railways, and the cargo operator Railion (Nederlandse Spoorwegen et al., 2003) have participated in this project. The main problem is how to facilitate the ever expanding railway traffic on the limited infrastructure. The starting point of the project is to better utilize the existing infrastructure, which is made feasible by means of small but smart infrastructure investments.

Homogenization of the railway system is one of the basic elements of “Benutten en Bouwen”. Although the main focus is on a limited homogenization where intercity trains stop at a few more stations of medium large size, full homogenization is also discussed and not excluded as a solution.

Other consequences of homogenization

Although reliability will increase when train services are homogenized, there are several other important characteristics to be considered both for passengers and operators. Homogenization can have its influence on many of those characteristics.

- *Travel time* for passengers is an important determinant of *service quality* in case of homogenization. The planned travel time may decrease for some passengers, but increase for others. The *number of passenger transfers* and the *transfer times* may also change. This requires a further mobility analysis, which falls outside the scope of this paper.
- *Infrastructural needs* can possibly change due to other train lengths, but also due to other locations for overtakings, and due to another way of coordinating trains at large transfer stations.

When the timetable is homogenized, the *rolling stock* can be standardized as well. The total required number of rolling stock units can also change.

Homogenization by one large operator may lead to *additional time-slots* in the timetable, which might be assigned to other operators. Evidently this would, in the end, lead to an increased SSHR and SAHR. Therefore, network wide cooperation is necessary for a beneficial introduction of a homogenized timetable.

Besides timetabling, the rolling stock and crew schedules can also have a considerable influence on delay propagation.

6. Conclusions

In this paper two timetable characteristics, the Sum of Shortest Headway Reciprocals (SSHR) and Sum of Arrival Headway Reciprocals (SAHR), were introduced. These measures are used for evaluating the heterogeneity of the timetable and for the prediction of the reliability. The SSHR can be applied to a whole section and has the desirable property that it decreases both when trains are spread better over the hour and when railway traffic is more homogeneous. The SAHR also has the property of decreasing when the trains are spread better over the hour. It lacks a direct link to the heterogeneity, but takes it into account implicitly.

The presented case shows a large reliability increase for homogenized services, which supports the hypothesis that we stated in section 3: *the heterogeneity resulting from the line plan and the timetable has a negative influence on the punctuality and the reliability of a railway system.*

In other words, when the SSHR and SAHR show large decreases, then there are usually also large decreases in delay propagation. Therefore, a relatively simple rule of thumb for timetable design is to minimize the SSHR and the SAHR. This may improve the reliability of the offered services.

Although homogenization may lead to a sizable increase in punctuality of the offered railway services, homogenization may also effect other features of the railway product, both for passengers and for operators and infrastructure managers. When homogenizing train services, these other consequences should also be considered. This is a subject for further research. The relationship between the consequences for the different operators and the infrastructure managers also stresses the importance of cooperation between these parties.

References

Bergmark, R.,1996. Railroad capacity and traffic analysis using SIMON. In: Brebbia, C.A., Allan, J., Hill, J., Sciutto, G., Sone, S. (Eds.), Computers in Railways V, Computational Mechanics, Southampton, pp. 183–191.

Carey, M., 1998. Optimizing scheduled times, allowing for behavioral response. Transportation Research Part B, 32 (5) 329–342.

Carey, M., 1999. Ex ante heuristic measures of schedule reliability. Transportation Research Part B, 33 (7), 473–494.

Carey, M., Kwieciński A., 1995. Properties of expected costs and performance measures in stochastic models of scheduled transport. European Journal of Operational Research, 83, 182–199.

De Kort, A.F., 2000. Advanced railway planning using (max,+) algebra. In: Allan, J., Hill, R.J., Brebbia, C.A., Sciutto, G., Sone, S. (Eds.), Computers in Railways VII, WIT Press, Southampton, pp. 257–266.

Fischer, H., Bolemant, M., Margaritov, A., Friedrich, G.-R., 1995. Das Projekt SABINE – Simulation und Auswertung des Betriebsablaufs im Netz. Simulation und Simulatoren für den Schienenverkehr, 451-462. In German.

Goverde, R.M.P., Hooghiemstra, G., Lopuhaä, H.P., 2001. Statistical Analysis of Train Traffic; The Eindhoven Case. Delft University Press, Delft.

Goverde, R.M.P., Odijk, M.A., 2002. Performance evaluation of network timetables using PETER. In: Allan, J., Hill, R.J., Brebbia, C.A., Sciutto, G., Sone, S. (Eds.), Computers in Railways VIII, WIT Press, Southampton, pp. , 89–108 , 89–108731–740.

Goverde, R.M.P., Soto y Koelemeijer, G., 2000. Performance evaluation of periodic railway timetables: theory and algorithms. Delft University Press, Delft.

Hansen, I.A., 2000. Station capacity and stability of train operations. In: Allan, J., Hill, R.J., Brebbia, C.A., Sciutto, G., Sone, S. (Eds.), Computers in Railways VII, WIT Press, Southampton, pp. 809–816.

Higgins, A., Kozan, E., 1998. Modeling train delays in urban networks. Transportation Science, 32 (4), 346–357.

Higgins, A., Kozan, E., Ferreira, L., 1995. Modeling delay risks associated with train schedules. Transportation Planning and Technology, 18, 89–108.

Hooghiemstra, J.S., 1994. An automatic timetable generator for the Dutch railway network. In: Brebbia, C.A., Murthy, T.K.S., Mellitt, B., Sciutto, G., Sone, S. (Eds), Computers in Railways IV (2), Computational Mechanics, Southampton, pp. 109–116.

Huisman, T., Boucherie, R.J., van Dijk, N.M., 2002. A solvable queueing network model for railway networks and its validation and applications for the Netherlands. European Journal of Operational Research, 142, 30–51.

Huisman, T., and Boucherie, R.J., 2001. Running times on railway sections with heterogeneous train traffic. Transportation Research Part B, 35, 271–292.

Hürlimann, D., 2001. Objektorientierte Modellierung von Infrastrukturelementen und Betriebsvorgängen im Eisenbahnwesen (Object oriented modeling of infrastructure elements and business processes in railways). PhD. thesis, ETH Zürich. In German.

Kaas, A.H., 2000. Punctuality model for railways. In: Allan, J., Hill, R.J., Brebbia, C.A., Sciutto, G., Sone, S. (Eds.), Computers in Railways VII, WIT Press, Southampton, pp. 809–816.

König, H., 2001. VirtuOS, Simulieren von Bahnbetrieb. Betrieb und Verkehr, 50(1-2), 44–47. In German.

Middelkoop, D., Bouwman, M., 2000. Train network simulator for support of network wide planning of infrastructure and timetables. In: Allan, J., Hill, R.J., Brebbia, C.A., Sciutto, G., Sone, S. (Eds.), Computer in Railways VII, WIT Press, Southampton, pp. 267–276.

Middelkoop, D., Bouwman, M., 2001. SIMONE: large scale train network simulations. In: Peters, B.A., Smith, J.S., Medeiros, D.J., Rohrer, M.W. (Eds.), Proceedings of the 2001

Winter Simulation Conference 2001, Institute of Electrical and Electronic Engineers, Piscataway, New Jersey, pp. 1042–1047.

Nederlandse Spoorwegen, 2002. Spoorboekje (Timetable), 15 december 2002 t/m 13 december 2003. Nederlandse Spoorwegen, Utrecht. In Dutch.

Nederlandse Spoorwegen, ProRail, Railion and ministerie van Verkeer en Waterstaat, 2003. Benutten en Bouwen, het plan van de spoorsector (Utilize and Build, the plan of the railway branch), Utrecht. In Dutch.

Schaafsma, A., 2001. Dynamisch railverkeersmanagement (Dynamic rail traffic management). PhD. thesis, Technical University Delft, Delft. In Dutch.

Soto y Koelemeijer, G., Iounoussov, A.R., Goverde, R.M.P. and Van Egmond, R.J., 2000. PETER, a performance evaluator for railway timetables. In: Allan, J., Hill, R.J., Brebbia, C.A., Sciutto, G., Sone, S. (Eds.), Computers in Railways VII,. WIT Press, Southampton, pp. 405–414.

Subiono, 2000. On classes of min-max-plus systems and their applications. PhD. thesis, Technical University Delft, Delft.

Van Egmond, R.J., 2000. An algebraic approach for scheduling train movements. Presented at the 8th International Conference on Computer-Aided Scheduling in Public Transport, Berlin.

Vromans, M.J.C.M., Dekker, R., Kroon, L.G., 2003. Reliability and heterogeneity of railway services. ERIM Research Series 2003 – 090, Rotterdam.

Wahlborg, M., 1996. Simulation models: important aids for Banverket's planning process. In: Brebbia, C.A., Allan, J., Hill, J., Sciutto, G., Sone, S. (Eds.), Computers in Railways V (1), Computational Mechanics, Southampton, pp. 175–181.