Automated real-time railway traffic control
an experimental analysis of reliability, resilience and robustness

Corman, Francesco; Quaglietta, Egidio; Goverde, Rob M.P.

DOI
10.1080/03081060.2018.1453916

Publication date
2018

Document Version
Accepted author manuscript

Published in
Transportation Planning and Technology

Postscript of paper published in: *Transportation Planning and Technology*, 41(4), 421–447, 2018.
https://doi.org/10.1080/03081060.2018.1453916

Automated real-time railway traffic control: An experimental analysis of reliability, resilience and robustness

Francesco Corman (1), Egidio Quaglietta (2), Rob M. P. Goverde (3)

(1) Institute for Transport Planning and Systems, ETH Zürich, Zurich, Switzerland;

(2) TRENOLab srls, Gorizia, Italy;

(3) Department of Transport and Planning, Delft University of Technology, Delft, The Netherlands.

francesco.corman@ivt.baug.ethz.ch; e.quaglietta@trenolab.com; r.m.p.goverde@tudelft.nl

Abstract

Railway transportation provides sustainable, fast and safe transport. Its attractiveness is linked to a broad concept of service reliability: the capability to adhere to the timetable also in presence of delays perturbing traffic. To counter these phenomena, real-time rescheduling can be used, changing train orders and times, according to rules of thumb, or mathematical optimization models, minimizing delays or maximizing punctuality. In literature different indices of robustness, reliability and resilience are defined for railway traffic. We review and evaluate those indices applied to railway traffic control, comparing optimal rescheduling approaches such as Open Loop and Closed Loop control, to a typical First-Come-First-Served dispatching rule, and following the timetable (no-action). This experimental analysis clarifies the benefits of automated traffic control for infrastructure managers, railway operators and passengers. The timetable order, normally used in assessing a-priori reliability in CBA, systematically overestimates unreliability of operations that can be reduced by real-time control.

Keywords: reliability, robustness, railway traffic, scheduling, closed loop control.

Acknowledgments

The research contained in this paper is partly supported by the EU FP7 project ON-TIME (www.ontime-project.eu), and the State Key Laboratory of Rail Traffic Control & Safety (contract No RCS2015K003), Beijing Jiaotong University.

Francesco Corman did this research while being with the Department of Maritime and Transport Technology, Delft University of Technology, Delft, The Netherlands; and at the Center for Industrial Management Katholieke Universiteit Leuven, Leuven Belgium. Egidio Quaglietta did this research while being with the Department of Transport and Planning, Delft University of Technology, Delft, The Netherlands; and at the group Control, Command & Signalling, Network Rail, The Quadrant, Elder Gate, Milton Keynes, United Kingdom.

1 Introduction and Background

Railway transportation is playing an increasing mobility role, thanks to its high capacity, low emissions, and high safety levels. Nevertheless, the attractiveness of the railway transport mode is linked to its ability in mitigating the propagation of delays that extend actual train travel times beyond those planned. In fact, railway operations are affected by unforeseen disturbances (e.g. extensions of dwell times at stations, unplanned stops at red signals) that induce deviations from the timetable and thereby reduce performances (e.g. punctuality). This issue is particularly relevant for those networks that run under a strong economic pressure and increase efficiency by squeezing train paths into a limited infrastructure capacity. Policy rules and white papers on transport are strongly suggesting the direction of increasing resource efficiency of the network (European Commission, 2011). To this end, relevant examples are countries like Switzerland, the Netherlands and Japan, that have a network utilization, in terms of train km, or passenger km, that is the highest in the world. Differently, when looking at the reliability of complete transport chains (Rietveld et al 2001), networks that are over-utilised result very easily in delays and unreliable operations.

To counter such delays, (deviations of trains from the schedule), typical possibilities include offline and online actions (Hansen and Pachl, 2014). Offline actions mostly relate to robust timetabling, i.e. defining service plans which are able to absorb small statistical variations in train operations (e.g. extensions of dwell and/or running times) (Dewilde et al 2011). In fact, this assumes that trains might run delayed but their orders are kept as scheduled; this kind of action is termed re-timing (Takeuchi and Tomii 2005). When larger perturbations affect traffic, time allowances in the timetable are not enough to absorb service deviations, and large-scale delay propagation (snowball effect of trains delaying each other) is experienced.

To avoid or reduce delay propagation, it is necessary to update train services online (rescheduling). A rescheduling plan involves retiming and/or, reordering trains, and is currently determined on the basis of rules-of-thumb or the experience of the dispatcher, with the aim of restoring the original timetable as soon as possible. These plans can be however ineffective or counterproductive due to the limited view that the human dispatcher has of downstream traffic behaviour. Lately, several approaches have been proposed to automatically solve the rescheduling problem by using mathematical models. A main stream of railway rescheduling models is given by Törnquist et al. (2007), Corman et al. (2011), Pellegrini et al. (2014), Lamorgese and Mannino (2013), Meng and Zhou (2014), Caimi et al (2012). Such approaches formulate the rescheduling problem in different ways adopting diverse objective functions and algorithms to solve it. A wider review on these models can be found in Cacchiani et al. (2014), Corman and Meng (2014), Narayanaswami and Rangaraj (2011).

Practitioners are still sceptic about using automated rescheduling for optimal traffic control, since this has only been tested in laboratory environments but not in real operations. Only scarce real-life installations can be mentioned, that apart from the Lötschberg base tunnel in Switzerland (Metha et al. 2010) go hardly beyond pilot tests (e.g. Mazzarello and Ottaviani, 2007, Mannino and Mascis, 2009, Lamorgese and Mannino, 2013). Many relevant aspects are therefore still unclear when these tools interact with real traffic phenomena and daily stochastic disturbances to operations. In fact, the main drawback of the majority of the works proposed in literature is that they do not consider dynamics of uncertainty, i.e. information on disturbances is perfect, immutable, and completely available beforehand. Uncertain information and unknown disturbances are instead the actual source of unreliable and/or non-robust operations in railway traffic control. Neglecting these factors, and the reaction of the system to uncertain events, constitutes a clear gap in the literature. It is unclear how these factors affect the robustness and the reliability of optimal plans.

The goal of this paper is the comprehensive analysis of these factors from the point of view of the benefits in terms of performance indicators and metrics, for a variety of stakeholders (infrastructure managers, railway operators, passengers). In fact, despite robustness and reliability of operational traffic in relation to the actual rescheduling plan is a hot topic in railway traffic, no agreement is found in literature about these concepts, yet. We aim at covering many of the definitions put forward by the academic community and in practice for these two concepts. In general reliability is intended as the capability of a plan to achieve acceptable traffic performance (such as punctuality, or generalized cost of passengers) also when stochastic disturbances affect traffic. The most recent technologies implemented in the railway world, in urban areas, typically impact reliability of operations under delays, rather than free-flow travel time (Van Oort et al 2015, Goverde et al 2014). Concerning robustness, concepts are more fuzzy. Meng and Zhou (2011) for instance consider robustness as the capability of a rescheduling plan to remain unchanged when it is implemented to traffic conditions which are stochastically known.

The present paper makes a comprehensive analysis of how uncertain factors in railway traffic influence robustness and reliability of service plans. We build upon the experience acquired in the ON-TIME project, using the same framework to analyse how uncertain and unknown information on traffic affect optimal railway plans, depending on a variety of parameters and modelling decisions. This paper is thus a complementary work to the framework definition, introduced in Corman and Quaglietta (2015), and the overall results of the projects in real-life pilots, presented in Quaglietta et al (2016). Compared to those two papers, the contributions are the focus of analysing and understanding the implications for an implementation point of view (i.e. for dispatchers), operations point of view (i.e. for infrastructure managers, and railway operators), and realised operations (i.e. for passengers). In particular, the latter point of view is also put in context of policy and planning decisions, regarding valuation of reliability of planned transport systems. The quantitative methods proposed and analysed in this paper complement the offline approaches of capacity planning, and determination, and infrastructure access (see for instance Burdett and Kozan, 2005), timetable design (Ke et al 2015) and their integration (Lindfelt, 2011) with the perspective of online control. In this sense, this paper addresses many of the issues raised in Watson (2000).

The rest of the paper is outlined as follows. Section 2 presents a broad overview of concepts of robustness, reliability and resilience as defined in literature. Section 3 reports on the approach and framework that we built up to analyse the different railway traffic control schemes. Section 4 and 5 present the metrics actually used, and the test case examined. Section 6 shows obtained results, while conclusions are provided in Section 7.

2 Related research on Railway Robustness, Resilience, and Reliability

In literature different definitions of robustness, reliability and resilience of train service plans have been given. The scientific community did not achieve yet a generally agreed and shared meaning for these concepts. The main reason is that authors have a different background and provide definitions from different points of view, namely: a social-economic view (that considers passengers perception, and impacts for social benefit and policy makers), a planning perspective (which considers the expected benefits of a transport system yet to be built), a control perspective (i.e. how to setup a system into practice, under which conditions the operators will be using it positively), and an operational analysis perspective (is the control system actually improving the operations?). We go through these different standpoints as follows.

2.1 Social-economic perspective: passenger perception

Dewilde et al. (2011) exploit a passenger-centric approach and refer to the average travel time that a passenger faces as the main robustness indicator for a timetable. This is thus related to a statistical average of all passenger travel times under small perturbations that do not require re-scheduling, but only retiming. This concept has also a strong relation with the perception of the passengers in terms of minimum travel time, the generalized cost and the value of time of different activities performed.

Similar performance indicators are considered in the general literature in delay management (see Dollevoet et al. (2012) for an overview) while they are not always explicitly called robustness or reliability. In general the average travel time, or the average deviation between the planned travel time and the realized travel time, is considered. Normally, small delays only are considered, which reduces strongly the set of rescheduling solutions, and allows tractability of the problem. Schöbel and Kratz (2009) introduce explicitly the term robustness to indicate a timetable that despite train delays is able to keep planned connections without re-ordering.

OECD (2010) refers to many definitions of reliability, which mostly relate to the variance of performance, i.e. a reliable system is one such that extreme deviations from expected operations are minimized. A similar study is reported by Rietveld et al (2001).

Börjesson and Eliasson (2011) report on the valuation of delays by the customers, and especially the relation among reliability, the extent of delay, and the risk (i.e. the probability of occurrence) of delay. They claim that the average delay is not a meaningful performance of train service, when referring to the passengers' perspective. Passengers are indeed afraid of large delays which usually have a lower probability. The average delay instead systematically underestimates the value of large delays with low probability since these are combined with small delays with higher probability. A correct valuation of delays must be based on statistical indicators which attribute a higher value to those large delays with a lower probability level. The benefits from reliable operations are quantified by van Oort et al (2015) for a light rail link; they account for a large amount of the social benefits of the transportation infrastructure. Similar studies on the impact of user perception to evaluate the robustness of a transit system are presented e.g. in Parbo et al (2014). The current work is basically reporting on reliability at the level of system, rather than at the level of single users.

2.2 Operations planning perspective: system design, planning, timetabling

In literature a number of definitions of robustness are given for the offline timetable. Goverde and Hansen (2013) differentiate between stability, robustness and resilience. Stability refers to the possibility of a timetable to absorb initial and primary delays so that delayed trains return to their scheduled train paths. Robustness refers to the possibility of a timetable to cope with process time deviations due to design errors, parameter variations, and changed operational conditions, excluding any real-time rescheduling actions. A key limitation is that the timetable order is considered as fixed according the timetable design, i.e. only retiming actions could be evaluated, and not reordering. The latter is reflected in resilience, which is defined as the flexibility of a timetable to reduce secondary delays using rescheduling (retiming, re-ordering, re-routing, cancel connections, cancel trains, etc.).

Takeuchi and Tomii (2005) attribute a robustness level to each different real-time control measure. The higher the degree of freedom of the control measure, the higher is the robustness

level. For instance, retiming is described as a control measure with robustness level of 0, while reordering has a higher robustness level (1).
The research on timetable robustness starts from the work of Carey (1999) who studies the conditions under which a timetable can absorb delay, or avoid delay propagation, when small disturbances affect operations. Based on the rationale that a minimum headway distance between trains is the critical factor initiating propagation of delay, many researchers proposed robustness indices that are all (partially) non-linear functions of the headways between trains. There is relatively large literature on this topic. Example are given by Vromans et al. (2006), who consider the reciprocal of headways, Carey (1999) who refers to the minimum headway, Dewilde et al. (2014) who define a composite function where the headway appears at the denominator; Andersson et al. (2013) adopt a composite function of the minimum headway margin.
Large scale delay propagation is studied by Büker and Seybold (2012), who describe the probability function of delays as trains interacting in the network, following the planned orders of the timetable. They measure a variety of performance indicators that include punctuality, mean delay, and mean variance. Goverde (2011) studied large-scale delay propagation to analyse robustness to delays. He also assumed fixed train orders and relate the delay propagation to timetable stability and realisability, resulting in either stable delay propagation or into structural delays, periodic delays or delay explosions.
Typical simulation software packages (Kaminsky et al, 1996; Nash and Huerlimann, 2004; Janecek and Weymann, 2010) have been used to study the robustness and reliability of timetables under stochastically perturbed traffic and straightforward dispatching rules.
The interested reader can refer to Dewilde et al. (2011) and Andersson et al. (2013) for other definitions of robustness and/or robustness-related indicators used for timetable planning.

2.3 Control perspective: delivering a service under uncertainty

From the control perspective, a rescheduling plan must result in an acceptable level of service and must be insensitive to dynamic variations of traffic over time and/or unknown/unforeseen disturbances to operations (i.e. disruptions, erroneous or missing train information).

Salido et al (2008) refers to a well-known concept of control theory and defines robustness as the ability to resist to imprecision. The same author introduces two measures of plan robustness. The former relates to the percentage of disruptions that a plan can accommodate without modifying the plan, i.e. by retiming only. The latter considers the settling time, i.e. the amount of time required to recover planned operations after having introduced a delay bounded in time. This same concept is called stability by Goverde and Hansen (2013).
Meng and Zhou (2011) define robustness of rescheduling actions as the ability to take decisions under incomplete information. A recursive stochastic programming is used to compute optimal dispatching actions for solving disruptions whose duration is only known in statistical terms.
From a control point of view, the concept of "stability" also assumes a relevant importance. Lyapunov defines as stable a dynamic system whose state is always close to a point of equilibrium (Liberzon, 2005). In the field of railway traffic, the concept of stability assumes different definitions. For Goverde and Hansen (2013) the equilibrium is the timetable and stability is its ability to absorb delays so that delayed trains return to the timetable on their own. In the view of the authors, stability relates to the ability to withstand exogenous entrance delays, while robustness is the ability to withstand internal deviations such as dwell times and running

time extensions to keep close to the timetable. For larger deviations or delays traffic control is required where resilience of the timetable is essential.

Quaglietta et al. (2013) consider stability as the capability of rescheduling plans to be insensitive to dynamic changes of traffic over time, to unplanned/unknown stochastic disturbances as well as to incomplete or erroneous traffic information. Such a study has mainly implications towards nervousness and acceptance factors in human-machine interactions.

Robustness and/or reliability is otherwise linked to the deviations between simulated and planned train paths, when stochastic perturbations are input to the simulation. Many researchers define as reliability of a plan the performance level achieved by traffic when it follows the plan under disturbed conditions. For instance, Delorme et al. (2009) refer to the sum of secondary delays in a station as a reliability indicator of railway service. This is a way to assess the capacity available at stations or junctions, based on different possible orders of trains. Medeossi et al. (2011) propose to identify reliability and robustness of a timetable based on a probabilistic description of the block occupation times, which allows the estimation of conflict probability. The conflict probability is also studied by Liu and Kozan (2010) as a function of duration of railway operations. Corman et al. (2010) study the sensitivity of rescheduling plans to disturbances in simulated operations. A larger scale study of robustness of real-time control actions is presented by Larsen et al. (2014) who examine the impact of stochastic factors on dispatching actions. Those actions are computed only based on expected train information, thus neglecting any influence of errors and variability in measured traffic states.

A general review of real-time rescheduling in situations of incomplete or erroneous traffic information has been recently proposed by Corman and Meng (2014).

3 Evaluation and Metrics used

The different control schemes and the resulting plans of operations are evaluated in terms of the generic concepts of robustness and reliability by using several metrics. The metrics adopted are those defined by several authors in literature, and can be grouped into three main categories, which are then traced back further to the perspective/stakeholders behind them. The general picture and the direct link with the literature is reported in Table 1. The three columns refer to the three perspectives of section 2, namely operations, planning, perceptions. The rows refer to a general perspective, one considering only retiming; one considering retiming and reordering. The bottom row reports on the experimental contributions of the paper along the three perspectives. We next describe the metrics used, for each measure/ perspective considered.

Operations control (2.3)	Analysis, Plans (2.2)	Perception (2.1)	
Variation operations/plan, conflicts	Delays average, maximum delay punctuality	Statistics Distribution, variance, tail	
Liu & Kozan 2010 Medeossi et al 2011 Delorme et al 2009	Vromans et al 2006 Carey 1999 Dewilde et al 2011 Andersson et al 2013 Van Oort et al 2015 Parbo et al 2014	OECD 2010 Börjesson & Eliasson 2011 Rietveld et al 2001	A-priori and general concepts
Corman et al 2010 Larsen et al 2014 In general: Goverde & Hansen,	Büker & Seybold 2012 Schobel & Kratz 2009 Dollevoet et al 2012 Goverde 2011		Retiming only

Takeuchi & Tomii 2005			
Meng & Zhou 2011 Quaglietta et al 2013 Salido et al 2008 In general: Goverde & Hansen, Takeuchi & Tomii 2005	Kaminsky et al 1996 Nash & Hürlimann, 2004 Janecek & Weymann 2010		Reordering
Stability / NRR **Deviation in actions** **Deviation from paths**	**Punctuality, average delays, consecutive delay**	**Delay cost; delay risk; delay variance**	**This paper**

Table 1. Summary of literature, contributions of the paper, metrics used.

Stability measures the sensitivity of a traffic control algorithm or the resulting rescheduling plans with respect to traffic dynamics, unknown random disturbances, partial or missing information (Quaglietta et al. 2013). Similarly to the classical Lyapunov's definition in control theory, rescheduling plans are stable when they do not change (from an equilibrium point), also if they are computed with respect to different traffic conditions, affected by random disturbances, with limited or missing information. According to Quaglietta et al. (2013), we can use the following metrics for stability:

Number of Relative Reordering (*NRR*). This metric describes for a certain CheckPoint location *CP* the similarity in terms of ordering between two plans computed at consecutive stages. Considering the plan given at stage *s*, we assume that a train is reordered if it is scheduled before some train that was preceding it, in the plan provided at stage *s*-1. The value of *NRR* is then calculated by counting all reordered trains.

The *average NRR* over all the stages gives a measure of how stable in terms of reordering are the optimal plans provided by the scheduler. The lower this average the higher is the plan stability. A condition of full stability is achieved when plans computed at consecutive stages are all the same, i.e. when the average NRR is zero.

Weighted NRR gives a higher weight to train reordering happening in the close future and a lower weight to train reordering occurring farther away in the future. This metric directly translates the fact that the far future has more variability, and short term variability is more important, since it requires more immediate actions from the dispatchers. Short term variability causes also discomfort to passengers who experience unexpected changes in their trips. In detail, the amount of trains (in percentage) that are scheduled to be reordered (in the future) is divided into very urgent reorderings (within 0 and 5 minutes from current time), quite urgent reorderings (between 5 and 10 minutes from current time), and less urgent reorderings (more than 10 minutes from current time).

Assessing the stability of operations is a clear indicator of the degree of acceptability by the control system operator, i.e. the dispatcher. It is thus related to operational control perspective, the **design of the control system**, and the control actions allowed and used. This involves to which extent disturbances can be absorbed just by shifting trains in time and thus propagating their delay (i.e. by *retiming*) or including rescheduling actions (*reorder* of trains). The former coincides with the concept of robustness defined by Goverde and Hansen (2013): the ability of the timetable to absorb delays by itself, without the need of any re-ordering or rerouting action. When instead disturbances are dealt with by means of rescheduling actions changing the order of trains, this refers to the *resilience*: the ability of a timetable to absorb delays if some rescheduling action is taken.

Variation in operations can be analysed by the *deviation in time and space* between the timetabled and actual train paths. This is obtained by measuring at a given location the difference between the scheduled and actual passage time. Vice versa at a given time we measure the difference between the scheduled and the actual passing location. We also consider the variance of both these deviations, i.e. how predictable they happen to be.

Reliability in planning is a concept closely related to what several authors define and analyse as quality of plans. Specifically it relates to the capability of plans to keep acceptable levels of perceived traffic performance by customers also when stochastic disturbances affect operations. We adopt the following metrics to assess reliability of service for the different control schemes:

The Average total arrival delay, or *Average delay* in short, is the average of the total arrival delay (i.e. difference between scheduled and actual arrival time) over all delayed trains reaching their final station.

The *Average consecutive delay* at a station is the average over all delayed trains of the delay that has been propagated from other delayed trains. This metric gives a measure of how much trains are hindered during their run by the presence of other conflicting trains. For each train the consecutive delay is obtained by subtracting from its total arrival delay at a station the unavoidable delays (i.e. the sum of entrance delays and dwell time disturbances cumulated at the previous stations). The *Max Consecutive Delay* is the maximum value of the consecutive delay over all trains reaching their final station.

The *Punctuality* at the final station with respect to a threshold of 5 minutes (P_{5min}). This number gives the percentage of trains whose total arrival delay at the final station is less than 5 minutes. 5 minutes is a typical threshold used in operations (see Hansen & Pachl, 2014). If a null threshold is considered, we just consider the *amount of traffic delayed*.

Reliability in perceptions relates to a more detailed understanding of the precise impact of the delay experienced towards the users. This refers to a *cost of delay*, related to some value of time, associated to the precise *distribution of arrival delays*, i.e. considering the whole observed distributions of arrival delays at every station of the network. Based on these distributions, it is possible to perform a statistical analysis of the *risk of delay* as defined by Börjesson and Eliasson (2011). This might take into account a nonlinear disutility for larger delays by the users, the characteristics of the *tail* of the distribution, as well as the *variance* of the delay.

4 Control approaches

To analyse robustness and reliability of different railway systems, depending on the control schemes used, we built up a framework which enables to accurately describe each one of the schemes. The framework considered allows to introduce stochastic disturbances to operations (e.g. entrance delays, extensions of dwell times) as well as missing or erroneous information on traffic (e.g. measurement errors on entrance delays or train positions). For each control scheme, it is possible to test the impact of stochastic disturbances and uncertain or missing information as well as the performance achieved within perturbed traffic conditions. The two most important indicators are thus the degree to which this stochasticity is considered (i.e. the degree of including updated information when those are available); and the possibility for optimizing operations based on the expectation of the future.

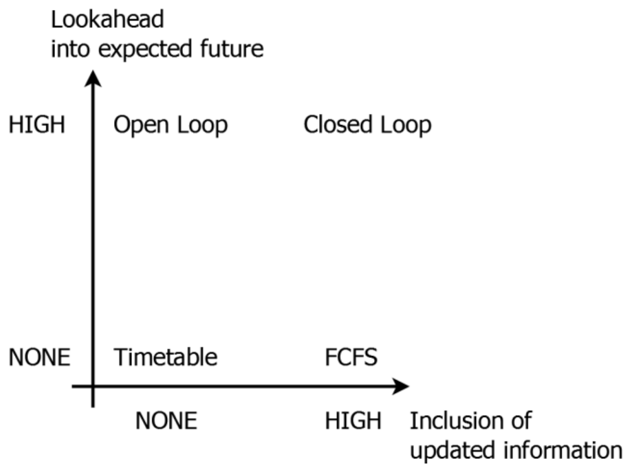


Figure 1. Qualitative differences between the 4 control schemes considered

We now formally detail all the four approaches. With **Timetable order** we mean that no real-time rescheduling action involving change of train order is considered, but trains keep on following the scheduled orders also under perturbed traffic conditions.

The **First-Come-First-Served (FCFS)** is a common dispatching strategy where trains pass a location (e.g. station, junction) in the same order they arrive.

In the **Open Loop** scheme the scheduler is run only once on the basis of the only expected train entrance delays. An optimal plan is computed once and for all, which solves all track conflicts detected over the whole horizon of traffic control. Such a plan is implemented at the beginning of traffic control horizon, and followed by trains for all its duration; no updated information or deviation from the plan is considered.

The **Closed Loop** scheme considers optimal plans, which are regularly updated on the basis of current traffic information.

The four control approaches considered (timetable order; FCFS; Open Loop and Closed Loop) are quantitatively analysed in Figure 1, along those two complementary evaluations: inclusion of updated information (x-axis) and lookahead into expected future (y-axis). Including updated information is performed by FCFS and Closed Loop approaches. The lookahead into expected future refers to the possibility of considering proactively the future to take better decisions, and is used by Open and Closed Loop.

The framework is composed of two main interacting modules that are an optimal *scheduler* of train services and an accurate simulator of railway operations (called *simulated operations*). For all control schemes, we manage traffic by retiming and reordering while considering train routes as those scheduled. Figure 2 functionally describes the four control approaches, in terms on how the modules and the input of our framework interact. Arrows represent information sharing, causal relation, input-output relations between the modules. Dotted arrows refer to approximated inclusion of effects and information. Inputs of the scheduler and the simulator are all the characteristics regarding the infrastructure (e.g. block sections, speed limits, track length, gradients), the rolling stock (e.g. mass, length, number of coaches, tractive-effort speed curve), the signalling and the safety systems such as ATP and interlocking. Train entrance delays are known in their realised value only by the simulator, while only their expected value is known by the scheduler. Moreover random dwell time extensions are considered in the simulator, whose realised values are unknown to the scheduler (which is only aware of their scheduled values). These assumptions reproduce what happens in real-life operations where the traffic control centre has only limited or even missing information on delays and traffic disturbances.

Figure 2 (a) considers the timetable order, where is no need to have a scheduler; the only way to manage the disturbances, set as input to the simulated operations, is by means of delay propagation. In other words, disturbances are dealt with only by retiming trains, i.e. shifting in time the scheduled arrival, departure and/or passing times, while keeping orders fixed. For this reason it is expected that the robustness and the reliability of this approach are in general poor with respect the other control schemes.

FCFS (Figure 2 (c)) does not consider any scheduler, priority rule, or policy to take order decision. Instead, decisions are taken in a myopic way, just looking at the immediate order of request of shared resources. For large stations with complex interlocking, there can be no guarantee that a feasible solution is found. This is a purely reactive strategy, which is able to incorporate changes to operations on the next decision to be taken. To apply the FCFS the only information needed from the simulated operations is the current position of trains. However this information has no impact on the next decision to be taken in the future given that FCFS does not adopt any kind of prediction of future traffic.

Figure 2 (b) refers to Open Loop. This setup is unable to adjust the optimized plan to current traffic conditions, since the plan is computed and implemented only once, based on the expected value of entrance delays. Most approaches presented in the literature have been considered and evaluated in an open loop structure, and further effects of uncertainties, traffic dynamics, and modeling errors are neglected.

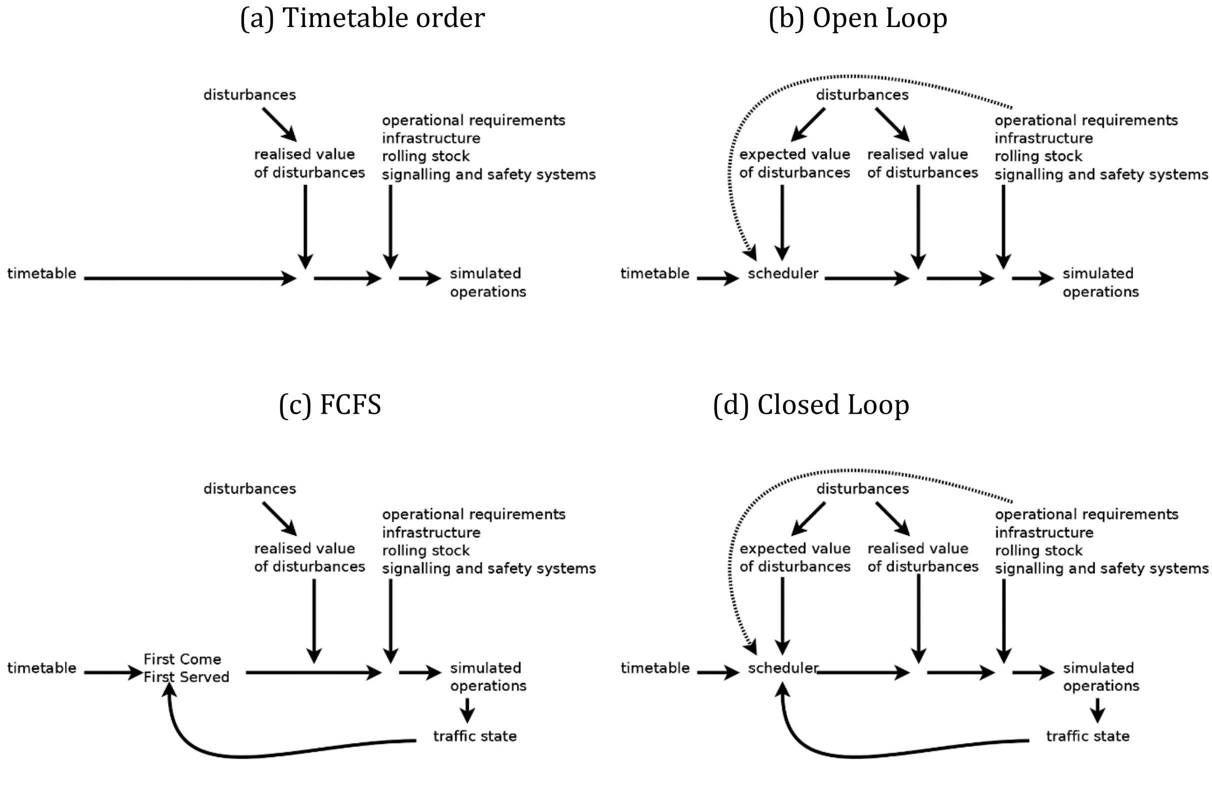


Figure 2. Architecture of the control structures considered

The Closed Loop scheme (Figure 2 (d)) is the most refined form of control, where implemented optimal plans are regularly updated on the basis of current traffic information. This means that at regular time intervals current train information (i.e. current positions and speeds) is collected from the simulator and set as input to the scheduler together with the expected values of (future) train entrance delays. The scheduler predicts future traffic operations in order to detect and solve all track conflicts forecasted over a prediction horizon. Optimal plans are produced that are successively implemented and followed by traffic. The essence of the closed-loop is represented by the arrow that transmits traffic information from the simulated operations to the scheduler. Differently from the FCFS, current traffic information collected in the closed loop has an impact on control decisions to be taken in the future. In the closed loop computed optimal plans take into account also dispatching decisions taken in the past, to prevent that suggested control measures are then not congruent with real operations, hence unfeasible. For this reason, traffic predictions have memory of past actions.

The scheduler used in our framework is the tool ROMA (Railway Optimization by Means of Alternative Graphs), which is based on the modelling paradigm of Alternative Graphs and a job-shop scheduling problem. The simulator of railway operations is the stochastic microscopic model EGTRAIN (Environment for the desiGn and simulaTion of RAIlway Networks). EGTRAIN is considered as a realistic simulation model since it has been validated by verifying that simulated train running times were congruent with those scheduled in reality, within undisturbed traffic conditions. A detailed description of ROMA and EGTRAIN can be found respectively in Corman et al., (2011) and Quaglietta (2011); and the general setup of the overall closed loop setup, including parameters, and functional description of the interface modules is described in Corman and Quaglietta (2015). The procedure goes iteratively along *stages*. In each of them, the simulator sends position and speed of trains to the scheduler; the information is used to determine a current traffic state and forecast future train paths to detect potential track conflicts over a *prediction horizon* (PH). The scheduler relies on deterministic traffic predictions where train running and dwell times are considered as deterministic. Track conflicts are detected as overlaps between the blocking times for all involved block sections (Hansen and Pachl, 2014). Detected conflicts are then solved by formulating the scheduling problem as a job-shop model with no-store constraints, and using a truncated version of a Branch and Bound algorithm (D'Ariano et al, 2007), yielding a new conflict-free plan minimizing delay propagation. The output is a set of advisory orders at given locations, which are implemented in the simulation core of EGTRAIN, after a control delay representing the communication to the field. According to the chosen orders, the traffic is microscopically simulated (using a time-driven and synchronous approach) The scheduler and the simulation model interact with each other according to a rolling horizon scheme, each stage being performed after a *Rescheduling Interval* (RI).

5 Case Study and parameters used

The experimental analysis of this paper is conducted on the railway corridor between Utrecht (Ut) and Den Bosch (Ht) in the Netherlands. This has a length of more than 48 km with 6 intermediate stations: Lunetten (Ln), Houten (Htn), Houten Castellum (Htnc), Culemborg (Cl), Geldermalsen (Gdm), and Zaltbommel (Zbm). The detailed layout is presented in Figure 3, together with the locations in which trains can overtake each other and a reordering is possible (those places are called CheckPoints, or CP in short: CP1, CP2, CP3). The network is equipped with a fixed-block signalling system and the traditional Dutch Automatic Train Protection ATB system. The hourly periodic timetable schedules 4 intercity trains (IC) per hour per direction between Ut and Ht without intermediate stops; and 4 local trains, two of which are limited between Ut and Gdm, while the other two run all the way till Ht. No freight trains are taken into account in the study. For the sake of simplicity, only trains running along the Ut-Ht direction are considered, as in this double-track corridor there is no interaction between trains running in opposite directions. The total horizon of traffic control corresponds to 2 hours of operations. The closed-loop setup considers the best parameters found in Corman and Quaglietta (2015), i.e. rescheduling interval *RI* =120 seconds, prediction horizon *PH*= 60 min, and a control delay of 10 s, i.e. the plans are actually implemented after 10 s from their computation.

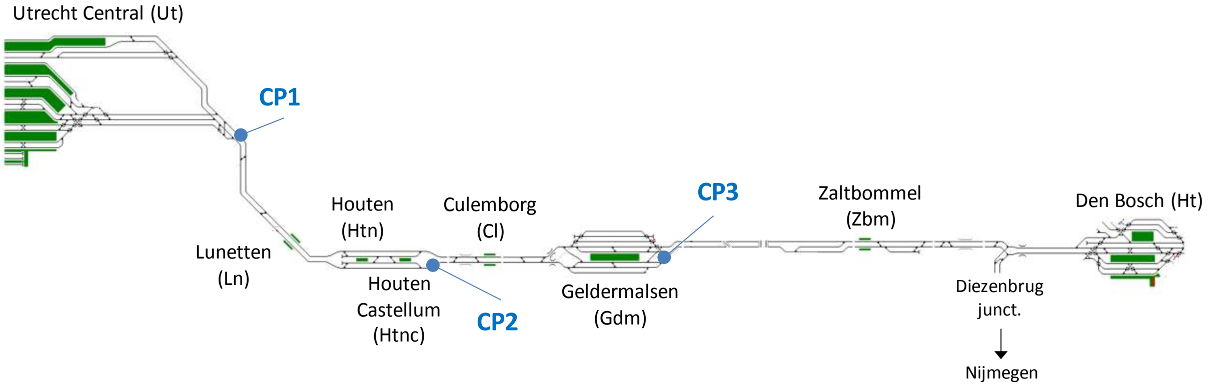


Figure 3. Detailed layout of the Utrecht - Den Bosch corridor, with the locations (CP1, CP2 and CP3) of the three checkpoints in which train reordering is considered.

The analysis is performed over 30 different perturbed scenarios in a typical Monte Carlo setup. Each scenario is generated by randomly sampling entrance delays, measurement errors on entrance delays and disturbances to dwell times at stations. Specifically entrance delays are drawn from a Weibull distribution fitted to recorded data (as in Corman et al 2011) with scale, shape and shift parameters that are different for ICs and local trains. Errors on entrance delays follow a Gaussian distribution with zero mean and a standard deviation equal to 20% of expected entrance delays. Those parameters have been defined in the projects requirements (Corman and Quaglietta 2015, and Quaglietta et al (2016)) as realistic values concerning the error of a prediction error, which is itself a stochastic dynamic process which has not been characterized properly in the literature. Station dwell times are drawn from a Weibull distribution, fitted to recorded data, as in Quaglietta et al. (2013). In both cases, we restrict our analysis to actually delayed operations, where rescheduling can prove its value, and moreover insert extra deviations that result in additional delays. Thus, the performance of the approaches should not be compared straightforward to real operations but rather used in a comparative manner.

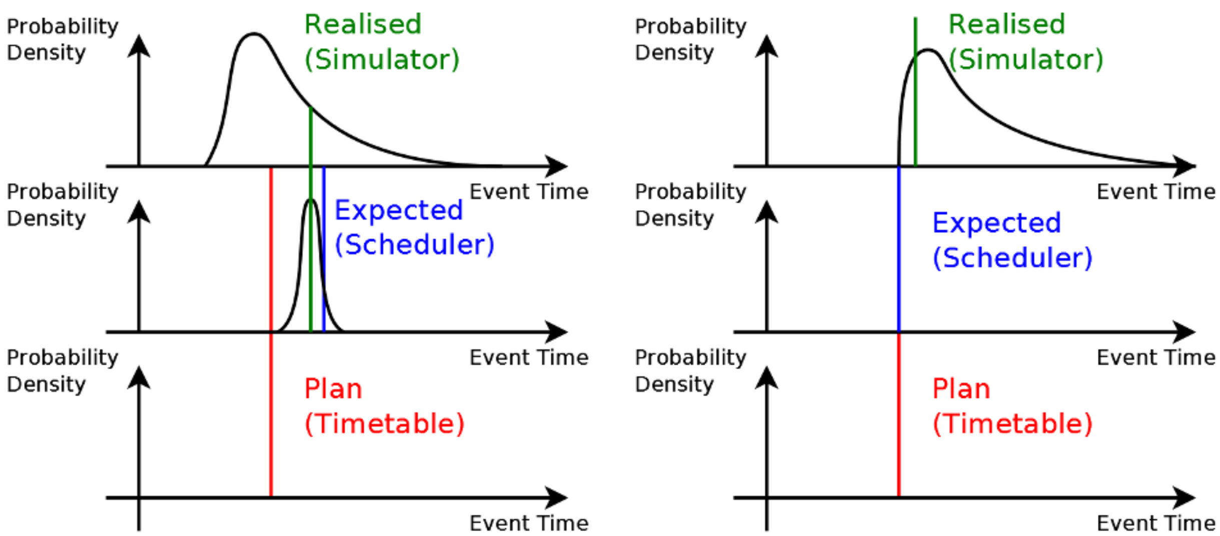
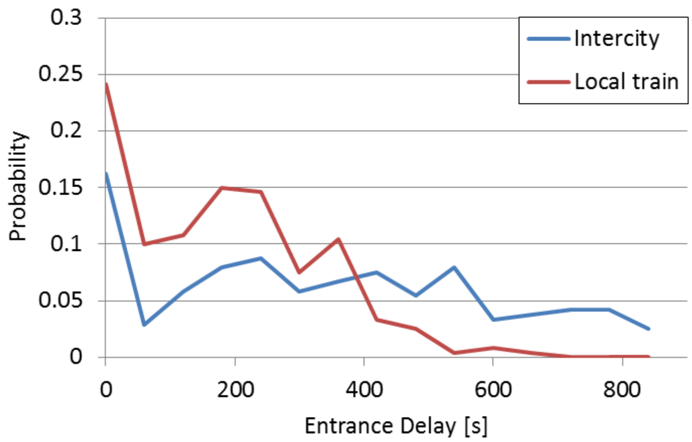
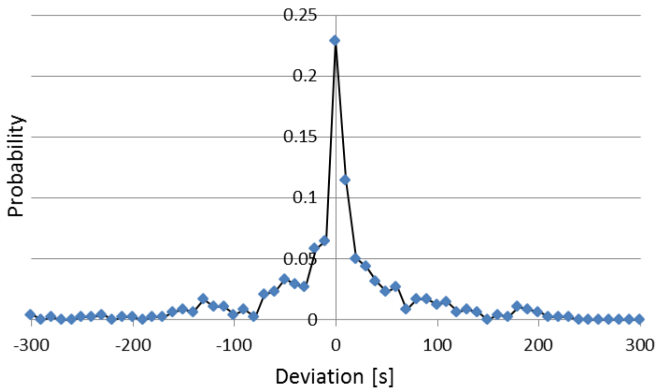


Figure 4. Process of information update for entrance delay (left) and dwell time (right)

Figure 4 reports the distributions, and the process by which entrance delays (on the left-hand side) and dwell times (on the right-hand side) are known by the scheduler and the simulator respectively. We also report the entrance times and dwell times as planned in the timetable (at the bottom). In the simulator (top left), trains enter the network with a Weibull-distributed delay with respect to the timetable entrance time (red vertical line). A realised entrance delay (green vertical line) is sampled from this distribution. Before a train enters the network, the realised entrance delay is transmitted to the scheduler (middle left) affected by a Gaussian-distributed measurement error (blue vertical line). After the event has occurred, i.e. the train has entered the network, the scheduler is finally aware of the actual entrance delay realised in the simulation.

For dwell times, we make the assumptions depicted in the right-hand side of Figure 6. In the simulator, trains will experience dwell times at stations that are distributed according to a Weibull probability. Realised dwell times (green vertical line, top right) differ from those planned (red vertical line, bottom right). The scheduler will be unaware of these differences since it expects trains dwelling according to the scheduled dwell time given by the timetable (blue vertical line). After that the event happens, i.e. the train has finished dwelling at a station, the control schemes that exploit a form of feedback from operations can include the realised value of the event time in their process (FCFS) or optimization (Closed Loop).





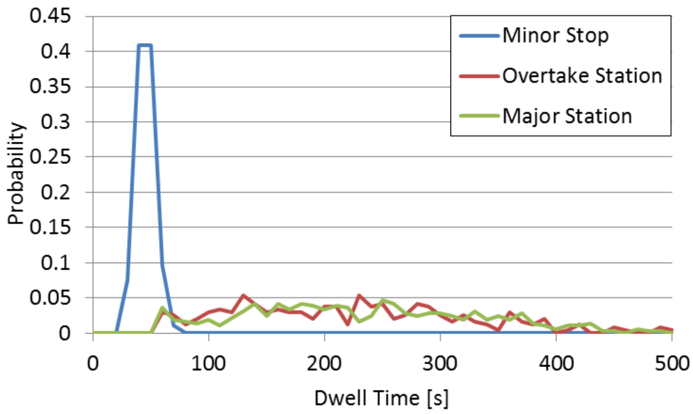


Figure 5. Probability distributions used for (top) entrance delay, (middle) errors in entrance delay, (bottom) dwell time.

Figure 5 reports the sampled probabilities for the 30 sampled scenarios for the realised entrance delays (at the top), the measurement errors on entrance delays (in the middle) and the realised dwell times at stations (at the bottom). Entrance delays are divided between intercity and local trains. No train can depart early from stations, thus the negative tail of the shifted Weibull is reported to 0. The average entrance delay for the samples considered is around 300 seconds. The measurement errors on entrance delays result in having a mean 0 and a maximum deviation of 300 seconds. In half of the cases, these errors are in the order of 20 seconds. We report the dwell time distributions for small stops where planned dwell times are less than a minute; for major stations where planned dwell times are about two minutes; and for the station of Geldermalsen, where local trains have longer planned stops, since they are scheduled to be overtaken by Intercity trains.

6 Evaluation

In this section we evaluate the metrics discussed in Section 3 for the different control schemes: Timetable, FCFS, Open Loop and Closed Loop. The values of the metrics are considered as the average over the 30 disturbed scenarios, when not stated otherwise.

6.1 Operational perspective: Stability and variations

Stability is a concept that makes sense only for the closed loop, since this is the only scheme where plans are regularly updated over time, based on current traffic conditions. For FCFS, there is no such a concept of a traffic plan, as the individual train orders are the same as their arrival orders at a location. We evaluate the number of relative reordering NRR to understand how the computed plans are sensitive to traffic dynamics, stochastic disturbances and uncertainties in information. A large NRR relates to instability of the plans, i.e. different train orders are computed at each stage. Plans that do not change at all have a NRR equal to 0. This would describe a theoretical situation of full stability of the plans.

Figure 6 illustrates the trend of NRR over time. Each time that a plan is computed (i.e. each RI=120 s) NRR gives the amount of train orders that differ from the plan computed at the previous stage (i.e. 120 s before). In the first 10 minutes, NRR is very low, because the amount of trains running is small enough to observe limited stochastic phenomena of delay propagation. As the disturbances start progressing over the network, the rescheduling plans become more unstable and vary over time. The reason of such instability is that the propagation of disturbances induces a deviation between actual and predicted train trajectories, altering from time to time the conflicts detected by ROMA and the corresponding plans. In general NRR varies over time in an erratical pattern, oscillating around an average of 0.09 (Average NRR) and reaching a maximum peak of 0.30.

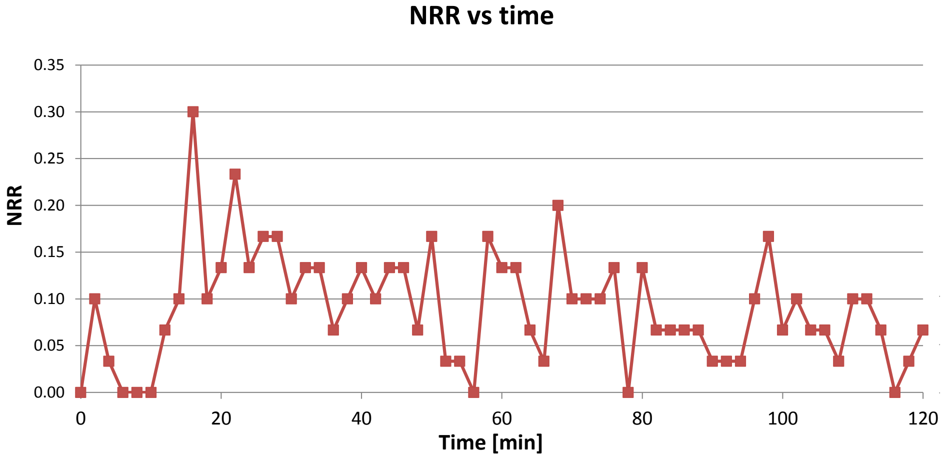


Figure 6. Number of Relative Reordering with respect to time.

We also evaluate the Weighted NRR as defined in Section 4, to determine a time dynamic for this instability. Figure 7 shows that only 11% of the train orders must be changed most urgently (i.e. in the first 5 minutes ahead from current time). About 25% of train orders are required to be changed between 5 and 10 minutes in the future. Most orders (64%) should be changed more than 10 minutes ahead of current time. This means that computed plans are stable in the short term (i.e. the first 5 minutes), and instability is due to prediction errors (hence inaccuracy in the plans) that obviously are larger, the farther away the operations to be predicted.

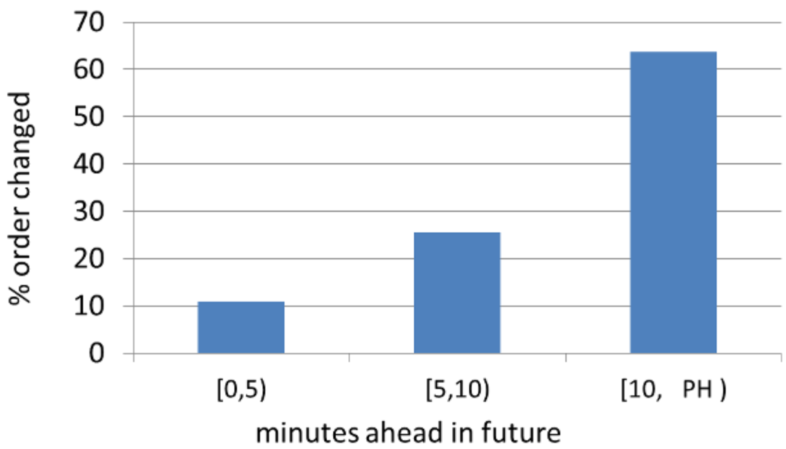


Figure 7. Number of orders changed, divided in the distance between the time of prediction, and the time of the order changed.

Table 2 reports the amount of retiming and reordering performed by each control approach. The retiming has been calculated as the average deviation in time between the train paths in the timetable and in realised operations. The amount of reorderings is calculated as the average number of train orders that differ from the timetable order. It is evident that the Timetable order solution has a retiming that is about 28% larger than the other approaches, as a result of not changing train orders. The other control schemes have a retiming that is practically equivalent, while the amount of reordering is quite different. By definition the Timetable order keeps the order of the timetable, and thus has 0 reorderings. FCFS is the solution with the largest amount of reorderings, due to the myopic nature of the approach. The open loop has a similar amount of reorderings, that is symptomatic of not very accurate traffic predictions when using very long prediction horizons (in this case 2 hours). Prediction errors increase when the operations to predict are farther away. In this case conflict detection is more inaccurate and the resulting updated plan will not completely match actual traffic. The closed loop instead has the lower amount of reorderings, highlighting its capability in adjusting the plans to better fit actual traffic conditions. On average 5.1 orders are changed, that means that one third of the 16 running trains will not follow the timetable order.

Control scheme	**Retiming** time deviation wrt planned timing [s]	**Reordering** order deviation wrt planned orders [-]
Timetable Order	377.2	0
Open Loop	294.8	6.3
FCFS	293.5	6.5
Closed Loop	295.4	5.1

Table 2. Retiming and reordering exploited by the 4 approaches considered.

As a relevant metric of reliability we also examine the deviation in time and space between the actual and planned train paths. Figure 8 reports the deviations in time (on the y-axis) along the whole network (on the x-axis; the intermediate stations are reported by their respective label). For each control scheme these deviations are the average over the 16 running trains and the 30 disturbed scenarios.

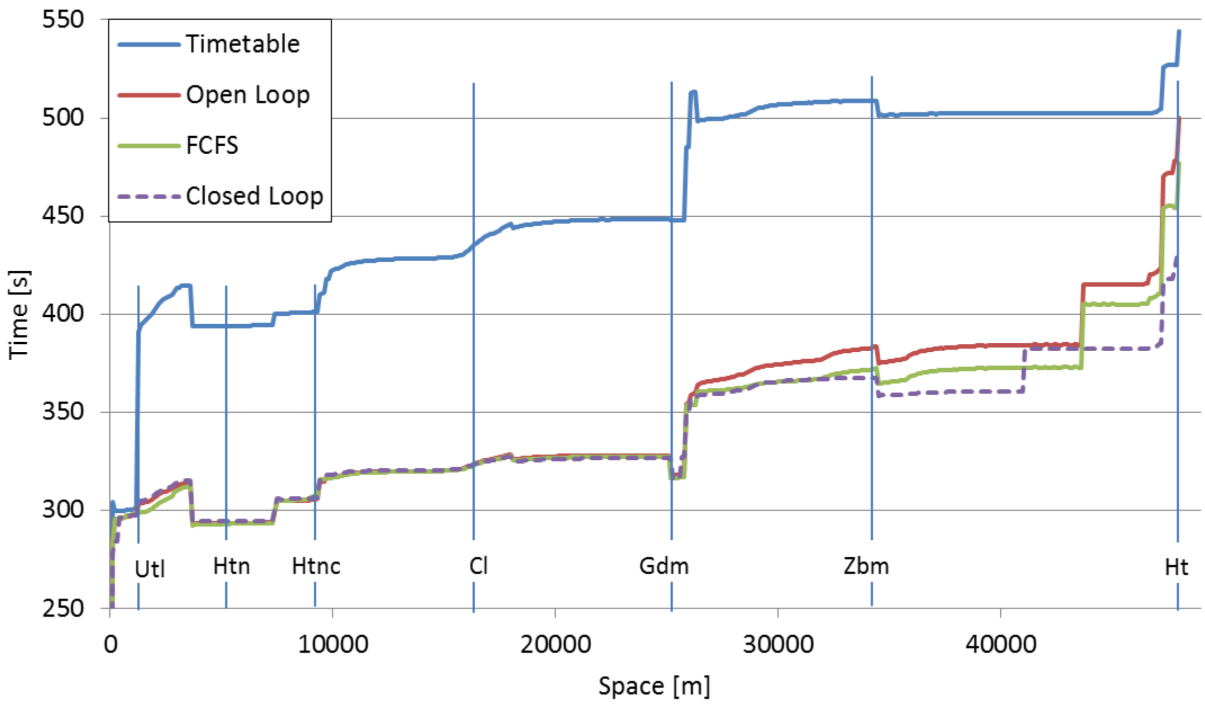


Figure 8. Average time deviations from scheduled train paths

It is evident that the delay of 300 s in Utrecht (at 0 metres) corresponds to the average entrance delay set as input in our experiments. From there on, deviations build on at stations (due to extra dwell time disturbances), at merging points (due to the necessity of holding trains to implement reordering) and along the line (due to hindrances with conflicting trains and restricted signal aspects).

Keeping the timetable order results in deviations that are about 100 seconds larger, than the other schemes. The peak of delay observed immediately after Gdm is significant, since it indicates suboptimal rescheduling actions, which force trains to wait a long time with respect to the scheduled order. The time deviations observed for the other schemes are very similar until Gdm, as can be seen by their diagrams that practically overlap. Optimal orders that significantly improve traffic performance are implemented in Gdm. This can be seen from the fact that from Gdm on, the time deviations of Open Loop, FCFS and Closed Loop follow different trends. The Closed Loop is the scheme that gives the lowest time deviations from the planned train paths.

We should also note that, despite time supplements in the timetable, delays never reduce; this is due to the sampled disturbances in dwell time, which are larger than time supplements. Such an effect of systematically small headway is relatively common in congested railway networks (Dewilde et al 2014). This kind of diagram can identify the bottlenecks of the network, where there is a sudden increase in time deviations. This also helps in determining the time supplements along the lines to increase the robustness of a timetable, as explained in Vromans et al. (2006). In our case study, we observe that when keeping the timetable order, the bottleneck (largest deviation) is located in Utl. The other control schemes reduce and shift the bottleneck, suggesting to place time supplements in Ht or Gdm instead of Utl. In terms of magnitude, the Timetable might result in overestimating the amount of buffer required by 25%.

In Figure 9 we report the space deviation (on the y-axis, in m) between actual and planned train paths over the whole horizon of traffic control of 2 hours (on the x-axis, in s). These deviations represent the distance at a given time between the planned and the actual position of a train. In other words, this represents (for each time point, resolution of one second) the span between the position where the train should be according to the plan, and its actual position, with a resolution of one meter. Deviations are computed for each of the 16 running trains (i.e. the 16 lines reported on each plot) as average over the 30 disturbed scenarios. Blue lines represent space deviations for the 8 Intercity trains while dark and light green lines depict deviations for the 8 local trains. Space deviations are always non positive which means that actual train paths are always behind the schedule.

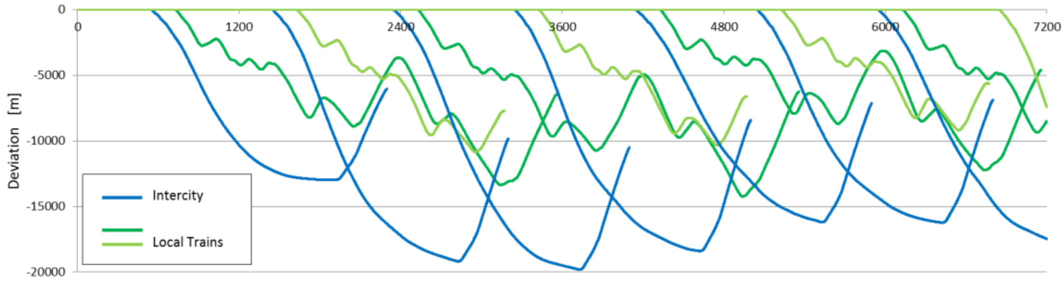
For each control scheme, the trend of the deviations looks similar even if they differ in value. Space deviations progressively increase over time for Intercity trains until a maximum is reached at around two thirds of their paths. For instance for the first intercity, departing at 600 s, the max deviation is reached at 1800s (close to Gdm station), and the train ends its service at 2400s, catching up some deviation. In general, intercity trains suffer from much large deviations on average than local trains. The Timetable scheme performs the worst, with a max deviation for the Intercities of about 20km. This is in line with the resulting delay that has been analysed so far. For local trains the maximum deviations in the Timetable solution is around 12 km while for the other schemes this value goes just beyond 5 km. The FCFS and the Closed Loop show the smallest space deviations. Closed Loop results in smaller max deviations of Intercity trains (18 km versus 19km of the FCFS), the FCFS results in limiting the max deviation for local trains (8 km versus 6.5 km). This latter is a consequence of the property of Closed Loop of considering future operations (that for intercity trains might be further away in time); the myopic approach of FCFS looks only at the current entrance time of trains, neglecting their future evolution.

Table 5 reports the average and variance of time and space deviations, across time horizons, trains running, delayed instances. The 2nd row (respectively 4th) reports to which extent the realised train paths are close to the planned ones, in terms of time (resp. space). Row Three and Five report the variance of those deviations. The average deviation is a measure of the expected bandwith of deviation of the train paths compared to their plans. A smaller bandwith means operations closer to the plan which allows a direct increase in the capacity of the infrastructure; this is a crucial goal of railway traffic control (Luethi 2009). The variance is a measure of how good the bandwith of train operations can be predicted. A small variance allows reducing delay consistently by inserting strategically buffer times and optimizing their distribution in time and space by robust timetabling (Vromans et al 2006, Dewilde et al 2014).

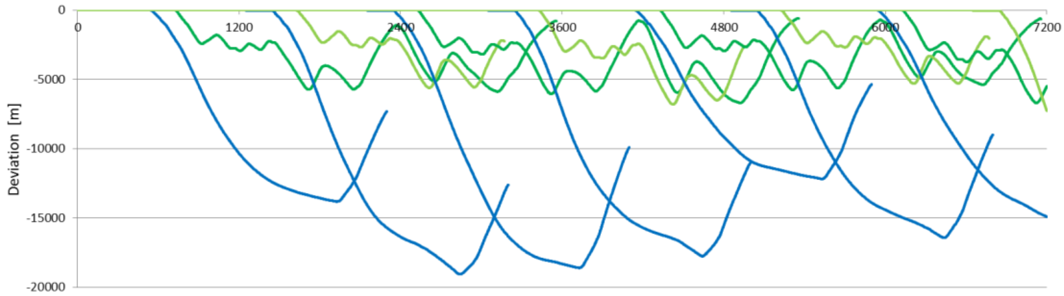
The ranks between the figures are all increasing from timetable to FCFS to open loop control to closed loop control, with a few exceptions. The timetable scores the smallest variance for deviation in time, i.e. all traffic is delayed consistently; a delayed train propagates a similar delay to all traffic afterwards. All other figures favour closed loop control, decreasing the extent and variability of the deviation, in time and space. The relative performance of FCFS and Open Loop sometimes favour one or the other approach. Open loop always has less variance than FCFS.

	Timetable	FCFS	Open Loop	Closed Loop
Time Deviation[s]	458	343	348	339
Variance Time	86	149	143	118
Space Deviation [m]	3107	2405	2393	2374
Variance Space	4790	4286	4189	4075

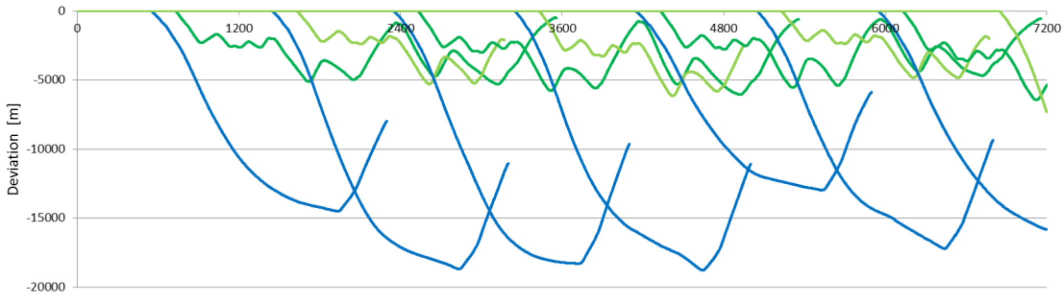
Table 5. Average deviations between planned and actual train paths, in time and space.



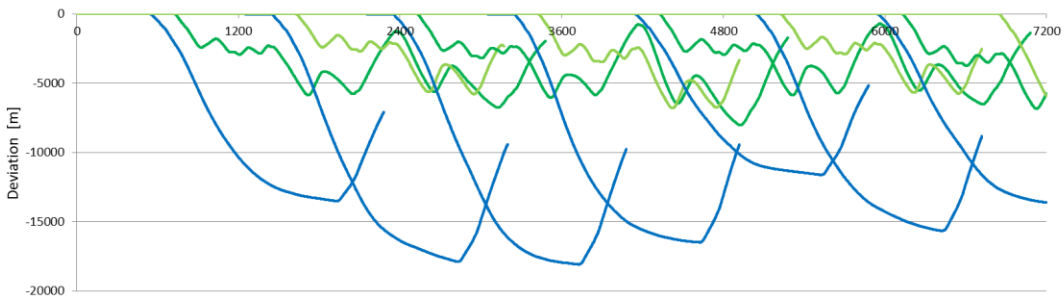
(a) Timetable



(b) Open Loop



(c) FCFS



(d) Closed Loop

Figure 9. Space deviations between actual operations and the scheduled time distance path.

6.2 Operational analysis and planning : delay performance

The four control schemes are now evaluated in terms of the average delay at all stations, the average consecutive delay, the max consecutive delay, the share of trains which are running delayed (the lower the better), the punctuality at 5 minutes (the higher the better) (Table 3).

Control scheme	Avg Delay [s]	Avg Cons Delay [s]	Max Cons Delay [s]	Delayed trains [%]	Punctuality 5 min [%]
Timetable Order	423.3	130.1	515.3	96.8	35.7
Open Loop	322.6	65.5	377.4	89.4	59.7
FCFS	307.9	52.8	309.2	87.7	60.3
Closed Loop	294.2	30.2	176.0	91.2	62.3

Table 3. Quality indices for the different control schemes.

The benefits of changing orders when rescheduling traffic are immediately highlighted. The Open Loop scheme already improves strongly traffic performance with respect to the timetable scheme (where no rescheduling is performed changing the order of trains). All delay-related indicators reduce, respectively by 23% for the average total delay, 50% for the average consecutive delay, and 27% for maximum consecutive delay. Consistent gains are also achieved in punctuality since the number of punctual trains increases by 37%. When applying FCFS a larger improvement is obtained with respect to the timetable scheme, which reaches up to 60% for the average consecutive delay and 38% in punctuality. The amount of delayed trains is lowest with FCFS, but this means that trains running on time are given priority over delayed trains. This turn out to be suboptimal at system level: the average delay experienced by FCFS is almost double compared to the closed loop approach. FCFS and Open Loop both represent an improvement with respect to the timetable, and sometimes perform relatively similar, but each thanks to a different factor. FCFS due to the always updated information it can exploit, Open loop due to the optimization approach.

For some performance indicators, the stronger importance of one or the other factor might result in better overall performances. Closed Loop outperforms all other schemes for all measures of performance. The comparison with FCFS is the most interesting, since it underlines the improvements that an automatic closed loop traffic control can give with respect to the strategy generally used in real-life. With respect to FCFS, the closed loop reduces by 5% the average delay, by 43% the consecutive delays and even by 43% the max consecutive delay. Also the number of punctual trains is increased by 5%. From the point of view of the infrastructure manager, an increase in punctuality can be associated to a direct increase in revenues, by either reduced ticket compensation, or by extra quality performance benefits. Concerning the former, For instance, (Kroon et al 2009) reported that a 1.5-percent increase in punctuality is associated to an increase in revenues of 20 Million EUR/ year, for the all Dutch network. Keeping this cost factor would result in extra revenues for more than 30 Million EUR/ year, even when compared to an application of FCFS (which has been used in very limited context for automated railway traffic management, see Corman et al (2011).

6.3 Passenger perception: Delay cost, delay risk and tails

As a metric of perceived reliability we here analyse in detail a few metrics related to the delays. A direct quantification of the delay cost can be achieved by multiplying the expected delay by a suitable Value of Time multiplier, which can be for instance 9 EUR/hour (Kouwenhoven et al 2014). This multiplies linearly the figures of Table 3, and combines them with a given and fixed amount of passengers of the network. A more precise understanding is based on the distribution of arrival delays at their final station. For each control scheme, the statistical analysis is performed over 480 samples (i.e. 16 trains over 30 scenarios). The distributions are reported in Figure 10.

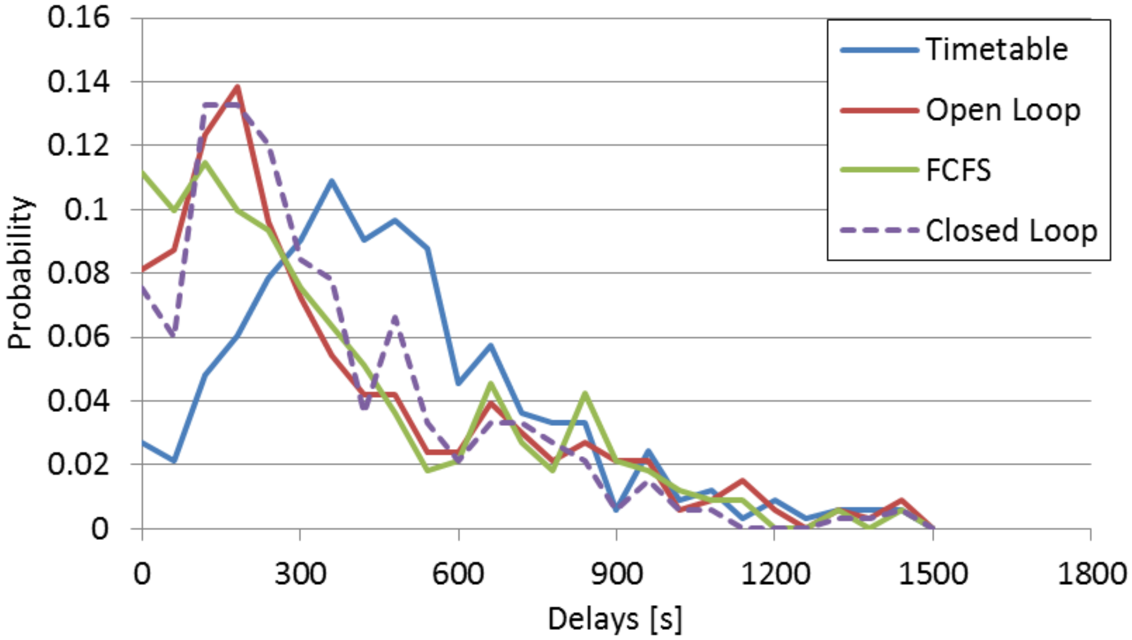


Figure 10. Statistic distribution of arrival delays at the final station.

When trains follow the timetable order there is a larger probability of experiencing delays between 300 and 600 s, as shown by the peak of the Timetable control scheme. For the other schemes, smaller delays of about 250-300 s are more probable, as shown by the peaks of their distributions. These values are similar in size to the average entrance delay used in our experiments (which is around 300 s), i.e. the delay propagation is strongly limited in these cases. FCFS exhibits a smaller peak (at 250-300 s) than the Open Loop and Closed Loop approaches, but has a higher probability of delays which are larger than 600 seconds. The maximum delay observed is about 1500 s, and it is almost the same for all the control schemes, as this depends on the maximum entrance delay.

	Timetable	Open Loop	FCFS	Closed Loop
Mean [s]	473.9	367.1	352.3	337.2
Median [s]	436.0	250.0	263.0	264.5
Variance [s^2]	79367.5	107775.2	98492.8	73886.4
Squared Delay [s]	563.5	500.6	482.2	444.7
Extreme values threshold[s]	1037.3	1023.6	979.9	880.8
Extreme values probability %	4.53	5.12	4.53	4.52

Table 4. Statistical measures of arrival delays at the final station.

The distributions of each control schemes are provided in terms of their significant statistical characteristics in Table 4. In particular, we report the mean of the total delay at the final station, its median, the variance and the squared delay, i.e. the Root Mean Squared Error (RMSE) of the delay. For this latter, the error is considered as the difference between scheduled and actual operations, and it is computed as the square root of the sum of the squared delays. The squared delay indicator weighs more larger delays, in line with the hypothesis of Börjesson and Eliasson (2011), which considers large delays with low probability (risk) as more relevant to the disutility of passengers. However the data considered in Börjesson and Eliasson (2011) showed no empirical evidence that a square function could describe this behaviour.

The last two rows show the extreme value threshold and the probability that a recorded delay is larger than the extreme value. The extreme value threshold is set at the mean plus twice the standard deviation, as in the recommendation from the OECD (2010). As can be seen the Closed Loop outperforms all the other control schemes, for any of the statistical measures considered, apart from the median. This latter is similar to the open loop and FCFS. For the Closed Loop, the arrival delays at the final station are on average smaller than the other schemes (lowest mean); less dispersed around the mean (lowest variance), with less large delays (lowest squared delay). In terms of extreme values we also observe that the closed loop has the lowest extreme value threshold for the same probability of 4.52%. According to the works cited on users' and economical perspective, this is positively valued by passengers. Still considering the extreme values, the Open Loop performs the worst since it has an extreme value that is just slightly less than the largest one (the one by the Timetable), but has the highest extreme values probability. In fact, the Timetable approach has a probability of extreme values that is in line with the other approaches, even though the threshold is much larger.

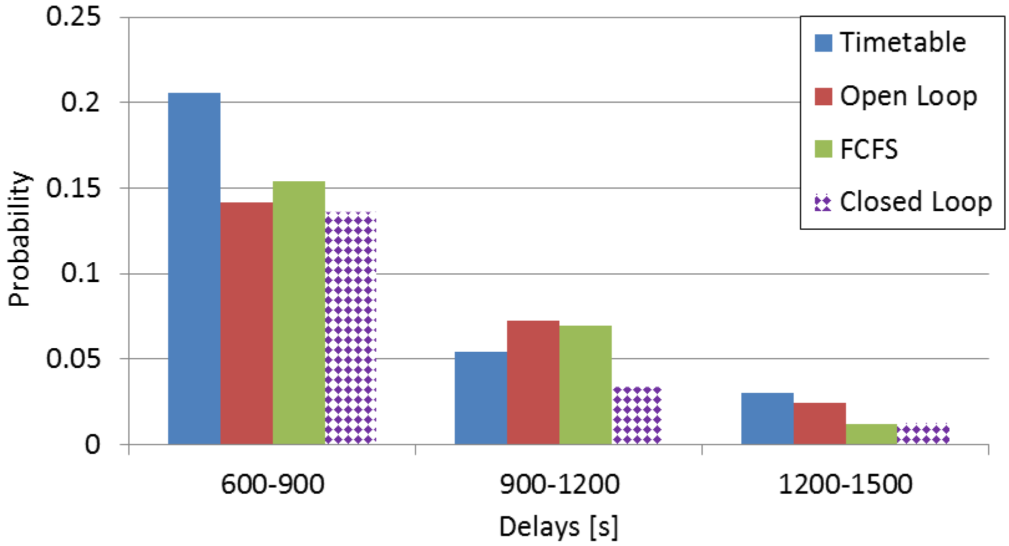


Figure 11. Analysis of the tails of the delay distributions observed at the final station

The tails of the delay distributions are further analysed in Figure 11 where we focus on delays larger than 600 s. Specifically we report for each control scheme the probabilities that arrival delays at the final station fall in the three intervals: 600-900 s (i.e. 10-15 min), 900-1200 s (i.e. 15-20 min) and larger than 1200 s (i.e. between 20 min and the max, which is about 25 minutes). Again the Closed Loop shows the smallest probability of experiencing delays belonging to each of the three intervals. The benefit of the Closed Loop can be traced back, at least partially, to the regular update of traffic information, which avoids large prediction errors in the scheduler. This results in more accurate plans that better fit to actual traffic conditions.

7 Conclusions

This paper presents an extensive experimental analysis of railway traffic control schemes and evaluates the impact of stochasticity and uncertainty on robustness, stability and reliability of railway operations. We consider several metrics as defined by specific literature in this field. A key contribution is the consideration of detailed uncertain effects, by means of control structures that go beyond the simple ones normally used for timetable robustness at planning stages, and include railway traffic control, and updates from the field in case of unreliable operations. This paper is an analytical complement to the framework definition and setup presented in Corman and Quaglietta (2015), to which the reader is directed for further details on the architecture.

We argue that without including the effect of uncertain dynamics, like delays, missing or erroneous information, the reliability of railway operations can be quantified only to a certain extent. This research is thus a first step to have a thorough appraisal of many uncertain factors in railway systems, from a variety of stakeholders (control system design; control operations; planning; operation analysis, passenger perception).

The practical implications of this research relate to the comprehensive assessment of operations from a wide range of points of views. We consider measures of reliability (i.e. the capability of keeping acceptable traffic performance also during disturbed traffic conditions), robustness (concerning the degrees of freedom available to cope with unforeseen events, ability of plans to absorb stochastic disturbances, avoiding large delays, and a limited sensitivity of control actions to uncertain information), and resilience (the impact of real-time traffic control to decrease delays). The ranking between different control schemes is in general consistent, with Timetable order scoring the worst. The Closed Loop scheme outperforms all the other schemes that also include the FCFS generally used in real-life to dispatch traffic operations. The closed loop results as the best control scheme for all the metrics of reliability and robustness, considered. This stresses the possibilities of improved traffic control in practice, as sought since years (Kauppi et al 2006, Schaafsma 2005), when the current deployment of advanced technology such as ERTMS/ETCS would make available interfaces to/from running traffic.

The main policy implications of this work relate to the valuation of reliability in railway projects. We have shown that considering only the timetable order (as currently it is common) might result in a systematic overestimation of unreliability, when railway traffic control will be implemented. This overestimation of reliability remains even when the rescheduling approaches so far assessed in perfect and full information, are tested with extensive degrees of uncertainty and information availability compatible with realistic situations. Also, the best allocation of buffer times in location and size differs for the different control schemes.

Future research should address further the impact to passenger traffic, by for instance studying how passenger might react to unreliable traffic, under different rescheduling and control approaches. Also, characterizing in a more precise way the different sources of uncertainty can improve predictions and reduce instability and errors. How to include this extra information in the optimization via stochastic (or robust) optimization is an interesting open challenge. Will it be possible to determine in advance most likely rescheduling actions to be implemented, and disseminate them timely to passengers? What might be the impact of reliable/unreliable information provision? Is there a need to rely on advanced personalised travel planners routing systems to achieve the reliability levels shown in this paper? How to characterize the unreliability for a transport a system that is only planned? A real-life pilot would also be a natural follow up of this research study.

References

[1] Andersson, E., Peterson, A., Tornquist Krasemann, J. (2013) Quantifying railway timetable robustness in critical points. Journal of Rail Traffic Planning and Management 3, 95-110.

[2] Börjesson, M., Eliasson, J. (2011) On the use of ‘‘average delay’’ as a measure of train reliability, Transportation Research Part A 45, 171-184.

[3] Büker, T., Seybold, B. (2012) Stochastic modelling of delay propagation in large networks. Journal of Rail Transport Planning and Management 2, 34–50.

[4] Cacchiani, V. Huisman, D. Kidd, M., Kroon, L., Toth, P., Veelenturf, L., Wagenaar, J. (2014) An overview of recovery models and algorithms for real-time railway rescheduling Transportation Research Part B Methodological 63:15–37.

[5] Carey, M. (1999) Ex ante heuristic measures of schedule reliability, Transportation Research Part B 33 (7).

[6] Caimi G., Fuchsberger M., Laumanns M., Luethi M. (2012) A Model Predictive Control Approach For Discrete-Time Rescheduling In Complex Central Railway Station Areas, Computers & Operations Research, Vol. 39, pp. 2578-2593.

[7] Corman, F., D’Ariano, A., Longo, G., Medeossi, G., (2010) Robustness and delay reduction of advanced train dispatching solutions under disturbances, Proceedings of the 12th World Conference on Transport Research, WCTR 2010, (pp. 1–12), Lisbon, Portugal.

[8] Corman F., D’Ariano A., Pranzo M., Hansen, I.A. (2011) Effectiveness of dynamic reordering and rerouting of trains in a complicated and densely occupied station area. Transportation Planning and Technology 34(4), pp. 341-362.

[9] Corman F., Meng, L., (2015) A Review of Online Dynamic Models and Algorithms for Railway Traffic Control. IEEE Transactions on ITS 16(3) pp 1274-1284.

[10] Corman F., Quaglietta, E. (2015) Closing the loop in railway traffic management. Transportation Research C 54, pp 15-39.

[11] D’Ariano A., Pacciarelli D., Pranzo M.(2007) A Branch And Bound Algorithm For Scheduling Trains In A Railway Network, European Journal of Operational Research, 183(2), pp. 643–657.

[12] D'Ariano A., (2009) Innovative Decision Support System for Railway Traffic Control IEEE Intelligent Transportation Systems Magazine Vol. 1 (4), pp. 8-16.

[13] Dewilde, T. Sels, P., Cattrysse, D., Vansteenwegen, P. (2011) Defining Robustness of a Railway Timetable. Proceedings of the 4th International Seminar on Railway Operations Modelling and Analysis (IAROR).

[14] Dewilde, T., P. Sels, D. Cattrysse, P. Vansteenwegen (2014) Improving the robustness in railway station areas. European Journal of Operational Research 235, pp. 276–286.

[15] Delorme, X., Gandibleux, X., Rodriguez, J. (2009) Stability evaluation of a railway timetable at the station level. Eur. J. Oper. Res. 195 (3), 780–790.

[16] Dollevoet, T.A.B., Huisman, D., Schmidt, M.E., Schöbel, A. (2012). Delay Management with Rerouting of Passengers. Transportation Science, 46(1), 74-89.

[17] European Commission, (2011) White paper Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system. COM/2011/0144.

[18] Goverde, R.M.P., (2011). A delay propagation algorithm for large-scale railway traffic networks. Transportation Research Part C, Vol. 18(3), pp. 269–287.

[19] Goverde, R., Hansen I.A. (2013) Performance Indicators for Railway Timetables. Proceedings of the 1st IEEE International conference on Intelligent Rail Transportation, Beijing, China.

[20] Hansen I.A., Pachl J. (2014) Railway Timetable and Traffic, Eurailpress, Hamburg.
[21] Janecek, D., Weymann, F. (2010). LUKS: Analysis of lines and junctions. Proceedings of the 12th World Conference on Transport Research (WCTR).Lisbon, Portugal.
[22] Kaminsky, R., Hauptmann, D., Radtke, A. (1996) Integrated planning-system for railways – A tool to improve the planning process. World Congress on Railway Research (WCRR) 1996, Colorado, Colorado, USA Volume 1
[23] Kauppi, A., J.Wikström, B. Sandblad, and A. W. Andersson (2006). Future train traffic control: control by re-planning. Cognition, Technology & Work, 8(1):50–56.
[24] Ke, B.R., C.L. Lin, H.H. Chien, H.W. Chiu & N. Chen (2015) A new approach for improving the performance of freight train timetabling of a single-track railway system, Transportation Planning and Technology, 38:2, 238-264, DOI: 10.1080/03081060.2014.959357
[25] Kozan, E., Robert Burdett (2005) A railway capacity determination model and rail access charging methodologies, Transportation Planning and Technology, 28:1, 27-45, DOI: 10.1080/0308106052000340378
[26] Kouwenhoven, de Jong, Koster, van den Berg, Verhoef,Bates, Warffemius (2014) New values of time and reliability in passenger transport in The Netherlands, Research in Transportation Economics, 47, 37-49.
[27] Kroon, L.G., D. Huisman, E. Abbink, P.J. Fioole, M. Fischetti, G. Maróti, A. Schrijver, A. Steenbeek, The new Dutch timetable: The OR revolution, Interfaces 39 (2009) 6—17.
[28] Lamorgese, L., Mannino, C. (2013) The track formulation for the train dispatching problem. Electronic Notes in Discrete Mathematics(41) 559-566.
[29] Larsen, R., Pranzo, M., D'Ariano, A., Corman, F., Pacciarelli, D. (2014) Susceptibility of Optimal Train Schedules to Stochastic Disturbances of Process Times, Flexible Services and Manufacturing Journal 26(4) 466-489.
[30] Liberzon D. (2005), Lyapunov functions and stability in Control theory, *Automatica*, Vol. 41 (12), pp. 2183-2184.
[31] Lindfeldt, O. (2011) An analysis of double-track railway line capacity, Transportation Planning and Technology, 34:4, 301-322, DOI: 10.1080/03081060.2011.577150
[32] Liu, S., Kozan, E. (2010) Scheduling Trains with Priorities: A No-Wait Blocking Parallel-Machine Job-Shop Scheduling Model. Transportation Science.
[33] Luethi M. (2009) Improving the efficiency of heavily used railway networks through integrated real-time rescheduling, PhD thesis, ETH Zurich.
[34] Mannino C., Mascis A. (2009) Optimal Real-Time Control in Metro Stations, Operations Research, Vol. 57(4), pp.1026-1039.
[35] Mazzarello M., Ottaviani E. (2007) A Traffic Management System For Real-Time Traffic Optimisation In Railways, Transportation Research Part B, 41(2), pp. 246-274.
[36] Medeossi, G., Longo, G., de Fabris, S. (2011) A method for using stochastic blocking times to improve timetable planning. Journal of Rail Transport Planning and Management 1 (1), 1–13.
[37] Meng L., Zhou X. (2011) "Robust Single-Track Train Dispatching Model Under A Dynamic And Stochastic Environment: A Scenario-Based Rolling Horizon Solution Approach", Transportation Research Part B, Vol. 45, pp.1080-1102.
[38] Meng, L. Zhou, X. (2014) Simultaneous train rerouting and rescheduling on an N-track network: A model reformulation with network-based cumulative flow variables. Transportation Research Part B, 67, 208-234.
[39] Metha, F., Roessiger, C., Montigel, M.,(2010) Latent energy savings due to the innovative use of advisory speeds to avoid occupation conflicts. *Computers in Railway XII*, WIT Press, pp. 99-108.

[40] Narayanaswami, S. Rangaraj, N. (2011). Scheduling and rescheduling of railway operations: A review and expository analysis, Technology Operation Management, vol. 2, no. 2, pp. 102–122.

[41] Nash, A., D. Hürlimann (2004) Railroad Simulation using OpenTrack, Proc. of the 9th International Conference on Computer in Railways (Comprail IX), eds. J. Allan, C.A. Brebbia, R.J. Hill, G. Sciutto & S. Sone, pp. 45-54, Dresden.

[42] OECD (2010). Improving reliability on surface transport networks. Paris, Organization for Economic Co-operation and Development.

[43] Parbo, J, Nielsen, O.A., Prato, C. G. (2014) User perspectives in public transport timetable optimization. Transportation Research Part C, 48, pp 269-284.

[44] Pellegrini, P., Marlière, G., Rodriguez, J., (2014). Optimal train routing and scheduling for managing traffic perturbations in complex junctions. Transport. Res. B 59, 58–80.

[45] Quaglietta E. (2011) A Microscopic Simulation Model For Supporting The Design Of Railway Systems: Development And Applications, PhD thesis, University of Naples Federico II, Naples.

[46] Quaglietta E., Corman F., Goverde R.M.P. (2013) Stability analysis of railway dispatching plans in a stochastic and dynamic environment, Journal of Rail Transport Planning and Management, Vol. 3 (4), pp.137-149

[47] Quaglietta, E., Pellegrini,P., Goverde,R.M.P., Albrecht, T., Jaekel, B., Marlière, G., Rodriguez, J., Dollevoet, T., Ambrogio, B., Carcasole,D., Giaroli, M., Nicholson, G. (2016) The ON-TIME real-time railway traffic management framework: A proof-of-concept using a scalable standardised data communication architecture. Transportation Research C 63, 23-50.

[48] Rietveld, P., Bruinsma, F.R., van Vuuren, D.J. (2001). Coping with unreliability in public transport chains: a case study for Netherlands, Transportation Research A, 35 (6), pp.539-559.

[49] Salido, M.A., Barber, F., Ingolotti, L. (2008) Robustness in railway transportation scheduling. In: 2008 7th World Congress on Intelligent Control and Automation. Chongqing, China, pp. 2880–2885.

[50] Schaafsma, A. A. M (2005). Dynamic traffic management – innovative solution for the Schiphol bottleneck 2007. Proceedings of the 1st International Seminar on Railway Operations Modelling and Analysis, Delft, The Netherlands.

[51] Schöbel, A., Kratz, A., (2009) A bicriteria approach for robust timetabling, In: Ahuja et al. (eds.), Robust and online large-scale optimization, LNCS, vol. 5868, 119– 144.

[52] Takeuchi Y., Tomii, N. (2005) Robustness indices for train rescheduling, Proceedings of Rail Delft – 1st International Seminar on Railway Operations Research.

[53] Törnquist J. (2007) Railway Traffic Disturbance Management-An Experimental Analysis Of Disturbance Complexity, Management Objectives And Limitations In Planning Horizon, Transportation Research Part A, Vol. 41, 249-266.

[54] Van Oort, N. van Leusden R. (2012) Quantifying benefits of enhanced service reliability in public transport. Proceedings of the 12th Conference on advanced systems for public transport, CASPT 2012, Santiago, Chile.

[55] Vromans, M., Dekker, R., Kroon, L. (2006) Reliability and heterogeneity of railway services, European Journal of Operational Research, 172:647-665.

[56] Watson, R. (2000) Prospects for computer aided railway scheduling: perspectives from users and parallels from mass transit, Transportation Planning and Technology, 23:4, 303-321, DOI: 10.1080/03081060008717655