A Survey on Robustness in Railway Planning

Lusby, Richard Martin ; Larsen, Jesper; Bull, Simon Henry

Published in:
European Journal of Operational Research

Link to article, DOI:
10.1016/j.ejor.2017.07.044

Publication date:
2018

Document Version
Peer reviewed version

Citation (APA):
Lusby, R. M., Larsen, J., & Bull, S. H. (2018). A Survey on Robustness in Railway Planning. *European Journal of Operational Research*, *266*(1), 1-15. https://doi.org/10.1016/j.ejor.2017.07.044

A Survey on Robustness in Railway Planning

Richard Lusby* Jesper Larsen† Simon Bull‡

October 9, 2017

Abstract

Planning problems in passenger railway range from long term strategic decision making to the detailed planning of operations.Operations research methods have played an increasing role in this planning process. However, recently more attention has been given to considerations of robustness in the quality of solutions to individual planning problems, and of operations in general. Robustness in general is the capacity for some system to absorb or resist changes. In the context of railway robustness it is often taken to be the capacity for operations to continue at some level when faced with a disruption such as delay or failure. This has resulted in more attention given to the inclusion of robustness measures and objectives in individual planning problems, and to the providing of tools to ensure operations continue under disrupted situations. In this paper we survey the literature on robustness in railway planning problems, considering how robustness is conceptualized and modelled for the individual problems of railway, the degree to which an overall railway robustness concept is present, and consider the future directions of robustness in railway planning.

Transportation, robustness, railway, optimization

1 Introduction

Operations Research (OR) plays a large and increasing role in the planning and execution of railway operations. As methods and approaches improve, and as rail utilization increases, it is increasingly important that solutions are not only of good quality during normal operations, but also perform well when encountering unexpected situations or true realizations of estimated parameters. OR can provide the tools to both assess and quantify planning problem solutions under uncertainty, and to find solutions that perform well despite uncertainty. Such well behaved plans may be considered "robust", and the quantification of performance may be a measure of robustness.

However, given the large scope of the different problems in railway planning and operations, there are many different interpretations and definitions of robustness. Some railway planning is carried out with a view to the long term rather than immediate operations, and consequently any robustness considerations would be different to robustness considerations in planning the immediate operations. Within the scope of rail, however, we aim to examine the different approaches to robustness and in particular consider the degree to which approaches to robustness in rail can be unified by their common application, despite being

*Department of Management Engineering, Technical University of Denmark, email: rmlu@dtu.dk
†*Corresponding author*, Department of Management Engineering, Technical University of Denmark, email jesla@dtu.dk
‡McKinsey & Company, Ved Stranden 14, 1061 Copenhagen, Denmark

different in scope and methodology. We therefore consider relevant, and include, not only research that defines and measures the robustness of solutions to railway planning problems, but also research which, through algorithmic implementation, attempts to optimize robustness measures. Research is categorized based on the particular planning problem it addresses, and the paper unfolds in a way that is consistent with how these problems appear in practice, starting with strategic planning problems and ending with more operational level planning. As such any paper addressing robustness aspects within the railway planning process is considered relevant. To our knowledge, we are the first to give such an all encompassing review on robustness on railway planning problems. A survey on literature that addresses passenger perspectives in railway timetabling, one aspect of robustness, can be found in Jensen et al. [2016]. Finally, this work primarily focuses on passenger railway but in some specific cases refer to relevant or related work in similar application areas of freight railway.

The remainder of the paper is organized as follows. In Section 2, we provide an introduction into robustness in general, while Section 3 gives an overview of the different railway planning stages, introducing railway robustness perspectives and the related concept of disruption management. Section 4 to Section 7 focus on specific levels in the railway planning process. Due to the vast amount of literature on railway timetabling, this section is subdivided into literature measuring robustness and literature optimizing robustness. Conclusions and future perspectives are given in Section 8.

2 Robustness

Robustness is a concept that exists in many fields with similar general interpretations, but at a detailed level robustness may be significantly different. Certainly, in cases where robustness is defined by some application-specific metric the definition has little use outside the application area. However, generally, robustness refers to how a system or plan behaves in the presence of uncertainty. This can for example be that some plan is created using estimates of parameters, but operated with a realization of those parameters that can differ from the estimates. A plan that behaves well under a wide range of realizations of parameters is considered robust, whereas a plan that behaves very differently or even fails under different realizations is less robust.

In the field of optimization, robustness has been considered in different ways. A common framework is termed *Robust Optimization*, which is surveyed by Ben-Tal and Nemirovski [2002]. This is strictly and clearly related to the conceptual idea that robustness is expressed in the context of an uncertainty of parameters or data for a problem. The following restrictions (among other) are imposed:

- The data are uncertain/inexact;
- The constraints must remain feasible for all meaningful realizations of the data.

More formally, the nominal (non-robust) definition of a Linear Programme (LP) can be stated as:

$$\min_x \{c^T x : Ax \leq b, x \in \mathbb{R}^n\}, \tag{1}$$

where x is a vector of decision variables, $c \in \mathbb{R}^n$ is a *known* vector of objective coefficients, $A \in \mathbb{R}^{m \times n}$ is a *known* matrix of constraint coefficients, and $b \in \mathbb{R}^m$ is a *known* vector of right-hand side values. In reality, parameters (c, A, b) are not known precisely, but lie in a so-called uncertainty set, U. This set contains all possible realizations of the uncertain input data. In robust optimization, this uncertainty can addressed by defining a set of LPs having a common structure.

$$\{\min_x\{c^T x : Ax \leq b, x \in \mathbb{R}^n\} : (c, A, b) \in U\},$$

Solving the following *robustness counterpart* then ensures that the obtained solution is the best solution that is "immune" to the uncertainty in the input parameters.

$$\min_{x,t}\{t \geq c^T x, Ax \leq b \;\forall (c, A, b) \in U, x \in \mathbb{R}^n\}$$

A potential problem with such an approach is that solutions can be somewhat pessimistic, by ensuring feasibility for any unlikely realization of uncertain parameters. Also, specifying the uncertainty set via linear constraints relating different uncertain parameters to each other may be difficult due to a lack of such data, whereas a lack of sufficient relationships between different uncertain parameters may result in solutions that are "robust" by being feasible in the case of particularly problematic but, in reality, impossible realizations of combinations of parameters.

Bertsimas and Sim [2004] offer a less conservative approach and define a "price of robustness" probabilistically, related to traditional robust optimization outline above but of a probabilistic nature. *... the trade-off between the probability of violation and the effect to the objective function of the nominal problem, ... is what we call the price of robustness*. The proposed approach assumes that data uncertainty is only present in the constraint matrix A. In other words, each element has a nominal value a_{ij} but due to uncertainty, in a real instance could be anywhere within the range $[a_{ij}, a_{ij} + \hat{a}_{ij}]$. The authors argue that it is extremely unlikely that all input parameters will simultaneously assume their worst case value. Instead, the authors reason that the number of "modified" coefficients in any constraint of the problem is bounded. Thus, any solutions that would be feasible given this restriction are probabilistically protected from uncertainty dependent on the probability that at most the specified number of coefficients is modified. More formally, in its robustness counterpart, and assuming the nominal problem has a set of decision variables N (with $n = |N|$), each constraint of the nominal problem outlined in (1) becomes:

$$\sum_{j=1}^{n} a_{ij} x_j + \beta(x, \Gamma_i) \leq b_i,$$

where

$$\beta(x, \Gamma_i) = max_{S \subseteq N : |S| \leq \Gamma_i} \sum_{j \in S} \hat{a}_{ij} x_j$$

gives the level of protection against the uncertainty that at most Γ_i coefficients in constraint i assume their worst case value, i.e, it states the largest increase in the left-hand side of constraint i at the optimal solution in such a case.

The price of itself is the increase in objective value over that given by solving only the nominal problem. Here, there is the explicit recognition that requiring robustness "costs" something, when compared with a solution that is apparently feasible with no consideration for robustness. In a problem (such as rail) where the nominal problem defines the most likely realization of parameters, and variation is unlikely but potentially problematic, this cost is something that negatively affects the normal operation for a benefit in the unusual situation. Assessing the value of such solutions, and whether or not the "cost" is worthwhile, requires detailed probabilistic knowledge of the likelihood of such realizations, and some idea of the "cost" incurred when a nominal, non-robust, solution would encounter such realizations. On the other hand, when compared with the robust optimization described above, solutions are not necessarily so pessimistic.

Fischetti and Monaci [2009] define the widely used "light robustness" framework. The authors introduce this as the "flexible counterpart" to robust optimization. The approach attempts to find a robust solution that is feasible for the nominal problem and which is within some threshold of the optimal nominal objective value, z^*. Slack variables are introduced on each of the constraints in the robust counterpart to guarantee feasibility, and the objective function is then to minimize the weighted sum of these variables. Using the robust optimization approach of Bertsimas and Sim [2004] as an example, the authors introduce slack variables $\gamma_i \geq 0$ $(\forall i = 1, \ldots, m)$, where γ_i indicates the magnitude of violation in the i-th constraint of the robust counterpart. The light robust counterpart therefore includes the following three types of constraints:

$$\sum_{j=1}^{n} a_{ij} x_j + \beta(x, \Gamma_i) - \gamma_i \leq b_i \qquad \forall i = 1, \ldots, m, \tag{2}$$

$$\sum_{i=1}^{n} c_j x_j \leq (1 + \delta) z^*, \tag{3}$$

$$\gamma_i \leq \beta(x, \Gamma_i) \qquad \forall i = 1, \ldots, m, \tag{4}$$

Constraints (2) are the now soft constraints from the robust optimization approach of Bertsimas and Sim [2004]. Constraints (3) restrict the deviation from the optimal nominal objective value, where the parameter $\delta \geq 0$ controls the level of deviation permitted. Finally, Constraints (4) enforce the maximum permitted slack on each constraint. The chosen values ensure that the optimal solution to the robust counterpart is also feasible for the nominal problem. The authors present railway timetabling as one possible application of light robustness. More recently, Schoebel [2014] generalizes this idea of light robustness, allowing arbitrary optimization problems and arbitrary uncertainty sets to be considered.

Stochastic programming provides a different framework for finding solutions to optimization problems in the presence of uncertainty. Kall and Wallace [1994] provide a textbook on stochastic optimization. Here, probability distributions for unknown parameters can be used to find a plan or policy that has high expected objective value, and is feasible for all realizations. Generally speaking, stochastic programming defines a problem in two or more *stages*, in which decisions may be made as uncertainty is revealed. If possible realizations are discrete in nature, this results in a tree structure of decision making with assigned possibilities. This tree may be fully enumerated, or, for tractability reasons, sampled as a limited set of discrete scenarios (see for example Kleywegt et al. [2002]). Rather than a full, fixed plan for operation, a solution is an initial plan dependent only on known problem parameters, and many contingency plans to be implemented as knowledge is revealed. In practice, two stage stochastic programmes are the most common. The first stage requires decisions to be made without knowledge of the values of the unknown parameters. In the second stage, the values of the unknown parameters are revealed, and recourse actions can be taken. With the such problems, the aim is to minimize the sum of the first stage costs and the expected value of all second stage costs.

Recoverable robustness (Liebchen et al. [2009]) is a new approach to robustness that, briefly, considers a nominal problem with scenarios and recovery possibilities; a solution to the nominal problem is *recoverable robust* if for any scenario, it can be recovered to feasibility using the given recovery possibilities. The authors give two rail applications as examples where recoverable robustness may be applicable. The authors contrast recoverable robustness with both robust optimization and stochastic optimization. Robust optimization can result in unnecessarily pessimistic solutions, because solutions must accommodate for every uncertainty simultaneously and without change. Stochastic programming instead permits some aspects of the solution to be fixed and solves the uncertain ones later in a second step; a 2-stage expansion in the scenarios. However, the authors claim that the complexity of their presented train timetabling problem results in problems that are too large and complex to approach with stochastic programming. Recoverable robustness, however, is claimed to combine “the flexibility of stochastic programming with the performance guarantee and the compactness of models found in robust optimization”.

As a short comparison, generally robust optimization is a worst-case planning method ensuring feasibility in all circumstances. As an advantage it does not require probabilistic information, which may be difficult to capture, with the associated risk that created plans are robust against extremely unlikely circumstances. If plans must be kept as originally specified then this guarantee of feasibility is more reasonable than in problems where reactive modifications are possible during operation. The concept of light robustness formulates robustness as a maximization of the level of “protection” in changes of problem parameters. This then permits less pessimistic solutions that robust optimization, although again not including probabilistic information. Stochastic optimization, in contrast, is probabilistic in nature, which has the potential downside of requiring the gathering or estimation of many parameters, generally related to a discrete number of disruption scenarios. Stochastic optimization can be a more natural modelling method if changes are to be made to plans as uncertainty is revealed, and a fixed unchanged plan is not a requirement.

3 Railway Planning

Passenger railway provids rail connections between stations as part of an overall public transportation system. It includes a range of operations such as commuter trains, metro systems, light rail (trams), and regional and intercity trains. Passenger railway interacts with other public transport systems such as buses and ferries. Passenger railway decisions take into account other modes of transport such as freight railway, car traffic and buses, due to both shared infrastructure (e.g., freight rail and passenger rail on shared tracks; light rail, cars and buses on roadways), and also to passengers moving between modes.

The railway planning problems differ between countries but in general there are several stages of planning from the most long term decision making to the day-to-day operations, and these stages are planned relatively independently and in sequence rather than as a single unified planning problem. The size and complexity of each individual problem alone tends to mean that only a sequential approach can be applied in practice. A typical sequence of planning problems may be:

- *Network design*
  Determining the topological structure of the rail network; station locations and capacities; see Magnanti and Wong [1984] for an overview of network design in general transportation.

- *Line planning*
  Determining the subsets of stations and a route between them that constitute train lines; deciding on frequencies, speeds, and rail stock types for each line. See Schöbel [2012] for a survey on line planning in public transport.

- *Timetabling*
  The determination of exact times for events that should occur for rail units such as driving between stations, the dwell times at platforms, and which specific station infrastructure is used. This planning level also addresses the issue of routing trains through dense, complex stations. A survey of railway track allocation, which includes timetabling, is given by Lusby et al. [2011]. A tutorial on non-periodic timetabling and platforming can be found in Cacchiani et al. [2015].

- *Rolling stock planning*
  The problem of allocating rolling stock units to the timetabled trips in circulations, considering

compositions and depot parking. An example of a nominal model for the problem is given by Fioole et al. [2006].

- *Crew scheduling*
  The problem of creating rosters of duties to be performed by a single crew person over a time horizon. There are similar problems in airline crew scheduling (which may be termed pairing), and likely many similar problems in other application areas. A following problem is assigning these rosters to specific crew members. Caprara et al. [1998] describe modelling the problem, with an application to Italian railways. Kohl and Karisch [2004] survey the literature on similar and well-studied airline crew rostering problems.

These stages may be further divided up, as for example crew scheduling often consists of two steps: creating anonymous crew schedules followed by crew rostering. Stages may also be combined, such as integrated line planning and timetabling (e.g., Schöbel [2015]). However there are clear differences and distinctions between the stages; network design for example is rarely changed, and decisions must stay valid for potentially many years, whereas crew rosters may be different every day. See Figure 1 for a suggested overview of how the different problems in rail planning may be carried out, indicating some estimate of the time they are considered before the day-of-operation (but not to scale). This sequence is commonly seen in the rail industry, and figures very similar to Figure 1 appear in other work (e.g., Liebchen and Möhring [2007]; Lusby et al. [2011]). We also indicate recovery but there are in fact several problems considered in the present, during operations, related to dispatching and managing delay and disruptions. However we do not focus on these present operational problems here.

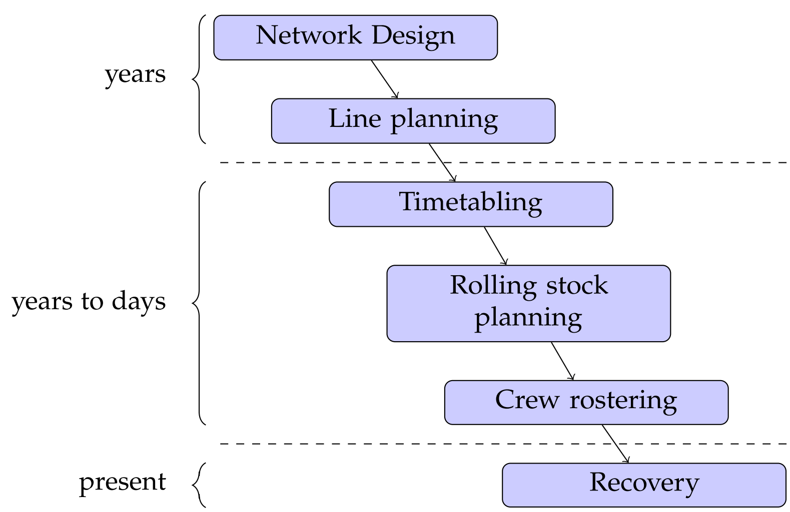

Figure 1: Overview of the possible steps of rail planning, indicating their relative time horizons.

The passenger data required for the strategic planning steps of network design, line planning, and timetabling is mostly based on transportation demands given by an origin-destination matrix (OD-matrix). Each entry in an OD-matrix gives the number of passengers that want to travel from one station in the network to another. It is specified over a fixed time horizon and therefore presents aggregated numbers for e.g. a day, the morning rush hour, or on an hourly basis. This data therefore only represents a snap shot and, in addition, information like service or price level, is not taken into account. Consequently, it is questionable how OD-matrices actually reflect real travel demand. Nevertheless, public transportation companies routinely develop and use OD-matrices despite such deficiencies.

3.1 Robustness perspectives

Figure 2 shows different considerations for a railway plan, and the trade-offs operators must generally make in creating a plan. High robustness can preclude (being close to) optimality, having high capacity utilization, and having a heterogeneous plan. In the general context of rail operations, capacity and heterogeneity do not necessarily have an obvious interpretation but capacity can be interpreted in relation to usage of infrastructure up to some maximum (and likely most fragile) level, and heterogeneity may be interpreted as a measure of the diversity of variety of rolling stock unit, or the variety in running times of trains on the same track, or the variety in the types of planned crew schedules. Heterogeneity has robustness implications for both small disruptions (e.g., dissimilar running times lead to delay propagation) and during large disruption scenarios (e.g., dissimilar unit types may not be usable as replacements). Other factors may also be included; for example speed (as in Salido et al. [2012]) may be reduced in a network, which also affects capacity and optimality with implications for robustness.

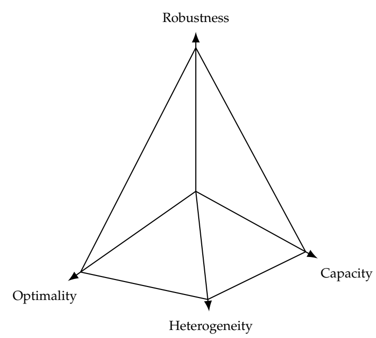

Figure 2: Considerations for a railway plan and their relationships (reproduced from Salido et al. [2008])

In the context of rail, robustness is a concept that is intuitively "obvious" or understandable, but there is no single clear definition. Further, any such proposed definition could not capture all elements of the intuitive "robustness" as understood by different stakeholders in public transport. There are in fact many differing definitions of robustness in the literature on rail planning, and indeed in other areas, but the core concepts are similar. In rail, as in most real-world problems, there are inherent uncertainties in aspects of the problem due either to inaccurate data, unpredictable occurrences, or stochastic processes.

Specifically referring to *passenger* robustness, De-Los-Santos et al. [2012] argue that generally agreed measurements of robustness do not exist. A passenger-robust timetable, the authors state, is one in which the travel time for passengers is not excessively increased when a particular link fails, and they define passenger-robust measurements indices. These are:

- the ratio of best-case travel time and the worst case (failure) for that connection;
- the ratio of best-case travel time and average case (failure) for that connection.

These can be calculated in the *with bridging* and *without bridging* cases, where bridging refers to the deployment of alternative routes by the operator (e.g., buses) rather than relying on passengers to re-route themselves or wait for the disruption to end. Dewilde [2014] investigates a definition of the robustness of a timetable, stating:

> A railway system that is robust against the daily occurring small disturbances minimizes the real weighted travel time of the passengers.

Dewilde [2014] considers the robustness of station areas. The author argues that as the purpose of rail networks is to serve passengers, a definition of robustness should (at least) consider passengers, while also arguing that there is no common measure of robustness in the literature. The weighted travel time assessment for passengers mentioned above, when the system is subjected to small disturbances, is argued to be the most relevant to users as it is a passenger-centric measure.

When solving railway planning problems, and when defining and quantifying robustness of operations, some authors take a point of view that indirectly benefits both the operator and the passenger (such as attempting to provide operations with little delay), while others explicitly take a viewpoint of the operator or of the passenger. While it is almost always true that it is the operator that makes decisions relating to robustness, the robustness itself can be viewed from the point of view of the operator or the point of view of the passenger, and the different perspectives do not always coincide. For example, a small delay to a train may not have a large affect on other operations, but the small delay may cause passengers to miss a transfer connection who may view the missed transfer as a symptom of poor robustness. Or, in contrast, to address an unexpected rolling stock problem an operator may be forced to use a different rolling stock unit than planned for, incurring additional operating cost, shunting and dead-head movements, but having no effect on passengers.

3.2 Disruption management

Operations are not always able to be carried out as planned, perhaps due to some unforeseen occurrence, an accident, or closure of track. In these situations plans must be modified or *recovered* to create a plan that can be carried out. There may be several recovery stages undertaken in sequence following the original stages of the tactical planning itself; for example a temporary track closure due to maintenance may necessitate a new timetable, then requiring modifications or recreation of the rolling stock and crew plans. It may be that the changes are known long enough in advance that planning can be done as comprehensively as for the normal case, or only known during operations and therefore recovery must be done quickly. During the planning stages, contingency plans may be created for some possible disruptions, and recovery may consist of finding an appropriate contingency plan and adjusting it as necessary to fit the exact situation. Decision support tools that assist in recovery situations may be explicitly related to the previously defined planning problems above (i.e. network design, line planning, timetabling etc.) or may span several (e.g., both rolling stock and crew).

In a small scale delay scenario, there can be decisions related to train waiting such as deciding whether a non-delayed train wait so that planned coordination between the trains is maintained. This problem, for example, is studied by Schöbel [2007] for public transport vehicles (e.g., buses or trains).

In contrast, during operations there may be instances of track closure or blockage that require significant modifications to plans for a possibly unknown duration. As some examples, Louwerse and Huisman [2014], Veelenturf et al. [2016], Vansteenwegen et al. [2016] study the problem of adjusting a passenger rail timetable. In Louwerse and Huisman [2014], the authors propose an approach to determine a new temporary timetable, and in doing so, which train series will be cancelled. However, they do not consider the transitions from the original timetable to the temporary timetable, and back again. The work of Veelenturf et al. [2016] extends that of Louwerse and Huisman [2014], and focuses on providing a complete rescheduling tool, with emphasis on the transition phases. The work of Vansteenwegen et al. [2016] deals with planned infrastructure maintenance and presents an algorithm to reroute, reschedule, and possibly cancel affected train services, while keeping the offered passenger service as high as possible.

4 Robustness in network design and line planning

In network design or line planning there has been little research into robustness, but some authors have attempted to define or incorporate robustness into line plans or even network design. The lack of a timetable at these stages make robustness measures related to exact passenger or train timings difficult to consider.

Goerigk et al. [2013a] investigates line plans or line "concepts", and their effect on resulting timetable quality and robustness. They question not whether the line plan *itself* is robust, but whether a resulting timetable for the line plan is or is not robust. The authors use two robustness measures: the relative increase in travel time for passengers under a disruption scenario, and the proportion of connections that would be missed in a disruption scenario. Their assessment is made via simulation, with a tool that both creates a timetable for a given line plan and provides the delay assessment. The authors examine two German intercity rail problems, with some provided and some estimated data and parameters, and the authors first find a line plan and then a timetable in an iterative way. Four different objectives or approaches for the line planning problem are used. Finally, the robustness of the resulting timetable is assessed. One conclusion is that when using an objective focused on the passenger, line plans offer nominally good solutions for passengers at the expense of robustness. One justification is that a good passenger solution can be considered one which provides many direct connections between stations, which is achieved with a variety of lines and tight transfers. There is then the potential for more missed transfers. In contrast, they observe that the (operator) cost oriented line plans have worse nominal passenger performance, but are more robust. The potential robustness of a resulting timetable is therefore very dependent on the goals of the line planning model.

Schöbel and Schwarze [2006] present a game-theory model for line planning, where individual lines seek to minimize their probability of delay, which depends on the presence and frequencies of other lines sharing infrastructure. The result is an overall line plan with an equally distributed probability of delay, which could be considered more robust than a plan where the delay probability is not equally distributed. As an exploratory numeric study the authors consider German intercity train lines, with some simplified details of the line planning problem, and show how their method is able to find equilibria for the problem.

A game theoretic approach is also adopted by Laporte et al. [2010], who address the problem of designing a robust railway transit network design in the presence of railway segment (or link) failures. For the duration of the failure it is assumed an alternative transport mode is available. The aim of the problem is to build a network which reacts well to link failure,i.e, a network that is still operative for the most number of passengers. The authors demonstrate that problem can be expressed as a non-cooperative two-player zero-sum game with perfect information.

Kontogiannis and Zaroliagis [2008] present a different view of robust line planning based on a network manager and independent line operators. The line operators compete to operate on shared infrastructure resources, managed by the single network manager. The incentive functions of the operators are the unknown parameters of the problem, and a robust line plan is one which maximises the aggregate utilities of the operators while being resilient to the unknown incentives of the line operators. The authors present a method for finding such robust solutions when certain conditions hold; however, the authors do not present a case study of the method.

Marín et al. [2009] consider robustness in both the line planning and network design problems. They consider robustness from two perspectives; the passenger and operator point of view, in the context of connection failures. Here, user robustness means that a failed connection should minimally affect travel time, and the passengers require geographically distinct routes but with similar travel times. Operator robustness, however, is measured in the cost of additional vehicles used to address such failures. The method developed is a heuristic that integrates both the network design and line planning problems. As a case study the authors consider the construction of a new high speed railway network in Andalucia (in southern Spain), and show how their method may be applied. They observe that both robustness measures (operator or passenger based) result in the same robust railway network but different line plans.

Michaelis and Schöbel [2009] present a heuristic integrating line planning, timetabling and vehicle scheduling. The focus is not on robustness, but of including robustness implicitly with slack times for the timetable, and the authors state that such additional time can provide robustness. However the slack times are not distributed in some way guided by robustness concerns; it is simply implied that the presence of slack times provides robustness.

Bull et al. [2016] consider a line planning problem with a nominal passenger focus, but some mention is made of metrics that may be relevant for robustness of a subsequent timetable. The robustness of the line plan *itself*, however, is not defined or given mention; only the degree to which the line plan influences the potential robustness of a timetable. The case study considered is from the S-tog network Copenhagen.

There is not a clear picture of what exactly robustness in line planning or network design should be, or what uncertainty it is that plans should be robust against. This may be explained in part by the fact that robustness in line planning and network design have not been extensively studied. One natural concern with a line plan, from a robustness perspective, is whether or not a robust timetable can subsequently be created. This is also a consideration for optimality, disregarding robustness, and both inspire a more integrated approach. Railway timetabling, examined in the following section, is the most studied planning problem from a robustness point of view, and naturally robustness viewpoints in earlier problems look to the timetable to assess robustness. Nonetheless some approaches do consider the earlier problems' inherent robustness, such as the approach of Marín et al. [2009] addressing network failure by requiring geographically distinct connections. Elements of the network design could also influence robustness of subsequent plans, especially taking a detailed view of track layout, crossings, and locations of overtaking sidings. However, generally, from an optimization point of view, such approaches to network design are not taken.

5 Robustness in timetabling

In the area of railway planning, timetabling is the area most studied for implementing robustness. This is perhaps natural due to timetabling being placed in a moderate planning stage, being not too far removed from the exact operations (as line planning and network design may be), but having large scope for changes that are not already very constrained (as rolling stock and crew scheduling may be). However, the planned timetable affects the robustness of the subsequent rolling stock and crew planning.

This section consists of three subsections. The first, Section 5.1, addresses literature that defines and measures robustness. The second, Section 5.2, surveys literature that attempts to optimize robustness measures. Non-exhaustive but representative sets of papers dealing with each of catergorizations are summarised in Table 1 on page 15 and Table 2 on page 21, respectively. Due to its interdependence with timetabling, Section 5.3 addresses robustness perspectives for the related station routing problem.

5.1 Defining and measuring robustness

Some work is concerned with the defining of robustness of a railway timetable, potentially for the assessment of different timetables, for the validation of methods producing robust timetables, or by defining metrics usable in models for creating railway timetables.

Many considerations of robustness relate to the distribution of buffer times. This is obviously more applicable to timetabling problems, where buffer times between trains can be included, and not directly to line planning or network design as exact times are generally not considered in those problems. Buffer times may also be included in rolling stock scheduling or crew scheduling though with a different interpretation: the buffer time between a particular unit finishing one service and beginning another. Buffer times mitigate the risk of the unknown precise *running times* and *dwell times* of trains. The running time refers to the time it takes a train to travel between two stations, while the dwell time indicates how long a train waits at a platform (after arrival) before departing. The latter, in particular, can be affected by the number of passengers that wish to board the train. Running times and dwell times are input parameters, but they are not necessarily known and are not fixed due to small delays. Buffer time supplementation is then an approach to create timetables that are still valid or operational for a range of uncertainties in input parameters. Up to some value over their anticipated duration, the timetable is still valid as the buffer time may be consumed. Smaller running times may also valid, if trains slow down or wait at stations. As pointed out by Dewilde [2014], much work on robustness in rail timetabling relates robustness to delay propagation, where a timetable that is unlikely to encounter propagating delay is more "robust", and the buffer time is almost always related to the absorption of delay and subsequent avoidance of delay propagation.

Buffer time can be considered a directly applicable or measurable metric for robustness, for some small deviation from expected running times. However for larger deviations the buffer times alone are not necessarily an adequate measure, as the role of train dispatchers is neglected. For example, dispatchers may decide to change the order of trains leaving a station due to a delay, choosing to change the plan to avoid inducing delay for other trains. Also, additional buffer time does not necessarily mean additional robustness, as the introduced buffer time may be excessive or introduced in the wrong part of the timetable. For example buffer time at the end of the day may be less beneficial than buffer time early in the day, as early delays can have longer lasting effect than delays late in the day. Deciding on the amount of buffer time to use is a decision on the trade-off between (potentially) increased robustness and a decrease in utilization of the rail network and a decrease in passenger service. More buffer can for example require fewer trains operating and greater waiting time for passengers transferring between different trains.

As a support for determining the propagation of delay and addressing of it in planning, Flier et al. [2009] present methods for the derivation of delay dependencies in actual delay data. The authors formulate and identify possibly timetabled dependencies between trains where one train may wait for another, and then quantify the correlation between relevant event times for those trains using data from certain points of the Swiss SBB network. The authors show that correlation alone, without the correct underlying model of the relationship, can fail to identify delay relationships between trains, while with a model and data relationships can be discovered.

Carey [1999] discusses "heuristic" measures of the reliability of a public transport schedule which can be calculated in advance, and distinguishes heuristic measures from analytic and simulation methods. It is stated that, for measuring schedule robustness, knock-on delays (where one train being delayed causes another to be delayed) are most relevant. The measures are divided into those that are probabilistic in nature, deriving measures of knock-on delay based on probability density functions for departure times of trains in the absence of delay, and those that are not probabilistic and are largely related to headway. The headway, or the interval between subsequent trains, is used by many authors as a measure of robustness or a target for robustness in various ways in other measures of robustness. In this case, the measures are not validated with real data and instances.

Considering the interval between events in a timetable, Goverde [2007] propose methods for analysing properties of railway timetables, including timetable robustness. Their robustness measure is related to slack times between events; the slackness between two events can be quantified by identifying a *critical path* which is the path relating those events with the least slack time of all such event paths. They also propose a delay propagation model to quantify the effect of some initial delays, though these measures are not explicitly equated with robustness by the authors. As a case study the authors use the Dutch national railway timetable, including both passenger and freight rail, and show how delay propagation may be quantified and show the identification of critical timetable elements. The identification shows *critical cycles* of dependent events that begin at some event and return the same event in a later timetable period, that consume all of the time of the timetable cycling period. Without considering planned excess time within events, such cycles have no capacity to absorb additional delay meaning initial delay in such a cycle will continue into later timetable periods.

Hofman et al. [2006] present a simulation study on the robustness of commuter train timetables in Copenhagen, Denmark. There is no single precise definition of what robustness is, but the study considers the addition of slack time, and different recovery strategies for late trains. The authors observe an upper limit to the effectiveness of additional buffer time, above which there is low effect. Early turn-around is one recovery method the authors consider, which is a late train stopping at a non-terminal station and becoming a train of the same line in the opposite direction, and the strategy is shown to be just as effective as the more drastic cancellation of late trains for some cases.

Vromans et al. [2006] propose methods to increase the *reliability* of railway services. Their approach is to increase homogeneity of the timetable, by reducing instances of trains running at different speeds in track sections. The authors are interested in small disturbances, and in fact state that *no* timetable is robust enough to handle larger disturbances (without major online adjustments). As a case study the authors consider a single line in the Netherlands, and compare a real "heterogeneous" timetable, and a "homogenized" version. They see a decrease in delay measure by simulation, and state that the improvement is predictable with the measures Sum of Shortest Headway Reciprocals (SSHR) and Sum of Arrival Headway Reciprocals (SAHR) that they introduce in this paper. For an ordering of trains running on a sequence of tracks, suppose train i precedes train $i+1$, and h_i^- is the minimal headway between trains i and $i+1$ in the track sequence (and final train N precedes train 1 in a cyclic manner).

$$\mathrm{SSHR} = \sum_{1}^{N} \frac{1}{h_i^-} \tag{5}$$

Claiming that headway at arrival in a section is more significant that headway later in the section, then authors define h_i^A to be the headway between trains i and $i+1$ at arrival in the track sequence.

$$\mathrm{SAHR} = \sum_{1}^{N} \frac{1}{h_i^A} \tag{6}$$

In the homogeneous case, where train speeds are all the same, both measures are equal and have minimal value when all trains are evenly spaced in the cyclic interval and has maximum value if they are "bunched".

Another train buffer related measure that appears in several works (such as Andersson et al. [2013], Fischetti et al. [2009], Kroon et al. [2007], Vromans [2005]) is Weighted Average Distance (WAD) of time supplements added to trips along a line, as a measure of how the supplements are distributed. The measure is intended to indicate the degree to which supplement or buffer is biased towards the start of the line, the end, or evenly distributed. If there are $N+1$ trips on a line, from trip $t=0$ to trip $t=N$, and each trip has supplement time s_t, then we may define WAD as follows:

$$\text{WAD} = \frac{1}{N} \frac{\sum_{i=0}^{N} i \cdot s_i}{\sum_{i=0}^{N} s_i} \tag{7}$$

Here, a WAD value of 0 would indicate that all buffer is allocated to trip $t=0$; a WAD value of 1 would indicate that all buffer is allocated to trip $t=N$, and a WAD of 0.5 would indicate that there is equal buffer in both the first half and second half of the trip.

Andersson et al. [2013] define a robust timetable as one which can maintain its planned operation, when subjected to some small delays, and also focus on the headways between trains. They define *critical points* in the timetable, that are a relationship between two trains at a location and time in the timetable, from which delay can propagate because the trains are running in the same direction on the same track, or where one is planned to overtake another. Critical points come from a timetable study by the same authors; Andersson et al. [2011]. The critical points are used to make timetable modifications such as headway increases. The authors measure the robustness of their original and modified timetables with other timetable-robustness measures from the literature. The comparative measures from literature are: WAD of runtime supplements from the start of a line (from Kroon et al. [2007], Fischetti et al. [2009], Vromans [2005]); SSHR (citing Salido et al. [2008], Vromans et al. [2006]). The measures of Andersson et al. [2013] are:

- Maximum runtime distance (MRD)
  Defined for parts of the network where the parts "are naturally bounded by the traffic structure"
- Total amount of runtime margin for each individual train (TAoRM)
- Robustness in critical points (RCP)
  A quantification of flexibility at the identified critical points, calculated for a pair of subsequent trains as the runtime margin between trains, the runtime margin for the earlier train to be earlier at the point, and for the later train to exit later from the point.

As a case study, the authors apply their methods to part of the Swedish mainline (a 200km double track in southern Sweden, between Malmö and Alvesta), and demonstrate the identification of potential modifications to increase robustness, showing how the different robustness measures may be interpreted in practice.

Landex and Jensen [2013] develop measures for analysing rail operations at stations. The authors suggest two methods that they specifically relate to robustness. The first is buffer time threshold, defined as a ratio of high risk conflicts, where buffer times between train are below some given threshold, to the total number of such potential conflicts. The second is and a complexity measure based on a probabilistic determination of trains delaying other trains. As an example application the authors apply the measures to Skanderborg station in Denmark, with a real timetable instance and two instances derived from the real timetable to achieve goals related to buffer times. The authors find that the modified timetable where small buffers are increased has the lowest value of timetable complexity (an indication of the interdependencies between train movements) and best level of robustness, as measured by the delay probability, and also deduce that that particular measure is the best for timetable comparison while noting that an analysis of more stations should be done for a better picture of overall timetable complexity.

Both Salido et al. [2008] and Salido et al. [2012] develop methods for comparing the robustness of timetables, where robustness is described as the ability *to absorb short disruptions*. Both papers address single railway line networks, and in both cases analytic approaches for measuring robustness are compared with simulation-based approaches In Salido et al. [2012], the quantification the authors present is of total delay encountered by other trains, given some initial input delay, and for both the case of homogeneous trains and heterogeneous trains the authors see agreement in total delay between the simulation and analytic approaches. The analytic methods consist of calculating the total delay of all trains for some initial delay in a corridor based on the absorption of the delay through headway, its propagation to subsequent trains, and by considering the possibility of overtaking. A timetable is said to be more robust than another if its total delay for the same initial input delay is lower. Salido et al. [2008] is somewhat similar; however, overtaking is not considered. Two analytic measures of robustness based on the size and location of buffer time and delay absorption are compared to an evaluation of robustness using simulation. A case study from Spain is used to validate the approaches. Generally speaking, analytic methods have the advantage of being computationally efficient compared to the simulation method and the authors suggest they are applicable in the generation of robust timetables.

Corman et al. [2014b] examine robustness of a train timetable as the performance of a timetable when faced with large scale disturbances (examples given are multiple train delays, speed restrictions, track blockages). The authors use different measures and compare different timetables under a range of disruption scenarios reporting values such as train delay and passenger travel time as a result of a disruption, taking those as indicating how robust the timetable was. The focus is not on the creation of timetables but rather of comparing timetables and responses to disruption with a robustness view. As a case study the authors use a section of the Dutch railway network, and compare the regular timetable and a shuttle timetable in which trains drive back and forward between pairs of major cities. For comparison the authors compare the disruption management methods of re-timing, where train order is maintained but times are altered; and rescheduling by changing train order using the method of Corman et al. [2014a]. The authors show results for both train delay and for passenger delay, showing the potentially conflicting experience of robustness passengers and the operator may have.

Takeuchi and Tomii [2006] (and Takeuchi et al. [2007]) define a robustness index for train timetables based on passenger disutility. The authors claim that any robustness measure should be defined from the passengers' perspective, not from the operator's perspective. The authors define passenger disutility as a weighted sum of congestion discomfort, number of transfers, waiting time at stations, and boarding time into trains (dependent on the number of other passengers boarding). The robustness index is the expected *increase* in this passenger disutility, given a sampling of delay scenarios, an estimate of the resultant train operations, and passenger re-routing given the new train operation. The authors give experimental results for an unspecified real urban train line, and compare the actual train timetable and a proposed modified timetable by re-allocating running time supplements. The authors show that their robustness index indicates better performance for the modified timetable. The method, however, is for the quantification of the robustness index and not for the modification itself.

Dewilde et al. [2011] present a claimed all-embracing, generic definition of the robustness of a railway timetable:

> "A railway timetable that is robust against small delays minimizes the real total travel time of passengers, in case of small delays. Limited knock-on delays and a short settling time are necessary but not sufficient conditions for a timetable to be robust. Furthermore, different weights can be assigned to different kinds of travel time prolongation."

Here the authors have a clear passenger focus; their robustness measure only relates to passengers; not to the increase in operator cost in the case of delays or in the nominal case to guarantee the robustness. The authors apply their method to the 2010 Belgian timetable, using both real and artificially modified timetables with increased running time supplement. The authors show that, while the timetables with additional running time supplement have fewer delays, they have worse robustness as calculated using their definition. The measure is also strictly related to small delays. In fact, many measures and metrics encountered in this section are implicitly or explicitly related to small delays only. However in the following section, applying different methods, more authors seek robustness as large-scale disruption robustness.

Table 1 summarises work defining or measuring railway timetable robustness, classifying whether they have a passenger or operator focus and noting whether there is a clear application to some real timetable problem instance.

Table 1: Summary of literature measuring or defining timetable robustness

Reference	Focus	Description	Application
Andersson et al. [2013]	Operator	Metric based assessment based on critical points from which delay can propagate.	Swedish single line
Carey [1999]	Mixed	Heuristic measures for timetable robustness.	-
Corman et al. [2014b]	Mixed	Analysis of train and passenger delay in the case of major disruption.	Southern Dutch network
Dewilde et al. [2011]	Passenger	Define robustness as realized passenger travel time, encountering small delays.	Belgian timetables
Flier et al. [2009]	Operator	Methods for the derivation of delay dependencies in actual delay data.	Several important points of the Swiss network
Goverde [2007]	Operator	Timetable assessment based on identification of critical paths and their relation to delay propagation.	Single dutch line
Hofman et al. [2006]	Operator	Assessment of timetables and recovery strategies; considers buffer times.	Copenhagen commuter timetable
Landex and Jensen [2013]	Operator	Station-centric robustness analysis; buffer time and probabilistic delay propagation measures.	Danish station
Salido et al., Salido et al. [2008, 2012]	Operator	Analysis based on delay propagation.	-
Takeuchi and Tomii [2006]	Passenger	Method for comparing timetable based on delay-induced passenger travel time.	Real, unspecified
Vromans et al. [2006]	Operator	Metric-based assessment, introducing two measures of heterogeneity and spreading of trains.	Dutch single line

5.2 Optimization with robustness

Cacchiani and Toth [2012] surveys the literature on the rail timetabling problem, in its regular and robust forms. Different approaches to optimizing timetable robustness are identified and these include the following: stochastic programming, light robustness, recoverable robustness, delay management, and bi-criteria approaches in that survey.

Cacchiani et al. [2012a] present a lagrangian relaxation heuristic for robustness problems, applying the method to the train timetabling problem. The method finds solutions that are approximately pareto optimal considering nominal efficiency and an estimate of robustness (against small delay). The solutions are validated using the tool of Fischetti et al. [2009], reporting average total delay. In the method itself, however, robustness is estimated as a dynamically weighted profit for buffer time for certain types of train event. The method is applied to corridor-focused Italian Railway instances, and also compared with results of Fischetti et al. [2009] showing better results in many cases found with less computation time.

Cicerone et al. [2008a] expand the concept of recoverable robustness, considering the *recoverability* of a timetable or more generally some schedule in the face of several recovery steps. This is applied to the train timetabling and delay management problem. Cicerone et al. [2009] consider the timetabling problem using the concept of recoverable robustness and quantifying the “price of robustness”. Several algorithms are proposed, which improve upon similar work and algorithms by Cicerone et al. [2008b]. The authors are explicit in stating the relationship between robust timetabling and delay management. They do not make an experimental study applying their methods to a real world problem, but note the value in doing so as future work. Outside the scope of passenger rail, Cicerone et al. [2007] apply similar concepts of recoverable robustness to the shunting of trains.

Delay management is the focus of Schachtebeck and Schoebel [2010]. The authors are primarily concerned with generating a so-called disposition timetable from the original timetable when delays occur, determining not only wait-depart decisions (i.e, which timetable connections should be kept, but also priority decisions concerning the order of trains on specific pieces of track. The authors are the first to consider the priority decisions for capacitated networks and present an integer linear programming formulaion of the problem. The methodology is tested on a railway network in Harz, Germany.

Goerigk and Schoebel [2014] define robustness in terms of a “recovery to optimality”; a solution is robust if, faced with the realization of uncertain parameters, can be recovered to an optimal solution with low recovery cost. Here, then, robustness is related to large scale disruption requiring intervention and recovery; not solutions that are still valid in the presence of small scale disruption. The authors state that their intention is to create a solution that is not strictly robust but rather one that can be recovered easily to an optimal, or good quality, solution for any disruption scenario. As an example the authors present an aperiodic timetabling application (an LP), with a specific created instance based on the German intercity rail system, and compare several different timetables:

1. A nominal (non-robust) solution
2. A strictly robust (robust optimization) solution
3. A lightly robust solution
4. A recoverable robust solution
5. A recoverability to optimality solution
6. A uniformly buffered solution (with 6% extra buffer time applied everywhere)

Testing the solutions using two different samplings of uncertainty, they show and conclude that their approach can find solutions that have good nominal objective values, and perform particularly well for a worst case recovery to optimality, and not as well simply recovering to feasibility. The authors suggest that their approach is effective in cases where a large-scale disruption is expected to occur, with the exact details unknown but with set of possible scenarios determined, such as knowledge that one of several disruptive construction plans will go ahead. Their approach leads to a solution that may be recovered or modified to a good solution in all cases, with low cost measured as difference from the nominal plan. However the authors note that generally, periodic timetables are sought and nominal periodic timetabling is (or can be) modelled as a Mixed Integer Programme (MIP), and not as an LP problem. They conclude that their method would require adaption.

Schöbel and Kratz [2009] use a bi-objective approach to robust aperiodic timetabling, using as one objective the nominal timetable quality, and as the other a robustness measure. The authors give three possible measures of robustness: the largest initial delay for which no passenger misses a transfer; the maximum number of passengers who could miss a transfer if all delays are bounded by some given value; the maximum accumulated passenger delay that could occur if all delays are bounded by some given value. The first would be maximized, and the second and third minimized. These, then, are examples of passenger-centric robustness approaches. The authors show how pareto-optimal solutions may be found for the problem by a formulation as a timetable problem with a robustness level parameter. Solutions to this problem, given some robustness level, are pareto optimal for the bi-objective formulation considering passenger travel time and robustness as objectives. The authors do not show results for a particular case study but rather consider formulations, and propose the idea of using a bi-objective for robustness in general.

A bi-objective approach is also described in Schlechte and Borndoerfer [2010] for optimizing timetable efficiency and robustness. On the one hand, the authors want to schedule as many trains as possible, while on the other, they would like to ensure that buffer times between trains are not too short. To find all efficient solutions of the bi-objective problem, a combined weighted sum and $\epsilon-$ constraint hybrid method utilizing column generation is outlined. Experiments are performed on part of the German long-distance railway network containing 37 stations.

Liebchen et al. [2010] introduce robustness into timetable planning by finding a periodic timetable with high delay resistance. The authors describe their robustness approach as an extension of light robustness. For the construction of timetables, the authors use a weighted sum of passenger travel time and expected deviation (delay) from the published time. The expected delay is computed in a simplified way in the constructive model using scenarios with given probability and input delays, and subsequent passenger delay is calculated using a no-wait policy where non-delayed trains all operate on time (which may be infeasible in practice). However, results are validated using a more detailed delay management model. Delay in the authors context is many, small delays. Their solutions are assessed with multiple input delays of at most 20 minutes, while in the creation of timetables delays of up to 40 minutes are considered. As an example, the authors apply their method to passenger railway lines in the Harz region of northern Germany and show timetables can be created that substantially reduce expected passenger delay with only minimal increase in nominal cost.

Vansteenwegen and Van Oudheusden [2006] use a passenger-centric approach to the cost of a timetable, and use a method to modify times of a given timetable to reduce expected passenger wait times when trains are delayed. The authors state that a robust timetable is one that performs well in non-ideal circumstances, and therefore their planning by including expected delay can be considered more robust than one that is created without considering the delay. The authors apply the method to passenger trains in a small part of the Belgian railway network, and their results show an improvement in overall waiting time although they also see an increase in passengers who miss their connection. The method uses arriving train delay probability, justifying their choice of distribution with delay data from the rail operator. The methodology is improved by Vansteenwegen and Van Oudheusden [2007] and applied to the full Belgian network. Their results show that the timetable created with their LP method outperforms the original timetable for several performance metrics, including passengers missing connections, and also show a reduction in long waiting times even in the no-delay case.

Fischetti et al. [2009] are concerned with the modification of railway timetables to introduce robustness, where robustness is related to the absorption of minor delays, and explicitly *not* delay requiring major adjustment. The robustness is validated with a separate validator that tests a solution on a number of delay scenarios, permitting limited adjustment but again stating that robustness is not related to major disruptions. Four methods are proposed: one effective but slow stochastic programming formulation; two "slim" stochastic programming methods, and a method based on light robustness. The full stochastic programming formulation minimizes delay time, similar to the validator, whereas the two slim formulations minimize weighted unabsorbed delay time (differing only in weighting). Instances are single-line problems and come from the Italian railway operator. As one measure, the authors use WAD and are able to draw a relationship between WAD and their measured robustness. The authors show that with a weighting putting more importance on earlier events than later events, both the light robustness formulation and the slim stochastic programming formulation perform well.

Kroon et al. [2008] consider the railway timetabling problem with stochastic disturbances, and equate robustness with the ability of the timetable to cope with those disturbances. They address the problem by a (re-)allocation of running time supplements, planning for certain events to take longer than their minimum required time. They state that their approach improves robustness to small delays only, and furthermore state that most robustness work in literature is focused on such small delay resilience. The work is an allocation of buffer time. Given a timetable, buffer time can best be allocated to achieve robustness to delay. The model itself is a two-stage stochastic optimization model, consisting of the creation and assessment of the timetable (considering average weighted delay of trains), by a sampling of possible delay scenarios. As a case study the authors apply their method to the train timetable for a northern section of the Dutch railway network, and compare the original to a timetable they create where running times of trains may change by only one minute, and another where running times may change by any (feasible) amount, and see improvement in (train) delay and punctuality over the original timetable even with only the small modification permitted. The authors sample input delay from a distribution that permits up to 10 minutes of delay.

While the solutions obtained by Kroon et al. [2008] are generally accepted by practitioners, the solution times exceeds what is allowed in an operational context. In Marti [2017] a branch and bound algorithm for solving the stochastic program of Kroon et al. [2008] is described. Computational tests are carried out on large real-life problem instances.

Also taking a stochastic optimization approach to railway timetabling for small disturbances, Kroon et al. [2007] study an optimal placement of running time supplements along a train line. One measure used is WAD, as defined above, and the authors observe that with different amounts of total slack allocated optimally, the distribution favours early buffer supplement (low WAD) when less total buffer supplement is available.

Meng and Zhou [2011] consider the problem of determining robust meet-pass plans when dispatching trains along a single track network. A two stage stochastic programming framework with recourse is devised to consider, by way of scenarios, the impact of probabilistic rail segment running times. The methodology is embedded in a rolling horizon framework in which a stochastic program combined with a multi layer branching strategy is used to select a robust horizon plan. Computational experiments focus on a 138km single track network connecting 18 stations in China. The robust solutions compare well to their expected value counterparts.

Shafia et al. [2012a] also address the problem of generating a robust timetable for single track networks. Like Meng and Zhou [2011], the robustness of the timetable is measured in its sensitivity to variations in rail segment times, and the aim is therefore to compute the necessary buffer times between successive train movements. The authors proposed a mixed integer programming approach and present two possible solution methods. The first is a branch-and-bound algorithm, while the second is a heuristic beam search method. Both approaches perform well on a set of 20 randomly generated numerical instances.

Hassannayebi et al. [2016] develop a robust stochastic programming approach to generate train timetables for rapid rail transit lines. Robustness in this context, is in line with that in Ben-Tal and Nemirovski [2002], and refers to the sensitivity of the solution to changes in the values of input parameters. The authors minimize the expected average passenger time as well as a cost variance and define two multi-objective stochastic programming formulations. Computational results from the Tehran underground railway indicate the potential of the methodology.

A tool for evaluating the robustness of a train schedule in the presence of stochastic disturbances is considered in Larsen et al. [2013]. The authors do not consider techniques for incorporating robustness when building schedules, but rather attempt to quantify a schedule's susceptibility to delay propagation. Monte Carlo simulation is used to assess the impact of variations in train running times and dwell times through a comparison with a deterministic variant in which no variation is assumed. Computational experiments are performed on a dense part of the Dutch railway network. The results indicate that a deterministic Branch-and-Bound approach for generating the train schedule generally provides the most robust solutions when stochasticity is considered. More recently, Khadilkar [2016], has addressed a similar problem, developing a delay propagation algorithm to evaluate schedule robustness. The proposed algorithm evaluates robustness using two metrics. The first focuses on individual trains and considers the potential for a train to recover its delay, while the second attempts to limit the impact of knock-on delays on the rest of the network. Empirical data from the Indian Railway network is used to confirm the applicability of the approach in practice.

Goverde et al. [2016] describe a three-level framework that integrates the construction and evaluation steps to ensure a high quality, robust timetable is obtained. The framework is comprised of a macroscopic timetabling module, a microscopic timetabling module, and a mesoscopic module to fine tune the running of trains on corridors between main nodes. The performance of a timetable is measured over several criteria, one of which is the timetable's "stability" (i.e., its ability to absorb delays). The authors also introduce an additional measure focusing on the energy efficiency of the timetable. The methodology is tested on a component of the Dutch railway network with promising results.

Sels et al. [2016] also address the problem of automatic, robust timetable construction and propose a MIP to achieve this. The proposed model minimizes expected passenger travel time in practice. Results on the full Belgian railway network can be obtained in approximately two hours. The obtained timetable is not only more robust, with a reduction in expected passenger travel time by 3.8%, but also more punctual than the original timetable. From a modelling perspective, the MIP model devised avoids issues caused by artificial upper bound supplements.

Burggraeve et al. [2017] consider an integrated approach to robustness of the line planning and timetabling problems, in a heuristic framework. The S-tog network of Copenhagen is again studied, and the line plans are modified by stopping pattern to alter the running times of lines using information from a timetable model that aims to build a robust timetable for the line plan. The overall aim is to create a robust timetable but, where this is not possible due to the line plan, modify the line plan to further facilitate timetable robustness. The claimed robustness improvement is related to buffer time in the created timetable, where, in creating the timetable, the goal is to maximize the smallest buffer time between trains.

The optimal allocation of buffer time placement when scheduling trains along single or double tracked corridors is considered by Jovanovic et al. [2017]. The authors model the problem as multidimensional knapsack problem, and solve the corresponding mathematical formulations using the Minizinc opensource integer programming solver. The methodology is tested on a busy mixed traffic rail corridor connecting the Hallsberg marshalling yard and Mjölby station in Sweden. Evaluation of the resulting train schedules via simulation demonstrates that the obtained schedules are more resistant to delay propagation.

Shafia et al. [2012b] also consider robust scheduling for mixed traffic, single track networks; however, the focus is on generating periodic timetables. The problem is formulated as a fuzzy periodic job shop scheduling problem and solved using a tailored simulated annealing metaheuristic. Robustness is measured in the timetables ability to absorb delays. Numerical experiments focus on a single track corridor connecting two of the most populated cities in Iran and results indicate that the proposed approach is able to quickly produce near optimal solutions.

Finally, considering the timetable, but entirely from a passenger's perspective, Goerigk et al. [2013b]

considers the concept of robustness in the *timetable information* problem, which is the problem of finding a passenger path, given a timetable, which minimises travel time and/or the number of transfers. A robust form of the problem is one which guarantees an unchanged, or minimally changed travel time, or likely unchanged travel time, given some set of disruption scenarios. The authors give a strictly robust formulation, where a path maintains validity in all delay scenarios, and a light robustness formulation where the path length is bounded to be close to the optimal path length, but additionally should have as few unreliable transfers as possible. Data instances are from the German rail schedule, and comprise either the full network or only high speed trains. One conclusion is that strict robustness is too costly to the passenger in terms of travel time to be used in practice, while in contrast the light robustness approach leads to solutions that are often as good in terms of robustness and suffer a much lower penalty in travel time.

Table 2 presents two ”reference papers” in each of the categories of bi-criteria approaches, recoverable robustness, light robustness, stochastic programming and delay management. In the table we give a summary of literature which optimizes robustness in railway timetables, classifying whether they have a passenger or operator focus and noting whether there is a clear application to some real timetable problem instance.

5.3 Station routing

When generating a railway timetable, typically a macroscopic network is considered. In such a network, stations are simply considered as nodes, and the routing possibilities within each station are ignored. Once arrival and departure times for each of the trains have been obtained, the microscopic network for each station is considered and routes must be determined for each train within each station. This section focuses on work that addresses robustness aspects for routing within station areas.

Caimi et al. [2005] consider the routing of trains through a station, where the timetable and station infrastructure layout are given but exact routings must be determined. The authors seek delay-resistant routings, measuring the maximum deviation from schedule an individual train may encounter without conflict if every other train is on time, and use a weighted sum over all trains as an objective to maximize in a local search heuristic method. The authors present results for application to Bern station in Switzerland and can find much improved solutions, as measured by their objective.

Caprara et al. [2010] also consider the problem of determining platforms and routes through the station infrastructure. The plans are assess using external delay and considering subsequent delay, when using different strategies to address the delay. The plans themselves, however, are created with one of two different ideas intended to introduce robustness. The first is to increase the robustness to small delay, by ensuring events have a large capacity to absorb delay. The second is a method of reserving both a platform and a backup platform for trains, as a contingency plan in the case of delay. In their Genova case study, this backup platform method is shown to be effective at reducing overall delay.

In an approach integrating routing and platforming with timetabling, Dewilde et al. [2013] consider the improvement of robustness at a single station. The authors argue that passenger service should be measured by realized total travel time, and that a robust railway system is one that minimizes real total travel time in the face of small, frequently occurring disturbances. The authors define a *weighted travel time extension* as

$$\frac{\text{realised passenger travel time} - \text{nominal travel time}}{\text{nominal travel time}}$$

This is for some system a measure of its robustness, and this may be compared to other systems by the percentage difference. They would say that system 1 is x-percent more robust than system 2 if the weighted travel time extensions of the systems compare in that way. In modelling the problem, the authors use a

Table 2: Summary of literature implementing robustness in timetable optimization

Reference	Focus	Description	Application
Bi-criteria approaches			
Cacchiani et al. [2012a]	Operator	Lagrangian heuristic for creating robust timetables for small delay with low cost to the nominal objective.	Italian "corridor" instances
Schöbel and Kratz [2009]	Passenger	Bi-objective approach to robustness, with application to aperiodic timetabling; theoretical. Gives passenger-based nominal and robustness definitions.	-
Recoverable robustness			
Cicerone et al. [2009]	Mixed	Study of the timetabling problem in terms of recoverable robustness; theoretical.	-
Goerigk and Schoebel [2014]	-	Finds aperiodic timetables that are recoverable to optimality in disruption scenarios; theoretical.	German Intercity System
Light robustness			
Fischetti et al. [2009]	Operator	Modify timetables to be more robust to small delay; present methods based on stochastic programming and light robustness.	Italian single line instances
Goerigk et al. [2013b]	-	Concepts of strict and light robustness for determining a best path for traveling in a public transportation network.	Complete German network and German high-speed network
Stochastic programming			
Kroon et al. [2008]	Operator	Stochastic-based modification to a timetable to improve expected punctuality in the face of input delay.	Northern Dutch network
Meng and Zhou [2011]	Operator	Robust time plan for a single track network	Track network in China
Delay management			
Liebchen et al. [2010]	Passenger	Light robustness (inspired) method to create timetables resistant to delays	North German network
Schachtebeck and Schoebel [2010]	Passenger	Priority decisions are added to the disposition timetable and heuristics and optimal methods are derived.	Rail network in the Hartz region of Germany

measure of the spreading of trains and assess their true defined robustness with a simulation method when a solution is found. The method itself consists of a routing module, a timetable module, and a platform assignment module in a heuristic framework. The authors study two Belgian station areas: Brussels containing three closely linked stations, and Antwerp, and in both cases show that they can find solutions that are 8% more robust than the original solutions.

Burggraeve and Vansteenwegen [2017] also address the problem of constructing a passenger robust timetable for dense railway networks. Like Dewilde et al. [2013], the authors integrate the timetabling problem with routing decisions. However, contrary to the conventional planning sequence of timetabling then routing, the routing problem is solved first. The authors attempt to first assign routes to trains that are sufficiently separate from a spatial perspective. To achieve this, the objective of the routing problem minimizes the maximum usage of any node in the network, where the use of each node is quadratically penalized. The timetabling procedure then uses the assigned routes and attempts to maximize a the minimum buffer time. The resulting timetable is simulated to analysis its robustness from a delay propagation perspective. Computational experiments are performed on the same Brussels case study from Dewilde et al. [2013], and the results suggest further improvements can be made in terms of robustness.

Finally, Jin et al. [2014] present work on the *resilience* of a metro network to disruptions, analogous to many definitions of robustness. Resilience is introduced by integrating the metro operation with another transport mode (buses), such that in the case of a metro disruption alternative capacity and routing may be provided by the unaffected alternate mode.

6 Robustness in rolling stock planning

Planning which rolling stock units from the available fleet will perform which timetabled trips is one area of railway planning in which the robustness introduced in the earlier planning problems, in particular line planning and timetabling, can be adversely affected. A rolling stock fleet is usually comprised of several different types of units, where all units of the same type have the same physical characteristics. Compositions of units – groups of units coupled together – are assigned to timetabled trips in such a way that the supplied capacity in terms of seat numbers matches the forecast passenger demand as closely as possible. As passenger numbers fluctuate during the course of a day, the composition assigned to a particular train is not usually static, and changes regularly occur. The (de)coupling activities associated with a composition change can introduce delay susceptibility to a schedule, particularly if these take longer than anticipated. Some common themes in robustness for rolling stock planning therefore include: minimizing composition changes, restricting composition changes, and ensuring homogeneous compositions. Papers addressing these issues are discussed below.

Abbink et al. [2004] present a model for the allocation of rolling stock units to timetabled trips, with an objective of minimising relative seat shortages. However, they do indicate the variety (i.e., lack of homogeneity) of rolling stock unit types as being indicative of a lack of robustness. Via constraints, the number of types of units allocated to a train line is limited, arguing that such limits lead to improved robustness at some cost to the nominal objective, although this cost of robustness is not quantified. The authors apply their model to an instance of regional trains in the Netherlands.

Alfieri et al. [2006] mention robustness of railway rolling stock operation briefly; they relate a minimum turning time requirement to the robustness of a rolling stock plan, which comes with a cost. They also briefly mention the impact of required shunting on robustness, where required shunting in a plan has a risk of delay, and is therefore less robust than a plan with less required shunting. However the work is more generally focused on rolling stock circulation without a clear theme of robustness, and is applied to an intercity line in the Netherlands as a case study.

Cadarso and Marín [2010] address the rolling stock routing problem for the rapid transit routing problem, and consider some robustness aspects. Robustness is included with two measures; the first is based on penalising propagated delay, which depends on expected arrival delay and assigned slack time to absorb the delay using historic arrival delay distributions. The second is by penalising certain operations which require additional crew. The authors apply their method to two instances of the Madrid suburban rail network; one of a single line, and another of two related lines; and show a reduction in expected delay when compared with the nominal solution.

Cadarso and Marín [2011] also consider the same two instances from the Madrid suburban rail network, considering the rolling stock assignment problem. Robustness here is said to be introduced by penalising solutions that require composition changes during rush hours, and empty train movements during rush hours, as both are claimed to risk leading to propagation of delay. The robust solutions presented do indeed reduce the number of empty movements at rush hour significantly, and the number of composition changes can be reduced, with a price for robustness paid for as increased operator cost but not in the service offered to the passenger. The same two instances from the Madrid suburban rail network are considered by Cadarso and Marín [2014], solving the rolling stock assignment and train routing problem for rapid transit simultaneously using a method based on a Benders' decomposition solved heuristically. Robustness considerations are introduced by penalising risky empty train movements and shunting movements, and expected delay, in a weighted objective with operator costs. Two other robustness considerations are a per-passenger-in-excess penalty on trains as very full arrivals at stations can lead to congestion, and finally restricting the number of unit types assigned per line. This is similar to Abbink et al. [2004]. In the case study of the two Madrid instances, the authors show improved solution quality over the nominal solution, and also find better solutions with less cost for including robustness over a sequential approach solving assignment and routing separately.

Cacchiani et al. [2012b] present an optimization model based on recoverability when faced with different disruption scenarios. It should be noted that in the nominal problem, robustness considerations may be included by a count of composition changes arguing that such events can lead to delay propagation. In their full formulation, robustness is considered by simultaneously solving the rolling stock circulation plan and a recovery plan (with an estimate of the recovery cost) for a number of disruptions. They also consider the solutions afterwards on a larger range of scenarios. The authors consider the worst case recovery scenario, rather than an expected value and therefore do not require probabilities for scenarios. The authors apply their method to an intercity line in the Netherlands, and when compared to a nominal optimal solution, can find solutions that experience fewer cancelled trips and have lower shunting costs for recovery when assessed with many scenarios. Their Benders' decomposition method is solved in a heuristic way, and estimates the true recovery cost for scenarios. The authors observe that the estimate is generally accurate, although note some cases where it makes a significant underestimate.

6.1 Rolling stock maintenance

Periodically rolling stock units need maintenance. Planning how units that require maintenance will transition from a planned schedule to the specified maintenance facility is an extremely important problem that may indirectly impact any robustness incorporated in the earlier planning stages. The problem of so-called maintenance routing has not been overly studied in the literature. Attempts can be found in, e.g. Maroti and Kroon [2005] and Wagenaar et al. [2017]. In Maroti and Kroon [2005], the authors model the problem as a variant of the "multi depot vehicle routing problem" and use the commercial solver CPLEX to solve the resulting integer programming formulation. Results on real-life instances provided by the

Netherlands Railways indicate the potenial for using the methodology in practice; however, the authors also highlight the possibility of generating feasible solutions to the model that are infeasible in practice. The work of Wagenaar et al. [2017] considers maintenance appointments when rescheduling rolling stock units after a disruption has occurred; the rolling stock schedule must be revised in a such a way that all maintenance appointments are retained. Three mixed integer programming formulations are developed and compared on a large number of practical instances provided by Netherlands Railways.

7 Robustness in crew planning

The size and complexity of crew scheduling and rostering problems for large railway companies calls for the application of sophisticated optimization algorithms. Some examples include Abbink et al. [2005], Caprara et al. [1998] and Freling et al. [2001]. Caprara et al. [1998] propose an integer programming formulation of the crew rostering problem and describe a heuristic algorithm, which utilizes a Lagrangian lower bound, as a possible solution method. Promising results are reported for an application arising at the Italian railways having approximately 1000 duties. Freling et al. [2001] consider the crew scheduling problem at Netherlands railways and present a heuristic branch and price algorithm. Computational results for instances with up to 1114 tasks can be solved to near optimality in reasonable time. Robustness is incorporated in the duty construction by penalizing changes from one train to another. However, this resilience to delay propagation is neither investigated nor quantified. Abbink et al. [2005] also address crew scheduling at Netherlands Railways. A set covering formulation is proposed and solved using the commercial solver TURNI. Impressive operational cost savings are reported.

With the exception of Freling et al. [2001] perhaps, robustness in crew scheduling has not been as thoroughly considered as robustness in other areas of railway. In contrast, in the more studied field of airline crew scheduling, robustness has been an area of study, such as the seminal paper of Ehrgott and Ryan [2002] which describes a bi-objective approach. The aim of this approach is to maximize the robustness of the schedule (from a delay propagation perspective), but also minimizes the operational cost.

Robustness aspects have also been considered for the related problem of scheduling bus drivers and vehicles by, e.g, Huisman and Wagelmans [2006]. The authors present a comparison of two different approaches to solve this problem, a static approach and a dynamic approach. For the static approach, the impact of fixed buffer durations between successive trips is analysed. While fixed buffers guard against delay propagation, the authors show, using realistic test cases from The Netherlands, that significantly more drivers and vehicles are needed as the buffer time is increased. The dynamic approach schedules drivers and vehicles dynamically, assuming the exact travel time for a trip is known a short time, i.e., 2 hours, prior to its departure. A sensitivity analysis is performed to analyse the impact of small variations in these trip times, and limited impact is reported. In addition to the static and dynamic variants, for each method the authors compare the traditional sequential approach of bus then driver scheduling with an integrated approach. Results suggest the integrated approach slightly outperforms the sequential methodology.

In the field of freight railway, Jütte et al. [2011] describe the modelling and implementation of a crew scheduling system for the DB Schenker German freight network. Robustness is not a focus of the work, but some robustness considerations are included. For example, buffer times are included when a crew person changes train, and the changes themselves are (optionally) penalised to avoid duties for a crew person that included multiple train changes. Another consideration is that employment rules dictate that a maximum of ten hours may be worked in a duty, and if during operation this would be exceeded for example due to delay. In that case, the remaining tasks must be reallocated which is costly and can cause further delay. The authors consider restricting the maximum number of work hours in a duty to be less than ten hours to reduce the risk of reallocation being necessary, showing results in ten minute increments for maximum work of between nine hours and ten hours. Robustness is improved as measured by a lower risk of disruption, based on historic delay data, with an increase in operator cost as the price for the robustness increases.

Potthoff et al. [2010] present a column generation algorithm for the rescheduling of crew in disruption cases with data from Netherlands Railways. The authors use a weighted objective with one consideration being robustness; they suggest penalising any change to the original schedule (as fewer drivers need to be informed of changes), and penalising any short transfers between consecutive tasks on different rolling stock units. However in the presented experiments the penalty for short transfers is set to zero. The authors show the results for their algorithm using ten instances and different problem parameters; however the exact effect of the robustness considerations are not exactly quantified or described.

Rezanova and Ryan [2010] consider the train driver recovery problem of rescheduling train driver duties in the case of disruption, using data from the Danish state railways, and also present a column generation based approach to solving the problem by expanding a neighbourhood of duties to consider changing. As one robustness consideration, the authors penalise changes to the original undisrupted schedule, and assign higher penalties to new duties based on the degree to which they differ from original duties. The authors also claim aim to ensure robustness by including a minimum idle time above the required minimum between duties, giving as an example 5–7 minutes for recovery duties. The authors show results for generated test instances based on historical data but the effect of the included robustness concerns are not quantified.

8 Conclusion

Robustness has become an increasingly significant factor in both research and application for railway planning problems. There has been a clear focus on robustness of the timetable, but more authors are applying similar ideas to other areas of railway planning such as rolling stock planning. Robustness work also appears in railway crew planning, and in the related field of airline crew planning robustness features.

Some work has appeared considering the integration of different planning problems of railway, and consequently some work features robustness in integrated problems. For example the integration of network design and line planning has been studied with robustness considered, as have the line planning timetabling problems.

There have been several metrics and measures for robustness derived for railway systems, mostly but not exclusively focusing on the timetable. However it is not evident that a single metric will ever be derived that satisfies all interested parties, as different metrics focus on different stakeholders and some are even contradictory.

Without integrating planning problems, robustness planned in the individual planning steps may be lost in the overall plan. For example a rolling stock plan designed to have recoverable-robustness, where in large scale disruption scenarios the rolling stock plan can always be recovered to a good plan, may in fact lose its robustness if planned crew schedules can not feasibly cover the required recovery train movements. A robust line plan may not facilitate robust timetables to be planned; or a robust timetable may lead to non-robust rolling stock and crew plans.

Measures of robustness indeed generally focus on a particular part of the plan such as the timetable or rolling stock plan. They are robustness of the timetable, or robustness measures of rolling stock; not robustness measures of passenger railway. While a completely integrated planning approach may not be achievable, an integrated metric of robustness may be more feasibly derived and provide insight into the possible loss of robustness from non-integrated planning. We do not yet see the conceptualization of robustness in railway planning problems in a holistic manner, but as metrics and concepts become more developed for the individual problems we may see more parallels between them.

We may expect to see more formulations that treat robustness as an additional objective function and therefore more multi-objective optimization approaches to planning with robustness. As robustness is generally not measured in units comparable to other costs, methods that provide several pareto-optimal solutions for manual planners to assess is important and will also be important in the future. Certainly this is an area attracting present research, and approximately half the works cited in this review are from the year 2010 or more recent.

Acknowledgement

The authors would like to acknowledge the constructive and thorough reviews made by the anonymous referees. Their comments on the text and their suggestions on additional references have led to a substantial improvement of the paper.

This work was supported the Innovation Fund Denmark under grant number 11-115893.

References

Erwin Abbink, Bianca Van den Berg, Leo Kroon, and Marc Salomon. Allocation of railway rolling stock for passenger trains. *Transportation Science*, 38(1):33–41, 2004.

Erwin Abbink, Matteo Fischetti, Leo Kroon, Gerrit Timmer, and Michiel Vromans. Reinventing crew scheduling at netherlands railways. *Interfaces*, 35(5):393–401, 2005.

Arianna Alfieri, Rutger Groot, Leo Kroon, and Alexander Schrijver. Efficient circulation of railway rolling stock. *Transportation Science*, 40(3):378–391, 2006.

Emma Andersson, Anders Peterson, and Johanna Törnquist Krasemann. Robustness in swedish railway traffic timetables. In *4th International Seminar on Railway Operations Modelling and Analysis, Sapienza-University of Rome, February 16-18, 2011*, 2011.

Emma V. Andersson, Anders Peterson, and Johanna Törnquist Krasemann. Quantifying railway timetable robustness in critical points. *Journal of Rail Transport Planning and Management*, 3(3):95–110, 2013. doi: 10.1016/j.jrtpm.2013.12.002.

Aharon Ben-Tal and Arkadi Nemirovski. Robust optimization–methodology and applications. *Mathematical Programming*, 92(3):453–480, 2002.

Dimitris Bertsimas and Melvyn Sim. The price of robustness. *Operations research*, 52(1):35–53, 2004.

Simon H Bull, Natalia J Rezanova, Richard M Lusby, and Jesper Larsen. An optimization based method for line planning to minimize travel time. Technical report, Technical University of Denmark, 2016.

Sofie Burggraeve and Pieter Vansteenwegen. Robust routing and timetabling in complex railway stations. *Transportation Research Part B: Methodological*, 101:228–244, 2017.

Sofie Burggraeve, Simon Henry Bull, Pieter Vansteenwegen, and Richard Martin Lusby. Integrating robust timetabling in line plan optimization for railway systems. *Transportation Research. Part C: Emerging Technologies*, 77:134–160, 2017.

Valentina Cacchiani and Paolo Toth. Nominal and robust train timetabling problems. *European Journal of Operational Research*, 219(3):727–737, 2012.

Valentina Cacchiani, Alberto Caprara, and Matteo Fischetti. A lagrangian heuristic for robustness, with an application to train timetabling. *Transportation Science*, 46(1):124–133, 2012a.

Valentina Cacchiani, Alberto Caprara, Laura Galli, Leo Kroon, Gábor Maróti, and Paolo Toth. Railway rolling stock planning: Robustness against large disruptions. *Transportation Science*, 46(2):217–232, 2012b.

Valentina Cacchiani, Laura Galli, and Paolo Toth. A tutorial on non-periodic train timetabling and platforming problems. *Euro Journal on Transportation and Logistics*, 4(3):285–320, 2015.

Luis Cadarso and Ángel Marín. Robust routing of rapid transit rolling stock. *Public Transport*, 2(1-2):51–68, 2010.

Luis Cadarso and Ángel Marín. Robust rolling stock in rapid transit networks. *Computers & Operations Research*, 38(8):1131–1142, 2011.

Luis Cadarso and Ángel Marín. Improving robustness of rolling stock circulations in rapid transit networks. *Computers & Operations Research*, 51:146–159, 2014.

Gabrio Caimi, Dan Burkolter, and Thomas Herrmann. Finding delay-tolerant train routings through stations. In *Operations Research Proceedings 2004*, pages 136–143. Springer, 2005.

Alberto Caprara, Paolo Toth, Daniele Vigo, and Matteo Fischetti. Modeling and solving the crew rostering problem. *Operations research*, 46(6):820–830, 1998.

Alberto Caprara, Laura Galli, Leo Kroon, Gbor Marti, and Paolo Toth. Robust train routing and online re-scheduling. *Openaccess Series in Informatics*, 14:24–33, 2010.

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

Serafino Cicerone, Gianlorenzo D'Angelo, Gabriele Di Stefano, Daniele Frigioni, and Alfredo Navarra. Robust algorithms and price of robustness in shunting problems. In *ATMOS*, pages 175–190. Citeseer, 2007.

Serafino Cicerone, Gabriele Di Stefano, Michael Schachtebeck, and Anita Schöbel. Dynamic algorithms for recoverable robustness problems. In *ATMOS*, 2008a.

Serafino Cicerone, Gianlorenzo DAngelo, Gabriele Di Stefano, Daniele Frigioni, and Alfredo Navarra. Delay management problem: Complexity results and robust algorithms. In *Combinatorial Optimization and Applications*, pages 458–468. Springer, 2008b.

Serafino Cicerone, Gianlorenzo DAngelo, Gabriele Di Stefano, Daniele Frigioni, and Alfredo Navarra. Recoverable robust timetabling for single delay: Complexity and polynomial algorithms for special cases. *Journal of combinatorial optimization*, 18(3):229–257, 2009.

Francesco Corman, Andrea D'Ariano, Dario Pacciarelli, and Marco Pranzo. Dispatching and coordination in multi-area railway traffic management. *Computers & Operations Research*, 44:146–160, 2014a.

Francesco Corman, Andrea DAriano, and Ingo A Hansen. Evaluating disturbance robustness of railway schedules. *Journal of Intelligent Transportation Systems*, 18(1):106–120, 2014b.

Alicia De-Los-Santos, Gilbert Laporte, Juan A. Mesa, and Federico Perea. Evaluating passenger robustness in a rail transit network. *Transportation Research Part C*, 20(1):34–46, 2012.

Thijs Dewilde. *Improving the robustness of a railway system in large and complex station areas*. PhD thesis, Arenberg doctoral school, Faculty of Engineering Science, KU Leuven, December 2014.

Thijs Dewilde, Peter Sels, Dirk Cattrysse, and Pieter Vansteenwegen. Defining robustness of a railway timetable. In *25th Annual Conference of the Belgian Operations Research Society*, pages 108–109, 2011.

Thijs Dewilde, Peter Sels, Dirk Cattrysse, and Pieter Vansteenwegen. Robust railway station planning: An interaction between routing, timetabling and platforming. *Journal of Rail Transport Planning and Management*, 3(3):68–77, 2013.

Matthias Ehrgott and David M Ryan. Constructing robust crew schedules with bicriteria optimization. *Journal of Multi-Criteria Decision Analysis*, 11(3):139–150, 2002.

Pieter-Jan Fioole, Leo Kroon, Gábor Maróti, and Alexander Schrijver. A rolling stock circulation model for combining and splitting of passenger trains. *European Journal of Operational Research*, 174(2):1281–1297, 2006.

Matteo Fischetti and Michele Monaci. Light robustness. *Lecture Notes in Computer Science), Lect. Notes Comput. Sci*, 5868:61–84, 2009.

Matteo Fischetti, Domenico Salvagnin, and Arrigo Zanette. Fast approaches to improve the robustness of a railway timetable. *Transportation Science*, 43(3):321–335, 2009.

Holger Flier, Rati Gelashvili, Thomas Graffagnino, and Marc Nunkesser. Mining railway delay dependencies in large-scale real-world delay data. In *Robust and online large-scale optimization*, pages 354–368. Springer, 2009.

Richard Freling, Ramon M. Lentink, and Michiel A. Odijk. *Scheduling Train Crews: A Case Study for the Dutch Railways*, pages 153–165. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001. ISBN 978-3-642-56423-9.

Marc Goerigk and Anita Schoebel. Recovery-to-optimality: A new two-stage approach to robustness with an application to aperiodic timetabling. *Computers and Operations Research*, 52:1–15, 2014.

Marc Goerigk, Michael Schachtebeck, and Anita Schöbel. Evaluating line concepts using travel times and robustness. *Public Transport*, 5(3):267–284, 2013a.

Marc Goerigk, Marie Schmidt, Anita Schöbel, Martin Knoth, and Matthias Müller-Hannemann. The price of strict and light robustness in timetable information. *Transportation Science*, 48(2):225–242, 2013b.

Rob M P Goverde, Nikola Bešinović, Anne Binder, Valentina Cacchiani, Egidio Quaglietta, Roberto Roberti, and Paolo Toth. A three-level framework for performance-based railway timetabling. *Transportation Research. Part C: Emerging Technologies*, 67:62–83, 2016.

Rob MP Goverde. Railway timetable stability analysis using max-plus system theory. *Transportation Research Part B: Methodological*, 41(2):179–201, 2007.

Erfan Hassannayebi, Seyed Hessameddin Zegordi, Mohammad Reza Amin-Naseri, and Masoud Yaghini. Train timetabling at rapid rail transit lines: a robust multi-objective stochastic programming approach. *Operational Research*, 2016. doi: 10.1007/s12351-016-0232-2. Available online.

Mads Hofman, Line Madsen, Julie Jespersen, Jens Clausen, and Jesper Larsen. Robustness and recovery in train scheduling-a case study from DSB s-tog a/s. In *ATMOS*, 2006.

D Huisman and APM Wagelmans. A solution approach for dynamic vehicle and crew scheduling. *European Journal of Operational Research*, 172(2):453–471, 2006.

Jens Parbo Jensen, Otto Anker Nielsen, and Carlo Giacomo Prato. Passenger perspectives in railway timetabling: A literature review. *Transport Reviews*, 36(4):500–526, 2016.

Jian Gang Jin, Loon Ching Tang, Lijun Sun, and Der-Horng Lee. Enhancing metro network resilience via localized integration with bus services. *Transportation Research Part E: Logistics and Transportation Review*, 63:17–30, 2014.

Predrag Jovanovic, Pavle Kecman, Nebojsa Bojovic, and Dragomir Mandic. Optimal allocation of buffer times to increase train schedule robustness. *European Journal of Operational Research*, 256(1):44–54, 2017.

Silke Jütte, Marc Albers, Ulrich W Thonemann, and Knut Haase. Optimizing railway crew scheduling at DB schenker. *Interfaces*, 41(2):109–122, 2011.

Peter Kall and Stein W Wallace. *Stochastic programming*. Springer, 1994.

Harshad Khadilkar. Data-enabled stochastic modeling for evaluating schedule robustness of railway networks. *Transportation Science*, 2016. doi: 10.1287/trsc.2016.0703. Available Online.

Anton J Kleywegt, Alexander Shapiro, and Tito Homem-de Mello. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12(2):479–502, 2002.

Niklas Kohl and Stefan E Karisch. Airline crew rostering: Problem types, modeling, and optimization. *Annals of Operations Research*, 127(1-4):223–257, 2004.

Spyros C Kontogiannis and Christos D Zaroliagis. Robust line planning under unknown incentives and elasticity of frequencies. In *ATMOS*, 2008.

Leo Kroon, Gábor Maróti, Mathijn Retel Helmrich, Michiel Vromans, and Rommert Dekker. Stochastic improvement of cyclic railway timetables. *Transportation Research Part B: Methodological*, 42(6):553–570, 2008.

Leo G. Kroon, Rommert Dekker, and M. J C M Vromans. Cyclic railway timetabling: A stochastic optimization approach. *Lecture Notes in Computer Science*, 4359:41–68, 2007.

Alex Landex and Lars Wittrup Jensen. Measures for track complexity and robustness of operation at stations. *Journal of Rail Transport Planning & Management*, 3(1):22–35, 2013.

Gilbert Laporte, Juan A. Mesa, and Federico Perea. A game theoretic framework for the robust railway transit network design problem. *Transportation Research Part B-methodological*, 44(4):447–459, 2010.

Rune Larsen, Marco Pranzo, Andrea DAriano, Francesco Gorman, and Dario Pacciarelli. Susceptibility of optimal train schedules to stochastic disturbances of process times. *Flexible Services and Manufacturing Journal*, 26(4):1–24, 2013.

Christian Liebchen and Rolf H Möhring. The modeling power of the periodic event scheduling problem: railway timetablesand beyond. In *Algorithmic methods for railway optimization*, pages 3–40. Springer, 2007.

Christian Liebchen, Marco Lübbecke, Rolf Möhring, and Sebastian Stiller. The concept of recoverable robustness, linear programming recovery, and railway applications. In *Robust and online large-scale optimization*, pages 1–27. Springer, 2009.

Christian Liebchen, Michael Schachtebeck, Anita Schöbel, Sebastian Stiller, and André Prigge. Computing delay resistant railway timetables. *Computers & Operations Research*, 37(5):857–868, 2010.

Ilse Louwerse and Dennis Huisman. Adjusting a railway timetable in case of partial or complete blockades. *European Journal of Operational Research*, 235(3):583–593, 2014.

Richard M Lusby, Jesper Larsen, Matthias Ehrgott, and David Ryan. Railway track allocation: models and methods. *OR spectrum*, 33(4):843–883, 2011.

Thomas L Magnanti and Richard T Wong. Network design and transportation planning: Models and algorithms. *Transportation science*, 18(1):1–55, 1984.

Ángel Marín, Juan A Mesa, and Federico Perea. Integrating robust railway network design and line planning under failures. In *Robust and online large-scale optimization*, pages 273–292. Springer, 2009.

Gabor Maroti and Leo Kroon. Maintenance routing for train units: The transition model. *Transportation Science*, 39(4):518–525, 2005.

G. Marti. A branch-and-bound approach for robust railway timetabling. *Public Transport*, 9(1-2):73–94, 2017. doi: 10.1007/s12469-016-0143-x. cited By 0.

Lingyun Meng and Xuesong Zhou. Robust single-track train dispatching model under a dynamic and stochastic environment: A scenario-based rolling horizon solution approach. *Transportation Research Part B-methodological*, 45(7):1080–1102, 2011.

Mathias Michaelis and Anita Schöbel. Integrating line planning, timetabling, and vehicle scheduling: a customer-oriented heuristic. *Public Transport*, 1(3):211–232, 2009.

Daniel Potthoff, Dennis Huisman, and Guy Desaulniers. Column generation with dynamic duty selection for railway crew rescheduling. *Transportation Science*, 44(4):493–505, 2010.

Natalia J Rezanova and David M Ryan. The train driver recovery problema set partitioning based model and solution method. *Computers & Operations Research*, 37(5):845–856, 2010.

Miguel A Salido, Federico Barber, and Laura Ingolotti. Robustness in railway transportation scheduling. In *Intelligent Control and Automation, 2008. WCICA 2008. 7th World Congress on*, pages 2880–2885. IEEE, 2008.

Miguel A Salido, Federico Barber, and Laura Ingolotti. Robustness for a single railway line: Analytical and simulation methods. *Expert Systems with Applications*, 39(18):13305–13327, 2012.

Michael Schachtebeck and Anita Schoebel. To wait or not to wait-and who goes first? delay management with priority decisions. *Transportation Science*, 44(3):307–321, 2010.

Thomas Schlechte and Ralf Borndoerfer. Balancing efficiency and robustness a - bi-criteria optimization approach to railway track allocation. *Lecture Notes in Economics and Mathematical Systems*, 634:105–116, 2010.

Anita Schöbel. Integer programming approaches for solving the delay management problem. *Lecture Notes in Computer Science*, 4359:145–170, 2007.

Anita Schöbel. Line planning in public transportation: models and methods. *OR spectrum*, 34(3):491–510, 2012.

Anita Schöbel. Integration of line planning and timetabling. In *Proceedings of the Conference on Advanced Systems in Public Transport*, 2015.

Anita Schöbel and Albrecht Kratz. A bicriteria approach for robust timetabling. *Lecture Notes in Computer Science*, 5868:119–144, 2009.

Anita Schöbel and Silvia Schwarze. A game-theoretic approach to line planning. *Openaccess Series in Informatics*, 5, 2006.

Anita Schoebel. Generalized light robustness and the trade-off between robustness and nominal quality. *Mathematical Methods of Operations Research*, 80(2):161–191, 2014.

Peter. Sels, Thijs. Dewilde, Dirk. Cattrysse, and Pieter. Vansteenwegen. Reducing the passenger travel time in practice by the automated construction of a robust railway timetable. *Transportation Research Part B-methodological*, 84:124–156, 2016.

Mohammad Ali Shafia, Mohsen Pourseyed Aghaee, Seyed Jafar Sadjadi, and Amin Jamili. Robust train timetabling problem: Mathematical model and branch and bound algorithm. *Ieee Transactions on Intelligent Transportation Systems*, 13(1):307–317, 2012a.

Mohammad Ali Shafia, Seyed Jafar Sadjadi, Amin Jamili, Reza Tavakkoli-Moghaddam, and Mohsen Pourseyed-Aghaee. The periodicity and robustness in a single-track train scheduling problem. *Applied Soft Computing*, 12(1):440–452, 2012b.

Yoko Takeuchi and N Tomii. Robustness indices based on passengers utilities. *WCRR2006, Montreal*, 2006.

Yoko Takeuchi, Norio Tomii, and Chikara Hirai. Evaluation method of robustness for train schedules. *Quarterly Report of RTRI*, 48(4):197–201, 2007.

Pieter Vansteenwegen and Dirk Van Oudheusden. Developing railway timetables which guarantee a better service. *European Journal of Operational Research*, 173(1):337–350, 2006.

Pieter Vansteenwegen and Dirk Van Oudheusden. Decreasing the passenger waiting time for an intercity rail network. *Transportation Research Part B-Methodological*, 41(4):478–492, 2007.

Pieter Vansteenwegen, Thijs Dewilde, Sofie Burggraeve, and Dirk Cattrysse. An iterative approach for reducing the impact of infrastructure maintenance on the performance of railway systems. *European Journal of Operational Research*, 252(1):39–53, 2016.

Lucas P. Veelenturf, Martin P. Kidd, Valentina Cacchiani, Leo G. Kroon, and Paolo Toth. A railway timetable rescheduling approach for handling large-scale disruptions. *Transportation Science*, 50(3):841–862, 2016.

Michiel Vromans. *Reliability of Railway Systems*. PhD thesis, Erasmus Research Institute of Management (ERIM), Erasmus University Rotterdam, July 2005.

Michiel JCM Vromans, Rommert Dekker, and Leo G Kroon. Reliability and heterogeneity of railway services. *European Journal of Operational Research*, 172(2):647–665, 2006.

Joris C. Wagenaar, Leo G. Kroon, and Marie Schmidt. Maintenance appointments in railway rolling stock rescheduling. *Transportation Science*, 2017. doi: 10.1287/trsc.2016.0701. Available Online.