Norwegian School of Economics
Bergen, Fall 2019

Rescheduling of light rail trains during disruption

An optimization model for Bybanen in Bergen

Sturla Lie & Jonathan Sinnes

Supervisor: Mario Guajardo

Master thesis, Economics and Business Administration

Major: Business Analytics

NORWEGIAN SCHOOL OF ECONOMICS

Acknowledgements

This master thesis was written as a part of our master's degree at the Norwegian School of Economics (NHH). It has been an exciting and educational process, where we would like to thank those who has helped us along the way.

We would first of all like to express our appreciation to Johan Haveland, Chief of Passenger Transportation at Bybanen, for contributing to our understanding of the challenges dispatchers face when operating an urban rail system. The completion of this thesis would not have been possible without data access and invaluable discussions with the team at Bybanen. We also want to express our gratitude for introducing us to a fascinating topic, sparking our interest of how optimization models can be used in real-world applications.

We further express our sincerest gratitude to our supervisor Mario Guajardo for excellent guidance when developing the model and writing the thesis. Your engagement at meetings and mail correspondence inspired us to perform at our highest level until completion of the thesis. Thank you.

Finally, we would like to thank our family and partners for unprecedented support, and for showing interest in the project.

Norwegian School of Economics

Bergen, December 2019

Sturla Lie | Jonathan Sinnes

Abstract

When a disruption occurs in an urban rail system, it usually results in significant disturbances due to limited operational flexibility. In this thesis, we develop an optimization model that efficiently reschedules trains during partial blockage on a double-tracked light rail line. The rescheduled timetable is obtained by a mixed-integer linear programming model that minimizes the sum of delay at all stations by rescheduling trains through the opposite track using crossovers.

The numerical analyses are performed on three case studies based on real-world data from Bybanen light rail system in the city of Bergen. Our findings suggest that the proposed optimization model can safely reschedule train operations through crossovers located at their actual position in the network. Our findings also indicate that when minimizing delay at all stations instead of at the final stations, it contributes to more evenly distribution of passenger delay. This is demonstrated by comparing two different objective functions.

The results furthermore imply that by increasing frequencies, a crossover strategy will be harder to implement following larger density of trains. Changing from manual to automatic crossovers seems to have little effect on rescheduling of train operations. When expanding to double-tracked crossovers, however, the results indicate that punctuality and train operations are significantly improved. Finally, as the optimization model solves the most comprehensive case study in six seconds, the model can be applied by dispatchers in real-time decisions.

Contents

List of Figures

List of Tables

Introduction and scope of research

Bybanen has been crucial for public transportation in Bergen since the introduction in 2010 (Bybanen, 2019a), aiming to be the most punctual, reliable and cost-efficient light rail system in Europe (Bybanen, 2019b). These objectives should be met through competence development for dispatchers and operators, as well as ensuring reliable and accurate passenger information (Bybanen, 2019b). In order to handle deviation management, disruptions should moreover be solved by utilizing dispatcher support tools that can deliver reliable, high-quality solutions in real time.

Similar to comparable railway systems, dispatchers at Bybanen decide based on previous experiences and simple operating heuristics. However, following higher frequencies of train operations and more advanced infrastructure, there is a broad agreement in the literature that dispatchers need decision support systems (Gao, Yang, & Gao, 2017; Pellegrini, Marlière, & Rodriguez, 2016; Samà, D'Ariano, Pacciarelli, Pellegrini, & Rodriguez, 2018). With expected increases of passenger demand following line extensions of Bybanen (Miljøløftet, 2019a), instantaneous decision making is likely to become more complex in the future. With numerous rescheduling possibilities during disruption, it is therefore nearly impossible for dispatchers to decide optimally. Thus, only by providing dispatchers efficient support can high punctuality and reliability be guaranteed.

The main purpose of this thesis is therefore to develop an optimization model to assist dispatchers in real-time during disruptions. This is achieved through mathematical programming by creating a mixed-integer optimization model with objective of minimizing sum of delay in the network. The model interacts with dispatchers in three phases. Firstly, dispatchers provide an operating timetable as input together with safety considerations. Secondly, if an incident occurs, dispatchers specify location and timeframe of the incident. Lastly, dispatchers receive a new, optimal timetable with specific rescheduling instructions as to how trains should operate when passing an incident section.

The model can therefore be used by dispatchers to safely reschedule train operations in real-time during disruptions. It may also be used for stress testing the network by simulating different scenarios and identifying possible bottlenecks, thereby providing a better understanding of critical sections. Moreover, dispatchers can use the model to increase knowledge of the train operations in general, such as the location where trains ideally should switch tracks during disruption. In future planning of line extensions, the model can be used to determine ideal placements of stations and crossovers. In other words, reducing the number of difficult operating decisions, and increasing the likelihood of seamless train operations.

Structure of thesis

The remaining thesis is organized as follows. Chapter 1 presents background information about urban rail systems in general and the light rail system of Bybanen specifically. Chapter 2 provides a literature review of how the real-time railway traffic management problem has been solved in comparable systems, as well as our contributions to the literature. Chapter 3 continues with a description of the problems considered in this thesis, including explanations of the assumptions made when creating the optimization model. Chapter 4 describes our optimization model in detail, before chapter 5 presents input data based on the light rail system of Bybanen. Moreover, we specify the process of solving the model as well as presenting the case studies analyzed in this thesis. Numerical results are presented in Chapter 6. In Chapter 7, implications of the numerical results are discussed, in addition to how the optimization model can be used as an element in a dispatcher support system in the future. Finally, Chapter 8 summarizes our findings before conclusions are presented.

1 Background

This chapter include a short introduction to urban railway systems, and more specifically, how Bybanen as a transportation mode has become a central part of public transportation in Bergen. It also contains a description of the infrastructure and characteristics of Bybanen.

1.1 Urban railway

Urban rail consists of railway systems in urban and suburban areas, and including commuter railways, metros and light rail systems (UITP, 2018a, 2018c, 2019). However, since commuter railways share many of the same characteristics as mainline railways, urban rails will be referred to as metros and light rail systems in the remainder of this thesis.

Metros consist of transportation systems with exclusive right-of-way and capacity to efficiently transport large number of passengers (UITP, 2012, 2018c). They are completely separated from other traffic, often located underground, permitting higher operating speed than light rail systems (UITP, 2012). As completely separated tracks require large investments, metros are most commonly implemented in large cities where high capital costs can be justified (UITP, 2012). In contrast to metro systems, light rail systems often share infrastructure with other users and operate partly on line-of-sight (UITP, 2019). This is normally solved by providing light rail priority in junction signals at the expense of other traffic, thereby reducing external disruptions (UITP, 2016). With right-of-way implemented, light rails can operate nearly congestion free, at velocities of 20-30 km/h (UITP, 2016).

The concept of light rail encapsulates both trams, light rail transit and quasi-metro rapid transit, depending on the level of segregation and capacity of the system (UITP, 2016). The two last-mentioned transportation modes have relatively high capacity, only surpassed by metros and heavy rail (UITP, 2016). Light rails have five to eight times lower costs than metros (UITP, 2016). The system has consequently become a popular option for small to medium sized cities aiming to reduce congestion, improve air quality and reduce greenhouse emissions (UITP, 2019). In recent years, there has been developed a range of light rail systems similar to Bybanen in cities such as Aarhus, Casablanca, Algiers, and Florence (UITP, 2015, 2018b).

Light rail systems have traditionally had a central role for public transportation in larger European cities, and the networks in Budapest, Prague and Paris together had more than 1100 million passengers in 2018 (UITP, 2019). In recent years, there has also been a steady increase of light rail systems in Asia, which is expected to continue in the future following heavy investments in China (UITP, 2019).

1.2 Bybanen

1.2.1 Development of Bybanen as transportation mode

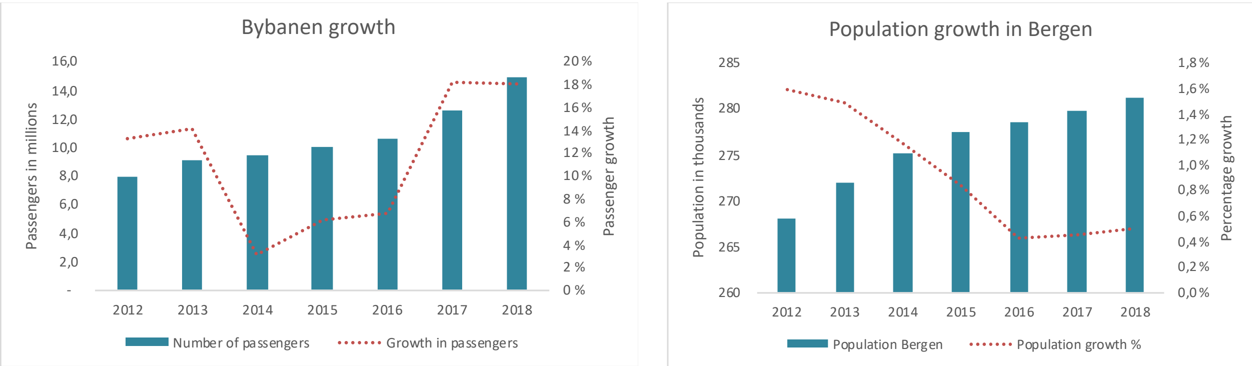
In 2012, the Norwegian Government presented a climate strategy with objective of meeting future transportation demands without increasing passenger travel by car (Norwegian Government, 2012). This should be accomplished by stimulating to increased use of public transportation, cycling and walking through state subsidies (Norwegian Government, 2012). The subsides are paid following an agreement between County Council and the State, hereafter named "*city growth agreement*", where County Council operate according to the National Transport Plan 2010-2019 (Norwegian Government, 2019). The Regional Climate Plan of Hordaland County Council for 2014-2030 states that in order to reach the national objectives in Bergen, increasing public transportation and reducing car traffic is crucial (Hordaland County Council, 2014). The change should consequently be stimulated by investing in infrastructure for public transportation and increasing car related costs.

In the most recent climate budget of Bergen County, there is correspondingly an objective of reducing passenger car traffic by at least 10 percent within 2020, compared to 2013 (Bergen County, 2019b). One of the main initiatives supporting this objective includes developing an integrated transportation network connecting Bybanen with other public transportation modes (Bergen County, 2019b). This is clearly stated through the city growth agreements from 2011-2014 and 2017-2023, where Bybanen is valued as the most important contributor to public transportation in Bergen (Norwegian Government, 2011, 2017).

The current network of Bybanen light rail is illustrated by solid red lines in Figure 1. The first section of Bybanen was built between Byparken and Nesttun, and opened in June 2010 (Bybanen, 2019a). It was further extended to Lagunen in June 2013, and in two stages to Bergen airport in August 2016 and April 2017 (Bybanen, 2019a). Since the start, there has been a consistent focus of supporting city development by creating network effects through integration of existing public transportation and Bybanen. There has for instance been created joint connecting points for Bybanen and buses, as well as bike paths along the tracks (Miljøløftet, 2019a).

Figure 1, Illustration of the current Bybanen network and future line extensions (Bergen County, 2019a)

Since the start of 2010, Bybanen has experienced a rapid growth of passengers. Figure 2a illustrates the passenger statistics since 2012. These figures show that Bybanen reached an all-time high of 14.9 million passengers in 2018, following 18% yearly growth two consecutive years. These numbers are even more impressive when including population growth in Bergen in the same period, illustrated by Figure 2b. The figures combined show a clear trend of increased use of public transportation in general, and Bybanen in specific. In fact, the number of passengers using public transportation in Bergen has doubled between 2010 and 2018 (SSB, 2019a), where Bybanen has been the major contributing factor (Engebretsen, Christiansen, & Strand, 2017).



(a) Passenger growth (b) Population growth

Figure 2, Passenger and population growth 2012 - 2018 (Skyss, 2015, 2016, 2017, 2019b; SSB, 2019b)

The development of Bybanen has further stimulated to population growth in areas with close proximity to the network, and more than a third of Bergen's population live within one km of the light rail (Engebretsen et al., 2017). This is consistent with comparable systems, where development of new infrastructure stimulates to increased activity and population growth (UITP, 2016). With the current development of a new track from the city center to Fyllingsdalen, the number of passengers within the proximity zone will further increase (Miljøløftet, 2019a). It has been decided to expand Bybanen with a new track from the city center to Åsane (Miljøløftet, 2019b), creating a network covering large parts of the Bergen area. Future line extensions of Bybanen are illustrated with stippled lines in Figure 1.

The increased focus on public transportation in Bergen has not only resulted in declining CO2-emissions from road traffic in recent years (Bergen County, 2019b), but also contributed to considerable improvements of air quality (Bergen County, 2019c). As a result of the upcoming line extensions of Bybanen, Bergen is expected to further decrease CO2-emissions, meeting the zero-growth objective of passenger transportation (Miljøløftet, 2019a).

1.2.2 Infrastructure and characteristics of Bybanen

The current route of Bybanen consists of 27 stations spread over nearly 20 km between BYP and FLE, illustrated in Figure 3. The figure illustrates abbreviated station names which will be used in the remainder of the thesis. Expanded station names are shown in Appendix A. The entire route is doubled tracked, meaning that each train can travel undisturbed at designated tracks in both directions. There are 16 crossover tracks throughout the system, divided in both manual and automatic crossovers.

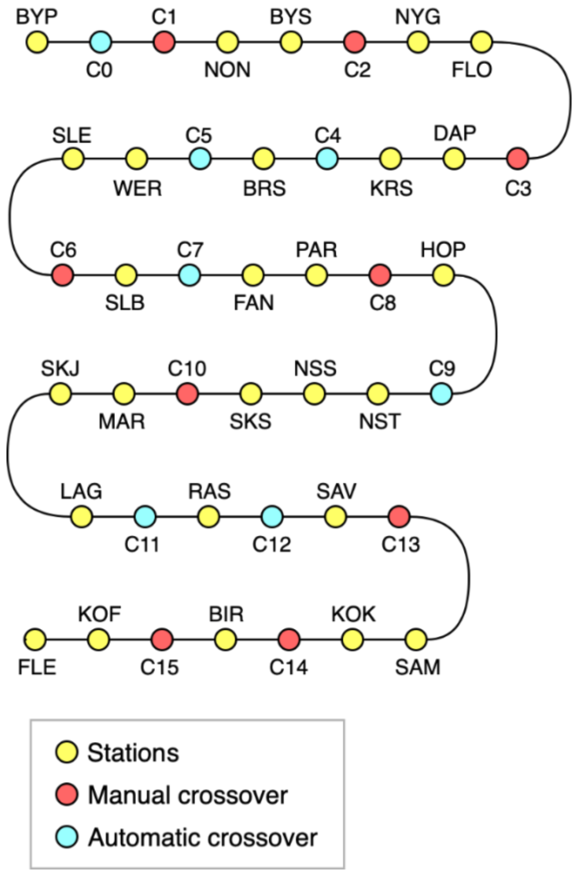


Figure 3, Schematic route Bybanen

Similar to other light rail systems (UITP, 2019), Bybanen is not a completely closed system, due to partly shared infrastructure with cars and pedestrians. The main operating rule is therefore to drive according to line-of-sight, except in tunnels where automatic operating systems are in place. Operators at Bybanen are therefore responsible for ensuring safe interactions with other traffic. In order to operate congestion free, Bybanen is prioritized at the expense of other traffic at intersections and traffic lights.

During normal operation, Bybanen operates according to desired frequency rather than a specific timetable. For ordinary weekdays, this frequency varies from five minutes in rush hour to 10 minutes in periods with lower demand (Skyss, 2019a). In order to meet the desired frequency during rush hour, 20 trains are currently in use. However, following increased passenger demand in recent years, utilization of 24 trains will shortly increase frequency to four minutes. In order to ensure that the required frequency is met, Bybanen measures performance according to several key performance indicators. The first indicator measures performance in terms of punctuality, by determining whether trains are operated according to the timetable (Bybanen, 2019b). Regularity is used to measure cancelled or abrupted train operations, while deviation management includes how deviations are handled and how passengers are informed during disruptions.

Bybanen generally performs well on all indicators, illustrated by a regularity of 99.45% and punctuality of 97.57% in 2018 (Bybanen, 2019a). That being said, Bybanen defines punctuality as deviation of less than three minutes from the timetable. A train can therefore be delayed in the view of passengers, but still be defined as punctual according to the operating plan. This is indicated by the customer satisfaction report from 2018 stating an experienced punctuality of 87% (Bybanen, 2019a). There is consequently a discrepancy between delay according to Bybanen and delay according to passengers.

A lower punctuality from the perspective of the passengers can be explained by the fact that delay is measured by the number of minutes a train is delayed, and not by passenger delay. If a train has to return to the origin station due to closed tracks, the current operating plan states that the original trip ends at the turnaround station, before a new trip starts when returning to origin. By contrast, passengers are likely to experience additional delay when they board a new train in order to reach their terminal station. In addition, as there are considerably more days with normal operation than disruptions, yearly punctuality doesn’t provide a clear picture of delay during disruptions. Thus, in order to measure performance of deviation management, a more thorough understanding of passenger delay during disruptions is needed.

2 Related work and contribution

This chapter provide an overview of relevant literature related to rescheduling of railway systems in general, and urban rail systems in specific. Contributions to research are further presented.

With increasing urbanization and passenger demand, well-functioning urban rail transit is important to efficiently transport large crowds (Chang et al., 2019). These systems are generally both safe and reliable. However, due to increasingly complex systems following the introduction of new technologies and equipment, managing normal operation is becoming more challenging (Chang et al., 2019). Consequently, when incidents occur, they often greatly influence normal train operations, including the safety of passengers (Chang et al., 2019). In general, there are three main contributing factors causing incidents; infrastructure failures, locomotive defaults and signal system failures (Chang et al., 2019; Xu, Li, & Yang, 2016).

When incidents occur, it is crucial for dispatchers to efficiently manage the situation to minimize delay and inconvenience for passengers (Chang et al., 2019). According to Xu et al. (2016), emergency responses to incidents depend on whether both tracks are affected. If the entire track is closed, passenger service is often suspended until the track is recovered. However, if only one track is closed, trains can be rescheduled in real-time using crossover tracks (Xu et al., 2016). Even when this is possible, dispatchers normally choose to wait for track recovery due to operational simplicity (Xu et al., 2016). This can be explained by lack of global considerations and that decisions are mostly based on experiences from previous situations and simple dispatching rules (Pellegrini et al., 2016; Yin, Tang, et al., 2017).

In other words, the main disadvantage of the experience-based rescheduling method is the lack of precision in complex situations (Gao et al., 2017). Only by incorporating global, systematic considerations can safety, service quality, and optimal operational costs be guaranteed (Gao et al., 2017). This is especially crucial for urban rail systems where passenger demand and departure frequencies are high (Yin, Tang, et al., 2017). Consequently, due to increasingly complex situations to manage and lack of efficient support systems, it becomes nearly impossible for dispatchers to correctly estimate decisions overall effect (Pellegrini et al., 2016; Samà et al., 2018; Xu et al., 2016).

With this in mind, providing dispatchers with useful tools has received considerable attention in recent years. In academic research it is common to differentiate between train scheduling and train rescheduling, where train scheduling is the process of creating a timetable in advance, and train rescheduling includes managing the timetable at an operational level (Sairong et al., 2019). As dispatchers deal with operational decisions, a model should be able to present information in real-time. The operational problems faced by dispatchers are therefore often defined as the real-time railway traffic management problem and involves managing disturbances that contribute to delays in the network (Samà, D′Ariano, Corman, & Pacciarelli, 2017).

Although there have been many attempts to solve this problem, few have been implemented in practice as most models solve very simplified problems that are only applicable for specific traffic situations rather than entire networks (D'ariano, Samà, D'ariano, & Pacciarelli, 2014). Ideally, an efficient system should provide dispatchers with a conflict-free schedule, minimize delay, safely deal with actual traffic conditions, and be solved in a matter of seconds (D'ariano et al., 2014; Samà et al., 2017). The balance of including both high level of details and solving the problem rapidly has been hard to manage (D'ariano et al., 2014), although advances in recent years have made it possible to deal with more effectively (Chang et al., 2019).

The majority of existing research on rescheduling problems are based on mainline railway systems (Chang et al., 2019; Gao et al., 2017; Yin, Wang, Tang, Xun, & Su, 2017), and utilized methods can broadly be classified as macro – and micro methods. The difference between the two perspectives is the level of granularity (Pellegrini, Marlière, & Rodriguez, 2014), macro methods describe the infrastructure based on groups of block-sections while micro methods define the infrastructure based on single block-sections (Samà et al., 2017). Macro methods consider an ideal or constant speed and optimize train operations only at stations, and micro methods also consider speed profiles and train movements throughout the network, thereby capturing a higher degree of details (Hangfei, Keping, & Paul, 2018). However, with more details it becomes harder to solve the problem within reasonable computation time (Samà, Pellegrini, D'ariano, Rodriguez, & Pacciarelli, 2016).

Two of the most popular methods for solving rescheduling problems in the literature are alternative graph models and mixed-integer linear programming (MILP) (Samà et al., 2016). Alternative graph models developed in a number of papers (D'Ariano, Corman, Pacciarelli, & Pranzo, 2008; D'Ariano & Pranzo, 2009; D'Ariano, Pranzo, & Hansen, 2007) have contributed to introducing the ROMA-system (railway traffic optimization by means of alternative graph). The system makes it possible to consider global information in a reasonable computation time, by utilizing blocking time theory for track occupation and alternative graphs for solving traffic control problems (D'Ariano, 2009). The system has later been verified by Corman and Quaglietta (2015) and is considered one of the most promising systems for mainline railways (Pellegrini et al., 2016).

Another line of studies have focused on how MILP-models can be utilized to improve rescheduling decisions. Törnquist and Persson (2007) proposed a MILP-model for rescheduling trains during small disruptions by considering possible track options. Moreover, Louwerse and Huisman (2014) considered both partial and complete line blockages during major disruptions. Zhan, Kroon, Zhao, and Peng (2016) rescheduled trains on a double-track high-speed railway when one track was unavailable, assuming that time of recovery is unknown beforehand and gradually updated.

Another promising MILP-approach for mainline railways is the RECIFE-MILP model developed by Pellegrini, Marlière, Pesenti, and Rodriguez (2015). The approach uses a heuristic algorithm based on the MILP formulation proposed by Pellegrini et al. (2014), handling the real-time railway traffic management problem in short computational time (Pellegrini et al., 2015). Samà et al. (2016) further improved the model by using ant colony optimization meta-heuristics to reduce the number of possible routes evaluated. This was done as the number of alternative routes affect problem size, and therefore heavily influence computation time (Samà et al., 2016). Recently, Pellegrini, Pesenti, and Rodriguez (2019) reformulated the MILP-model from Pellegrini et al. (2015) by exploiting inequalities to reduce the number of binary variables. This was done as previous models sometimes failed to deliver within the computation time required in real-world instances (Pellegrini et al., 2019).

The research on rescheduling problems for urban railways is limited (Chang et al., 2019; Gao et al., 2017), but due to increasing passenger demand in cities and better opportunities for automatic systems, the topic has gained more interest in recent years (Yin, Tang, et al., 2017). In comparison to mainline railway systems, urban railways have higher departure frequencies and shorter distances between stations (Gao et al., 2017). The station layout and infrastructure are also much simpler, and in normal operations trains are not allowed to meet or overtake each other (Chang et al., 2019; Gao et al., 2017). Consequently, disruptions often have larger consequences for the entire network due to greater interactions between trains and limited operational flexibility (Chang et al., 2019; Gao et al., 2017). The objectives and model formulation in urban railway systems are therefore somewhat different from mainline systems (Chang et al., 2019; Gao et al., 2017). The minimum headway between two successive trains in urban railway systems are normally between two and five minutes. However, with increasing passenger demand in recent years, headways of only two minutes are not uncommon in busy urban rail systems (Chang et al., 2019). This increases the need for real-time rescheduling of urban rail systems, to ensure passenger satisfaction and reduce operational costs (Gao et al., 2017).

Rescheduling in urban railways is often performed by adjusting headway, however, with increasing frequencies it becomes more complicated for dispatchers to respect headways during rescheduling decisions (Yin, Wang, et al., 2017). Different rescheduling strategies have therefore been proposed including deadheading, holding, stop-skipping and short turning. Deadheading was proposed as one of the earliest contributions in literature for urban railway systems where trains pass some of the stations empty at the start of the trip to reduce headways at later stations (Eberlein, Wilson, Barnhart, & Bernstein, 1998). Gao et al. (2017) recently presented a version of a holding strategy where a MILP-model was used to incorporate a real-time rescheduling strategy for an urban railway system, by utilizing information of fault handling to adjust run – and dwell time.

A stop-skipping strategy allows late trains to skip low-demand stations in order to return to normal schedule (Gao, Kroon, Schmidt, & Yang, 2016). Considering an overcrowded metro system in Beijing, Gao et al. (2016) demonstrated how a stop-skipping pattern increased circulation of trains and reduced the number of waiting passengers. According to Yin, Wang, et al. (2017), however, stop-skipping strategies are rarely allowed in practice as the wait time will increase for some passengers. To recover disruptions in a metro system, Yin, Wang, et al. (2017) proposed an alternative approach, where back-up trains located at depots or sidings increased transport capacity during disruption. Thereby contributing to faster return to normal operation. As most metro infrastructures includes several storage sidings, the method can easily be implemented in real-word applications (Yin, Wang, et al., 2017).

Chang et al. (2019) recently presented a short-turning strategy during complete blockage for a double-tracked urban railway, where trains can turn at intermediate stations in order to use tracks in the opposite direction. Similar to Yin, Wang, et al. (2017), this strategy also included back-up trains at depots (Chang et al., 2019). The case study on a subway line in Beijing demonstrated that a rescheduled timetable could be obtained within short computation time by adjusting runtime, dwell time, and rolling stock circulation by the use of crossover tracks (Chang et al., 2019). Despite this, since the strategy brings inconveniences for passengers who have to leave the train when changing direction, more research is needed to evaluate the effectiveness of the approach (Chang et al., 2019). Xu et al. (2016) considered an incident at one of the tracks on a double-track subway line, with objective of minimizing total delay. In the presented model, impacted trains during disruption are rescheduled using the opposite track through crossovers. The case study on a subway line between Beijing and Yizhuang indicated that the model is able to reschedule a large number of trains within short computational time (Xu et al., 2016).

The presented research about urban railway systems have in common that a macro perspective is used to model train operations. This can be explained by the high-frequency nature of urban rail systems, where dispatchers need to decide rapidly (Yin, Tang, et al., 2017). Due to differences in frequencies for urban – and mainline railways, what is considered a short computation time will also differ. As an example, one of the most successful MILP-models for mainline railways, RECIFE-MILP, does not manage to optimally solve complicated cases within three minutes (Pellegrini et al., 2019). This would clearly be problematic for more frequent urban rail systems (Chang et al., 2019). In general, a rescheduling plan should therefore be determined within one minute (Yin, Tang, et al., 2017).

In this paper, we focus on a double-tracked light rail system where one track is unavailable due to an incident, e.g. due to power loss. Moreover, during disruption, affected trains can utilize the opposite track through crossovers. Possible dispatching measures therefore consist of both adjusting run – and dwell times and utilizing crossovers to optimize rolling stock circulation. Similar to previous research of urban railways, this thesis considers a rescheduling problem at macro perspective, to balance the trade-off between accuracy and computation time. The purpose is to generate an optimal rescheduling plan where sum of delay is minimized at each station. This is done to minimize delay for all passengers, not only passengers travelling to the terminal station. This differs from previous research that minimize delay at the final stations (Pellegrini et al., 2015; Xu et al., 2016), although necessary for Bybanen where the majority of passengers do not travel the entire route (Bybanen, 2019a).

By contrast to Xu et al. (2016) who modelled crossovers at stations, we present a model where crossovers are located at their actual positions in the network. This contributes to safer and more accurate train positions when changing tracks. Summarized, we aim to make the following contributions to the study of rescheduling problems for urban railways:

- A MILP-model that can assist dispatchers in real-time during disruption by accurately rescheduling trains using crossovers. This is done by minimizing sum of delay at all stations, resulting in a conflict free, rescheduled timetable
- Most existing research for urban railways uses metro or subway systems as illustrative cases. As this thesis considers a light rail system with somewhat different characteristics, we demonstrate the applicability of rescheduling models for a new transportation mode
- Addressing real-world problems faced by dispatchers and operators at Bybanen light rail system. By providing knowledge of how train operations ideally should be managed, and an urban rail infrastructure developed, the optimization model can contribute to more effective train services

3 Problem description

This chapter provides a description of the problems faced by dispatchers when one of the tracks are closed due to an incident, including limitations of the current rescheduling method. Moreover, the alternative approach considered in this thesis is presented.

In the current system of Bybanen, dispatchers decide according to previous experiences and an operating manual based on simple dispatching rules, similar to comparable systems (Xu et al., 2016; Yin, Tang, et al., 2017). This is problematic as experiences and dispatching rules alone cannot account for every possible scenario, making it harder to guarantee optimality (Gao et al., 2017). On average, dispatchers at Bybanen use five minutes from an incident has occurred to a decision has been made. During this time frame, trains operate to their subsequent station and wait until further instructions. Consequently, a decision support model that can instantaneously reschedule trains is therefore likely to considerably reduce delay.

The main characteristics of Bybanen are illustrated in Figure 4, consisting of stations and crossovers. All stations include two platforms with dwell capacity of one train in each direction. Two trains travelling in the same direction cannot dwell at the same station simultaneously. Moreover, trains are not allowed to dwell at crossovers following safety considerations.

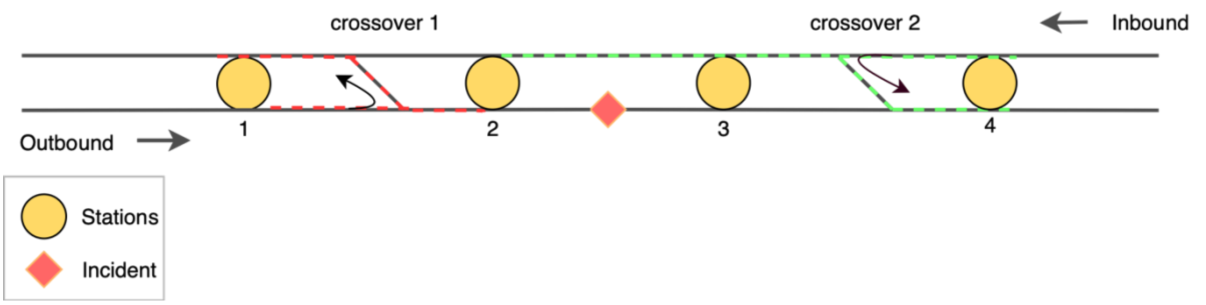


Figure 4, Emergency situation current system

Figure 4 illustrates a situation where trains change tracks through crossovers when outbound track is closed due to an incident. Based on the current operating rules of Bybanen, dispatchers normally choose between two options. Firstly, dispatchers estimate the expected incident duration time. If the problem can be solved within the decision-making time of five minutes, outbound trains wait, and inbound trains drive according to schedule. Secondly, if recovery time is unknown, dispatchers normally use a short-turning strategy similar to Chang et al. (2019), as illustrated in Figure 4. In the presented scenario, inbound trains will drive to station 2 and drop off passengers. They will then change operating direction and return to origin station after changing to outbound track at crossover 2. Outbound trains will similarly drive to station

2 and drop of passengers before returning to origin station after changing to inbound track at crossover 1.

In the operating handbook of Bybanen, this is only one of many possible solutions. Often, only some inbound trains will drive to station 2, while the remainder will drop off passengers at station 4 before changing tracks. This is done due to headway considerations and to reduce congestion. Consequently, some of the inbound passengers have to wait until the subsequent inbound train can transport them to station 2, before boarding a third train to station 1. This is problematic due to inconvenience for passengers when they have to change trains (Chang et al., 2019), and also contributes to considerable delays.

Outbound passengers will similarly experience additional delay when waiting for inbound trains to arrive at station 2. Consequently, delay in the view of passengers will be greater than measured per train, thus, overestimating the punctuality of train operations. As the dispatcher strategy causes both increased delay and number of transfers, passenger satisfaction is likely to be lowered.

With this in mind, the proposed method in this thesis attempts to reduce the number of transfers as well as overall delay experienced by passengers. The alternative approach is presented in Figure 5, illustrating a situation where the outbound track is closed due to an incident. During disruption, outbound trains are allowed to pass the incident by changing tracks at crossover 1, before operating in the opposite direction until crossover 2 where it returns to their designated track.

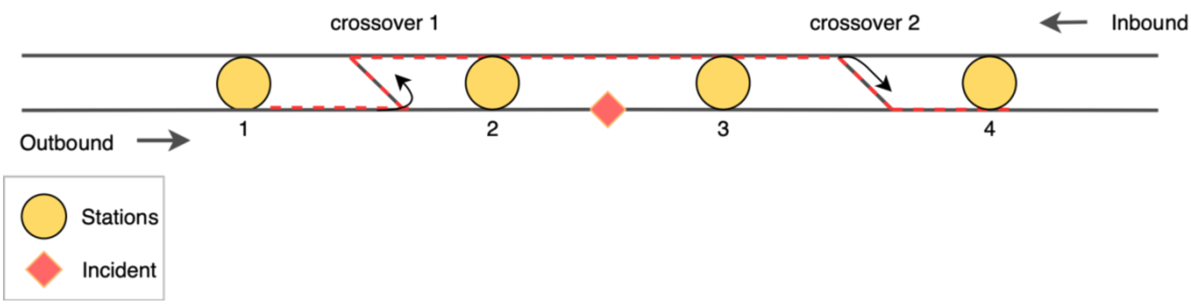


Figure 5, Proposed method during an incident

Obviously, in order to avoid head-to-head collisions, safety headways are crucial for such a strategy to be implemented. As a similar method has demonstrated effectiveness for subways (Xu et al., 2016), there is reason to believe that the proposed solution can be applied to Bybanen due to lower density of trains.

The alternative approach has two main advantages compared to the current dispatcher method. Firstly, as trains travel in their designated direction, the need for passengers to change trains is nonexistent. It will also be easier to include a passenger perspective, as passenger delay and train delay are identical. Secondly, with an optimization model supporting dispatchers, the average decision time of five minutes can be reduced. In addition, with global considerations in mind, more complex situations, like operating trains in the opposite direction, can be handled efficiently (Gao et al., 2017).

When dealing with delays, implementing a model where only directly affected trains change tracks is likely to be more efficient than a short-turning strategy where trains in both directions have to change tracks. In the remainder of this thesis, directly affected trains are defined as trains operating on a track where an incident has occurred. For the incident in Figure 5, inbound trains are for instance not directly affected by the incident on outbound track. Compared with the scenario in Figure 4 where all trains change tracks, inbound trains only have to adjust runtime to safely interact with outbound trains. If the interaction is feasible, inbound trains are expected to be considerably less delayed. Moreover, when disruption has recovered, returning to normal operation is likely to be faster when fewer trains have changed tracks.

To summarize, the proposed model in this thesis attempts to increase decision quality through an optimization model which takes global considerations into account and produces reliable and feasible solutions within short computational time. When formulating the model of characteristics of Bybanen when one track is closed due to an incident, we make the following assumptions:

- During disruption, all trains can be rescheduled before reaching a crossover section. Thus, trains cannot be trapped between two crossovers
- Time of recovery and location of incident section is known
- Communication systems between operators and dispatchers are of sufficient quality for trains to safely operate on the opposite track
- There are well-functioning systems to inform passengers of which platform trains will arrive after changing tracks
- In real-world operations, trains operate continuously in a loop by turning when reaching the end of track. For simplicity, in the proposed model a train terminates when reaching the final stations
- Passenger demand is equal at all stations. Thus, all stations should be prioritized equally

4 Model formulation

In this chapter we first provide a short introduction to MILP-models, before the optimization model used to reschedule urban rail systems are presented in full. We start by presenting sets, parameters and decision variables used in the model, before the objective function is described. Finally, constraints necessary to ensure safe train operations and to uphold the desired frequency are defined.

4.1 Mixed-integer linear programming model

An integer programming model is formulated as a problem where one or several of the decision variables have integer values (Lundgren, Rönnqvist, & Värbrand, 2010, p. 323). This includes both pure – and mixed integer programming models, where pure integer models only consist of integer variables, by contrast to mixed-integer models where both integer and continuous variables are defined (Lundgren et al., 2010, p. 325).

There are generally two reasons for introducing integer variables. Firstly, variables should be defined as integer when they are naturally integer values (Lundgren et al., 2010, p. 325), e.g. number of persons. The second reason is when logical or binary 0/1 variables are necessary (Lundgren et al., 2010, pp. 325-326), e.g. if a train utilizes crossovers or not.

To efficiently model train operations, we define a mixed-integer programming model with both continuous and binary variables, presented in full in the subsequent sections.

4.2 Mathematical formulation

4.2.1 Sets

The sets included in the optimization model are presented in Table 1. When running the model, we use sets for both inbound and outbound trains, links and routes. However, for ease of presentation in the model formulation, T, $LINKS$, and $LINKSC$ are defined as sets containing both inbound and outbound characteristics.

$LINKS$ displays the order of stations and crossover tracks in the network. From Figure 5, a link is defined as the section from station 1 to crossover 1. The links are important in order to define the correct order of events in the network, and to ensure that trains do not skip stations or crossovers. Furthermore, $LINKSC$ consists of links between crossover tracks, e.g. between crossover 1 and 2 in Figure 5, and are crucial to determine where trains switch tracks. It is important when specifying the additional time of utilizing crossovers.

Table 1, Sets

O	Outbound trains
I	Inbound trains
T	All trains ($O \cup I$)
EP	Set of end points in a link
S	Passenger stations
C	Crossover tracks
$Sall$	Set of start points in a link
$LINKSI$	Inbound links between stations, $(s, e) \in LINKSI: s \in Sall, e \in EP$
$LINKSO$	Outbound links between stations, $(s, e) \in LINKSO: s \in Sall, e \in EP$
$LINKS$	Links between stations ($LINKSI \cup LINKSO$)
$LINKSFI$	Inbound crossover links, $(a, b) \in LINKSFI: a \in Sall, b \in EP$
$LINKSFO$	Outbound crossover links, $(a, b) \in LINKSFO: a \in Sall, b \in EP$
$LINKSC$	Crossover links ($LINKSFI \cup LINKSFO$)
SI	Start of incident section
EI	End of incident section

4.2.2 Subscripts of sets

Table 2, Subscripts

t	Index of trains, $t \in EP$
e	Index of end points, $e \in EP$
s	Index of start points, $s \in Sall$
j	Index of stations, $j \in EP$
c	Index of crossovers, c $\in C$
s, e	Index of start and endpoints between stations and crossovers, $(s, e) \in LINKS$
a, b	Index of start and endpoints in crossover links, $(a, b) \in LINKSC$

4.2.3 Parameters

Parameters included in the model are presented in Table 3. Scheduled arrival and departure are given by the timetable for trains at all stations and crossovers. By contrast to mainline railways where runtime over crossovers often is neglected (Sairong et al., 2019), this is crucial to define for urban railways where margins are tighter due to shorter overall operating time. $f_{a,a,b}$ is therefore used to define utilization time of crossovers within a crossover link.

Table 3, Parameters

$a_{t,s}$	Scheduled arrival, $t \in T, s \in Sall$
$d_{t,s}$	Scheduled departure, $t \in T, s \in Sall$
$f_{a,a,b}$	Utilization time of crossover, $(a, b) \in LINKSC$

4.2.4 Constants

Table 4 presents constants defined in the model. We consider dwell time as constant in line with existing literature (Samà et al., 2018), although in practice there are normally some differences depending on passenger demand. Furthermore, headway for adjacent trains is defined to ensure that safety considerations are respected. As we are evaluating one incident at a time, incident occurrence and recovery are further defined as constants. Finally, we define a large M to be used in the binary relations in the model. In order to improve solution performance and speed up computation time it is important that this value is not too large (Pellegrini et al., 2015).

Table 4, Constants

w	Minimum dwell time
h_0	Minimum headway between two trains travelling in the same direction
to	Time when incident occurs
tr	Time when incident recovers
N	Number of trains in each direction
M	Sufficient large number

4.2.5 Decision variables

Decision variables used in the model are presented in Table 5 and include both continuous and binary variables. The first two variables are continuous and used to define actual arrival and departure of all trains. These are crucial in order to determine the tardiness of trains, as well as ensuring safe train operations. The binary variables are moreover used to define which trains utilize crossovers to pass an incident.

The binary variables together determine possible train operations for directly affected trains. Firstly, l_t states that only trains arriving at start of incident section before or at the time of recovery will be affected. Thus, trains arriving after incident recovery can drive according to normal operations. Secondly, k_t helps determine which trains arrive at end of incident section before incident occurrence, where only trains that are yet to arrive are affected. The two conditions are captured by the auxiliary variable lok_t which is only true when both l_t and k_t are true.

Table 5, Decision variables

$a1_{t,s}$	Actual arrival, $t \in T, s \in Sall$
$d1_{t,s}$	Actual departure, $t \in T, s \in Sall$
l_t	1 if directly affected trains arrive at start of incident section before or equal to incident recovery, 0 otherwise
k_t	1 if an incident has occurred and directly affected train has not reached end of incident section, 0 otherwise
lok_t	1 if directly affected trains arrive in incident section during incident, 0 otherwise

4.2.6 Objective function

The objective function minimizes sum of delay for all trains at all stations in the network. This is done to ensure that delay in the perspective of passengers is minimized, such that train delay equivalents passenger delay.

$$\sum_{t \in \mathrm{O}, j \in S} \left(a1_{t,j} - a_{t,j} \right) + \sum_{t \in \mathrm{I}, j \in S} \left(a1_{t,j} - a_{t,j} \right) \tag{1}$$

4.2.7 Constraints for normal operation

4.2.7.1 Runtime constraints

If a train is delayed, it cannot be recovered later in the system as the timetable reflects the ideal runtime between stations. This is consistent with existing research where initial delay cannot be recovered, and rescheduling decisions are limited to handling consecutive delay related to solving conflicts in the network (Samà et al., 2016; Shakibayifar, Sheikholeslami, Corman, &

Hassannayebi, 2017). Delay is accumulated throughout the system from the section where trains are initially delayed. Constraint (2) therefore states that runtime in the network is variable although restricted by a maximum runtime between stations and crossovers.

$$\mathrm{a}1_{t,e} - d1_{t,s} \geq \mathrm{a}_{t,e} - d_{t,s}, \quad t \in T, (s,e) \in LINKS \tag{2}$$

4.2.7.2 Blockage area constraints

To avoid dangerous situations, Bybanen has restrictions stating that only one train in each direction can be located in the same section simultaneously. Constraints (3) and (4) therefore state that within $LINKS$, there is a maximum capacity of one train.

$$\mathrm{a}1_{t+1,s} \geq \mathrm{a}1_{t,s} + (\mathrm{a}1_{t,e} - a1_{t,s}), \quad t \in T, (s,e) \in LINKS\colon t < N \tag{3}$$

$$\mathrm{a}1_{t,s} \geq \mathrm{a}1_{t-1,s} + (\mathrm{a}1_{t,e} - a1_{t,s}), \quad t \in T, (s,e) \in LINKS\colon t = N \tag{4}$$

4.2.7.3 Dwell time and order of events constraints

In the predefined timetable, a minimum dwell time is included to guarantee safe boarding of passengers as well as basic operational procedures like opening and closing of doors. Constraint (5) therefore states a minimum dwell time of w at every passenger station.

$$d1_{t,j} - a1_{t,j} \geq w, \quad t \in T\ j \in S \tag{5}$$

Moreover, as trains are not allowed to dwell at crossover tracks, constraint (6) is specified to ensure that departure always occur after arrival, thereby determining the correct order of events.

$$d1_{t,c} \geq a1_{t,c}, \quad t \in T, c \in C \tag{6}$$

4.2.7.4 Headway constraints for adjacent trains

Safety headways for adjacent trains are crucial both for ensuring a fixed frequency of train operations and for avoiding rear-end collisions. Constraints (7) - (14) therefore specify headways between trains travelling in the same direction. The first constraint states that the arrival of train $t+1$ at s should be larger or equal to arrival of train t at s, and the headway of h_0. Similarly, constraint (8) illustrates the relationship for departing trains. Constraints (9) and (10) are specified in order to avoid that more than one train, travelling in the same direction dwell at the same station or crossover. Since we are modelling a train service that terminates at the final station, the final train will not have a succeeding train. Constraints (11) - (14) are consequently specified.

$$a1_{t+1,s} \geq a1_{t,s} + h_0, \quad t \in T, s \in Sall: t < N \tag{7}$$

$$d1_{t+1,s} \geq d1_{t,s} + h_0, \quad t \in T, s \in Sall: t < N \tag{8}$$

$$a1_{t+1,j} \geq d1_{t,j} + h_0 - w, \quad t \in T, j \in S: t < N \tag{9}$$

$$a1_{t+1,c} \geq d1_{t,c} + h_0, \quad t \in T, c \in C: t < N \tag{10}$$

$$a1_{t,s} \geq a1_{t-1,s} + h_0, \quad t \in T, s \in Sall: t = N \tag{11}$$

$$d1_{t,s} \geq d1_{t-1,s} + h_0, \quad t \in T, s \in Sall: t = N \tag{12}$$

$$a1_{t,j} \geq d1_{t-1,j} + h_0 - w, \quad t \in T, j \in S: t = N \tag{13}$$

$$a1_{t,c} \geq d1_{t-1,c} + h_0, \quad t \in T, c \in C: t = N \tag{14}$$

4.2.8 Constraints for incident situations

4.2.8.1 Determination of which trains are affected during an incident

If one track is closed due to an incident, then during disruption, partial service using single track is accepted. Figure 6 provides an illustrative example of how trains are affected during disruption. Given that an incident occurs between two stations, the first station is defined as start of incident section and the second as the end of incident section, where only trains arriving to the section during disruption are affected. Moreover, as we are modelling the network within a macro perspective, the location of an incident is not determined with exact precision. In the model, an incident is therefore defined to occur somewhere between two members of $LINKS$, determined as incident section.

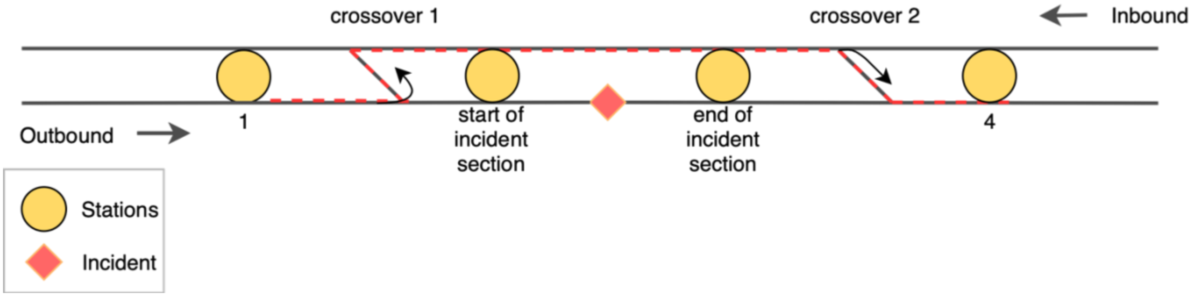


Figure 6, Incident section

With Figure 6 in mind, directly affected trains are trains travelling in the outbound direction, determined by constraints (15) - (21). Constraints (15) and (16) state that trains arriving to start of incident section before or at the time of incident recovery are affected. Moreover, constraints (17) and (18) state that trains which are yet to arrive at end of incident section when an incident occurs are affected. Finally, constraint (19) - (21) determine that a directly affected train must take action if it arrives at start of incident section before incident recovery and if it has not reached the end of incident section when an incident occurs. If either of these conditions are unfulfilled, trains will drive according to normal operation.

$$a1_{t,s} \leq tr + M(1 - l_t), \quad t \in T, s \in SI \tag{15}$$

$$a1_{t,s} \geq tr - M(l_t), \quad t \in T, s \in SI \tag{16}$$

$$to \leq a1_{t,s} + M(1 - k_t), \quad t \in T, s \in EI \tag{17}$$

$$to \geq a1_{t,s} - M(k_t), \quad t \in T, s \in EI \tag{18}$$

$$l_t + k_t \leq 1 + lok_t, \quad t \in T \tag{19}$$

$$lok_t \leq l_t, \quad t \in T \tag{20}$$

$$lok_t \leq k_t, \quad t \in T \tag{21}$$

4.2.8.2 Headway constraints opposite direction

It is relatively easy to ensure headways during normal operation where trains are operating according to a pre-specified timetable (Xu et al., 2016). However, when utilizing the opposite track, an additional headway needs to be specified in order to avoid front-to-front collisions. In Figure 7, the blue train is travelling in outbound direction on inbound track due to an incident. To avoid collision, the yellow train therefore has to wait for the blue train to finish the red-stippled section. The minimum safety headway necessary to avoid front-to-front collisions is consequently the green-stippled section. Similarly, there should be a safety headway when trains enter the opposite track, illustrated in Figure 8. Thus, the blue train cannot enter the red-stippled section through crossover 1 before the yellow train has passed crossover 1 in the inbound direction.

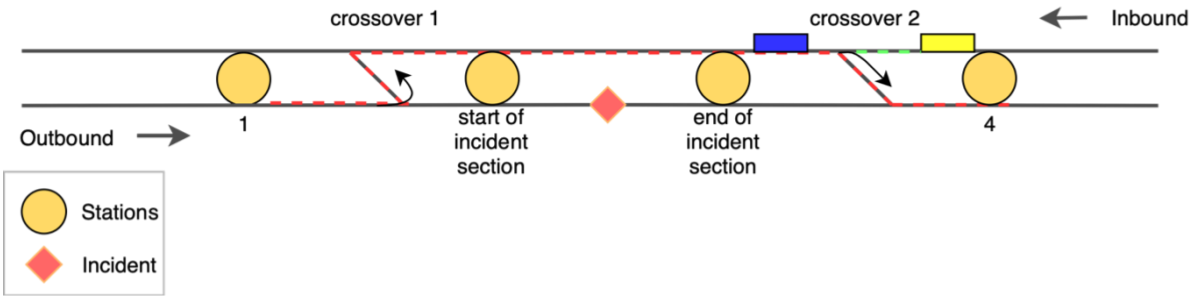


Figure 7, Headway when leaving opposite track

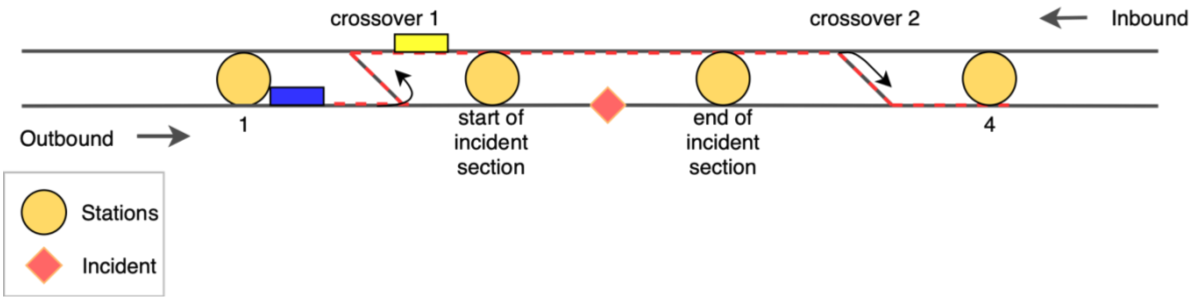


Figure 8, Headway when entering opposite track

Constraints (22) and (23) state the situation in figure 7 and figure 8 respectively, and are both written in the perspective of when the outbound track is closed. Thus, when disruption occur on inbound track, $t \in 0, u \in I$ is written as $t \in I, u \in O$, where u is defined as indirectly affected trains.

Constraint (22) consists of four parts. First, $a1_{u,b} - a1_{t,b}$ describes the situation where inbound and outbound trains arrive at crossover 2 on inbound track. $f_{b,a,b}$ is further included to define when the outbound train has reached crossover 2 on outbound track. To ensure safe train operations, the resulting headway should be equal or greater than actual runtime between crossover 1 and 2, divided by two. This condition is included following two reasons. Firstly, it guarantees that inbound trains will not be present at any station in the crossover section when outbound trains use inbound track. Secondly, it ensures enough time to safely decelerate and accelerate before and after reaching a crossover. It moreover includes the time it takes to drive across a crossover. Finally, the constraint is activated when directly affected trains utilize crossovers, encapsulated by lok_t.

Constraint (22) and (23) together account for the entire section between crossover 1 and crossover 2, and actual runtime is consequently divided by two for both constraints. If for instance the crossover section in Figure 7 takes 200 seconds to pass, a safety margin of 100 seconds is included from outbound trains has returned to outbound track. The safety margin is similarly included when inbound trains have reached crossover 1. Thus, making it possible to ensure safe train operations. Constraint (23) finally guarantees a headway when inbound trains depart from, and outbound trains arrive at crossover 1.

$$a1_{u,b} - a1_{t,b} - f_{b,a,b} + M(1 - lok_t) \geq \left(\frac{a1_{t,b} - d1_{t,a}}{2}\right), \quad t \in 0, u \in I, (a,b) \in LINKSC \quad (22)$$

$$a1_{u,a} - d1_{t,a} + M(1 - lok_t) \geq \left(\frac{a1_{t,b} - d1_{t,a}}{2}\right), \quad t \in 0, u \in I, (a,b) \in LINKSC \quad (23)$$

4.2.8.3 Added time when using crossover track constraints

Constraints (24) and (25) state that directly affected trains using crossovers will increase runtime according to switching procedures. The first constraint states how the additional time of $f_{a,a,b}$ is added when the first crossover within a link is utilized and is respected only when lok_t is 1. Similarly, the second constraint describes the relationship between arrival and departure at crossovers where trains switch back to their original track.

$$d1_{t,a} \geq a1_{t,a} + f_{a,a,b} * lok_t, \quad t \in T, (a,b) \in LINKSC \quad (24)$$

$$d1_{t,b} \geq a1_{t,b} + f_{b,a,b} * lok_t, \quad t \in T, (a,b) \in LINKSC \quad (25)$$

5 Computational implementation

In this chapter, we describe how input data is created, including timetables and other data necessary to ensure safe and reliable train operations. The architectural setup when solving the optimization model is thereafter presented, before case studies considered in this thesis are described. Finally, possible expansions currently under consideration at Bybanen are presented.

5.1 Data description

Data used in the optimization model is developed in cooperation with Bybanen and consist of both variable and fixed characteristics. Variable data is in this thesis defined as incident specific data related to incident duration and location. Fixed data on the other hand includes information about train services during normal operation, and utilization time of crossovers.

For normal operations, runtimes between stations are calculated according to speed profiles from the operating handbook of Bybanen. These include start – and maximum velocities between different sections in the network, as well as exact locations of stations. Thus, runtimes can be calculated based on the second kinematic equation (Johnson, 2001, p. 135).

$$\Delta x = \left(\frac{v + v_0}{2} \right) t \tag{26}$$

where Δx is the change in distance between two positions in the network, v the maximum velocity, v_0 the starting velocity, and t time.

When solving for t, it is therefore relatively straightforward to calculate runtimes. The accuracy of the calculations has been confirmed through comparisons with the operating handbook, where runtimes for the second – and third line extensions are included. Operating experts at Bybanen has confirmed the accuracy of calculated runtimes. This has been especially crucial for the section from Byparken to Nesttun, as runtimes were not calculated in the first building phase.

Most of the crossovers, however, are not included in the speed profiles. Thus, runtimes cannot be calculated according to the kinematic equation. That being said, the operating manual includes information about locations of crossovers for both inbound and outbound direction. It is therefore possible to calculate runtimes from stations to crossovers through the equation of:

$$R_{tsc} = \left(\frac{D_{sc}}{D_s}\right) R_{ts} \tag{27}$$

where R_{tsc} is the runtime and D_{sc} the distance from a station to a crossover, while D_s is the distance and R_{ts} the runtime between the two stations adjoining a crossover. Thus, runtime is calculated based on the percentage distance between stations and crossovers.

To provide an illustration, distance and runtime between two stations in outbound direction is 765 meters and 82 seconds respectively. Distance from the first station to the first crossover is 706 meters, resulting in a runtime of 76 seconds. In order to calculate runtime to the second station, R_{tsc} is thereafter subtracted from R_{ts}, resulting in a runtime of 6 seconds. By utilizing this method, it is possible to ensure correct placements of crossovers in terms of runtime during normal operation.

With considerations to dwell time and frequencies, calculated runtimes have been used to create a timetable. According to the current operation of Bybanen during rush hour, dwell time is defined as 20 seconds per station, while frequency between trains is 300 seconds. The resulting timetable is presented in Table 6, which illustrates arrival and departure for the first two trains operating in outbound and inbound direction respectively. It moreover displays arrival and departure for both stations and crossovers, where crossovers range from C0 to C15. Since Bybanen normally operates 20 trains to ensure a frequency of 300 seconds, a total of 10 trains in each direction are included in the optimization model.

Furthermore, outbound trains arrive at their terminal station at Flesland (FLE) after 2429 seconds while inbound trains arrive at their terminal station at Byparken (BYP) after 2486 seconds. Inbound trains therefore have a slightly higher runtime due to characteristics of the track and placement of stations. In real-world applications this difference is adjusted with respect to headway by regulating dwell time at terminal stations. However, due to the assumption of trains terminating at the final stations, these factors are not taken into consideration in this thesis. Table 6 shows that most of the crossovers are located in close proximity to stations, whereas crossover 0,11 and 12 are located at stations.

Table 6, Timetable first two trains

Outbound station	Arrival outbound	Departure outbound	Inbound station	Arrival inbound	Departure inbound
BYP	0	20	FLE	0	20
C0	20	20	KOF	98	118
C1	52	52	C15	129	129
NON	68	88	BIR	202	222
BYS	154	174	C14	252	252
C2	239	239	KOK	273	293
NYG	259	279	SAM	347	367
FLO	345	365	C13	413	413
C3	441	441	SAV	436	456
DAP	447	467	C12	555	555
KRS	541	561	RAS	555	575
C4	564	564	C11	659	659
BRS	616	636	LAG	659	679
C5	691	691	SKJ	751	771
WER	695	715	MAR	840	860
SLE	802	822	C10	872	872
C6	899	899	SKS	907	927
SLB	916	936	NSS	992	1 012
C7	1 012	1 012	NST	1 078	1 098
FAN	1 031	1 051	C9	1 108	1 108
PAR	1 126	1 146	HOP	1 204	1 224
C8	1 156	1 156	C8	1 293	1 293
HOP	1 238	1 258	PAR	1 299	1 319
C9	1 319	1 319	FAN	1 425	1 445
NST	1 326	1 346	C7	1 546	1 546
NSS	1 412	1 432	SLB	1 555	1 575
SKS	1 501	1 521	C6	1 596	1 596
C10	1 558	1 558	SLE	1 664	1 684
MAR	1 565	1 585	WER	1 789	1 809
SKJ	1 655	1 675	C5	1 815	1 815
LAG	1 753	1 773	BRS	1 857	1 877
C11	1 773	1 773	C4	1 941	1 941
RAS	1 858	1 878	KRS	1 943	1 963
C12	1 878	1 878	DAP	2 044	2 064
SAV	1 982	2 002	C3	2 073	2 073
C13	2 041	2 041	FLO	2 145	2 165
SAM	2 071	2 091	NYG	2 224	2 244
KOK	2 145	2 165	C2	2 273	2 273
C14	2 196	2 196	BYS	2 345	2 365
BIR	2 220	2 240	NON	2 419	2 439
C15	2 326	2 326	C1	2 458	2 458
KOF	2 331	2 351	C0	2 486	2 486
FLE	2 429	2 449	BYP	2 486	2 506

When passing a crossover section during disruption, utilization time depends on whether the involved crossovers are manual or automatic. The 16 crossovers are divided in nine manual and seven automatic crossovers. For the manual crossovers, train operators have to leave the train in order to switch tracks. According to the operating handbook of Bybanen, this takes 49 seconds on average, and is used as default in this thesis. The automatic crossovers can on the other hand be controlled from inside the train, thereby reducing delay compared to using manual crossovers.

The first crossover located at BYP is both doubled-tracked and automatic, resulting in normal runtime when utilizing the crossover. The other crossovers, however, imposes additional delay when utilized, due to being single tracked. The four-step process of passing a single-track crossover is illustrated in Figure 9, displaying the situation when a train travelling in the outbound direction has to switch tracks. Since trains at Bybanen are operated based on line-of-sight, train operators must always have a clear visual. Thus, removing the possibility of reversing. In order to pass the crossover, operators consequently have to walk to the other side of the train in step two. From the operator stops the train to having started driving in the correct direction, it normally takes 60 seconds. The same process must be repeated in step three where the operator must walk to the other side of the train to continue in outbound direction. Finally, since the train is operating in correct direction when returning to outbound track in step four, eventual additional time relates to whether the crossover is manual or automatic.

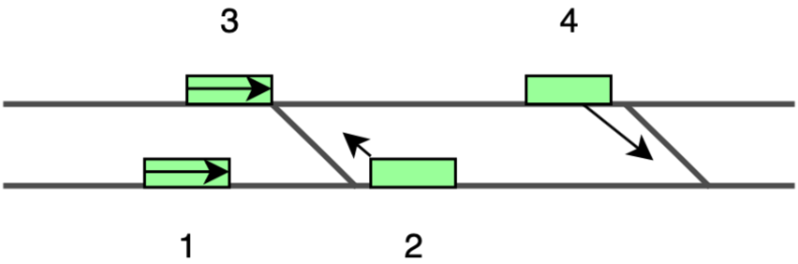


Figure 9, Switching tracks

Table 7 displays the four possible combinations when passing a crossover section. If both crossovers are manual, they cause an additional runtime of 218 seconds. Moreover, 169 seconds are added if one crossover is automatic, while 120 seconds are added if both are automatic. Consequently, the location of an incident may heavily impact overall delay, depending on the characteristics of the involved crossovers. This is based on the following two reasons. Firstly, as manual crossovers impose a greater delay on the system, the crossovers surrounding an incident will have an obvious impact. Secondly, as manual crossovers are more time consuming to utilize, fewer trains are likely to pass an incident through the opposite track due to headway considerations.

Table 7, Runtime for single-track crossovers

	Second manual crossover	Second automatic crossover
First manual crossover first	218 seconds	169 seconds
First automatic crossover	169 seconds	120 seconds

Headway for adjacent trains during normal operation equals the desired frequency set by the timetable. The minimum headway between two trains travelling in the same direction should therefore never be less than 300 seconds during rush hour. The large M is defined as 10 000, in order for binary relations to be valid also during major disruptions.

In real-world situations, trains are not allowed to operate on the opposite track for more than a short distance following the impracticalities to meeting train operations and passengers. Potential crossover links when passing an incident are therefore restricted. This increases the possibility of rapidly finding an optimal route, as smaller subsets positively impact computation time (Samà et al., 2016). Potential crossover links are determined by the number of stations a train passes when operating in the opposite direction. If a link causes a train to pass more than three stations, the crossover link cannot be used. This is a reasonable assumption, given the inconveniences operating in the opposite direction cause meeting train services, and for passengers having to embark on different platforms. Consequently, only the closest and likely the most optimal crossover links are evaluated, considerably reducing computation time.

Three new sets which are used to determine possible crossover links when running the model are therefore introduced, presented in Table 8. $RLINK$ displays stations and crossovers within a crossover link by connecting $LINKS$ and $LINKSC$. $RLINK$ between crossover 1 and 2 in Figure 5 therefore include three links consisting of crossover 1 to station 1, station 1 to station 2 and station 2 to crossover 2.

Table 8, Sets determining possible crossover links

$RLINKI$	Set of tuples s, e, a, b such that link $s, e \in LINKSI$ and $a, b \in LINKSFI$
$RLINKO$	Set of tuples s, e, a, b such that link $s, e \in LINKSO$ and $a, b \in LINKSFO$
$RLINK$	Set of tuples $(RLINKI \cup RLINKO)$

5.2 Implementation of optimization model

Figure 10 illustrates the process of solving the optimization model in AMPL, using the CPLEX 12.9.0 solver and data input from Bybanen. In order to specify variable data, running the model and presenting output in one location, the proposed method is centered around the open-source programming software of RStudio (RStudio, 2019). Running the model from R is made possible through AMPL API, which enables access to AMPL models and solvers, increasing stability and speed (AMPL, 2019). The R script used to run and process the optimization model is presented in Appendix B.

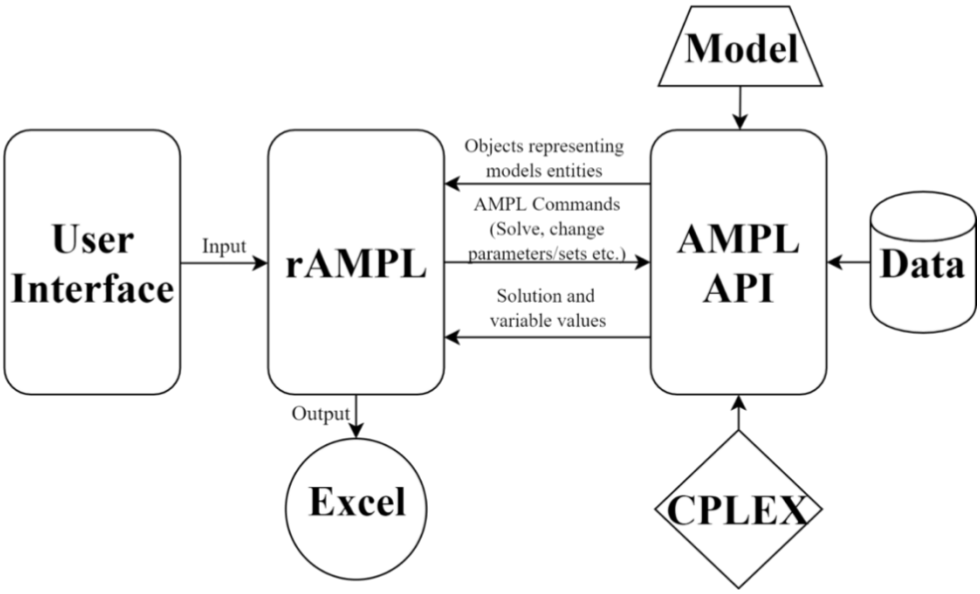


Figure 10, Dispatcher support system

Variable data is specified through the simple user interface in Figure 11, displaying an incident between WER and SLE in the outbound direction, with corresponding incident occurrence and duration. When running the model, variable data can easily be specified by dispatchers and other users. In this process, the specified values are assigned as updated parameter values when running the optimization model. The R script is furthermore used to iterate over all possible crossover links from $RLINK$, returning rescheduled timetables presented as Excel-files. For crossover sections with more than one possible crossover link, multiple files are therefore created. They are moreover named according to the utilized crossover link and objective function value, making it possible to easily determine the optimal rescheduling route in terms of safety and delay. In addition to a rescheduled timetable with specific instructions as to how trains should operate during disruption, the output also include train graphs and delay measurements.

```
Which frequency 4min,5min or 6min:
5
Incident inbound or outbound:
OUTBOUND
Enter start of incident section:
WER
Enter end of incident section:
SLE
Enter start time of incident:
1735
Enter end time of incident:
5206
```

Figure 11, Input specification in user interface

5.3 Case studies

Since the network varies both in terms of where crossovers are placed and density of stations, delay is likely to vary depending on where an incident occurs. In order to display the versatility of the model, three different case studies are considered, illustrated in Figure 12.

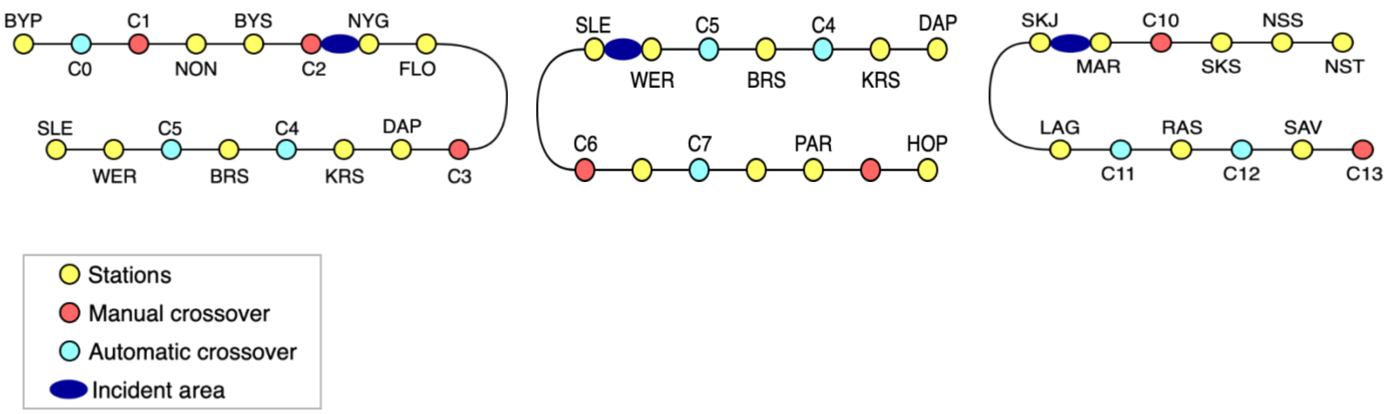


Figure 12, Case studies - incident 1-3

The first case study displays an incident occurring in the outbound direction between C2 and NYG, close to the city center. The area is characterized by large density of stations with few possibilities of passing the incident. In fact, there is only one possible crossover section consisting of three or fewer stations, where both crossovers are manual. There are, however, short runtimes between stations and crossovers, which increases the possibility of respecting headway considerations.

In the second case study, the incident occurs between WER and SLE in the outbound direction. It differs from the previous case as crossover links from C4-C6, C5-C6, and C5-C7 are possible, and that automatic crossovers can be utilized. An interesting question is whether it will be beneficial, in terms of minimizing delay, to choose a longer crossover section to reduce utilization time of crossovers. The situation might occur as C5 to C7 both consists of automatic crossovers, considerably reducing switching time. Finally, the third case study illustrates an incident between SKJ and MAR in inbound direction. The incident differs from the previous cases as it occurs on inbound track and because of lower density of stations. Moreover, only one crossover link can be used to pass the incident, consisting of both automatic and manual crossovers.

How incidents are handled greatly depends on incident duration. To demonstrate the effectiveness of the model, three different incident durations are consequently explored, ranging from minor to major incidents. In the proposed model, one day of train service consists of 5206 seconds. To ensure that trains in both directions are affected by an incident, incidence occurrence is defined to start one third into the day. Thus, an incident starts after 1735 seconds (5206*1/3) and is defined as incident occurrence for all case studies evaluated in this thesis.

For a major disruption, the incident lasts until the end of day at 5206 seconds, which entails a duration of 67%. A medium disruption is defined to last 30% of the day, until 3229 seconds. Finally, a minor disruption lasts 10% of the day, until 2256 seconds. All case studies use a fixed frequency of 300 seconds in order to display an incident situation during rush hour.

The case studies used for numerical analyses are summarized in Table 9.

Table 9, Case studies used for numerical analyses

	Case study 1	**Case study 2**	**Case study 3**
Start of incident section	C2	WER	SKJ
End of incident section	NYG	SLE	MAR
Direction of incident	Outbound	Outbound	Inbound
Incident occurrence	1735	1735	1735
Incident recovery, minor disruption	2256	2256	2256
Incident recovery, medium disruption	3229	3229	3229
Incident recovery, major disruption	5206	5206	5206

5.4 What-if cases

In addition to contributing to better decision-making during real-time train operations, an efficient optimization model can increase knowledge by estimating the impact of changes to the network. This includes both changes in the general infrastructure and more specific train operation characteristics such as frequency and headways. By utilizing this information, critical aspects of the network can be estimated, working as a foundation for discussion of future developments of the network. In this section, three possible network changes currently under consideration by Bybanen is therefore presented.

Finally, an alternative objective function which minimizes the sum of delay at the final stations is presented. This is done in order to compare previous literature (Pellegrini et al., 2015; Xu et al., 2016) to the results obtained from the proposed optimization model. During the comparative analyses for both network and model changes, we define an incident to occur and recover according to a major disruption.

5.4.1 Changes in frequency

By evaluating differences in punctuality and train interactions at different frequencies, it is possible to determine the robustness of the proposed model. Since Bybanen is planning to increase frequency in the future, the effect of increasing frequency to four minutes is consequently evaluated. This requires that additionally four trains are put to use, two in each direction, and that minimum headway is reduced to 240 seconds. The effect of reducing frequency to six minutes by removing two trains in each direction and increasing headway to 360 seconds is also evaluated. Following the changes in rolling stock, it is necessary to determine new incident time frames as length of day changes. These are calculated according to the same methodology as the numerical case studies. The resulting incident occurrence and duration for four – and six minutes frequencies are 1715-5146 seconds and 1675-5026 seconds, respectively.

5.4.2 Implementation of automatic – and doubled tracked crossovers

In the numerical case studies, at least one of the crossovers with closest proximity to the incident section are manual. This involves that operators have to leave the train in order to switch tracks. As this causes increased delay and extra efforts from operators, introducing automatic crossovers can be useful to ensure safe train operations and reduced delay. With an ambition of autonomous train operations within ten years (Mæland, 2019), automatic crossovers have to be in place. The effect of changing the involved crossover sections from manual to automatic is therefore evaluated.

Even with automatic crossover, operators still have a central role when changing tracks as Bybanen only have one double-tracked crossover. In order for trains to be operated centrally, a minimum requirement is therefore to introduce more double-tracked crossovers. Extending the infrastructure in terms of extra tracks induces additional costs as capital cost per kilometer for light rail systems is quite significant (UITP, 2016). However, if double-tracked crossovers prove to considerably improve train services, it may be argued for the development of new tracks both in the current network and in future line extensions. In the third what-if case, the analysis of automatic crossovers is therefore extended to also be double-tracked. Consequently, reducing utilization time of crossovers to zero.

5.4.3 Comparison of objective functions

In order to compare the objective function presented in this thesis with previous literature, an alternative objective function similar to Xu et al. (2016) is formulated, where sum of delay is minimized at the final stations. This is done in order to demonstrate how train interactions differ depending on where delay is minimized. The alternative objective function is presented in (28), where parameter n defines the number of stations in the network.

$$\sum_{t \in \mathrm{O}} \left(a1_{t,n} - a_{t,n} \right) + \sum_{t \in \mathrm{I}} \left(a1_{t,1} - a_{t,1} \right) \tag{28}$$

6 Numerical results

In this chapter, results from the case studies are presented with main focus on how safety regulations are handled. Moreover, how delay varies depending on when and where incidents occur is presented. The numerical analyses are presented according to size, ranging from minor to major disruptions. Finally, results from the what-if cases are presented as the second part of this chapter.

6.1 Analyses of minor disruptions

For dispatchers running the model, it is crucial that the optimal solution provides information about how trains interact during disruption. This includes information about how trains are affected by an incident, not only when using crossovers, but also when waiting for incident recovery. It is consequently crucial to know when and where trains arrive and depart at different sections throughout the network. This can be used to inform passengers about expected delay. In order to demonstrate that safety regulations are met, the rescheduled timetables are visualized through train graphs for every disruption. For each disruption type, one case study is presented through train graphs while the remaining are presented through numerical values.

6.1.1 Case study: SKJ - MAR

Figure 13 displays the rescheduled timetable during a minor disruption at inbound track between SKJ and MAR, where red and green colors respectively illustrate train services in inbound and outbound direction. Track blockage is illustrated by the blue rectangle between SKJ and MAR, occurring at 1735 seconds and recovering at 2256 seconds. When describing train graphs for specific inbound and outbound trains, inbound trains range from I1 to I10, while outbound trains range from O1 to O10.

The figure illustrates that trains arriving before incident occurrence operate according to schedule. Delay is minimized when later arriving inbound trains wait until incident recovery, while outbound trains operate according to plan. Thus, crossovers are not in use. This is clearly displayed where I4 waits between C11 and LAG until track blockage is recovered. Later arriving inbound trains are consequently delayed from LAG in order to respect headway considerations for adjacent trains. The figure indicates a short distance between waiting inbound trains at C11 and LAG. Despite this, the restriction where only one train can be in the same section is respected as RAS-C11 and C11-LAG are two separate links. This is true even if there is only a short distance between C11 and LAG.

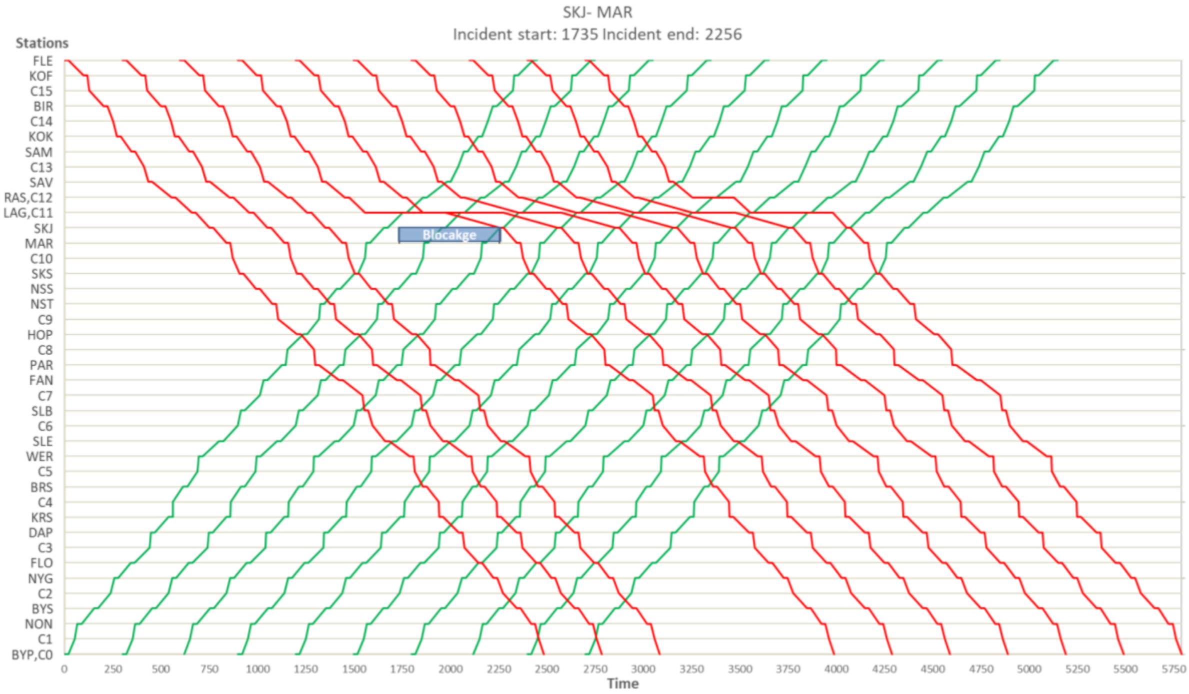


Figure 13, Rescheduled timetable SKJ-MAR

For the situation between SKJ and MAR during minor disruption, there are two closely connected effects contributing to delay. The first effect relates to I4 waiting until incident recovery, causing 10:05 minutes delay for passengers arriving between SKJ and BYP. The second effect is caused by respecting headway considerations for adjacent trains, where remaining inbound trains experience the same delay as I4 from SKJ to BYP. I5-I10 are in addition 06:37 minutes delayed at LAG due to restrictions of only one train at the same section simultaneously.

Maximum delay for passengers in inbound direction is consequently 10:05 minutes and is summarized with overall punctuality in Table 10. Punctuality is calculated by dividing the objective function value with scheduled arrival. Bybanen punctuality is moreover calculated with a margin of 3:00 minutes in order to reflect their current definition of punctuality. Table 10 states a perfect punctuality for outbound passengers, while inbound passengers arrive according to schedule with Bybanen punctuality of almost 92%. Absolute punctuality is in contrast 88%. It is important to remark that even if inbound trains do not arrive according to schedule, most passengers travelling from FLE to LAG and from SKJ to BYP will experience a nearly normal frequency. Passengers who experience abnormal frequencies and delays are largely those who have to pass the incident area, and passengers waiting for I4 to arrive from

SKJ to BYP. Passengers travelling with I5-I10 from SKJ will in contrast experience normal frequencies.

Table 10, Punctuality and maximum delay during minor disruption, SKJ -MAR

	Outbound trains	Inbound trains	All trains
Bybanen punctuality	100 %	91.57 %	95.68 %
Absolute punctuality	100 %	87.93 %	93.81 %
Maximum delay measured at station	0:00 minutes	10:05 minutes	10:05 minutes

6.1.2 Case study: C2-NYG

By contrast to the incident between SKJ and MAR, a minor disruption between C2 and NYG causes one train to use crossovers. The resulting train graph demonstrates that O6 utilizes C2 and C3 to pass the incident, while incoming inbound trains are only slightly affected by switching procedures. This occurs as the incident section is located far from the origin station of inbound trains, such that I1 only has to adjust runtime with a few minutes for O6 to pass the incident. The train graph further shows that O7 waits until disruption is recovered. The crossover effect caused by O6 switching tracks at C2 produces a delay of 2:49 minutes for passengers arriving at NYG and FLO. The delay further increases after returning to outbound track at C3, such that passengers arriving between DAP and FLE experience a delay of 3:38 minutes.

O7 on the other hand, is 3:37 minutes delayed at NYG while waiting for incident recovery. Delay is further adjusted to 3:38 minutes from DAP to respect headway for adjacent trains. Passengers travelling with later arriving outbound trains experience the same delay following headway effects. Finally, I1 adjusts runtime in order to ensure safe train operations when O6 switches tracks and passengers experience a delay of 3:07 minutes from BYS to BYP. The same is true for later arriving inbound trains where runtime is adjusted to respect adjacent headway.

Maximum delay and punctuality are presented in Table 11, displaying high punctuality for both outbound and inbound trains. The results show that maximum delay is just slightly higher than 3:00 minutes. Thus, based on Bybanen punctuality almost all trains arrive according to schedule.

Table 11, Punctuality and maximum delay during minor disruption, C2-NYG

	Outbound trains	Inbound trains	All trains
Bybanen punctuality	99.35%	99.95 %	99.65 %
Absolute punctuality	96.22 %	98.66 %	97.45 %
Maximum delay measured at station	3:38 minutes	03:07 minutes	3:38 minutes

6.1.3 Case study: WER-SLE

Similar to SKJ-MAR, crossovers are not in use when a minor disruption occurs between WER and SLE. Delay is minimized when outbound trains wait until the track blockage is recovered. The situation arises as O4 passes the incident section shortly before incident occurrence, and because O5 arrive almost halfway through disruption. Thus, O5 does not experience major delays by waiting. Incident occurrence and location results in several inbound trains being located in the incident section. Consequently, creating difficulties of respecting headway for meeting trains without considerably adjusting runtime.

When O6 waits for disruption to be recovered it results in 06:01 minutes delay for passengers arriving between WER and FLE. Due to headway effects for adjacent trains, passengers of later arriving outbound trains therefore experience identical delays. The resulting punctuality is presented in Table 12, where the perfect punctuality of inbound trains is achieved at the expense of outbound trains. Despite this, outbound punctuality is high even when some trains experience a delay of 06:01 minutes.

Table 12, Punctuality and maximum delay during minor disruption, WER-SLE

	Outbound trains	Inbound trains	All trains
Bybanen punctuality	97.01%	100 %	98.51 %
Absolute punctuality	94.03 %	100 %	97.04 %
Maximum delay measured at station	6:01 minutes	0:00 minutes	6:01 minutes

For all case studies the computation time of solving the optimization model is averagely two seconds per crossover link. When solving for the most comprehensive incident section between WER and SLE, the optimal solutions are therefore presented in six seconds.

6.2 Analyses of medium disruptions

6.2.1 Case study: C2-NYG

Figure 14 illustrates the rescheduled timetable when track blockage occurs at outbound track between C2 and NYG at 1735 seconds and recovers at 3229 seconds. The figure shows that O6 and O7 passes blockage through C2 and C3, while later arriving outbound trains wait until disruption is recovered. It shows that I1 adjusts runtime between DAP and C3 to respect safety headways for meeting trains. This is clearly illustrated by the figure, as O6 and O7 arrive at C3 before inbound trains, ensuring safe train operations. In this process, inbound trains wait between DAP and C3.

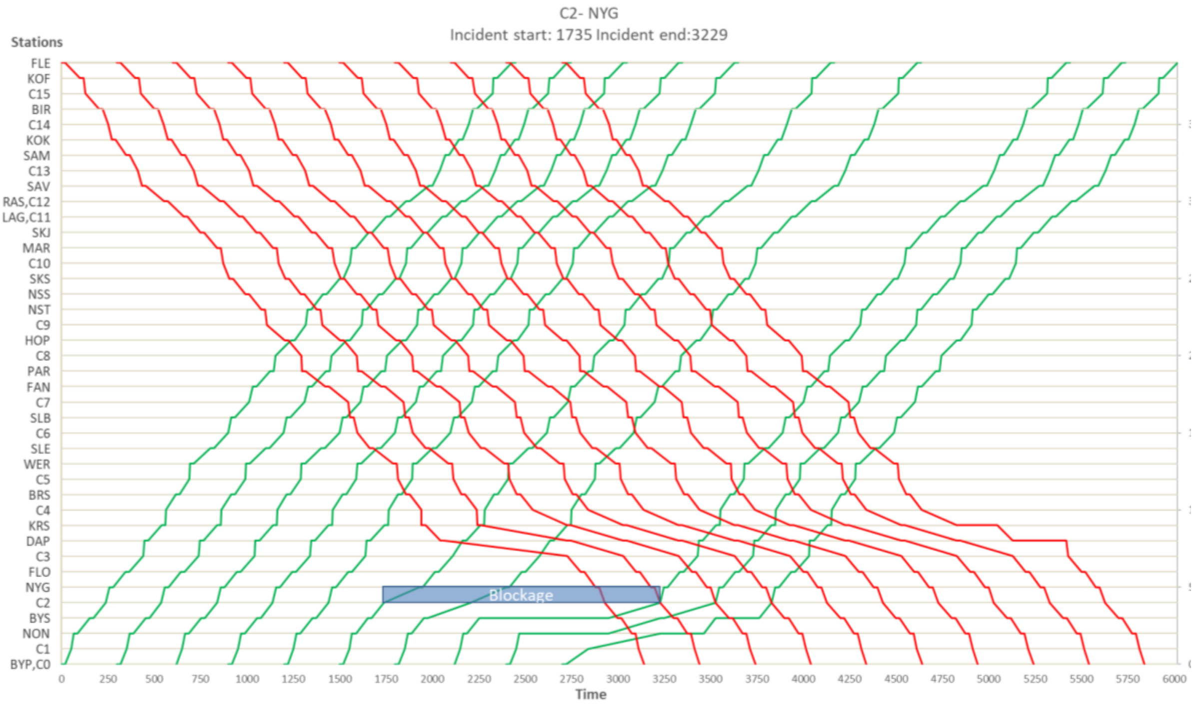


Figure 14, Medium disruption between C2 and NYG

The train interactions illustrated in Figure 14 contribute to crossover effects, headway effects for meeting trains, and headway effects for adjacent trains. When switching tracks, passengers travelling with O6 experience a delay of 2:49 minutes at NYG and 3:38 minutes from DAP to FLE. The succeeding outbound train is 5:38 minutes delayed at NYG and 6:27 minutes from DAP to FLE. This occurs as O7 has to respect both headway for adjacent trains at C2, and utilization time of crossovers. Thus, delay caused by the headway effect is 2:49 minutes. Passengers travelling with O8-O10 are in addition 14:50 minutes delayed from NYG to FLE.

Headway effects for meeting trains contribute to 10:56 minutes delay for all inbound passengers travelling from FLO to BYP. I2 is delayed by 06:25 minutes at DAP, while I3-I10 are 03:06 minutes delayed at KRS. The resulting punctuality in Table 13 illustrates that increased blockage duration decreases punctuality compared to minor disruption. However, punctuality is still high with more than 94% absolute punctuality for inbound trains and 88% for outbound trains.

Table 13, Punctuality and maximum delay during medium disruption, C2-NYG

	Outbound trains	Inbound trains	All trains
Bybanen punctuality	91.57 %	96.33 %	93.97 %
Absolute punctuality	88.36 %	94.61 %	91.51 %
Maximum delay measured at station	14:50 minutes	10:56 minutes	14:50 minutes

6.2.2 Case study: WER-SLE

Despite increased blockage duration between WER and SLE, overall delay is still minimized when outbound trains are waiting until incident recovery. Passengers travelling in outbound direction therefore experience a noteworthy increase of delay compared to minor disruption. Following a wait time of 22:14 minutes for O6, all succeeding outbound trains will experience the same delay from WER to FLE. Passengers travelling with later arriving outbound trains will in addition experience delay at earlier stations, and O10 is delayed already at DAP. This occurs following headway effects and restrictions of the number of trains that can wait in the same section.

The resulting punctuality is presented in Table 14, where absolute punctuality of outbound trains is reduced from 94% during minor disruption to 77% during medium disruption. Maximum delay has also increased considerably.

Table 14, Punctuality and maximum delay during medium disruption, WER-SLE

	Outbound trains	Inbound trains	All trains
Bybanen punctuality	80.52 %	100 %	90.33 %
Absolute punctuality	77.37 %	100 %	88.77 %
Maximum delay measured at station	22:14 minutes	00:00 minutes	22:14 minutes

6.2.3 Case study: SKJ-MAR

In contrast to a minor disruption between SKJ and MAR, a medium disruption causes I4 to switch tracks. This is done through C11 and C10, while O1 correspondingly adjusts runtime to ensure safe train operations. Later arriving outbound trains experience similar runtime adjustments when respecting headway for adjacent trains. In order to respect meeting train operations, I5-I10 wait until disruption is recovered, causing delay for passengers arriving between SKJ and BYP.

Due to C11 being automatic, crossover effect for I4 is only 2:49 minutes in total. This is divided in 2:00 minutes delay from LAG to MAR and an additional 0:49 minutes from SKS to BYP. I5 is further 21:18 minutes delayed at SKJ after waiting for incident recovery. Following headway effects, later arriving inbound trains are equally delayed. I6-I10 are in addition delayed at earlier stations, and I10 experiences delays already at SAV. Outbound trains are on the other hand 08:09 minutes delayed from MAR to FLE, while O2-O10 also experience minor delays at earlier stations. Maximum delay and corresponding punctuality are presented in Table 15, indicating a doubling of maximum delay compared to the minor disruption. Absolute punctuality is, however, still more than 85% when measured for all trains.

Table 15, Punctuality and maximum delay during medium disruption, SKJ-MAR

	Outbound trains	Inbound trains	All trains
Bybanen punctuality	95.42%	81.31 %	88.31 %
Absolute punctuality	92.54 %	77.68 %	85.05 %
Maximum delay measured at station	08:09 minutes	21:18 minutes	21:18 minutes

6.3 Analyses of major disruptions

6.3.1 Case study: WER-SLE

Figure 15 illustrates a major disruption at the outbound track between WER and SLE, occurring at 1735 seconds and recovering at 5206 seconds. The train graph indicates that the consequences of blockage are considerably more severe than for previous cases. During blockage, the optimal solution states that O5-O7 change tracks between C5 and C6, while remaining outbound trains wait until incident recovery. The optimal solution indicates that even if C4 and C6 both are automatic, delay is minimized when operating through the crossover section of the nearest proximity to the incident. When O5-O7 utilize the opposite track, inbound trains correspondingly have to wait in order to respect headway considerations.

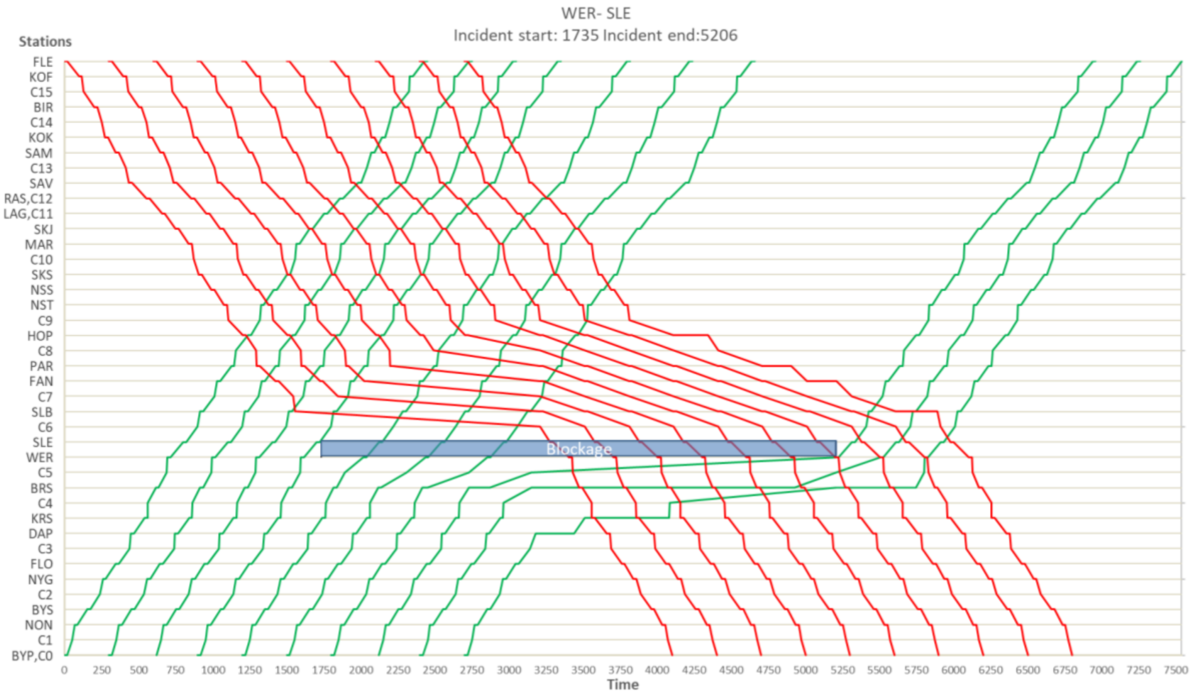


Figure 15, Major disruption between WER and SLE

Outbound trains operating on inbound track are delayed according to the crossover effect induced by utilizing crossovers. O5 is therefore 2:00 minutes delayed at WER and 2:49 minutes delayed from SLB to FLE. Delay further increases by 2:00 minutes for every train using crossovers following headway effects, and O7 is 06:49 minutes delayed from SLB to FLE.

Later arriving outbound trains are, however, considerably more delayed as O8 waits 40:11 minutes at WER. O9 and O10 are in addition delayed at earlier stations. All inbound trains are 26:56 minutes delayed from SLE to BYP as I1 waits until O6-O8 have passed the incident. Later arriving inbound trains are also delayed at earlier stations, and I8-I10 have delays already from HOP.

The major disruption causes large delays for all passengers arriving after incident occurrence, where punctuality and maximum delay are presented in Table 16. The table states that absolute punctuality is reduced from almost 89% at medium disruption to 75 % during major disruption, and maximum delay has doubled. Maintaining a steady frequency is challenged to a higher degree compared to less comprehensive disruptions where passengers travelling with O7-O10 experience considerable delay. O1-O6 and inbound trains by contrast, experience a more normal frequency.

Table 16, Punctuality and maximum delay during major disruption, WER-SLE

	Outbound trains	Inbound trains	All trains
Bybanen punctuality	80.39 %	76.79 %	78.57 %
Absolute punctuality	77.38 %	73.55 %	75.45 %
Maximum delay measured at station	40:11 minutes	26:56 minutes	40:11 minutes

6.3.2 Case study: C2-NYG

Compared to the medium disruption between C2 and NYG, a major disruption causes additionally two trains to switch tracks. Thus, four trains in total utilize crossovers to pass the incident. Following the same analogy as in previous cases, manual crossovers induce 3:38 minutes delay from DAP to FLE, whereas headway effects result in additionally 2:49 minutes delay between each train at C2. As the final train to use crossovers, O9 is therefore 12:05 minutes delayed from DAP. In order to ensure safe train operation when O6-O9 operate on inbound track, meeting operations wait before C3, resulting in 26:34 minutes delay from FLO. Furthermore, O10 waits until disruption is recovered, causing passengers travelling from DAP to FLE a delay of 37:47 minutes.

Table 17 demonstrates that absolute punctuality is still high, and equally distributed between outbound and inbound trains. A longer disruption has consequently caused less impact on punctuality than for WER-SLE. However, maximum delay has increased considerably compared to minor and medium disruptions, and both outbound and inbound trains experience significant delays.

Table 17, Punctuality and maximum delay during major disruption, C2-NYG

	Outbound trains	Inbound trains	All trains
Bybanen punctuality	88.56 %	86.98 %	87.76 %
Absolute punctuality	85.36 %	85.06 %	85.21 %
Maximum delay measured at station	37:47 minutes	26:34 minutes	37:47 minutes

6.3.3 Case study: SKJ-MAR

Similar to the other case studies, a major disruption between SKJ and MAR causes several trains to use crossovers. During blockage, I4-I7 operate on outbound track through C11 and C10. Based on the analogy for medium disruptions, trains utilizing crossovers causes 2:49 minutes delay for passengers from SKS to BYP. In addition, headway effects increase delay by 2:00 minutes at SKS for every train that utilizes crossovers. In order to respect headway considerations for meeting train operations, outbound trains wait 29:10 minutes at LAG for inbound trains to safely switch tracks.

I7-I10 on the other hand wait between C11 and LAG until blockage is recovered. The results show a delay of 39:15 minutes for passengers travelling between SKS and BYP. I9 and I10 are in addition considerably delayed at RAS and LAG, while I10 is marginally delayed already from SAV. The resulting punctuality is presented in Table 18, displaying an absolute punctuality of 73%. Outbound trains are punctual less than 70% of the time, while inbound trains experience the longest delay. Major disruptions between SKJ and MAR result in considerable challenges in terms of upholding a desired service level.

Table 18, Punctuality and maximum delay during major disruption, SKJ-MAR

	Outbound trains	Inbound trains	All trains
Bybanen punctuality	73.01 %	79.65 %	76.36 %
Absolute punctuality	69.60 %	76.01 %	72.83 %
Maximum delay measured at station	29:10 minutes	39:15 minutes	39:15 minutes

6.4 Analyses of what-if cases

6.4.1 Sensitivity of frequency changes

Figure 16 illustrates how absolute punctuality varies during major disruptions at different frequencies. The figure shows that punctuality is relatively stable for all frequencies, although slightly higher when fewer trains are in the system. In addition, higher frequencies cause decreased punctuality compared to base case, except for C2-NYG where punctuality increases.

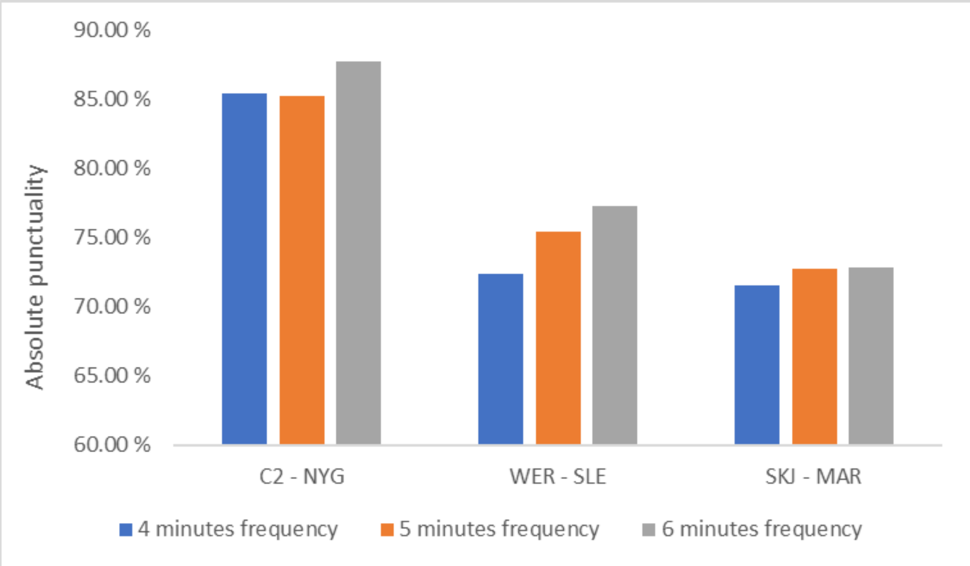


Figure 16, Absolute punctuality at different frequencies

Differences in punctuality are results of timing and mode of train operations at incident sections. The resulting train graphs for different frequencies show that punctuality increases with lower frequencies due to fewer directly affected trains waiting for incident recovery. As maximum delay in the base cases were all caused by waiting trains, fewer trains therefore increase overall punctuality. Increased rolling stock similarly produces decreased punctuality for WER-SLE and SKJ-MAR.

In contrast, waiting trains for C2-NYG are identical during the two most frequent train services. However, there are more trains operating according to schedule at four minutes frequencies, due to one additional train passing before incident occurrence. The resulting punctuality is therefore higher. The results also show that changing frequencies only impact utilization rate of crossovers for SKJ-MAR and WER-SLE. For the case study between SKJ-MAR, crossovers are used less at frequencies of six minutes, while identical with base case at four minutes. WER-SLE, on the other hand, experience the same utilization rate of crossovers during five – and six minutes frequencies. During higher frequencies, one additional train utilizes crossovers.

All things considered, changes in frequency only slightly influence how train services are operated during disruptions. The case studies indicate that reducing density of trains cause fewer conflicts and increase overall punctuality.

6.4.2 Implementation of automatic crossovers

By introducing automatic crossovers utilization time of crossovers is reduced by 98 seconds for C2-NYG, and 49 seconds for the remaining case studies. This is explained by the fact that both crossovers surrounding C2-NYG are manual, while WER-SLE and SKJ-MAR both are surrounded by one automatic crossover. The explicit effect of introducing automatic crossovers is consequently larger for the first case study.

The resulting timetable of C2-NYG is presented in Figure 17 and illustrates reduced delay for outbound trains. This occurs as automatic crossovers causes O10 to switch tracks instead of waiting for incident recovery.

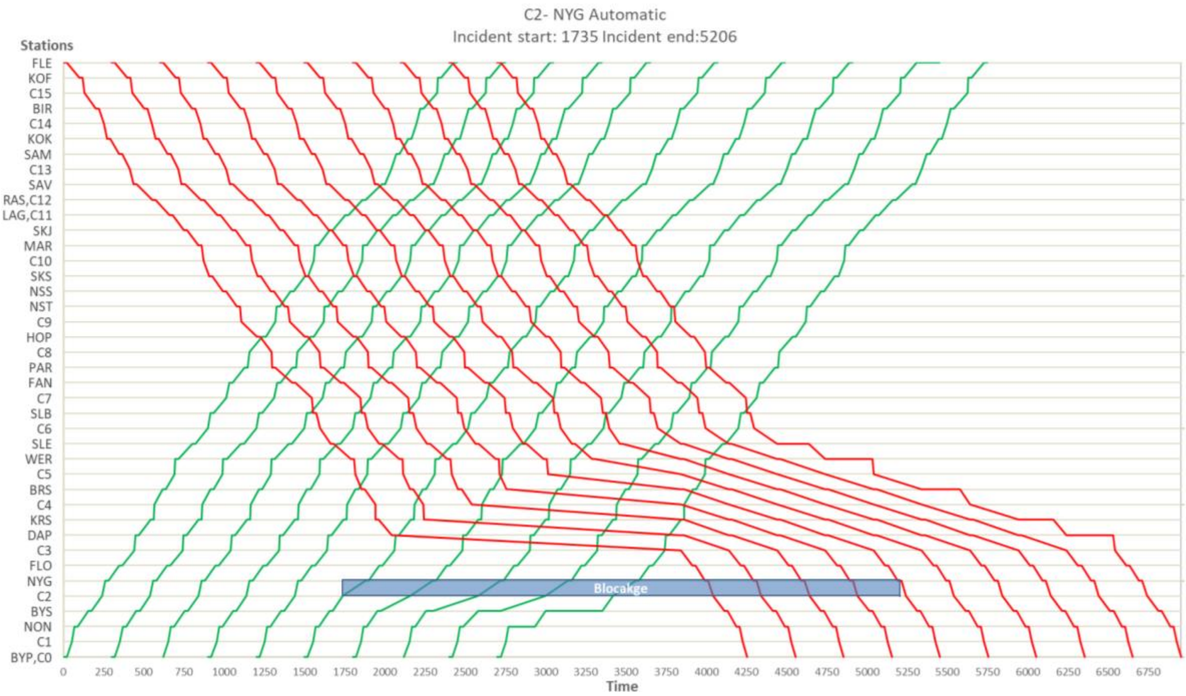


Figure 17, Train graph for automatic crossovers, C2-NYG

Maximum delay for outbound trains is reduced from 37:47 minutes to 10:00 minutes. Inbound passengers are by contrast more delayed following increased wait time at C3. The resulting maximum delay is therefore increased from 26:34 to 29:29 minutes. Despite this, absolute punctuality is considerably improved compared to base case, as the overall effect of delay reduction for outbound trains is larger than the increase for inbound trains. Corresponding punctuality is presented in Table 19, showing an overall punctuality of more than 88% when utilizing automatic crossovers.

Table 19, Absolute punctuality for automatic crossovers, C2-NYG

	Outbound trains	Inbound trains	All trains
Automatic crossovers	93.71 %	83.10 %	88.36 %
Base case	85.36 %	85.06 %	85.21 %

In contrast to C2-NYG, introducing automatic crossovers for the other case studies have no impact on the number of trains switching tracks. The overall effect on punctuality is marginal, following the reduction of 49 seconds when utilizing crossovers and corresponding headway effects. Absolute punctuality for the two case studies are presented in Table 20 and 21, displaying an increase in overall punctuality of less than one percent.

Table 20, Absolute punctuality for automatic crossovers, SKJ-MAR

	Outbound trains	Inbound trains	All trains
Automatic crossovers	70.54 %	76.49 %	73.54 %
Base case	69.60 %	76.01 %	72.83 %

Table 21, Absolute punctuality for automatic crossovers, WER-SLE

	Outbound trains	Inbound trains	All trains
Automatic crossovers	77.74 %	74.43 %	76.07 %
Base case	77.38 %	73.55 %	75.45 %

6.4.3 Implementation of double-tracked crossovers

Since utilization time of crossovers is zero when double-tracked, automatic crossovers are in use, delay caused by disruptions only consist of headway effects for adjacent and meeting train operations. Figure 18 illustrates the rescheduled timetable when automatic, double-tracked crossovers are used to pass an incident between SKJ and MAR. By comparison to base case, two additional trains utilize crossovers, resulting in considerable delay reduction. This is clearly indicated by the figure where six trains change tracks, and only I10 have to wait for incident recovery. Outbound trains correspondingly experience increased delay, although only marginally compared to the improvements of inbound trains.

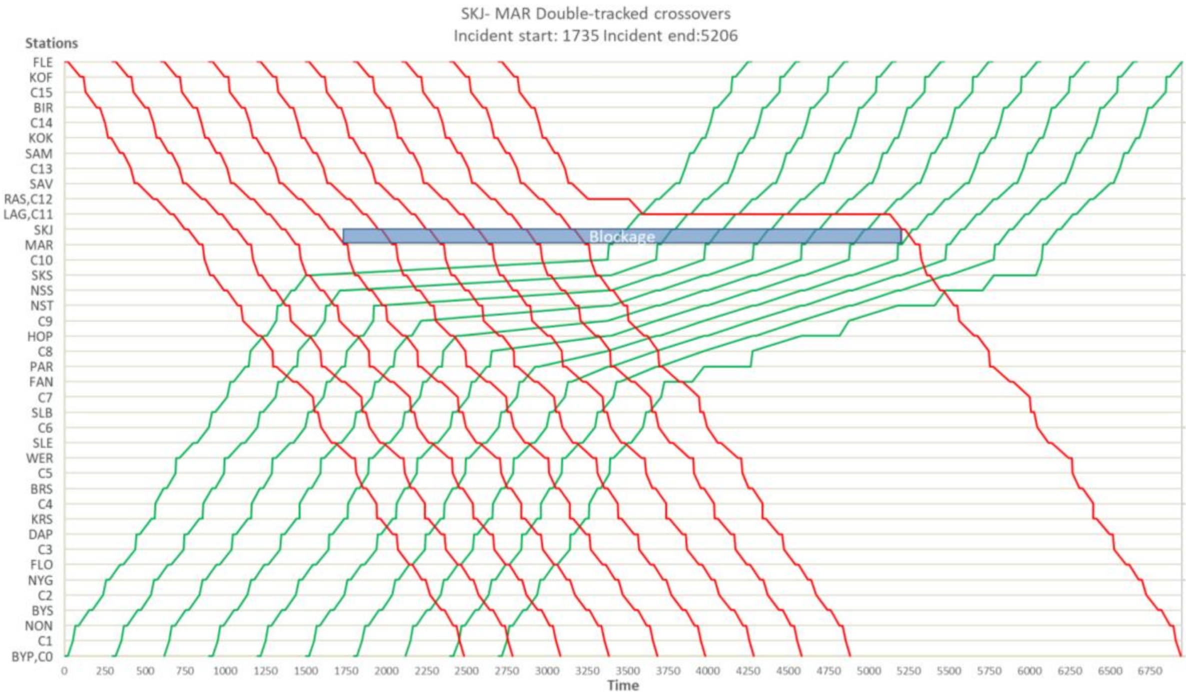


Figure 18, Rescheduled timetable when double-tracked crossovers, SKJ-MAR

Maximum delay has increased to 30:21 minutes for outbound trains as compared to 29:10 minutes in base case. For inbound trains, on the other hand, maximum delay is reduced from 39:15 to 29:15 minutes. The resulting absolute punctuality is presented in Table 22, displaying a considerable improvement in overall punctuality. By introducing double-tracked crossovers, inbound trains accomplish a punctuality of 95% compared to 76% in base case. Outbound trains experience a reduced punctuality due to increased wait time. Since wait time for outbound trains only marginally increases, overall punctuality is improved from 73% to 82%.

Table 22, Absolute punctuality for double-tracked automatic crossovers, SKJ-MAR

	Outbound trains	Inbound trains	All trains
Double crossovers	68.24 %	95.12 %	81.79 %
Base case	69.60 %	76.01 %	72.83 %

The rescheduled timetable for WER-SLE produces similar results to SKJ-MAR, as two additional trains switch tracks and only one train waits until incident recovery. Changes in train operations decrease maximum delay for outbound trains while slightly increasing for inbound trains. The corresponding punctuality is presented in Table 23, showing the same trends as for SKJ-MAR. Trains that are switching tracks considerably increase punctuality, while meeting operations perform worse. The effect on overall punctuality is therefore positive, improving from 75% in base case to 82% when double-tracked crossovers are used.

Table 23, Absolute punctuality for double-tracked automatic crossovers, WER-SLE

	Outbound trains	Inbound trains	All trains
Double crossovers	94.97 %	70.11 %	82.45 %
Base case	77.38 %	73.55 %	75.45 %

Finally, the case study between C2-NYG displays the same train operation effects as for the introduction of automatic crossovers. Thus, the final outbound train utilizes crossovers instead of waiting. However, double-tracked crossovers considerably improve punctuality for both outbound and inbound trains. In the rescheduled timetable, outbound trains achieve a perfect punctuality as automatic, double-tracked crossovers allow runtimes according to normal operation. Inbound trains, on the other hand, experience additional delay when waiting for outbound trains to pass. Compared to base case, maximum delay for inbound trains are reduced from 26:34 to 19:29 minutes. The overall punctuality therefore shows a considerable improvement, increasing from 85% to 95%.

Table 24, Absolute punctuality for double-tracked automatic crossovers, C2-NYG

	Outbound trains	Inbound trains	All trains
Double crossovers	100 %	89.54 %	94.73 %
Base case	85.36 %	85.06 %	85.21 %

Overall, introducing double-tracked crossovers therefore have significant impact for all case studies, not only in terms of improved punctuality, but also in terms of ensuring safe train operations.

6.4.4 Comparison of objective functions

When minimizing delay for the final stations, the resulting train graphs illustrate that unaffected trains are prioritized at the expense of affected trains. As an example, Figure 19 displays train operations during major disruption between C2 and NYG when the alternative objective function is used. The figure shows that inbound trains operate almost according to plan, while the majority of outbound trains wait. By comparison, only one outbound train had to wait in base case.

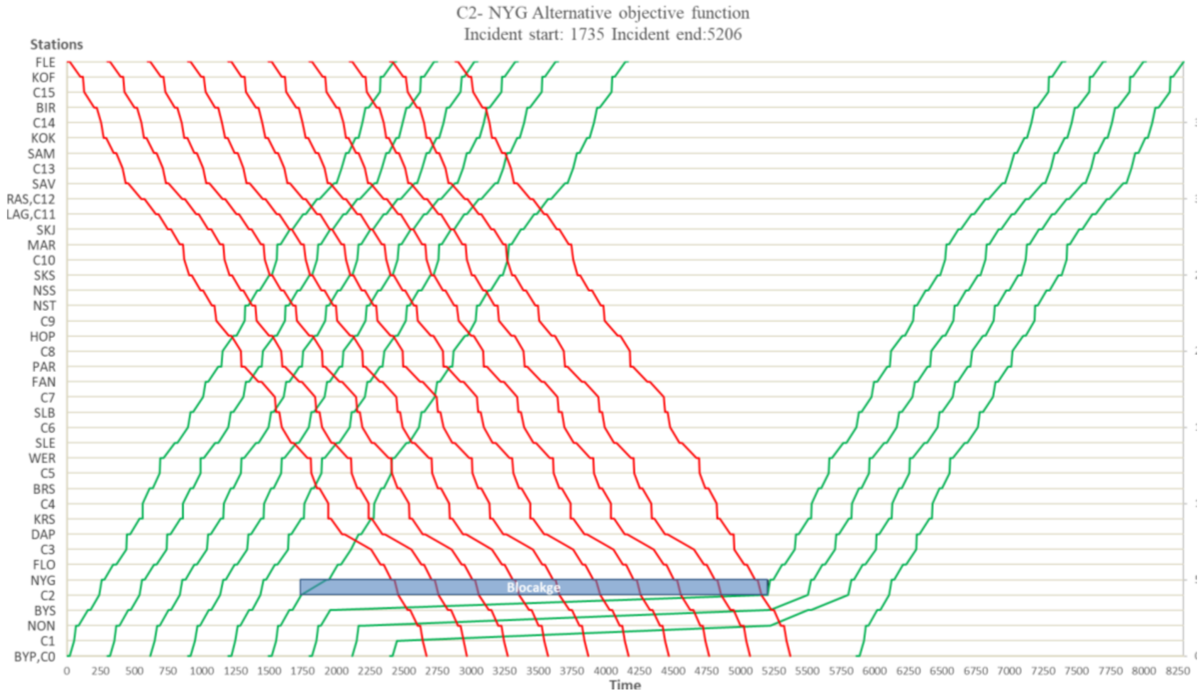


Figure 19, Train graph C2-NYG with alternative objective function

The resulting absolute punctuality is presented in Table 25, divided in base case and the alternative objective function. Since the objective functions differ in terms of where delay is minimized, direct comparison of overall punctuality is inaccurate. It is therefore more interesting to analyze distribution of punctuality between outbound and inbound trains. As shown in Table 25, punctuality is considerably skewed towards inbound trains when minimizing delay at final stations. In fact, only 66% of outbound trains arrive according to schedule. When minimizing delay at all stations, there is almost an equal distributed punctuality for passengers in both directions.

Table 25, Absolute punctuality for alternative objective function, C2-NYG

	Outbound trains	Inbound trains	All trains
Base case all stations	85.36 %	85.06 %	85.21 %
Alternative final stations	65.90 %	95.13 %	80.62 %

Moreover, the alternative objective function causes all outbound trains to wait when a major disruption occurs between WER and SLE. This differs from base case where three outbound trains were allowed to use crossovers. The corresponding increase of maximum delay for outbound trains is therefore considerable, from 40:11 to 55:11 minutes. Inbound trains on the other hand, operate according to schedule. In the case study between SKJ and MAR, two additional trains have to wait instead of utilizing crossovers. Maximum delay consequently increases from 39:15 to 49:15 minutes for inbound passengers, while outbound passengers experience less delay. The corresponding punctualities are presented in Table 26-27, indicating the same trend as for C2-NYG. Thus, by utilizing the alternative objective function, punctuality is skewed towards unaffected trains. This is especially true for WER-SLE where outbound trains experience low punctuality at the expense of perfect punctuality for inbound trains. SKJ-MAR on the other hand, show a more equal distribution of delay, although noticeably more unbalanced than base case.

In general, the alternative objective function contributes to more unbalanced train services in comparison to minimizing delay at all stations.

Table 26, Absolute punctuality for alternative objective functions, WER-SLE

	Outbound trains	Inbound trains	All trains
Base case all stations	77.38 %	73.55 %	75.45 %
Alternative final station	47.43 %	100 %	73.91 %

Table 27, Absolute punctuality for alternative objective function, SKJ-MAR

	Outbound trains	Inbound trains	All trains
Base case all stations	69.60 %	76.01 %	72.83 %
Alternative final station	75.93 %	60.29 %	68.05 %

7 Discussion

In this chapter, the implications of our findings and how the optimization model performs in terms of safety and punctuality are discussed. Moreover, we discuss how the model performs compared to existing literature. Potential sources of error related to the assumptions made in this thesis are thereafter discussed. Finally, possible practical implications of the model as part of a real-time dispatcher support system is presented.

7.1 Implications from numerical analyses

In the numerical analyses, we studied three incidents located at different sections in the light rail system. This was done in order to evaluate how delay and train interactions vary depending on the location and timeframe of an incident. Results were furthermore presented in train graphs and corresponding performance indicators. Case studies for minor disruptions indicated that delay was minimized when few trains utilized crossovers, which can be explained by incident duration. When the expected incident duration is short, utilizing crossovers seem to do more harm than good as crossover operations cause additional delays following both crossover – and headway effects. This is illustrated by a punctuality of more than 93% in all case studies even when affected train services had to wait.

In the case study where crossovers were utilized, meeting operations had barely reached the incident section. Thus, crossover operations could easily be performed without inflicting major delays. It is therefore likely that crossovers would have been omitted if the incident occurred later in the day. That being said, as the case study is characterized by short runtimes between stations and crossovers, respecting headways for meeting operations is easier compared to the other case studies. This is true even if the crossover section in question include manual crossovers.

Consistent with existing research (Chang et al., 2019; Xu et al., 2016), increased incident duration caused larger delays for all case studies due to an increased number of affected trains. In contrast to traditional urban rail systems where operational flexibility is limited to adjusting run – and dwell times (Gao et al., 2017), our proposed model created increased flexibility as crossovers were used to minimize delay. This was demonstrated during medium disruptions where additional trains, compared to minor disruptions, utilized crossovers to pass the incident. Thus, train services were improved in terms of punctuality when using crossovers. Our findings indicated that incident location impacts prioritization of train operations. To provide an illustration, even with several automatic crossovers surrounding the section WER-SLE, the optimal solution proposed that affected trains should wait until incident recovery. This can be explained by operating time between links in the crossover section, and the timeframe when conflicting train operations arrive at the start of the section. This information is highly relevant for dispatchers not only when planning where crossovers should be located, but also during real-time instances. By apprehending information about handling of critical sections, the light rail system might be improved in the future.

The analyses of major disruptions further demonstrated a correlation between delay and disruption length. The findings also showed that even during severe track blockage, utilization of crossovers can efficiently reschedule train operations. This was indicated by train graphs displaying safe train interactions, as well as obtaining a reasonable punctuality of more than 72% in the most affected case study. Similar to optimization models for other railway systems (Gao et al., 2017; Samà et al., 2017), the presented optimization model provided conflict-free timetables during disruption while minimizing delay.

In contrast to previous research (Pellegrini et al., 2015; Xu et al., 2016), we minimized delay at all stations to ensure more balanced train services in terms of punctuality. The resulting punctuality demonstrated a relatively proportionate distribution for all case studies, with the largest difference between outbound and inbound trains being 7% during major disruption. The applicability of the proposed objective function was determined through comparison to an alternative objective function where delay is minimized at final stations, similar to Xu et al. (2016). The results indicated that when minimizing delay at all stations, punctuality was more evenly distributed between outbound and inbound passengers.

For the case study between C2-NYG there was a skewness of punctuality by almost 30 % with the alternative objective function, and less than one percent when minimizing for all stations. Additionally, all case studies demonstrated that fewer trains changed tracks with the alternative objective function. In order to force a more balanced train service, both Xu et al. (2016) and Pellegrini et al. (2015) introduced weighting of train operations, either in the objective function (Pellegrini et al., 2015) or as an additional constraint (Xu et al., 2016). In light of this, minimizing delay at all stations seems preferable as it provides an evenly distributed train service while avoiding unnecessary restrictions when exploring the optimal rescheduling route.

Even with high overall punctuality, the optimal solutions have potential for improvement as several passengers experienced considerable delay. In real-world operations it is nearly impossible to avoid that passengers are delayed during extensive track blockages. That being said, the presented what-if cases demonstrated several possibilities to improve rescheduling of train operations. Similar to Xu et al. (2016), our analyses showed that delay decreased with lower frequencies. This occurred as fewer trains and longer headways made it easier to respect safety considerations, thereby decreasing headway effects for meeting operations.

Moreover, the analyses indicated that reducing train frequency from five to four minutes only marginally impacted overall punctuality. This can be explained by the fact that reducing minimum headway correspondingly reduces delay caused by headway effects. The extra delay due to an increased number of waiting trains was almost compensated by reduced headway effects. Following this analogy, a possible dispatcher solution during disruption could involve reduction of the required headway while maintaining the same rolling stock. Consequently, shorter distances between trains would be allowed, reducing delay caused by headway effects in both directions.

The what-if cases indicated that introducing automatic crossovers only was effective when both crossovers surrounding an incident section were manual beforehand. According to this, the cost-benefit of implementing automatic crossovers in itself might not have a positive effect. The findings suggest that further expanding the model to include double-tracked crossovers can significantly improve train operations. These results are caused by considerable delay reduction following crossover effects, where affected trains safely can switch tracks without causing large disturbances to the system. This was clearly illustrated by punctuality improvements ranging from 7-9%, as well as significant reduction of maximum delay for all case studies.

As opposed to the short-turning strategy presented by Chang et al. (2019), double-tracked crossovers would produce larger improvements for our proposed model, as a short-turning strategy implies that train operators have to change driving direction. At Bybanen, this process currently takes 60 seconds. Thus, as the proposed solution entails driving in the designated direction, introducing double-tracked crossovers would reduce operating time by 60 seconds compared to the current dispatcher method of Bybanen.

In light of our findings it might therefore be beneficial to increase the number of double-tracked crossovers both in the current system and future line extensions. However, as capital cost per kilometer of track is relatively high (UITP, 2016), and Bybanen is already an expensive project (Norwegian Government, 2011, 2017), it is not realistic with only double-tracked crossovers.

The proposed optimization model can, however, provide an understanding of the location of double-tracked crossovers. Determining the effect of such crossovers could be used as basis of negotiation with Hordaland County Council when determining future budgets of Bybanen, as well as in applications for national funding.

The combined findings suggest that the proposed optimization model can be used to safely reschedule train operations during disruption. It also shows applicability when minimizing delay at all stations and improving service level for passengers. Rescheduling trains was done through utilization of crossovers, located at their actual position in the network, and can be seen as new contributions to the urban railway literature. As the proposed model is able to optimally calculate solutions within seconds, it can also be used by dispatchers during real-time disruptions. In keeping with this, the model may be integrated as part of a dispatcher support system, assisting with real-time decisions.

7.2 Sources of error

Although the findings suggest that urban rail operations can efficiently be rescheduled through crossovers, it is important to note that the applicability of the proposed model is highly dependent on the characteristics of the system. In comparison to busy metro systems, Bybanen has considerably longer frequency intervals. Since the results suggest that increased frequency cause reduced punctuality, crossover operations are likely to be more challenging for metro systems. That being said, for metro systems with double-tracked crossovers, these procedures might prove to be efficient.

The validity of the results greatly depends on the assumptions made when modelling the network. Although optimization models with macro perspective have proven to be efficient in the past (Chang et al., 2019; Yin, Tang, et al., 2017), they create uncertainties as trains cannot always be located with exact precision. Nor can incident location be determined with high level of granularity. Although these considerations are included in terms of safety margins, it may be easier for dispatchers with more exact locations.

In the proposed model, headway for meeting operations were moreover determined by a proxy to ensure safe train operations. This was necessary following lacking information of crossover operations, likely to overestimate headways. Before implementing the model in real-world instances, it would therefore be necessary to define crossovers with higher granularity. This entails exact acceleration and deceleration when using crossovers, as well as the exact reverse time when operating passed it.

In the current model, train services terminate when reaching their final stations. This is a reasonable assumption when demonstrating applicability, however, not sufficient in real-world instances. During normal operations, train services operate in loops by changing operating direction at final stations. Rescheduling of looping train operations is therefore likely to cause different optimal solutions. These operations also have to consider delay when returning to origin station, compared to the proposed model where only one direction is considered. A probable outcome would be for affected and unaffected trains to take turns in passing the incident, resulting in better distribution of delay between inbound and outbound trains. Overall punctuality would most likely be reduced in short term perspective but be more evenly distributed throughout the day.

7.3 Development of an efficient dispatcher support system

Since the proposed optimization model can optimally reschedule trains within seconds, it demonstrates potential for an integrated dispatcher support system. At the current state, however, the model is most suited to increase knowledge through scenario analyses and stress testing of the network. It is also suitable when planning future line extensions and determining characteristics of the infrastructure. Thus, the model could be used as a basis for knowledge to expand to double-tracked crossovers.

In order to develop an integrated system, data should first of all be specified and generated in the same interface. This would make it easier for dispatchers to process data, thereby decreasing decision-making time. A possible solution could be an integrated application where the main page is used to specify input. The resulting output could be presented through a user-friendly interface consisting of real-time train positions, as well as suggested rescheduling operations. In addition, relevant train graphs and displays for safe interaction between trains could be presented.

The application should also include necessary software to run the optimization model, including R, AMPL and CPLEX. As AMPL and CPLEX both require licensing fees, a cost-benefit analysis should therefore be conducted before a system is developed. However, the costs are not likely to be substantial compared to alternative support systems.

In order to develop the network of Bybanen, an integrated system could include sensors throughout the light rail system, to determine the exact location of an incident. This could provide a more autonomous system, reducing the necessary tasks by dispatchers. Sensors could also be used to provide real-time information to passengers when crossover procedures are initiated, informing of when and where trains will arrive. Informing passengers of train arrivals is especially crucial as the optimization model reschedules trains to arrive at platforms in the opposite direction. The information system should further ensure clear communication not only between dispatchers and operators, but also between operators at different train services. This is necessary to ensure that operators adjust runtimes corresponding to the rescheduled timetables provided by dispatchers, ensuring safe and reliable train services.

8 Conclusions

The purpose of this thesis was to develop an optimization model that efficiently reschedules trains during disruptions. This was achieved through a mixed-integer optimization model that minimizes sum of delay at all stations while respecting safety considerations. Following subset reduction of crossover links, the model was able to reschedule trains within short computation time.

The numerical analyses were performed on three case studies from Bybanen light rail system, during various disruption timeframes. The results suggest that the proposed optimization model can safely reschedule train operations through crossovers located at their actual position in the network. Minimizing delay at all stations may also contribute to a more even distribution of punctuality compared to minimizing delay at final stations. These combined findings are highly relevant for comparable urban rail systems as utilization of crossovers rarely have been researched or tried in practice.

With increasing rolling stock in the network, implementing a crossover strategy is problematic due to increased frequency. Our findings further indicate that introducing double-tracked crossovers could significantly improve urban rail operations in terms of punctuality and service level. Consequently, similar urban rail systems could utilize the proposed model to optimally plan infrastructure and crossover sections.

In the proposed optimization model, delay was minimized with an assumption of equal passenger demand at all stations. In reality, there are considerable differences in demand depending on where a station is located. Further development of the model could include minimized delay at stations weighted according to expected passenger demands. To improve service level, the model could also include train operations turning at the terminal stations, as well as higher granularity of crossovers. Refining the model with these conditions may contribute to more efficient crossover operations and evenly distributed delay. In future research it would be interesting to evaluate the effect of adding back-up trains to reduce delay for passengers succeeding the incident area. This could increase the frequency of train operations during disruption, thereby possibly improving service levels.

9 References

AMPL. (2019). AMPL R API. Retrieved from https://ampl.com/api/nightly/R/index.html

Bergen County. (2019a). Bybanens fem byggetrinn. Retrieved from https://www.bergen.kommune.no/hvaskjer/tema/reguleringsplaner-for-bybanen/bybanens-fem-byggetrinn

Bergen County. (2019b). Klimabudsjett. Retrieved from https://pub.framsikt.net/2020/bergen/bm-2020-hop20-23/#/generic/summary/climatesummary

Bergen County. (2019c). Rekordlav luftforurensing i 2018. Retrieved from https://www.bergen.kommune.no/hvaskjer/tema/luftkvalitet/nyhetsarkiv/rekordlav-luftforurensing-i-2018

Bybanen. (2019a). *Bybanen annual report 2018*. Retrieved from https://www.bybanen.no/wp-content/uploads/2019/07/Bybanen_årsberetning_2018.pdf

Bybanen. (2019b). PRAKT. Retrieved from https://www.bybanen.no/vi-er-bybanen/prakt/

Chang, Y., Niu, R., Wang, Y., Luan, X., D'Ariano, A., & Sama, M. (2019). *Train Rescheduling for an Urban Rail Transit Line under Disruptions.* Paper presented at the 8th International Conference on Railway Operations Modelling and Analysis - RailNorrk¨oping 2019.

Corman, F., & Quaglietta, E. (2015). Closing the loop in real-time railway control: Framework design and impacts on operations. *Transportation Research Part C: Emerging Technologies, 54*. doi:10.1016/j.trc.2015.01.014

D'Ariano, A. (2009). Innovative Decision Support System for Railway Traffic Control. *Intelligent Transportation Systems Magazine, IEEE, 1*, 8 - 16. doi:10.1109/MITS.2010.935910

D'Ariano, A., Corman, F., Pacciarelli, D., & Pranzo, M. (2008). Reordering and Local Rerouting Strategies to Manage Train Traffic in Real Time. *Transportation Science, 42*, 405-419. doi:10.1287/trsc.1080.0247

D'Ariano, A., & Pranzo, M. (2009). An Advanced Real-Time Train Dispatching System for Minimizing the Propagation of Delays in a Dispatching Area Under Severe Disturbances. *Networks and Spatial Economics, 9*, 63-84. doi:10.1007/s11067-008-9088-1

D'Ariano, A., Pranzo, M., & Hansen, I. A. (2007). Conflict Resolution and Train Speed Coordination for Solving Real-Time Timetable Perturbations. *Intelligent Transportation Systems, IEEE Transactions on, 8,* 208 - 222. doi:10.1109/TITS.2006.888605

D’ariano, A., Samà, M., D’ariano, P., & Pacciarelli, D. (2014). Evaluating the Applicability of Advanced Techniques for Practical Real-time Train Scheduling. *Transportation Research Procedia, 3*(C), 279-288. doi:10.1016/j.trpro.2014.10.007

Eberlein, X. J., Wilson, N. H. M., Barnhart, C., & Bernstein, D. (1998). The real-time deadheading problem in transit operations control. *Transportation Research Part B, 32*(2), 77-100. doi:10.1016/S0191-2615(97)00013-1

Engebretsen, Ø., Christiansen, P., & Strand, A. (2017). Bergen light rail – Effects on travel behaviour. *Journal of Transport Geography, 62*, 111-121. doi:10.1016/j.jtrangeo.2017.05.013

Gao, Y., Kroon, L., Schmidt, M., & Yang, L. (2016). Rescheduling a metro line in an over-crowded situation after disruptions. *Transportation Research Part B, 93*(PA), 425-449. doi:10.1016/j.trb.2016.08.011

Gao, Y., Yang, L., & Gao, Z. (2017). Real-time automatic rescheduling strategy for an urban rail line by integrating the information of fault handling. *Transportation Research Part C, 81*, 246-267. doi:10.1016/j.trc.2017.06.005

Hangfei, H., Keping, L., & Paul, S. (2018). Real-time energy-saving metro train rescheduling with primary delay identification. *PLoS ONE, 13*(2), e0192792. doi:10.1371/journal.pone.0192792

Hordaland County Council. (2014). *Klimaplan for Hordaland 2014-2030: Regional klima- og energiplan*. Retrieved from https://www.hordaland.no/globalassets/for-hfk/plan-og-planarbeid/regionale-planar/a4_klimaplan14-30_web-bokmerke-og-navigasjon.pdf

Johnson, K. (2001). *Physics for you : revised national curriculum edition for GCSE*. Cheltenham: Nelson Thornes.

Louwerse, I., & Huisman, D. (2014). Adjusting a railway timetable in case of partial or complete blockades. *European Journal of Operational Research, 235*(3), 583-593. doi:10.1016/j.ejor.2013.12.020

Lundgren, J., Rönnqvist, M., & Värbrand, P. (2010). *Optimization*. Lund: Studentlitteratur.

Miljøløftet. (2019a). Bybanen til Fyllingsdalen. Retrieved from http://miljoloftet.no/prosjektliste/kollektiv/bybanen-til-fyllingsdalen/

Miljøløftet. (2019b). Bybanen til Åsane. Retrieved from https://miljoloftet.no/prosjektliste/kollektiv/bybanen-til-asane/

Mæland, P. A. (2019). Om ti år er Bybanen førerløs. Retrieved from https://www.bt.no/nyheter/lokalt/i/kJPXPj/om-ti-aar-er-bybanen-foererloes

Norwegian Government. (2011). *Avtale mellom Samferdselsdepartementet og Hordaland fylkeskommune og Bergen kommune om belønningstilskudd til bedre kollektivtransport og mindre bilbruk 2011 – 2014*. Retrieved from https://www.regjeringen.no/contentassets/66644bf4b3e642acaf10bea324af42b8/bergen-20111-2014.pdf

Norwegian Government. (2012). *Meld. St. 21 (2011–2012) Norsk klimapolitikk*. Retrieved from https://www.regjeringen.no/no/dokumenter/meld-st-21-2011-2012/id679374/sec1

Norwegian Government. (2017). *Byvekstavtale mellom Bergen kommune, Hordaland fylkeskommune og Staten v Samferdselsdepartementet og Kommunal- og moderniseringsdepartementet*. Retrieved from https://www.regjeringen.no/contentassets/66644bf4b3e642acaf10bea324af42b8/byvekstavtale-bergen-2017-2023.pdf

Norwegian Government. (2019). Belønningsordningen, bymiljøavtaler og byvekstavtaler. Retrieved from https://www.regjeringen.no/no/tema/transport-og-kommunikasjon/kollektivtransport/belonningsordningen-bymiljoavtaler-og-byvekstavtaler/id2571977/

Pellegrini, P., Marlière, G., Pesenti, R., & Rodriguez, J. (2015). RECIFE-MILP: An effective MILP-based heuristic for the real-time railway traffic management problem. *IEEE Transactions on Intelligent Transportation Systems*, *16*, 1-11. doi:10.1109/TITS.2015.2414294

Pellegrini, P., Marlière, G., & Rodriguez, J. (2014). Optimal train routing and scheduling for managing traffic perturbations in complex junctions. *Transportation Research Part B*, *59*(C), 58-80. doi:10.1016/j.trb.2013.10.013

Pellegrini, P., Marlière, G., & Rodriguez, J. (2016). A detailed analysis of the actual impact of real-time railway traffic management optimization. *Journal of Rail Transport Planning & Management*, *6*(1), 13-31. doi:10.1016/j.jrtpm.2016.01.002

Pellegrini, P., Pesenti, R., & Rodriguez, J. (2019). Efficient train re-routing and rescheduling: Valid inequalities and reformulation of RECIFE-MILP. *Transportation Research Part B*, *120*, 33-48. doi:10.1016/j.trb.2018.12.008

RStudio. (2019). About RStudio. Retrieved from https://rstudio.com/about/

Sairong, P., Xin, Y., Hongwei, W., Hairong, D., Bin, N., Haichuan, T., . . . Ruijun, T. (2019). Dispatching High-Speed Rail Trains via Utilizing the Reverse Direction Track: Adaptive Rescheduling Strategies and Application. *Sustainability, 11*(8), 2351. doi:10.3390/su11082351

Samà, M., D'Ariano, A., Pacciarelli, D., Pellegrini, P., & Rodriguez, J. (2018). Applications of train routing selection methods for real-time railway traffic management. 72-77. doi:10.1109/ITSC.2018.8569555

Samà, M., D′Ariano, A., Corman, F., & Pacciarelli, D. (2017). A variable neighbourhood search for fast train scheduling and routing during disturbed railway traffic situations. *Computers and Operations Research, 78*, 480-499. doi:10.1016/j.cor.2016.02.008

Samà, M., Pellegrini, P., D'ariano, A., Rodriguez, J., & Pacciarelli, D. (2016). Ant colony optimization for the real-time train routing selection problem. *Transportation Research Part B, 85*, 89-108. doi:10.1016/j.trb.2016.01.005

Shakibayifar, M., Sheikholeslami, A., Corman, F., & Hassannayebi, E. (2017). An integrated rescheduling model for minimizing train delays in the case of line blockage. *Operational Research*. doi:10.1007/s12351-017-0316-7

Skyss. (2015). Kollektivmeldinga 2014. Retrieved from https://www.hordaland.no/globalassets/for-hfk/rapportar-og-statistikk/andre-rapportar---pdf/kollektivmeldinga2014_150dpi_2.pdf

Skyss. (2016). KOLLEKTIVSTRATEGI FOR HORDALAND ÅRSRAPPORT 2015. Retrieved from https://www.skyss.no/globalassets/strategiar-og-fagstoff/statistikk-og-fakta/kollektivstrategi-arsrapport-2015-oppslag.pdf

Skyss. (2017). KOLLEKTIVSTRATEGI FOR HORDALAND ÅRSRAPPORT 2016. Retrieved from https://www.skyss.no/globalassets/strategiar-og-fagstoff/strategiar-og-handlingsprogram/arsrapport/endelig-arsrapport-2016-enkle-sider.pdf

Skyss. (2019a). Bybanen Rutetabell. Retrieved from https://www.skyss.no/globalassets/bussruter/hostruter-2019/bergen-sentrum/1_bybanen.pdf

Skyss. (2019b). KOLLEKTIVSTRATEGI FOR HORDALAND ÅRSRAPPORT 2018. Retrieved from https://www.skyss.no/globalassets/strategiar-og-fagstoff/strategiar-og-handlingsprogram/arsrapport/skyss_arsrapport_2018_08042019.pdf

SSB. (2019a). KOLLEKTIVTRANSPORTUTVIKLING 2010-2018: Mer kollektive byer. Retrieved from https://www.ssb.no/transport-og-reiseliv/artikler-og-publikasjoner/mer-kollektive-byer

SSB. (2019b). Statistikkbanken - befolkning. Retrieved from https://www.ssb.no/statbank/table/01222/

Törnquist, J., & Persson, J. A. (2007). N-tracked railway traffic re-scheduling during disturbances. *Transportation Research Part B, 41*(3), 342-362. doi:10.1016/j.trb.2006.06.001

UITP. (2012). *Metro, light rail and tram systems in Europe*. Retrieved from http://www.uitp.org/sites/default/files/cck-focus-papers-files/errac_metrolr_tramsystemsineurope.pdf

UITP. (2015). *Light rail in figures. Statistics brief. Worldwide outlook*. Retrieved from https://www.uitp.org/sites/default/files/cck-focus-papers-files/UITP_Statistic_Brief_4p-Light%20rail-Web.pdf

UITP. (2016). *Knowledge brief. Light rail: A tool to serve customers and cities*. Retrieved from https://www.uitp.org/sites/default/files/LRT_Knowledge%20Brief_LQ.pdf

UITP. (2018a). *Commuter railway landscape*. Retrieved from http://www.uitp.org/sites/default/files/Commuter%20Teaser.pdf

UITP. (2018b). *New urban rail infrastrcture*. Retrieved from https://www.uitp.org/sites/default/files/cck-focus-papers-files/Statistics%20Brief%20-%20NEW%20URBAN%20RAIL%20INFRASTRUCTURE%202018.pdf

UITP. (2018c). *World metro figures*. Retrieved from https://www.uitp.org/sites/default/files/cck-focus-papers-files/Statistics%20Brief%20-%20World%20metro%20figures%202018V4_WEB.pdf

UITP. (2019). *The global tram and light rail landscape*. Retrieved from http://www.uitp.org/sites/default/files/cck-focus-papers-files/Statistics%20Brief%20-%20World%20LRT_web.pdf

Xu, X., Li, K., & Yang, L. (2016). Rescheduling subway trains by a discrete event model considering service balance performance. *Applied Mathematical Modelling, 40*(2), 1446-1466. doi:10.1016/j.apm.2015.06.031

Yin, J., Tang, T., Yang, L., Xun, J., Huang, Y., & Gao, Z. (2017). Research and development of automatic train operation for railway transportation systems: A survey. *Transportation Research Part C, 85*, 548-572. doi:10.1016/j.trc.2017.09.009

Yin, J., Wang, Y., Tang, T., Xun, J., & Su, S. (2017). Metro train rescheduling by adding backup trains under disrupted scenarios. *Frontiers of Engineering Management, 4*, 418. doi:10.15302/J-FEM-2017044

Zhan, S., Kroon, L. G., Zhao, J., & Peng, Q. (2016). A rolling horizon approach to the high speed train rescheduling problem in case of a partial segment blockage. *Transportation Research Part E, 95*, 32-61. doi:10.1016/j.tre.2016.07.015

10 Appendix

A Station names

Abbreviation	Full names
BYP	Byparken
NON	Nonneseter
BYS	Bystasjonen
NYG	Nygård
FLO	Florida
DAP	Danmarks plass
KRS	Kronstad
BRS	Brann stadion
WER	Wergeland
SLE	Sletten
SLB	Slettebakken
FAN	Fantoft
PAR	Paradis
HOP	Hop
NST	Nesttun terminal
NSS	Nesttun sentrum
SKS	Skjoldskiftet
MAR	Mårdalen
SKJ	Skjold
LAG	Lagunen
RAS	Råstølen
SAV	Sandslivegen
SAM	Sandslimarka
KOK	Kokstad
BIR	Birkelandsskiftet
KOF	Kokstadflaten
FLE	Flesland

B R-file

```
#Installation

#install.packages("Rcpp", type="source")
#install.packages("https://ampl.com/dl/API/rAMPL.tar.gz", repos=NULL)
#-------------------------------------------------------------------------
------
library(rAMPL)
ampl <- new(AMPL, new(Environment, "C:/Users/xx"))

setwd("C:/Users/xx")
# Decide which frequency to use
con <- if (interactive()) stdin() else file('stdin')
message('Which frequency 4min,5min or 6min:')
frequency <- scan(file=con,what = numeric(), nlines=1, quiet=TRUE)

if (frequency == 4){
  ampl$read("close240.mod")
}else if (frequency == 5){
  ampl$read("close.mod")
}else if (frequency == 6){
  ampl$read("close360.mod")
}

if (frequency == 4){
  ampl$readData("240head.dat")
}else if (frequency == 5){
  ampl$readData("300head.dat")
}else if (frequency == 6){
  ampl$readData("360head.dat")
}

#Decide solver cplex
ampl$setOption("solver","cplex")

#If the incident is located in outbound or inbound direction
con <- if (interactive()) stdin() else file('stdin')
message('Incident inbound or outbound:')
incident <- scan(file=con,what = character(), nlines=1, quiet=TRUE)

#Insert incident section, start and end
message('Enter start of incident section: ')
incident.start <- scan(file=con,what = character(), nlines=1, quiet=TRUE)

message('Enter end of incident section: ')
incident.end <- scan(file=con,what = character(), nlines=1, quiet=TRUE)


#Insert incident time and recovery
message('Enter start time of incident: ')
incident.occurs <- scan(file=con,what = integer(), nlines=1, quiet=TRUE)
message('Enter end time of incident: ')
incident.resolves <- scan(file=con,what = integer(), nlines=1, quiet=TRUE)

#If statement with two choices, either outbound or inbound
if (incident=="OUTBOUND"){
  #Find possible crossover links with input as incident section
  RLINKO <- ampl$getSet("RLINKO")
  R <- RLINKO$getValues()
  datalist <- list()
```

```
  for (row in 1:nrow(R)){
    if ((R[row,1]==incident.start) && (R[row,2]==incident.end)){
      x <-R[row,3]
      y <-R[row,4]
      dat <- data.frame(x,y)
      datalist[[row]] <- dat # add crossoverlink to list
    }
  }
  All_crossover_links = do.call(rbind, datalist)

  #Find links within used crossoverlinks, creating a dataframe, and store
as a list
  list <- list()
  for(r in 1:nrow(All_crossover_links)){
    df <- data.frame()
    for (row in 1:nrow(R)){
      if ((R[row,3]==All_crossover_links[r,1]) &&
(R[row,4]==All_crossover_links[r,2])){
        x <-R[row,]
        df = rbind(df,x)
        list[[r]] <- df
      }
    }
  }



  library(openxlsx)
  #Solve for all possible crossover links, storing wait time and used
crossover link
  Wait_list <- list()
  New_times <- list()
  for (row in 1:nrow(All_crossover_links)){
    library(rAMPL) #needs to reset for every loop
    df <- data.frame()
    ampl <- new(AMPL, new(Environment, "C:/Users/xx"))
    setwd("C:/Users/xx»)
    ampl$setOption("solver","cplex")
    # Interpret the two files
    if (frequency == 4){
      ampl$read("close240.mod")
    }else if (frequency == 5){
      ampl$read("close.mod")
    }else if (frequency == 6){
      ampl$read("close360.mod")
    }

    if (frequency == 4){
      ampl$readData("240head.dat")
    }else if (frequency == 5){
      ampl$readData("300head.dat")
    }else if (frequency == 6){
      ampl$readData("360head.dat")
    }
    #The lines over resets so need to specify sets again
    #Determine correct values for sets and parameters
    ampl$getSet("SO")$setValues(incident.start)
    ampl$getSet("EO")$setValues(incident.end)
    ampl$getParameter("too")$setValues(as.integer(incident.occurs))
    ampl$getParameter("tro")$setValues(as.integer(incident.resolves))
```

```
    ampl$getSet("LINKSFO")$setValues(All_crossover_links[row,]) #determine
new set
    new_RLINKO = list[[row]]
    ampl$setData(new_RLINKO,4, "RLINKO") #determine new set
    ampl$solve() #solve
    total_wait_time <- ampl$getObjective("total_wait_time")
    #Present variables
    Var <- ampl$getVariables() #Get variables
    Outbound_times <- cbind(Var$ao1$getValues(),Var$do1$getValues()[,-1:-
2])
    names(Outbound_times)[names(Outbound_times)=="index0"] <- "Train"
    names(Outbound_times)[names(Outbound_times)=="index1"] <- "Station"
    names(Outbound_times)[names(Outbound_times)=="ao1.val"] <- "Arrival
Outbound"
    names(Outbound_times)[names(Outbound_times)=="Var$do1$getValues()[, -
1:-2]"] <- "Departure Outbound"

    Inbound_times <- cbind(Var$ai1$getValues(),Var$di1$getValues()[,-1:-2])
    names(Inbound_times)[names(Inbound_times)=="index0"] <- "Train"
    names(Inbound_times)[names(Inbound_times)=="index1"] <- "Station"
    names(Inbound_times)[names(Inbound_times)=="ai1.val"] <- "Arrival
Inbound"
    names(Inbound_times)[names(Inbound_times)=="Var$di1$getValues()[, -1:-
2]"] <- "Departure Inbound"
    cat(sprintf("Objective is: %g\n", total_wait_time$value()))
    New_times[[row]] <- cbind(Outbound_times,Inbound_times)
    Wait_list[[row]] <-
cbind(total_wait_time$value(),All_crossover_links[row,])
    #Make excel files
    if (frequency == 4){
      wb <- loadWorkbook('Excel240.xlsx')
    }else if (frequency == 5){
      wb <- loadWorkbook('Excel300.xlsx')
    }else if (frequency == 6){
      wb <- loadWorkbook('Excel360.xlsx')
    }
    total_wait_time <- total_wait_time$value()
    crossover <-
paste(All_crossover_links[row,1],All_crossover_links[row,2])
    writeData(wb,sheet = 1,New_times[[row]],startCol = 1, startRow = 1,
colNames = TRUE)
    saveWorkbook(wb, as.character(paste(total_wait_time,
crossover,".xlsx")),overwrite = TRUE)
    } #ends with storing total wait time and the used crossoverlink

  print("Look at Excel-files for the Rescheduled times")
} else if (incident=="INBOUND"){
  #Find possible crossover links with input incident section
  RLINKI <- ampl$getSet("RLINKI")
  R <- RLINKI$getValues()
  datalist <- list()
  for (row in 1:nrow(R)){
    if ((R[row,1]==incident.start) && (R[row,2]==incident.end)){
      x <-R[row,3]
      y <-R[row,4]
      dat <- data.frame(x,y)
      datalist[[row]] <- dat # add crossoverlink to list
    }
  }
  All_crossover_links = do.call(rbind, datalist)
```

```
  # Find links within used crossoverlinks, creating a dataframe, and store
as a list
  list <- list()
  for(r in 1:nrow(All_crossover_links)){
    df <- data.frame()
    for (row in 1:nrow(R)){
      if ((R[row,3]==All_crossover_links[r,1]) &&
(R[row,4]==All_crossover_links[r,2])){
        x <-R[row,]
        df = rbind(df,x)
        list[[r]] <- df
      }
    }
  }

  library(openxlsx)
  #Solve for all possible crossover links, storing wait time and used
crossover link
  Wait_list <- list()
  New_times <- list()
  for (row in 1:nrow(All_crossover_links)){
    library(rAMPL) #needs to reset for every loop
    df <- data.frame()
    ampl <- new(AMPL, new(Environment, "C:/Users/xx"))
    setwd("C:/Users/xx")
    ampl$setOption("solver","cplex")
    # Interpret the two files
    if (frequency == 4){
      ampl$read("close240.mod")
    }else if (frequency == 5){
      ampl$read("close.mod")
    }else if (frequency == 6){
      ampl$read("close360.mod")
    }

    if (frequency == 4){
      ampl$readData("240head.dat")
    }else if (frequency == 5){
      ampl$readData("300head.dat")
    }else if (frequency == 6){
      ampl$readData("360head.dat")
    }
    #The lines over resets so need to specify sets again
    #Determine correct values for sets and parameters
    ampl$getSet("SO")$setValues(incident.start)
    ampl$getSet("EO")$setValues(incident.end)
    ampl$getParameter("toi")$setValues(as.integer(incident.occurs))
    ampl$getParameter("tri")$setValues(as.integer(incident.resolves))
    ampl$getSet("LINKSFI")$setValues(All_crossover_links[row,]) #determine
new set
    new_RLINKI = list[[row]]
    ampl$setData(new_RLINKI,4, "RLINKI") #determine new set
    ampl$solve() #solve
    total_wait_time <- ampl$getObjective("total_wait_time")
    #Present variables
    Var <- ampl$getVariables() #Get variables
    Outbound_times <- cbind(Var$ao1$getValues(),Var$do1$getValues()[,-1:-
2])
    names(Outbound_times)[names(Outbound_times)=="index0"] <- "Train"
    names(Outbound_times)[names(Outbound_times)=="index1"] <- "Station"
```

```
    names(Outbound_times)[names(Outbound_times)=="ao1.val"] <- "Arrival
Outbound"
    names(Outbound_times)[names(Outbound_times)=="Var$do1$getValues()[, -
1:-2]"] <- "Departure Outbound"

    Inbound_times <- cbind(Var$ai1$getValues(),Var$di1$getValues()[,-1:-2])
    names(Inbound_times)[names(Inbound_times)=="index0"] <- "Train"
    names(Inbound_times)[names(Inbound_times)=="index1"] <- "Station"
    names(Inbound_times)[names(Inbound_times)=="ai1.val"] <- "Arrival
Inbound"
    names(Inbound_times)[names(Inbound_times)=="Var$di1$getValues()[, -1:-
2]"] <- "Departure Inbound"
    cat(sprintf("Objective is: %g\n", total_wait_time$value()))
    New_times[[row]] <- cbind(Outbound_times,Inbound_times)
    Wait_list[[row]] <-
cbind(total_wait_time$value(),All_crossover_links[row,])
    #Make excel files
    if (frequency == 4){
      wb <- loadWorkbook('Excel240.xlsx')
    }else if (frequency == 5){
      wb <- loadWorkbook('Excel300.xlsx')
    }else if (frequency == 6){
      wb <- loadWorkbook('Excel360.xlsx')
    }
    total_wait_time <- total_wait_time$value()
    crossover <-
paste(All_crossover_links[row,1],All_crossover_links[row,2])
    writeData(wb,sheet = 1,New_times[[row]],startCol = 1, startRow = 1,
colNames = TRUE)
    saveWorkbook(wb, as.character(paste(total_wait_time,
crossover,".xlsx")),overwrite = TRUE)
  } #ends with storing total wait time and the used crossoverlink

  print("Look at Excel-files for the Rescheduled times")

}
```