Incorporating Weather Impact in Railway Traffic Control

Ying Wang

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

Institute for Transport Studies

October, 2018

Intellectual Property and Publication Statements

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Chapter 4 is based on work published in:

Wang, Y., Liu R. and Kwan, RSK. 2017. Railway rescheduling under adverse weather conditions on a single-track line. *7th International Conference on Railway Operations Modeling and Analysis*, 2017 Lille.

And based on work currently being considered for publication in Transportation Research Part C (responding to second round review comments)

Wang, Y., Liu R. and Kwan, RSK., D'Ariano, A., and Ye, H. Train timetable rescheduling on single-track lines under adverse weather conditions.

Chapter 5 is based on work published in:

Wang, Y., Liu R. and Kwan, RSK. 2016. Railway rescheduling under adverse weather conditions using genetic algorithm. *14th World Conference on Transport Research*, 2016 Shanghai.

Chapter 6 is based on work published in:

Wang, Y., Liu R. and Kwan, RSK. 2018. Simultaneous rerouting and rescheduling on rail networks under weather impact. *Transportation Research Board (TRB) 97th Annual Meeting*, 2018 Washington, D.C.

The original work contained within these papers is all the candidate's own work with guidance provided by Kwan and Liu.

Acknowledgements

All work in this thesis has been undertaken by the candidate under the specific guidance of Professor Ronghui Liu and Professor Raymond Kwan. Their guidance has been encouraging and inspiring and is highly appreciated. Gratitude also goes to Dr. Hongbo Ye and Dr. Zhiyuan Lin who have provided suggestions in modelling and thesis editing.

Special thanks to Network Rail for supplying the data and industry standards and reports. Particularly thanks to Nadia Hoodbhoy, Richard Storer and Rhodri Lloyd in Network Rail and Dr. Ryan Neely in the School of Earth and Environment at University of Leeds for their help in acquiring the data and documents.

I also thank the Chinese Scholarship Council and University of Leeds for their financial support for my tuition fee and living expenses for three years, and thank British Railway Safety and Standard Board and Leeds for Life Award Committee for sponsoring me to attend conferences.

Of course, big thanks to my kind colleagues, supporting staff and tutors of the Institute for Transport Studies, especially Daniel, Deborah, Paul, Jennifer, Tatjana, Peter, Weiming, Tianli, Tamas, Jing, Jeff, Fangqing, Yao and Izza. They always provide unconditional help whenever I need.

Finally, I thank my family. This study would not have been possible without their enormous love and support.

Abstract

Abnormal weather events can have significant impacts on the safety and operational performance of the railways. In Great Britain, weather related train delays run into 1 to 2 million of minutes each year. With the rapid advances in weather forecasting and emerging information technology, the weather forecasting data can be utilised to improve the performance of train control models in dealing with weather events. In this thesis, the forecasted moving weather fronts are map in terms of their temporal and spatial coverage, as well as the corresponding speed restrictions and/or track blockages according to the severity of the weather fronts, onto the railway lines. This enables the control models to consider multiple disruptions in advance of them commencing, instead of dealing with them one by one after they have commenced. Then the proactive train control methods are proposed, i.e. mixed integer liner programming (MILP) and genetic algorithm (GA) for single-track rescheduling in adverse condition, and an MILP model for simultaneous train rerouting and rescheduling model, taking into account forecasted severe weather perturbations. In the models, the forecasted moving weather perturbations on different parts of the rail network are represented as individual constraints, whereby, trains travelling through the adversely impacted zones follow reduced speed limits and in the severely impacted zones where the tracks are blocked, trains need to be rerouted or wait until the blockage disappears. The case studies indicate: a) compared with existing control methods our rescheduling methods have shown to make significant reduction in total train delays (in the case studies examined, an average 21% reduction in delays); b) within the timescale considered, the further ahead the weather forecast information is considered, the less the overall delay tends to be; c) under severe weather disruptions (with track blockage), the proposed rerouting and rescheduling model is shown to be able to effectively and efficiently find a cost effective route and timetable.

Keywords:

Forecasted adverse weather; Railway traffic control methods; Train timetable rescheduling; Simultaneous rerouting and rescheduling; Mixed integer liner programming; Genetic algorithm

Table of Contents

List of Tables

List of Figures

Acronyms

ATM	Air Traffic Management
CWAM	Convective Weather Avoidance Model
E 4-1	Experiment 4-1 using UISPR with MILP method
E 4-2	Experiment 4-2 using PCDPR with MILP method
E 4-3	Experiment 4-3 using a rolling PCDPR with MILP method
E 5-1	Experiment 5-1 using a PCDPR method with GA
FAA	Federal Aviation Administration
GA	Genetic Algorithm
MILP	Mixed Integer Liner Programing
MIP	Mixed Integer Programing
MP	Mathematical Programming
PCDP	Predictable, Compound and Dynamic Perturbations
PCDPC	Predictable, Compound and Dynamic Perturbations Control
PCDPR	Predictable, Compound and Dynamic Perturbations Rescheduling
PCDPRR	Predictable, Compound and Dynamic Perturbations Rerouting and Rescheduling
PIP	Pure Integer Programming
RTC	Railway Traffic Control
TOCs	Train Operation Companies

Chapter 1 Introduction

1.1. Context and motivation

Every year, abnormal weather (including adverse weather and severe weather) causes a massive amount of financial loss for the railway industry. For example, as shown in Figure 1-1, in the Great Britain, the infrastructure manager Network Rail, pays tens of millions of pounds each year to train operating companies for weather related delay and cancellation compensations. From April 2006 to March 2014, the least payment is about £25 millions in year 2011-12, while the most is £95 million in year 2013-14. Wind, snow and flood are the three main hazards contributing to the payment among the listed nine different weather hazards.

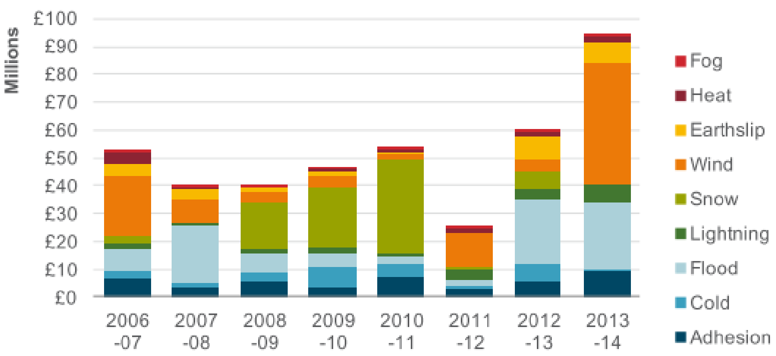


Figure 1-1: Weather attributed compensation payments by Network Rail to train operators between April 2006 and March 2014 (source: Network Rail, 2014).

We conduct further investigation on how the weather hazards lead to delay and cancellation. Table 1-1 lists many potential consequences of weather hazards summarised from Network Rail (2011). As shown in Table 1-1, different weather hazards cause a variety of consequences: flooding may result in obstructions on the line, extremely high temperature may result in railway buckles, etc.. To ensure safety, railway operation authorities mandates detailed mitigation strategies such as emergency speed limitation on tracks where adverse weather will happen, and service suspensions (track blockages) on the tracks where severe weather will happen. Detailed mitigations are reviewed in Section 2.4.2.

Table 1-1: Consequences of weather hazards

Weather Hazards	Consequences
Flooding / High Seas / Heavy Rain	bstructions on the line; scour action; land-slide, slope failure or washout; inundation (flooding), including equipment failure; sea spray; erosion.
High Wind Speeds	overhead line damage; structural damage, including station roofs and canopies; fallen trees (or parts thereof); leaf fall (includes railhead contamination and loss of track circuit detection); shifted load or loose sheeting.
Railhead Contamination	station over-run; low rail adhesion; loss of track circuit detection (Wrong side failures); rail / wheel defects.
Extremes of Temperature	rail buckles; track circuit failures; point failures through loss of detection, especially switch diamonds; overhead line sag; overheating relay rooms.
Thunderstorms / Lightning	failure of electrical and electronic equipment; structural / tree damage; lineside fires.
Fog / Mist / Low level Cloud Cover	signal passed at danger; level crossing collision.
Snow / Hail / Ice / Frost, including Freezing Rain and Freezing Fog	points failures; signal failure; structural / tree damage; ground heave during extended periods of low temperatures; icing of electrical supply equipment, including conductor rails and OLE; icicles, including in tunnels; signal passed at danger; level crossing collision; platforms and walkways covered by snow or ice; track circuit failures at level crossings caused by applications of road salt.
Long Periods of Dry Weather	embankment settlement through internal collapse or shrinkage; lineside fires; fires on land / premises adjoining the railway; fires resulting from the operation of steam locomotives.
Any of the Hazards above	obstruction of the line; stranded trains; severe disruption and delays to train services.

However, these mitigations are neither combined with the existing real-time railway traffic control (RTC) mechanisms in industry, nor in research. In practice, when the local controllers obtain the forecasted abnormal weather information and the mitigation notes, the corresponding speed limitation marks or signals will be placed alongside the railway lines. The train drivers then apply the actions when acknowledging the information. This entire process does not include any proactive control, i.e. rescheduling or rerouting in advance, but includes only reactive responses from the train drivers' side. The speed limits or track blockages will further develop to perturbations propagating in the network and the controllers will start the real-time control when significant delay is observed.

There is a large body of literature on railway traffic control which considering reduce the impact of stochastic perturbations, which weather impact is classified in (Nielsen et al., 2012; Pender et al., 2013;). Such stochastic perturbations are *Unpredictable*, *Independent* and *Static Perturbations* (UISP); in other words, they are unknown or unpredictable in advance and unrelated to each other, and they occur at static time moments and static locations, i.e. their impact areas do not change with time. The impact of each perturbation is usually modelled as a fixed amount of delay to a prescribed set of trains or a fixed track blockage for one period. In response to an UISP event, the railway traffic control (UISPC) model will be triggered only when a delay or a blockage is detected; consequently, the initial delays are irreversible. As such, future perturbation events, such as the forecasted abnormal weather conditions, are not considered in UISPC.

Though in some countries such as the Netherlands, there are thousands of pre-prepared emergency timetables for the condition of severe weather impact, they are not weather specific and the utilisation of railway resources is not optimised.

With the advances in weather forecasting technologies, the forecasting accuracy is very high. The UK Met Office's Global and Regional Ensemble Prediction System (MOGREPS) produces UK weather forecast (on temperature, pressure, wind and humidity) for the next 54 hours on a forecast grid of 2.2km by 2.2km (Golding et al., 2014). The Met Office's most detailed forecast model applies to a 1.5km-by-1.5km grid inner domain (Met Office, 2017). The forecast accuracy is 95.1% for the next-day's temperature, and 93.6% for wind speed, while a 0.562 Equitable Threat Score (a verification index for rain versus no rain) of three-hourly weather is correctly forecasted for rain (MetOffice, 2016).

With the accurately forecasted weather data, weather-related perturbations can be (at least partially) predicted. We can then design train control models which utilise these forecasted perturbations, so that they can further help the controllers make well-advised decisions, which can minimise the delay and the impact globally from both temporal and spatial dimensions.

1.2. Aims and objectives

Instead of controlling trains after the occurrence of weather events, rescheduling and rerouting can be conducted in advance of these events while taking their impact into account. In this thesis, we consider adverse weather impacts to be: (a) accurately *predictable* for the rescheduling horizon, e.g. the next day or the next 3-4 hours depending on the availability and accuracy of the weather forecasting data; (b) *compound* in the sense that they have both time and space dimensions and can be considered as a group; and (c) highly *dynamic* as the impacted areas can change with time as the weather fronts move, and the trains being affected can vary with the rescheduling plan. In summary, with an accurate weather forecast, disturbances due to abnormal weather can be treated as *Predictable, Compound* and *Dynamic Perturbations* (PCDP).

Though weather hazards may result in tens of different consequences, according to the industry mitigation guidance, the mitigations are of two types: emergency speed limitations and track blockage. We map the PCDP into train control models, so that a train scheduled to go through a speed limitation zone would follow the reduced speed limit of that zone. While a train scheduled to go through the track blockage zone, would have to wait until the track blockage is cleared or to divert to other tracks.

By treating the weather impacts as PCDP, we propose new proactive railway traffic control (PCDPC) methods in rescheduling and rerouting to mitigate the weather-related perturbations to train services. In particular, we focus on a) a rescheduling model which deals with adverse weather conditions whose impact is merely speed restrictions, so that the system delay could be minimised; b) a simultaneous rerouting and rescheduling model which deals with severe weather conditions whose impact is both track blockages and speed restrictions, so that the controllers could alter the train route considering the different penalty cost of using backup lines and of using tracks that normally serve trains travelling in opposite directions; and c) an efficient algorithm so that for large networks the model can give feasible solutions in short computation time.

To clarify, the models are based on the following assumptions and simplifications:

a) The weather forecast is accurate. This is justified by the significantly improved technology in atmospheric modelling and weather forecasting, which is briefly described in Section 1.1.
b) Crew and rolling stock scheduling is considered separately and outside the scope of this research, and is assumed always rectifiable after any railway rescheduling.
c) For simplification, the railway is assumed to operate a "moving-block" signalling system under which each track segment can be used by any number of trains as long as they can maintain the safety headways between each other, and the average speed is used for calculating travel time on segments.

1.3. Contributions

The contributions of this thesis are listed below:

First, we propose the concept of predictable, compounded and dynamic perturbations (PCDP) for forecasted abnormal weather and propose a method to model weather impact from temporal and spatial dimensions.

Second, we adopt weather forecasting data in an optimal rescheduling model and propose a novel and proactive train rescheduling method dealing with adverse weather. This model can deal with multiple weather impacts happening at different locations at the same time and different levels of weather impacts happening at the same place at different time.

Third, to improve the computation efficiency for PCDP, we design a modified GA. We introduce a new concept, i.e. a conflict resolution matrix (chromosome) in which the solution for each potential conflicted of train pair is an element (gene).

Fourth, we propose a simultaneous rerouting and rescheduling MILP model with the modified track occupation constraints so that the unidirectional track could change to bidirectional tracks and be temporarily used by the opposite trains under severe weather impact. This can help the controllers make better routing decisions considering the penalties of using backup lines and of borrowing tracks from opposite trains.

1.4. Outline

The thesis is organised as follows.

In Chapter 2, we first review the railway service planning process. Second, we review the existing control models including reactive and proactive control, control objectives, rescheduling and rerouting methods, and the approach for modelling perturbations. By comparing the strengths and limitations of the existing research, we build our models in the later chapters.

Chapter 3 considers the weather as predictable perturbations in the railway system and focuses on transferring the weather data into train control models. Specifically, we introduce suitable infrastructure models, including assumptions made in the models. Then we introduce the control models considered in this thesis, including the control actions, constraints and solution approaches. This is followed by the introduction of how we incorporate the weather data into the railway system control models. It includes the analysis of differences between UISP and PCDP, how the weather is abstracted from temporal and spatial dimensions, and the justification methods of train speed. The chapter ends with an illustrative case study of comparing UISP and PCDP.

In Chapter 4, we consider MILP formulation for single-track rescheduling under weather-related emergency speed restrictions. The most related single-track rescheduling models are first reviewed, followed by the mathematic formulation considering weather constraints. Several experiments are conducted to test the performance of PCDP rescheduling (PCDPR) and UISP rescheduling, the rolling PCDPR with partial information, and the sensitivity of the PCDPR model.

Chapter 5 proposes a GA to efficiently solve the PCDPR problem. The genetic formulation and the process of updating chromosomes are introduced after the literature review of the GA. This is followed by the rescheduling processes with each chromosome. A case study is conducted to compare GA and the MILP regarding their solution quality and computation efficiency.

Chapter 6 proposes a simultaneous rerouting and rescheduling model considering the severe weather impact which causes not only speed restrictions, but also track blockages. We first introduce the train rerouting and rescheduling problem. Then we formulate this problem as MILP with improved track occupation constraints. Finally, we apply this model on both a small and a large network to demonstrate how this model will help controllers in making route choice decisions.

Finally, Chapter 7 summarises the work in this thesis, highlights the contributions and discusses the potential future research directions.

Chapter 2 An overview of railway traffic planning and control

The complex railway system is composed of elements such as infrastructure, movable devices and crews that interact in a regulated manner to deliver services to passengers and freight. In an average railway system, there are hundreds if not thousands of train services on any given day that use these elements in planned orders. To ensure efficiency and safety, railway authorities plan and manage railway services at various levels. This chapter will introduce the different levels of railway planning as well as the terminologies and models that are relevant to understanding the research problems and our methodologies.

Though worldwide railway systems obey universal physical limitations, there are some differences in infrastructure requirements, protocols, terminology definitions, etc. In this Chapter, we will first introduce the major important terminologies used in this research in Section 2.1. Then we introduce the three main planning processes, i.e., strategic level, tactical level, and operational level in Section 2.2; these will help to identify the research scopes in this thesis. In Section 2.3, we introduce the benchmarks for this research, i.e. the existing railway traffic control models, which aim to reduce delays or prevent conflicts. In Section 2.4, we introduce the weather impact on the railway and the operational guidance for weather events, which shows the practical needs and possibilities for considering weather in the traffic control model in Chapter 3. In Section 2.5, we reviewed the weather impact in other transportation modes, such as aviation and ground traffic.

2.1. Terminology

To avoid confusion, the key terminologies used in this thesis are defined below (Hansen and Pachl, 2014).

- Points: in this thesis stand for location points without specific instructions. They are the general terms for physical stations, loops, junctions in macroscopic level as well as joints and switches in microscopic level.

- Lines: refer to different meanings according to the context. i) They can refer to tracks between points. In this sense, lines are divided into two classes: running lines and sidings. Running lines are the tracks on which trains move through the network, including main lines and side lines. Sidings are lines used for assembling trains, storing vehicles and trains, loading and unloading, and similar purposes, but not for regular train movements. ii) A railway line which refers to the entire line consists of stations and segments between stations and provides the complete railway services.
- Nodes: are representations of arbitrary locations in a railway network modelling. In the macroscopic model, they represent the railway stations, loops, and junctions, while in the microscopic model, they can represent the switches on the tracks.
- Edges: are arbitrary non-directional representations of running lines, sidings or tracks in railway models.
- Links: are directional connections between two nodes in railway models.
- Capacity: The maximum number of trains that can be run through a certain area (station or open line) in a given period of time.
- Perturbations: include all the abnormal events that cause, or potentially cause, delays in the rail system. Perturbations are further categorised into disturbances and disruptions (Cacchiani et al., 2014) .
    - Disturbances: a disturbance happens when certain railway processes (e.g., moving from one station to another, or dwelling in a station) last longer than specified in the timetable. As a consequence, trains may depart and/or arrive later than planned. This can be handled by rescheduling the timetable only, without rescheduling the resource duties.
    - Disruptions: a disruption is a relatively large external incident, strongly influencing the timetable, and requiring the resource duties to be rescheduled as well. A disruption may be caused by a temporary blockage of the railway infrastructure, for example by malfunctioning infrastructure or rolling stock, or by an accident. Due to a blockage, a number of trains may incur large delays, or a number of trips in the timetable must be cancelled.
- Railway traffic control: is to minimise the negative impact of perturbations by adjusting railway services according to real-time conditions. The general control

actions include: detouring, cancellation, skipping stations, rerouting, rescheduling, etc. In this thesis, we consider only rerouting and rescheduling, which are defined as follows (Meng and Zhou, 2014) .

- Rescheduling: includes (1) changing arrival and/or departure times, and (2) changing arrival and/or departure orders.
- Rerouting: includes (1) using a different track, and (2) using a different route on a network.

2.2. Rail service planning process

To identify the research scope, the railway service planning process will be firstly reviewed. Most railway authorities divide planning and management of the railway service into three levels: strategic, tactical and operational (Hansen and Pachl, 2014). The strategic level is mainly concerned with matching the traffic demand with service supply, e.g., deciding how many services, tracks and rolling stocks are needed to cover the target demand. At a tactical level, the infrastructure is usually fixed, while the movable resources are adjusted in terms of quantity, quality and intensity of operation. Operational level covers pre-operations resource allocation and operations management.

2.2.1. Strategic level planning

The strategic level is also known as “advanced timetable development and capacity planning”. It typically includes two main activities: network design and line planning, projecting 5 to 15 years ahead.

- Network design consists of the construction of new or change of current railway infrastructure, due to changes in travel requirements, increased or decreased demand, and implementation of new technologies or standards. The relevant authorities, e.g., government and railway operators normally have different objectives at this level (Hooghiemstra et al., 1999). Because the construction or revision of infrastructure costs millions of pounds, the design will be in execution and revision for several years, before being established and approved.
- Line planning consists of designing train lines that are defined as itineraries between two designated stations and some intermediate stations traversed by

trains. The frequency, desired schedules of the trains, and types of the required rolling stock are also defined in this procedure. The quality indicators for a line plan are direct connections between lines, total travel time for passengers, and so on.

2.2.2. Tactical level planning

Upon the start date of an annual railway service plan, there are five main tasks: maintenance planning, timetabling, capacity allocation, rolling stock planning and crew scheduling. These need to be performed in a period of five years.

- Maintenance planning: It plans maintenance activities needed to maintain operation; this includes all preventive maintenance activities and time slots reserved for possible corrective maintenance activities. A maintenance activity is normally a set of actions performed for retaining or restoring a system, or an item, so that it can perform its required function, modifying and constructing new infrastructure also can be part of maintenance planning.
- Timetabling: This is known as the Train Timetabling Problem or Train Scheduling Problem, and has attracted wide attention in research. In this activity, each Railway Undertaking (RU) submits their desired schedule to the Infrastructure Manager (IM), and the IM is responsible for solving any possible incompatibilities and then producing an integrated timetable which meets all RU’s requests.
- Capacity allocation: In this activity, track routes and station platforms are allocated to each train according to their schedules. In case of allocation impossibility, the IM will negotiate with RUs for other options. It is also known as track allocation problem, train routing problem, train path allocation problem and, in some cases, train platforming problem.
- Rolling stock planning: This consists of finding and making assignments of rolling stock to the scheduled services and also includes scheduling of empty rides and shunting movements. The objective is normally to minimise the number of vehicles, or the total cost, necessary to meet the requirements of the timetable.

- Crew scheduling: RUs are responsible for their crew scheduling, i.e., generating crew duties for each of their train services at minimal cost, with the precondition of meeting all work regulations and operational requirements.

2.2.3. Operational level planning

The Operational level concerns short-term plans (normally between the start of an annual service plan and several days before the operation day) for unexpected requirements and management during daily operations. The operational level mainly deals with the abnormal events that cause perturbations to the timetable during daily operation. In a dense timetable, any perturbation like a signal failure, or severe weather can easily cause significant delay to services. It may hinder the subsequent services that are scheduled over the same resource, e.g., railway infrastructure, rolling stock, or crew, or may cause conflicts between trains. The operational-level planning includes two main types of activities: pre-operations resource (re)allocation and operations management.

- Pre-operations resource (re)allocation could be triggered by short-term supply or demand changes, known from several months to several days before the operation day. The examples include extra passenger services due to sports events or changes in the crew rotations due to strikes.
- Operations management is responsible for overseeing, managing and coordinating the train traffic in daily operation. With unforeseen events, such as signal or infrastructure failure or abnormal weather that occurs within the railway system, the operation management team is required to provide feasible solutions to avoid conflicts or reduce delays by rescheduling or rerouting trains or even cancelling services. However, the computing time is normally very limited due to the real-time feature. Much of the literature has focused on developing advanced methodologies to improve the solutions and reduce the computation time to deal with these unforeseen events, and these methodologies will be reviewed in Section 2.3.

The RTC is applied in the operational control level. We further propose a new way with proactive traffic management and control to reduce delays caused by one of the special events, abnormal weather. More specifically, we introduce a new procedure at the operation level by taking account of improvements in weather forecasting technology. This would allow rescheduling one day, or even several hours before the operation day in case of predicted adverse weather impact. Related to this, the existing research on weather impact to railway operation will be reviewed in Section 2.4 and the proposed method to convert the weather data into a suitable format for timetable rescheduling will be explained in Section 3.3.

2.3. Railway traffic control models

As introduced in Section 2.2.3, there might be many unexpected events, including weather impact, during daily operations. Railway traffic control aims to minimise the negative impact of them by adjusting railway services according to real-time conditions. It is also referred to as Train or Railway Dispatching Problem, Railway Dynamic Traffic Management, or Railway Traffic Control in the literature (Corman and Meng, 2015).

Assuming the following information are known: the topological structure and the physical characteristics of a railway network, the set of train routes and associated passing/stopping times at each relevant point in the network, and the position and speed of trains at the given starting time, RTC is defined as meeting the following requirements (Hansen and Pachl, 2014).

a. Solve all potential conflicts between trains;
b. Does not result in deadlock situations (trains that are all waiting for each other, making any planned movement impossible);
c. Compatible with the initial positions of all trains;
d. The selected train routes are not blocked;
e. The speed profiles are acceptable;
f. No train appears in the network before its expected entrance time (including the entrance delays);
g. No train departs from a relevant point before its scheduled departure time; and
h. Train arrives at the relevant points with the smallest possible knock-on-delay.

There are varieties of actions to adjust the services while satisfying the above conditions, such as detouring, cancellation, skipping stations, reordering, rerouting, rescheduling, and so on. In this thesis, we consider the most common RTC methods: train rescheduling and rerouting.

In this subsection, the reactive and proactive RTC will be reviewed first in Section 2.3.1. The way of modelling perturbations is reviewed in Section 2.3.2, which is also a fundamental difference between the UISP and PCDP. Then the two control actions used in this thesis, i.e. rescheduling and rerouting will be reviewed in Section 2.3.3 and Section 2.3.4, respectively. Last but not the least, the control objective is reviewed in Section 2.3.5.

2.3.1. Reactive and proactive railway traffic control

Railway traffic control approaches are further distinguished between *reactive approaches*, which do not take account of the future traffic conditions when making decisions and *proactive* approaches, which take account of the perturbations and the prognosis of future statuses of the network.

Most of the current operational traffic management methods are mostly reactive, which UISP belongs to. In reactive control, local traffic controllers (called "controllers" hereafter) can update orders and routing decisions within an area of limited geographical size (called "dispatching area" hereafter). Traffic controllers have very limited knowledge of the current status of the railway network, mostly limited to the block section where the train is at present. They have no precise information on the train's position, speed or acceleration. For this reason, dispatchers can only update the plan when a considerable delay has accumulated.

In proactive control, each train driver receives an advisory travel time or speed to maintain. Proactive traffic management requires the following:

a. Precise monitoring of current train positions;
b. Predicting train speed profiles or running times in a defined geographical area and for a defined time window;
c. Detecting the effects of perturbations to train traffic conflicts;
d. Rescheduling trains in real time, so that consecutive delays are minimised, by adjusting orders, routes set, times; and
e. Communicating the advisory location-time-speed targets to train drivers.

Proactive control is commonly used, for example, in maintenance activities. The proposed control methods for weather-induced PCDP fall into this category.

2.3.1.1. Reactive control

In some studies on reactive control, a perturbation is represented as a single delay occurred to one of the trains in the timetable (Chen et al., 2010; Corman et al., 2014; Larsen et al., 2014; Tornquist and Persson, 2007). For example, Tornquist and Persson (2007) proposed a dispatching approach for an n-track network when a disturbance occurred on one train.

On the other hand, other research considered perturbations either on all trains passing one prescribed failure location, or directly on a set of prescribed trains (D’Ariano et al., 2007; D’Ariano, 2008; Pellegrini et al., 2015).

Jacobs (2004) designed the asynchronous disposition method of asynchronous traffic regulation to identify and resolve conflicts on large sub-networks, which can produce conflict-resolution proposals or suggest a new train-regulating schedule. The method will act when the deviations are detected and once for all.

Törnquist (2012) designed an effective algorithm for fast dispatching under disturbances. It considered three different types of disturbances: (i) a single train with a certain delay at one section; (ii) a train having a “permanent” malfunction resulting in increased running times on all line sections it is planned to traverse; and (iii) a speed-limit reduction on a certain section, which results in increased running times for all trains running through that section. Among them, case (iii) is most similar to weather impact, i.e. the perturbation is modelled on one fixed location. But the delay applied to all the trains passing that location, i.e. all these delays are not able to be reduced by dispatching.

In the reactive control, disturbances are considered as UISP, which cause fixed initial delays on specific trains that cannot be reduced or eliminated by dispatching no matter whether the perturbations are stochastic or deterministic, and the perturbations in the control time window are not considered in the dispatching procedure.

2.3.1.2. Proactive control

In the proactive category, some of the literature deals with perturbations during daily operations. Boccia et al. (2013) described two heuristic approaches to solve the optimal real time train dispatching problem, based on a MIP formulation. The disruption cases considered are intervals called maintenance of way windows defined by the start and end points and the time. The objective function is to minimise the weighted cost of delay on each edge, the deviation for trains at nodes required adherence and terminals, time spent on unpreserved edges.

Dollevoet et al. (2017) proposed an iterative rescheduling framework considering timetable, rolling stock and crew, which led to an overall feasible solution for all resources. They first used a timetable rescheduling method in Veelenturf et al. (2016) to get an optimised timetable for the rolling stock composition capacity; the objective was to minimise the total duration of cancelled train services. Second, they used an approach in Nielsen et al. (2012) to allocate rolling stock compositions to trips many as possible. If any trips are not covered by compositions, they will be cancelled and it goes back to a timetable rescheduling process to generate a new timetable. Third, new duties are assigned to crew members following Veelenturf et al. (2012) with the objective of covering as many tasks as possible. Again, if any trips are not covered in this step, it will go back to timetable rescheduling processes to generate a new timetable and start a new iteration.

Another big branch of proactive control is track maintenance scheduling. The literature on railway maintenance can be classified into three types: (i) scheduling both trains and maintenance events (Albrecht et al., 2013; Forsgren et al., 2013); (ii) scheduling maintenance while considering fixed train schedules as constraints (Cheung et al., 1999; Higgins, 1998; Jardine et al., 2006; Lake et al., 2010; Peng et al., 2011; Santos et al., 2015); and (iii) scheduling trains while considering fixed maintenance schedules as constraints (Diego, 2016). Type (iii) is similar to the scenario of interest in this thesis, but it attracts very little attention in the literature.

Specifically, Lidén and Joborn (2017) allowed the trains to pass the maintenance work sites with a reduced speed limit, and optimised the schedules of maintenances and trains simultaneously; however, the meet/pass constraints for conflict avoidance were not considered.

Diego (2016) scheduled the trains with several fixed maintenance activities which can lead to both track closure and reduced speed limits. This consideration is similar to weather impact to railway. They claimed they designed the first microscopic model in the literature to tackle maintenance while considering specific factors such as temporary speed limitations. In their model, a same track segment can be impacted by several maintenance activities in different time slots (similar to the impact of different adverse weather events), but these different time slots have to follow the same level of reduced speed limits. This is however not the situation of the adverse weather: even for the same location, different adverse weather events may lead to different levels of speed limits, due to the different types and severity of the weather events.

In the existing literature, the information considered in proactive approaches includes part or all of: current status of infrastructure, train positions and speeds, precise prediction of delay characteristics and expected time of future events. The events here mean arrival or departure of trains at certain key points, such as stations and junctions.

To our best knowledge, the weather forecasting data is not considered in any real-time control models. This research considers predictable future disturbances associated with tracks (and over the predicted time periods) rather than with any specific trains, and the trains disturbed are not fixed but determined by the control decision, in that some of the initial effects on trains can be avoided by active control. Moreover, the proposed algorithms in this thesis are aiming at obtaining the optimal control plan taking account of all the forecasted disturbances. We will further introduce how the weather forecasting data is abstracted to the railway system in Chapter 3 and how it is built in the control models in Chapter 4, Chapter 5 and Chapter 6.

2.3.2. Modelling perturbations in railway traffic control

Among the literature, two types of perturbation are mainly considered, which are single-train perturbation and multiple-train perturbations. In this section, the focus will be on how the existing perturbations are considered.

2.3.2.1. Single-train perturbation

Some researchers focus on dealing with a single perturbation in the whole network and some of them generated their test cases by one specific distribution. Corman et al. (2014) tested their model with 50 delay cases generated by Weibull distributions which fitted to the historical data of real-life operations. Chen et al. (2010) tested their model by some specific trains' arrival delays generated by a normal distribution. Larsen et al. (2014) generated 1,000 Monte Carlo trials to evaluate the susceptibility of optimal train schedules. These trials are the result of, for example, measurement errors, coarse-grained train detection data, additional unexpected disturbances, overcrowding, and additional delays at station platforms.

In other times, the disturbances were generated randomly without any specific distribution. Corman and Quaglietta (2015) designed a framework to reproduce the interactions between an automatic rescheduling tool and railway operations under random disturbances such as the unplanned extension of trains' running times and/or dwelling times at stations. Tornquist and Persson (2007) considered one random train malfunctioning temporarily while studying disturbance propagation and rescheduling algorithms during disturbances. Yang et al. (2010) applied the stochastic-length-disturbances on the leading train when studying the movement model of a group of trains.

2.3.2.2. Multiple-trains perturbations

The rescheduling algorithms for single-train perturbation might not be suitable for multiple-train perturbations. Törnquist (2012) designed a greedy rescheduling algorithm and applied it to three different categories of disturbances. Among these categories, the infrastructure failure is most related to weather impact and leads to increased running times for all trains running through the affected section. Corman et al. (2014) also studied the scenario of speed reduction for a railway section of about 10 km under adverse weather condition. Although the impact was described as happening on the tracks, it was transferred directly to all trains, i.e. all the trains pass through this impacted section were delayed, and the corresponding delays are modelled and solved sequentially.

The dispatching support tool Railway traffic Optimisation by Means of Alternative (ROMA) reacts to various types of disturbance, such as multiple delayed trains and dwell time perturbations (Corman et al., 2010; D'Ariano, 2008; D'Ariano et al., 2007). In these studies, the disturbances to trains either happen at the same track location and apply to all the trains passing that place, or directly happen on multiple trains, i.e. the affected trains and their initial delay were determined and will not be changed by rescheduling decisions. In other words, all the disturbances are considered as UISP, which are modelled as fixed delays on specific trains, no matter whether the disturbances are stochastic or deterministic, and the future disturbances are not considered in the rescheduling procedure.

2.3.3. Rescheduling models

The two most commonly-used rescheduling models are mathematical programming (MP) models and simulation-based models. The former formulates and solves the rescheduling problems via MP and can obtain an optimal solution but might be difficult to compute. The latter is driven by simulation based on either uniformly sampled time points (Zhou and Mi, 2013) or discrete events such as a train's arrival or departure at a station (Li et al., 2008). We will review these two categories and their pros and cons to choose a suitable method underpin this research.

2.3.3.1. Simulation models

Simulation offers a powerful method for modelling the complex operation of a railway system and the dynamic interactions among train scheduling, railway signalling and speed controls, and train movements. Discrete time models and discrete event models are the two main branches of the simulation models for scheduling (Zhou and Mi, 2013).

In discrete time models, the time span is divided into equal-length intervals. Caimi et al. (2012) propose a closed-loop discrete-time control framework for a fixed-block railway network rescheduling. The evolution of rail system is first forecasted by operational data from the physical layer, then the potential resource conflicts are detected and resolved by the forecasted results, and at the last stage the control loop is closed by forwarding disposition decisions to the physical layer. Yang et al. (2010) consider the discrete time model under the stochastic disturbance condition, where the train will perform the braking operation when a stochastic disturbance occurs.

In discrete event models, train movements are driven by events such as one train's arrival at a station. Dorfman and Medanic (2002; 2004) proposed a local greedy travel advance strategy driven by the earliest event in the coming period. In their research, a capacity check algorithm is used to prevent deadlock. Based on these researches, Li et al. (2008; 2014) proposed an advanced travel advance strategy that considered the network global information. Moreover, they introduced additional events (i.e. acceleration and deceleration) and proposed a less conservative deadlock check algorithm

The simulation models mimic the real world operation, but the rescheduled results are not necessarily optimal, as the decisions are made according to the limited local information in each time step. In the next subsection, we will review the mathematical programming model, which is able to achieve optimal solutions.

2.3.3.2. Mathematical programming models

Pure integer programming (PIP) and mixed-integer programming (MIP) are two commonly used MP models in railway rescheduling. In the IP models, the decision variables such as the priority of two trains, connection maintenance, sequences of trains and the assignment of resources are represented as binary variables, while the departure, arrival and delay times are non-binary integers but are commonly represented as discrete time intervals (see a review in Fang et al., 2015). Schachtebeck and Schöbel (2010) proposed an IP formulation of the delay management problem, which determines which train is allowed to pass the track first.

In the MIP models, the departure, arrival and delay times are continuous decision variables, while the binary decision variables are similar to those in the IP model (Fang et al., 2015).

Higgins et al. (1996) proposed a non-linear MIP model for a single-track line to minimise the train delays and train operating costs. Their model structure and constraints were quite clear and used by many later researches. Li et al. (2014) improved the model in Higgins et al. (1996) with more constraints, such as loading and unloading constraints and station capacity constraints. The station capacity constraints effectively prevented deadlock and capacity shortage in the stations, which made the rescheduling model more realistic. The key differences between the above two pieces of work and this thesis are listed in Appendix.

Using a case study based on the Dutch Railway network, Narayanaswami and Rangaraj (2013) studied the single-track rescheduling algorithm under disturbance. In their model, a disturbance is modelled by its location and time of occurrence and incorporated in a MIP model. However, only one disturbance was considered in their paper and the model was not able to deal with multiple disturbances happening at different times and locations.

2.3.3.3. Heuristic solution algorithms

The mathematical programming models can be solved using standard solvers to deliver optimal results. However, for large-scale rescheduling problems, this approach may be computationally challenging and time consuming. To address this issue, many researchers designed heuristic approaches, such as greedy algorithm, tabu search algorithm, genetic algorithm, customised algorithms for special problems, and so on.

Cai and Goh (1994) designed a greedy algorithm to tackle conflicts between trains running on the opposite directions and on the same direction. For each conflict, if stopping one train was less costly than stopping another, the algorithm will choose to stop the former. The computational results showed that the method could deliver a feasible solution very quickly though it is not globally optimal.

Higgins et al. (1997) compared several heuristics in a single line scheduling problem. Among a local search heuristic, a genetic algorithm, a tabu search algorithm and two hybrid algorithms, the genetic and hybrid algorithms were able to generate near optimal solution for at least 90% of the test cases when computation time is not limited.

Boccia et al. (2013) designed fix routes heuristic and fix trains heuristic for a MILP multitrack territories problem. In the first heuristic, they assigned the most promising route for each train before invoking the MILP solver. While in the second heuristic, they fixed the routing variables and meeting variables of the solution in the previous round to the next round of the MILP model. Computational results showed that their algorithms are better than other approaches.

Mladenovic et al. (2016) combined three classes of heuristics: bound heuristic, which is to limit the domains of decision variables and the objective function to increase the search efficiency; separation heuristic, which aimed to separate and simultaneously schedule only activities which affect each other; and search heuristics which is corresponding to find a feasible solution for real-time train rescheduling problem.

Xu et al. (2018) designed a genetic algorithm for last passenger train delay management. The objectives were maximising connecting passengers and minimising average waiting time. A chromosome was composed of actual section running time and dwell time of the last train of each line.

The probability of explorating the search area in GA is comparable to other heuristic algorithms. It could also use parallel computing to save computation time. Besides, Higgins et al. (1997) had showed the genetic algorithm had a higher chance to get a better result than the tabu search method for small to median scale train rescheduling problems. In Chapter 5, we apply GA as a way of efficiently generating feasible solutions for our proposed train rescheduling problem.

2.3.4. Rerouting models

Carey (1994) developed an optimisation model which can assign not only departure and arrival times, but also platforms and tracks to trains. He proposed to decompose the complex network into a set of subnetworks in order to reduce the complexity. Each of the sub problems consisted of pathing one train while fixing the sequence of all already pathed trains on all links, but not the times. The advantage was the number of binary variables would not increase if more trains are introduced in the subproblems but they do not intersect with the current pathing train.

Caimi et al. (2004) proposed two algorithms for finding train routes through railway stations for a given timetable. Based on an independent set model, the first algorithm searched for a feasible solution by using a fixed-point iteration method. Then the second algorithm amended the initial solution so that the time interval during which a train can arrive is increased. Results showed that the arriving time interval is doubled, so that the routes are more robust.

Corman et al. (2010) investigated the effectiveness of a tabu search scheme using different neighbourhood searching strategies for train rerouting: (1) directly reroute the train with the largest consecutive delay, (2) reroute another train j which has precedence on train I and contributes to its delay, and (3) anticipate the arrival time of train j at the conflict point with train i by rerouting another train k which has precedence on j before the conflict point with i. Experiment showed that for small instances, the new tabu search algorithms are able to find optimal solutions. For large instances, the solutions generated by the new algorithms after 20 seconds of computation are up to more than 15% better than those achieved within 180s by the previous methods.

Mu and Dessouky (2011) proposed MIP formulations for both FixedPath and FlexiblePath models in scheduling freight trains on complex networks. The FlexiblePath model might achieve a significant reduction of total delay compared to the FixedPath model. However, the computation time increased significantly due to the increase of additional binary variables regarding nodes occupation of each train, the sequence of each train pairs in non-predetermined nodes, as well as arrival and departure times at each non-predetermined node.

In order to reduce the computational time while maintaining the solution quality, the following four heuristic algorithms were proposed. (1) LtdFlePath, similar to FlexiblePath, but only reasonable candidate paths were allowed. (2) Genetic+Fixed Path, used genetic algorithm to evolve the population of the candidate paths to generate better results based on the Fixed Path. (3) Decomp algorithm, which included horizontal and vertical decompositions, decomposed the network into several smaller sections and cluster trains into groups. (4) Parallel algorithm, similar to the Vertical decompositions, trains were firstly decomposed into clusters, however each cluster was independent from each other, so that all the sub-problems were be solved in parallel.

For moderate size networks, the GA+Fixed Path algorithm generates the best schedules which balance the quality of the solution and the computational time; while for larger networks, the Decomp algorithm performs the best.

Most of the existing researches consider rescheduling and rerouting separately and thus can only get local optimum. Meng and Zhou (2014) developed an innovative simultaneous rerouting and rescheduling model for the multiple-track train dispatching problems. The route choices and arrival and departure timings are combined by several groups of constraints so that the system optimal result could be generated. We follow their principles when the rerouting is involved. The differences between the above work and this thesis are also listed in Appendix.

2.3.5. Control objective

The objective of railway traffic management is to improve performances of running traffic. Minimise total delays, for all or some specific trains/stations, are mostly used as the objective. For example, Chigusa et al. (2012) tried to minimise the passengers' arrival delay time at their destinations.

Some other researchers use the delay cost as an optimal objective. Andersson (2014) built the delay cost formula of each train t as the sum cost of each passenger type, which is the product of the delay of train t in hours, the number of passengers onboard train t, share of passenger type α at train t in percent and the value of reliability for passenger type α. In Li et al. (2014), the weight of delay costs is related to the train type and the proportion of the delay time among the total travel time. While Gatto et al. (2004) only weighted delay costs by the passenger number of each train path. In these delay costs optimisation models, the values of delay minutes are not connected very comprehensively to the real world.

In addition to the delay time and delay cost, Sato et al. (2013) proposed to use the inconvenience to passengers as the optimal objective, which consists of the travelling time on board, the waiting time at the platforms and the number of transfers. A bi-objective conflict detection and resolution algorithm were studied in Corman et al. (2012), in which minimising train delays and missed connections are objectives.

Different objectives will result in different rearranged timetables, even though the original timetable and perturbations are the same. As our research is to propose a new way of considering weather impact, the objective function will not affect our result in illustrating the effectiveness of the new models, as long as we use the same objective function in the new methods and the benchmark methods. We choose the most commonly used total delay at the destination as the objective in the rescheduling model, and the weighted total cost in the rerouting model considering the penalties of delay at the destination and of using alternative tracks and opposite tracks.

2.4. Weather impact on railway and operational guidance for weather events

2.4.1. Research regarding weather impact

Academic studies on weather impact to the railways have so far been limited to statistical analysis of the causal factors influencing train operation performances under different types and severity of weather conditions.

By analysing major disruptive events on the Dutch railway network between 2011 and 2013 and the historical weather data, Yap (2014) concluded that vehicle breakdowns, switch failures and signal failures occurred significantly more frequently during snowy days than in normal winter days; the frequencies of the above-mentioned three types of disruptive events on a snowy day are 11, 46 and 11 per week, respectively, compared to 5, 4.7 and 4.4 per week respectively on regular winter days.

Xia et al. (2013) analysed the role of weather condition to 424,768 disruptions to infrastructure in the Netherlands from 2001 to 2008. The result shows that: (a) train cancellations are almost always due to disruption in railway infrastructure; (b) train punctuality is negatively impacted by snow, falling leaves on tracks, high temperatures and large variation in temperature; and (c) cancellations and punctuality are both directly and indirectly impacted by gusts, precipitation and low temperatures.

Brazil et al. (2017) analysed the impact of weather conditions on the performance of metropolitan commuter rail in the Dublin Area Rapid Transit. They found that rain was the primary factor for poor train performance. Interactions between wind and rain, as well as that between wind/rain conditions and the month in which a journey took place, were also observed to be significant and resulting in delays to services.

The literature indicates the bad weather (snow, high temperature, gusts, low temperature, rain, etc.) has big impacts to railway systems in many other countries in addition to the UK. The primary weather factors impacting the system might be different in different regions and periods.

2.4.2. Operational guidance for weather events

The rail industry generally provides synthesised guidelines to reduce weather impact. Jaroszweski et al., (2014) suggested long-term planning and short-term actions which can be implemented before, during and after a weather event. These include actions for improving the resilience of physical infrastructure to specific weather conditions, learning from past events and dealing with affected passengers. During service disruption, Virgin Trains (2015) provided guidance to all key staff on their roles and in particular providing additional support and information to customers. However, none of these guidelines deal with the timetable adjustment.

The Secretary Delay Attribution Board (2015) in Britain has developed a range of "delay code guidance" for various weather conditions, in the form of flow charts indicating the organisations to involve and the actions to take. Here we introduce the guidance for extreme high wind (Table 2-1) and high temperature extreme conditions (Table 2-2 and Table 2-3) as examples.

Table 2-1: Actions triggered by wind conditions (source: Network Rail, 2014)

Element	Wind Speed	Action
Wind 1	Forecast of gusts up to 59mph	No action
Wind 2	Forecast of gusts from 60mph to 69mph (not sustained)	Be aware of the possibility of 'Wind 3' being reached
Wind 3	Forecast of frequent gusts from 60 to 69mph (sustained over 4 hours+)	50 mph speed restriction for all trains in the affected Weather Forecast Area
Wind 3	Forecast gusts 70mph or over	50 mph speed restriction for all trains in the affected Weather Forecast Area
Wind 3	Forecast gusts 90mph or over	All services suspended in the affected Weather Forecast Area

Table 2-1 shows the mitigations triggered by wind conditions, with the wind gust increase, the speed restrictions becomes lower. When the forecast gusts reaches 90mph or over, all services suspended in the affected Weather Forecast Area, i.e. the tracks are blocked in the area.

Table 2-2 and Table 2-3 illustrate the temperature related mitigations. High temperature may result in track buckles, so when temperature reaches to a certain level, emergency speed restrictions (ESR) will be imposed to the affected sites to avoid derailing. The level of speed restrictions is set according to the calculated critical rail temperature, when the exceptionally hot weather restrictions accrue, the speed restriction will be applied for several hours.

Table 2-2: Mitigations triggered by critical rail temperature[1] (CRT) levels (source: Network Rail, 2013)

CRT level	Watchmen on site[2]	Watchmen not on site[3]
CRT(W)	Watchmen only	Impose 30/60mph ESR[4]
CRT(30/60)	Impose 30/60mph ESR	Impose 20mph ESR[4]
CRT(20)	Impose 20mph ESR	Impose 20mph ESR[4] on affected line and adjoining lines

Notes:

1 For methods of calculation CRT, please refer to Network Rail, 2013.

2 The watchman must be able to continuously observe the length of affected track. Where the watchman cannot do so, or a watchman cannot be provided, then the requirement for 'watchman not on site' shall be applied.

3 If there is no watchman on site, an alternative means of determining the actual rail temperature on site will be required to enable the staged measures described to be applied at the right time.

4 The restrictions shall not be removed until the track has received a visual examination. If 20mph ESR is not imposed on the adjoining lines, the affected line shall be blocked.

Table 2-3: Mitigations triggered by exceptionally hot weather restrictions (source: Network Rail, 2013)

Forecast air temperature	Restriction[1]	Period of restriction[2]
> 36℃	Impose 45/90mph ESR	1200hrs to 2000hrs
> 41℃	Impose 30/60mph ESR	1400hrs to 1800hrs

Notes:

1 These ESRs will be imposed by Route Control offices on the basis of weather forecasts provided on the previous day.

2 These precautions can be reviewed on the actual day and may be withdrawn if forecast air temperatures are not occurring.

When the abnormal weather is expected, a pre-prepared emergency timetable would normally be put in place (The Secretary Delay Attribution Board, 2015).

However, such a generic timetable is usually not directly linked to the specific abnormal weather conditions, and certainly not with the levels of details in terms of the temporal and spatial weather impacts. However, even when specific speed-restriction guidance for abnormal weather events is available (such as that in Table 2-1), the spatial and temporal range of the responding actions are only vaguely defined and often conservative.

In practice, in case of bad weather, the speed limitations and track blockages will be notified to train drivers. These will further develop to perturbations propagating in the network. The controllers then will conduct the real-time railway traffic control when significant delay is observed. These processes wasted the time which could be used to control the perturbations before the delays happened.

The rapid advances in weather forecasting technologies provide high accuracy forecasting. Having an accurately forecasted abnormal weather conditions means knowing “unique” traffic information of a rail network. When such weather information is translated into more localised and timely speed reduction and track blockages information according to the operation guidance and further considered by the railway control models, more railway capacity could be utilised and less delay will occur in the system.

Therefore, efficient control methods, automatically taking into account weather forecasts and the related speed regulations and track blockages in the near future, is beneficial for generating effective timetables to reduce cost under abnormal weather conditions.

2.5. Weather impact to other transport systems

Weather may also result in perturbations to other transport systems. In order to have a wider prospective of the potential methods in managing the weather impact, we also survey the weather impact to other systems, such as aviation system and road traffic system.

2.5.1. Weather impact to aviation system and the corresponding management methods

According to US Federal Aviation Administration (2017), the largest cause of delay in the US National Airspace System is weather, which caused 69% of system impacting delays (> 15 minutes) over the six years from 2008 to 2013.

These delays mainly from the weather impact on both terminals and enroute flights. The weather impact on terminals mainly causes airport capacity reduction, which requires slot scheduling to reduce congestions (Balakrishnan, 2007; Zografos et al., 2017). The impact on enroute flights mainly requires scheduling, so that the flight can avoided the dangerous areas. Compared to railway system, the former corresponds to the capacity utilization in train stations, which beyond the research scope of this paper. While the latter is more closed to the retiming and rerouting of trains in terms of management mechanism and methodologies, which will be mainly focused on in this sub section.

2.5.1.1. The framework of the air traffic management (ATM) integration concept

Flathers et al. (2013) reported integration of weather into air traffic management decision-making processes is the chief goal of the Generation Air Transportation System in US. Federal Aviation Administration (FAA) ATM and weather communities formed several of the core weather integration concepts consisted of four elements:

a) Weather Information, i.e. the sources of most of the meteorological data.
b) Weather Translation, which turns weather information as constraint or threshold through filters such as safety regulations, operating limitations and standard operating procedures.
c) ATM Impact Conversion, which transforms the constraint or threshold into an impact or state change by identifying the individual aircraft making up the projected demand, calculating the aircraft-specific, weather-constrained capacity, etc.
d) ATM Decision Support, which mitigates the impact of weather constraint by taking the impact information and developing solutions.

During these years, the US FAA and the aviation industry were achieved to some degree in these concepts. This is the underpin supporting our proposed framework in modeling weather impact in railway system. The specific adaptations according to the railway features will be introduced in Chapter 3. In next sub section, we will introduce the weather management methodologies on enroute flight.

2.5.1.2. Weather management methodologies for enroute flights

The management of the weather impact on enroute flights normally required to minimise the fuel and time cost subject to a wide range of capacity constraints by airports, airspaces and human factors. When bad weather appeared in air sectors, the pilot will need to decide to deviate or fly through. The dispatchers will need to schedule or reroute other flight when the impact propagated.

In the management literature, some research model the impact of weather as a deterministic forbidden area. Chan, et al. (2007) proposed a method which calculate the accuracy of the model predicting the pilots' decision in deviating or flying through the convective zones.

Campbell and Delaura (2011) extended the scope of the Convective Weather Avoidance Model (CWAM), which provides the likelihood of pilot deviation due to convective weather in a given area, to include low-altitude flights. A database with 309 deviation cases in nearly 1000 encounters with convective weather was studies. The new introduced low altitude CWAM performed better in both accuracy and decisiveness.

Enayatollahi and Atashgah (2018) studied the impact of headwind and tailwind on arrival times to the waypoints for all flights in the terminal area using cellular automata. A 1D array of identical rectangular cells is used to model the arrival phase and the wind is modelled as a change in aircraft speed at each time step. The verification studies showed that the model is 3-15% accuracy with about 2.9 seconds run time for a 2-hour operation.

Lim and Zhong (2018) studied a reroute mechanism based on the cellular automaton model to avoid the prohibited area, restricted area, danger area as well as the bad weather. The model is suitable for dynamic properties of weather, i.e. transiting, growing and deforming.

Some other researches tend to develop the air traffic management models with weather uncertainty using a serious of different methods.

Balakrishnan and Chandran (2014) presented an integer programming approach for solving deterministic large-scale air traffic flow management problems and extended it to stochastic scenarios, i.e. bad weather conditions, which were represented by probabilistic scenario trees. A tree was grouped by a series of continues flight events. An event had some paralleled sub events with a probability to become materialized.

Yang (2018) designed a 4-dimentional strategic air traffic management formulation and solution in dealing with the system uncertainty such as convective weather. In his paper, he used probability distributions to depict the uncertainties of convective weather, and the probabilistic chance constraints to state the impact of convective weather. The trade-off between the safety and operation efficiency is captured by risk tolerance index, which convert the stochastic chance constrained programming model into a deterministic programming model.

These methodologies in the management for enroute flights are not very suit the weather management of railway system as the system structure, capacity constraints are not the same, especially in managing the railway overtaking and meet events between trains. Railway tracks are the essential part and limitations causing the conflicts in the networks. Rail traffic is strictly constrained by the tracks they have to run on (no overtaking and relatively little routing flexibility), while the flights have more flexibility, which can fly in all three dimensions. Besides, rail traffic follows a much stricter timetable than air.

2.5.2. Weather impact to road traffic and the corresponding management researches

Similar to railway and aviation systems, weather will reduce the road capacity, result in lower running speed or sometimes cause deviation to vehicles. In this subsection, we review the researches in weather impact to road traffic and the corresponding management researches.

Some researchers studied the weather impact to traffic flow characteristics of urban transport by empirical data analysis or simulation. Akin et al. (2011) studied the speed and volume in freeway under weather impact using Remote Traffic Microwave Sensor data. They concluded rain may reduce an average of 8 to 12% on vehicular speeds and 7-8% on road capacity; and light snow will result a significant reduction in traffic volume. Snelder and Calvert (2016) first reviewed the researches focusing on quantifying the impact of bad weather conditions and the adaptation measures on road network, such as the reduction amount of rain and snow on capacity and speed. They further conducted a case study of Rotterdam using a combination of models to analyse the most vulnerable links under the local weather impact by simulation.

Some others studied the evaluation of transport networks after disaster happened.

Researchers are not only focused on studying how the weather is impacting the road traffic, but also, actively looking for the mitigation methods to improve the safety.

The intelligent transportation system technology are widely used in providing weather mitigation solutions in the developed countries (Dey et al., 2015). For example, Colorado Department of Transport installed a warning system with sensors measuring the pavement surface friction and information board giving advisory speed and warning messages on State Highway 82 in Snowmass Canyon. It resulted in no winter crash in the first year after the system was functioned (Goodwin, 2012).

Some other researches consider weather impact in providing route guidance. Lin et al. (2015) proposed a dynamic real-time route guidance system which aimed to reduce the traffic efficiency and mitigate traffic congestions. In addition to real-time traffic information, they also consider the bad weather and incidents. The weather is modelled as a factor impacting the state of one specific segment of road, but without the dimension of time.

As stated above, many researchers studied the characteristics of weather impact on road traffic and proposed many methods in improving the road safety against bad weather. We find very limited researches on system wide control methods which produce traffic route guidance to optimise system efficiency. This may be due to the large uncertainty in the system and individual participates do not aiming a network wide system optimal in making their decisions. The origins and destinations of vehicles are not acknowledged by a control centre and other users in the system, so the future status of the system is hard to be predicted to the traffic management authorities. Even by prediction, some advices will be given to vehicles to avoid congestion, individual vehicles may not follow the given traffic guidance exactly. This is because the road traffics are not centrally managed, which means the vehicles are not managed and controlled strictly by a control centre. Moreover, the vehicles have strong flexibility in speed and road choices, they can change running lanes and routs freely according to their preferences.

2.6. Summary

When perturbations occur and cannot be absorbed by the original timetable, RTC will be needed to make the system back to normal or to reduce system delay as much as possible. In this chapter, we first defined the terms involved in RTC, especially the infrastructure modelling terms such as nodes, edges, links as well as the different types of perturbations and the corresponding control methods. Second, we introduced the existing rail service planning processes and identified the research scope of this thesis, which is the RTC at the operational level.

Third, the railway traffic control models, including control types, perturbation modelling methods, rescheduling and rerouting models, solution methods/algorithms and control objectives are reviewed. Different from reactive UISPC, the proposed proactive PCDPC requires considering the perturbations in advance. By comparing the models used for rescheduling, the MIP models proposed by Higgins et al. (1996) and Li et al. (2014) were found most suitable as fundamentals for the PCDPR model, and the GA proposed by Dündar and Şahin (2013) are most suitable to further develop as a heuristic algorithm to get feasible solutions in short time. We further studied the different models in rescheduling and rerouting. Meng and Zhou (2014) proposed the idea of simultaneously rerouting and rescheduling trains, which resulted in less delay than sequential rerouting and rescheduling. Diego (2016) proposed the proactive microscopic model in train timetabling with fixed maintenance activities which lead to track closure and reduced speed limitations. Based on their ideas, we will propose our simultaneous rerouting and rescheduling method which incorporates weather impact.

Fourth, we discussed the existing weather related research and the operational guidance under abnormal weather impact in railway systems. According to the guidance, abnormal weather can be translated into sectional speed restrictions or blockages within defined periods and treated together at the same time by the control process instead of one at a time. This is assumed to have advantages to the control models which will be verified in the following chapters.

Last, we discussed the weather impact in some other transportation modes, i.e. aviation and road. Though the road traffic management and railway traffic control did not share many similarities, weather integration concepts of aviation system proposed in Flathers et al. (2013) underpinned our idea of translating weather data into railway traffic control models. However, the methodology in managing enroute flights is not very suitable for the management of railway system as rail traffic is strictly constrained by the tracks and rail traffic follows a much stricter timetable than air traffic.

Chapter 3 Methodology for modelling disturbances to railway traffic control

As reviewed in Corman and Meng (2015), the existing proactive control methods majorly focus on train maintenance and we seldom find any research considering the future weather as completely known when conducting control. In this chapter, relying on the high accuracy of weather forecasting technology, we consider the weather as predictable perturbations in the railway system so that the system delay could be minimised globally from both temporal and special dimensions. Our research effort in this chapter is focused on transferring the weather data into railway control models, and designing the corresponding weather constraints for rescheduling and rerouting.

This chapter is structured as follows. First, in Section 3.1 we introduce the basic infrastructure models used in this thesis, as well as the corresponding assumptions. Then in Section 3.2 we analyse how the problem is modelled and the elements of control models considered in this thesis. In Section 3.3, we describe how the weather is modelled in this thesis and the main differences to other works. This is followed by an illustrative case study to explain the differences in Section 3.4. Finally, the chapter is summarised in Section 3.5.

3.1. Infrastructure

3.1.1. Characteristics of a railway network

A railway network is a group of railway lines of different directions and destinations connected by stations or junctions. At the functioning level, a railway line is as considered in this thesis, to consist of (i) points, i.e. stations, loops, junctions and switches, and (ii) tracks connecting two adjacent points. Stations and loops are referred to as passing points (Oliveira, 2001). There are other very important features of railway lines, such as signals and blocks. However, as the main objective of this thesis is to compare the different ways of considering weather impact, the features of signalling and block sections are simplified and absorbed by the representation of track/line and run times in this thesis.

Different types of passing points have different functions. Stations are places where trains can stop to be loaded and unloaded, manoeuvre and change crew; at a loop a train can only stop or slow down in order to let another train pass. However, to resolve conflicts between trains, these passing loops will be considered in this thesis as special stations which does not allow passenger boarding and alighting. A passing point has a specified capacity limit in the number of trains it can hold at any one time and a conflict occurs when its capacity is exceeded. We set the value of capacity limit as non-directional and equal to the number of parallel tracks at the passing point. Two trains running in opposite directions or in the same direction are not allowed to occupy the same track at the passing point.

There are two different types of segments between two points: single-track segments which have only one piece of consecutive track between two nodes and multiple-track segments which have parallel pieces of tracks between two nodes. On a single-track segment, trains running in opposite directions need to use the line in the proper order. A conflict can occur when different-direction trains are planned to use the same track segment at the same time. On a multiple-track segment, trains running in opposite directions are separated into different pieces of tracks and the directions are normally fixed for each piece of track. For safety concerns, there are corresponding speed limitations applied to different parts of the segments, which trains must not exceed. If there are only single-track segments between stations (including loops) in one line, we call this single-track line, otherwise we call it multiple-track line.

3.1.2. Infrastructure models

There are three types of models in representing the above-mentioned characteristics of the infrastructure, i.e. microscopic model, macroscopic model and mesoscopic model. Figure 3-1 illustrates the different levels of modelling the infrastructures. The points and tracks are modeled as nodes and (directional or non-directional) links containing different levels of details.

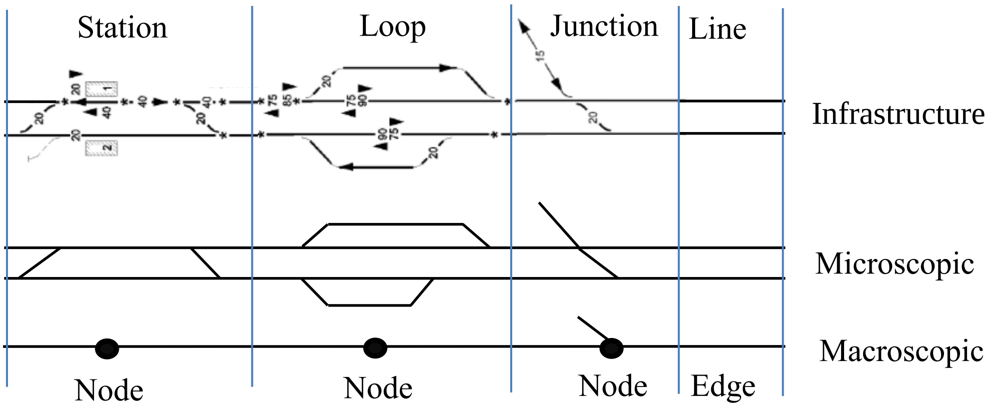


Figure 3-1: Microscopic and macroscopic representation of the railway infrastructure

Microscopic model contains the finest details on characteristics of points and segments, depending on the purpose. A typical microscopic infrastructure model contains all tracks with key information about the segments and the stations. The information includes but is not limited to, speed, gradient, signalling system (such as signals, block sections, release points) and some operational information (such as routes, alternative platforms, timing points and availability at each time stamp). A new node would be needed for any change in one of the attributes to split an existing link and to generate a new one. For example, a station consists of tracks and switches which connect tracks from different directions; in the microscopic models, it would be modelled as a group of nodes and links which represent the switches and tracks respectively.

Macroscopic model contains aggregated information on nodes and links. A station or a junction is represented as a single node (shown on Figure 3-1), which contains the following information: geographic attributes such as ID, coordinates and name, the type of node such as station or junction, and operational information, such as terminal and capacity. The segments are normally modelled as single edges with the length. It contains the information including: types such as passenger or freight, the number of parallel tracks, train availability such as electrification, average running time, and average capacity.

Mesoscopic model is a synthesis of microscopic and macroscopic model, which can contain features of both models according to the tasks. For example, to verify some strategic question, a simplified simulation model for complex networks might be used. For the key part relating to the strategic question, the microscopic representation needs to be applied, but for some other subsidiary part, the macroscopic model is sufficient.

In this research, we adopt a macroscopic representation of network infrastructure for the pure rescheduling model. The features represented include: stations, station capacity, track/segment between stations, and average running time along the tracks. The rescheduled train timetable is represented in terms of the stations to stop, and the arrival and departure times of trains at the stations.

For the problem involving rerouting, we need to model the topological structure of the network to find alternative train routes, so a mesoscopic model is used to describe the railway features. We consider every switch which can lead a train to more than one direction, and consider lines connecting the switches as node and edges. In other words, no matter on the single-track or the multiple-track lines, joints between two consecutive tracks are neglected but the topology in stations, loops and junctions are depicted. An entire segment between two passing points is also described as an edge. Although there are different speed limitations for different parts of a segment, we consider only one entire segment between two junctions and use the average speed to calculate the running time in control models and the average speed limitations as the maximum speed allowed for the entire segment. However, we do not consider details such as gradient, considering computational resource limitation.

3.2. Real-time traffic control models

The timetable is the basis of railway operation. It specifies the starting, passing and ending points of each train service and the associated departure and arrival times at these points. In this way a train performs tasks both when it is waiting for loading or changing crew at a station, and when it is traversing track segments between two stations.

When perturbations happen, the given desired timetable may not be strictly followed, leading to potential conflicts between trains. Conflicts may happen not only between trains running in opposite directions, but also between trains running in the same direction; for instance, when a faster train intends to overtake a slower one on a track segment, or when two trains are applying for the same track in a passing point.

To eliminate the conflicts, trains need to be controlled. In this section, we will introduce the actions and the control constraints in the control model.

3.2.1. Control actions

In practice, a controller might take a series of control actions including choice of time, speed, order, route and service to change the traffic to a certain desired state. In the research of control models, very few address control actions that are different in times, orders, and routes. Speed advice is normally provided in reducing energy consumption and not easy to model and solve within short computational time. Service adjustment normally includes cancelling or short turning trains, or adding or skipping stops, and the objectives are normally the global services quality, which belongs to another branch of control research.

In this thesis, we consider the control actions related to time, order and routes in adjusting train timetable and classify the control into two levels. The first level is rescheduling, in which only time and order of trains are adjusted. In this level, the weather has an adverse impact on railway system but not severe. Speed limitations are applied to some tracks, but no track blockages happen.

The second level is rerouting and rescheduling, in which not only time and order but also routes of trains are adjusted. In this level, the weather has severe impact to railway system, where track blockages happen on some rail tracks accompanying possible speed limitations to some other tracks. In the literature where both rescheduling and rerouting are involved, the majority of them considered rescheduling and rerouting sequentially. Generally, rerouting will be conducted first and followed by rescheduling, although sometimes the reverse process was also used. The drawback of the sequential methods is the possibility of missing out the globally optimal solutions. To address the issue, a simultaneous rerouting and rescheduling method was proposed (Meng and Zhou, 2014), which is also the fundamental of the rerouting and rescheduling model in this thesis.

3.2.2. Basic constraints for traffic control

For simplicity but without loss of generality, we omit some of the infrastructure features in building the traffic control model; the omitted features include track gradients, signals and block sections. We consider the acceleration and deceleration processes implicitly by considering the average speed over the entire segment.

We classify three types of constraints. The first type is constraints must be obeyed, which is due to safety needs, physical limitation, and operation limitation. The second type is constraints which may vary in different operation system. The third type is constraints which is allowed to be broken with certain compensation to allow the global optimisation.

Constraints must be obeyed:

For safety:

1) Speed constraints - These constraints enforce the maximum average speed limit for a vehicle on a particular track segment.
2) Headway constraints at segment - They specify the minimum distance separation between two trains which use the same track segment; they guarantee that, in case the leading train applies the emergency brake, the following train will have enough distance to brake to avoid collisions. This is often worked out as minimum time separation in constraints according the speed limitation.

For physical limit:

3) Conflicts-avoidance constraints - These constraints are for two kinds of purposes. The first guarantees that two trains travelling in the same direction will not overtake one another (headway constraints for following and overtaking). The second enforces that if there are two journeys on a line in opposing directions, these journeys will be conducted in turn (conflict-avoidance constraints between trains traveling on opposite directions).

Constraints that vary in different operation systems:

In real world, operation requirements in stations can vary due to the local restrictions, such as station structure or crew schedule. For the events of trains entering and exiting a station, some operation systems allow several entering and exiting events happening at the same station at the same time, while some can allow only one entering or one exiting event at a time. Similarly, for capacity restrictions on platforms, some stations allow multiple trains stopping on one same platform while some stations can allow only one train at a time. We list the following constraints which are optionally to be considered, according to the different operation systems.

4) Station entry constraints (headway constraints at stations) - Due to the operation restrictions, such as dispatchers or signalling system limitations, the number of trains a station can receive (train entering events) at a time is limited. In some cases, there needs a time headway after receiving a train until a station can receive another one, even the trains are from different directions or tracks. In some other cases, stations can allow multiple trains from different directions or tracks entering at the same time.

5) Station exit constraints - These are similar to the station entry constraints, but now for the exit situation.

6) Station exit & entry constraints - These are similar to the station entry constraints, but now for time headway of a train's entry after another train's exit or a train's exit after another train's entry.

7) Station capacity constraints - The station has limited number of tracks to allow trains loading and unloading. At some railway system, one platform can only be occupied by one train at a time, but in other systems, one platform can hold several trains at the same time.

Constraints that are allowed to be broken with some compensation:

In some of the operation system, the below constraints must not be broken. But in some other system, for the control purpose these constraints can be broken at some penalty cost to prevent a bigger loss. In the latter case, we put these constraints in the objective function with some penalty coefficients and let the program decide whether to break them or not by comparing the cost.

Passenger load

8) Dwell time constraints - These constraints enforce the minimum dwell time for a train at a station.

9) Original timetable constraints - These contains two sub-types of constraints: departure time constraints and arrival time constraints. Departure time constraints limit the rescheduled departure time to being not early than the planned time to ensure passengers will not miss their train. The arrival time of the train cannot exit the departure time plus the minimum running time on the segments.

10) Route constraints - trains need to pass certain key stations, including original and destination stations to load and unload passengers.

3.2.3. Solution approaches

Varieties of solution approaches for MIP models and their frequencies being used are analysed in Fang et al. (2015). These include rule-based approaches (e.g. first come first serve), heuristic approaches (e.g. Greedy), meta-heuristic approaches (e.g. Genetic Algorithm (GA) and Ant Colony Optimisation (ACO)), branch & bound (B&B), and standard solvers (e.g. CPLEX). Among them, B&B and standard solvers can generate optimal solutions. However, both rescheduling and rerouting are NP-hard problems and the computation time will increase greatly with the increasing problem scale. Using rule-based approaches, heuristic approaches or meta-heuristic approaches is a way to find feasible solutions quickly.

For large scale problems, researchers also tried different ways to reduce the solution space so that the computational complexity and solution quality can be balanced. One measure is to decompress the problem by spatial or temporal scale or groups of trains. The second is to fix some of the variables such as train orders and routes. Last but not the least, people also tried to divide the solution approach into different stages, where the first stage obtains a near optimal solution quickly, and the latter stages, improve the solutions.

As we use the mesoscopic model and the decision variables are not in a very big scale, so we will first use the standard solver to generate the optimal solution and compare it with result from the GA, one of the meta-heuristic approaches.

3.3. Incorporating weather forecasting data into the control system

In this section, we illustrate the differences of UISP and PCDP and explain how we transfer the weather data into the constraints into our traffic control models. And give a small example to show the advantage of our method.

3.3.1. A new way of considering weather as perturbations in railway

We consider all the predicted future weather perturbations on tracks rather than on trains, and some of the initial effects on trains can be avoided by active control.

Moreover, the proposed control algorithm is aimed at obtaining the whole time-span globally optimal taking account of all the forecasted disturbances.

The first key difference between UISPC and PCDPC lies in the way the weather perturbation is accounted for. In the traditional UISPC models, weather perturbations are normally modelled as delays on trains and they are unreducible. In our PCDPC model, the adverse weather leads to reduced speed limits or track blockages which vary with space and time. Trains would be affected only when they plan to go through the weather-affected area in the impacted period.

The second key difference is the utilisation of weather information and the time of conducting the control. In UISPC, the weather events are considered sequentially in time. The planner has no “prior knowledge” on or does not consider future perturbations, and control is conducted only when a train is delayed by a perturbation. Figure 3-2 (a)-(c) illustrate under UISPC how the control is conducted repeatedly in response to the occurrence of each perturbation.

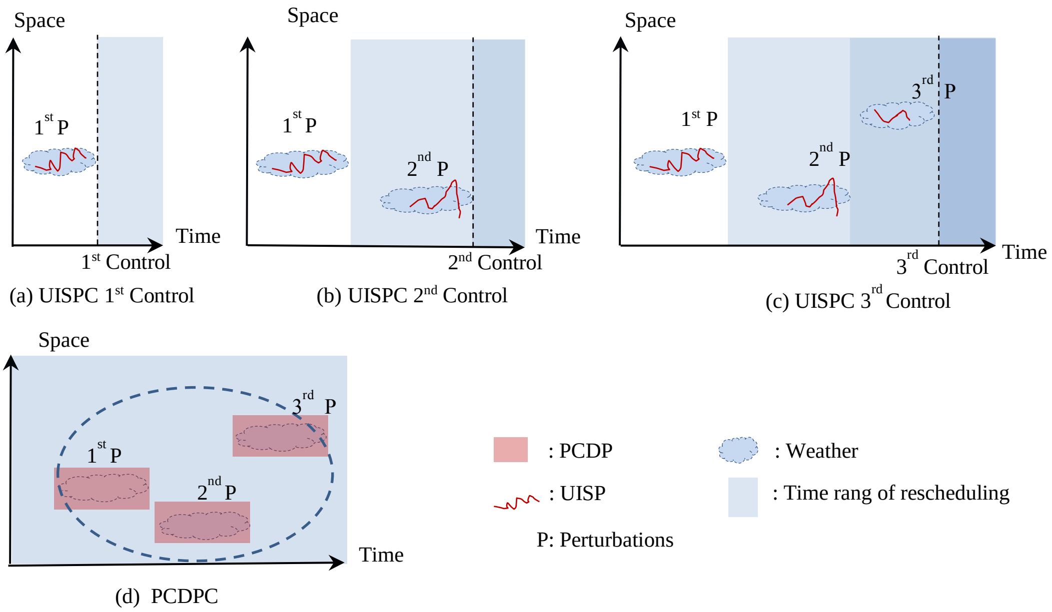


Figure 3-2: Comparison of UISPC and PCDPC: (a)-(c) sequential rescheduling considering one disturbance each time by UISPC; (d) PCDPC

Figure 3-2 (a) shows after the 1st perturbation happens, the 1st control will be conducted, Figure 3-2 (b) represents after the 1st control, the 2nd control happens again, will be 2nd control will be conducted, and the same for Figure 3-2 (c), the control will again be conducted after the third perturbation happens. On the other hand, as shown in Figure 3-2 (d) PCDPC simultaneously considers the spatiotemporal impacts of all the three perturbations over the entire rescheduling horizon in a single control problem.

The procedures of executing the UISPC and PCDPC are further illustrated in Figure 3-3. As shown in Figure 3-3(a), the UISPC is designed to be conducted after the occurrence of initial delays to minimise the subsequent secondary delays. When further external disturbances occur, the rescheduling will be activated again. Therefore, in UISP, initial delays are irreducible, and as the rescheduling is conducted sequentially after each disturbance, the overall delay achieved at the end of the operation period may not be globally minimal.

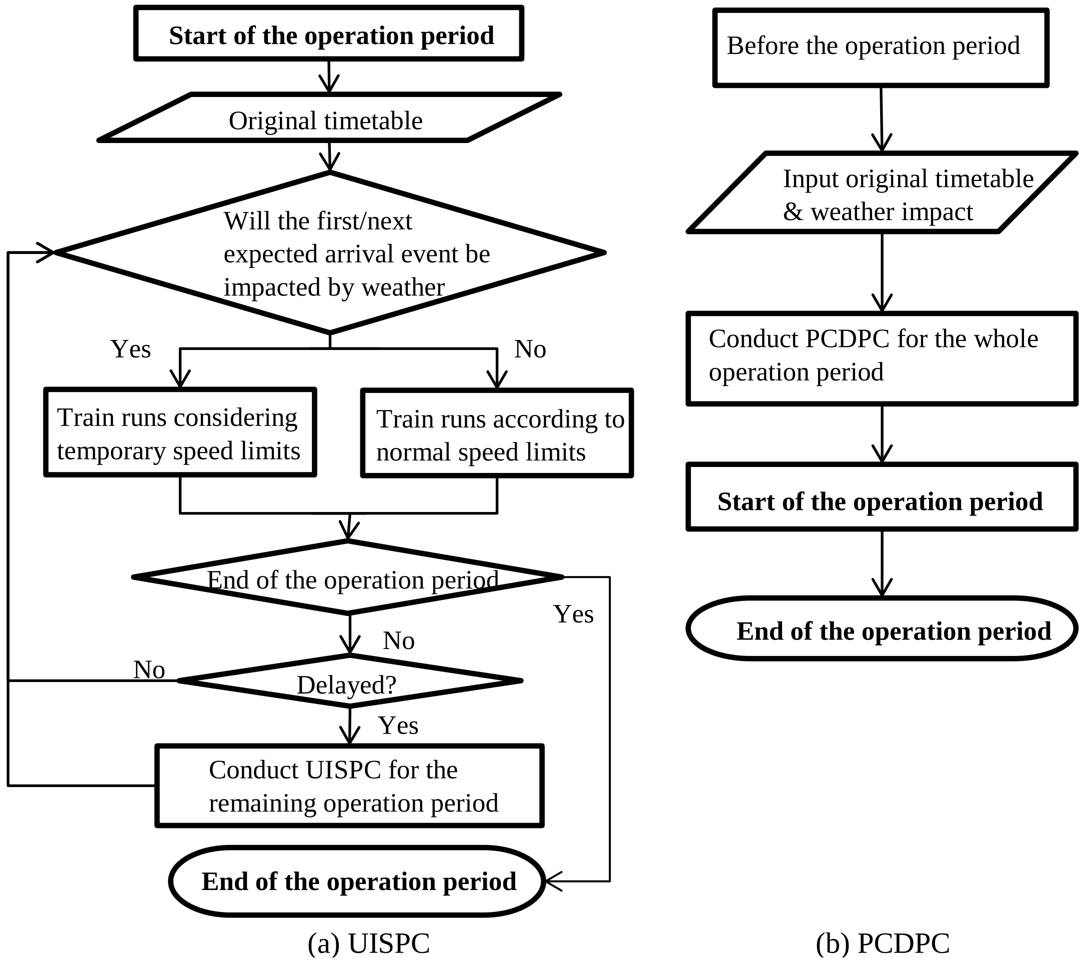


Figure 3-3: Flowcharts of UISPC and PCDPC

3.3.2. Mapping weather data onto the railway network

As the PCDPR relies on prescribed information of temporary speed limits over space and time during the rescheduling time horizon, a fundamental question here is then how to derive such temporary speed limit information from weather data. In the UK, the Met Office maps the weather conditions onto $g*g$ grids, and the location of each grid cell is expressed by the coordinate of its geographic centre. The weather condition of each track location can then be decided by referring to the grid cell that it is located in. The detailed process is described as shown in Figure 3-4.

Step 1: Obtain the coordinates of the weather grid cells and of the discretised points on the railway tracks;

Step 2: Map the railway points to the weather grid cells;

Step 3: Locate the weather cell that each railway point is in;

Step 4: For the time horizon that the rescheduling is concerned, obtain the weather forecast of each railway point in a time resolution of T;

Step 5: Based on the rail industry weather management criterions, estimate the weather impact on segment l during w^{th} period, p_{w}^{l}=[$b_{\mathrm{w}}^{l}, e_{\mathrm{w}}^{l}, v_{\mathrm{w}}^{l}$], where b_{w}^{l} and e_{w}^{l} are the starting time end time of the period, respectively, and v_{w}^{l} is the temporary speed restrictions.

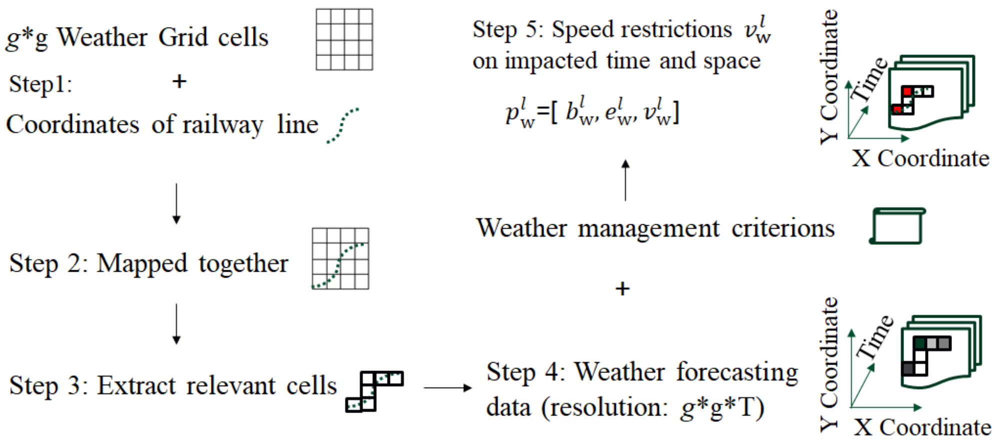


Figure 3-4: The process for mapping the adverse weather events as PCDP

To reduce the computation space, we aggregate the weather impact for each segment. As long as the impacts to a segment are continued for several continuous periods, we consider them as one single piece and record the entire start and end time. The p_{w}^{l}=[$b_{\mathrm{w}}^{l}, e_{\mathrm{w}}^{l}, v_{\mathrm{w}}^{l}$] is transferred to p_q=[l_q, b_q, e_q, v_q], which means the qth

weather impact in the system happened on segment l and start from and end at time b_q and e_q, respectively, and the corresponding speed limitation is v_q.

3.3.3. Modelling the impact of adverse weather on the minimum running time

The minimum link running time depends on weather impact on each link. Consider a train k travelling on link c with the length of x_c. The normal speed limit on the link without weather impact is $v_{c,k}$ and the corresponding standard minimum travel time is $t_{c,k}$. A weather-related disturbance p_q occurs on the link during time period $[b_q, e_q]$, leading to a restricted speed limit of v_q during that period. Then, depending on whether train k's planned trajectory overlaps with the impact zone or not, i.e. depending on $a'_{c,k}$ and $d'_{c,k}$, the planned entering time and leaving time on link c, its minimum running time can be derived according to the following two scenarios as shown in Figure 3-5.

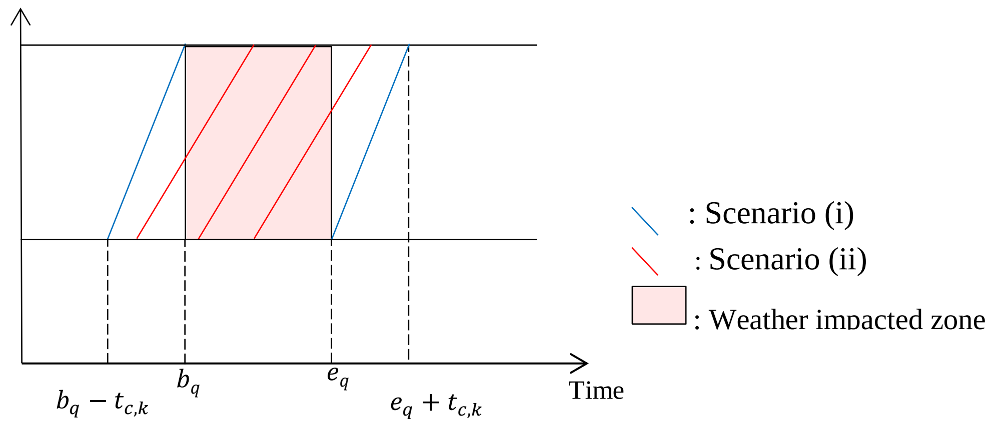


Figure 3-5: Illustration of train trajectories in different situations

Scenario (i): Train not disturbed by the weather:

If under the rescheduling plan, train k would leave the link before the start of the weather impact, or enter after the end of the impact, i.e.

$$b_q \geq a'_{c,k} \tag{3-1}$$

or

$$e_q \leq d'_{c,k} \tag{3-2}$$

then the temporary speed limit is not effective on the train and thus its minimum travel time through the link would be:

$$t'_{c,k} = \frac{x_c}{v_{c,k}} \tag{3-3}$$

Scenario (ii): Train disturbed by the weather:

If the train enters the link within period $[b_q - t_{c,k}, e_q]$, i.e.,

$$b_q - t_{c,k} < d'_{c,k} < e_q \tag{3-4}$$

then it will be disturbed. In this case, we assume that trains can only receive speed restriction information at particular control points such as stations; therefore, for safety concerns, even if only part of the train journey on this segment encounters the abnormal weather, the train will run under the temporary speed restriction for the entire segment.

When the weather impact is adverse, i.e. the non-zero temporary speed limitation is invoked, the corresponding minimum link travel time is:

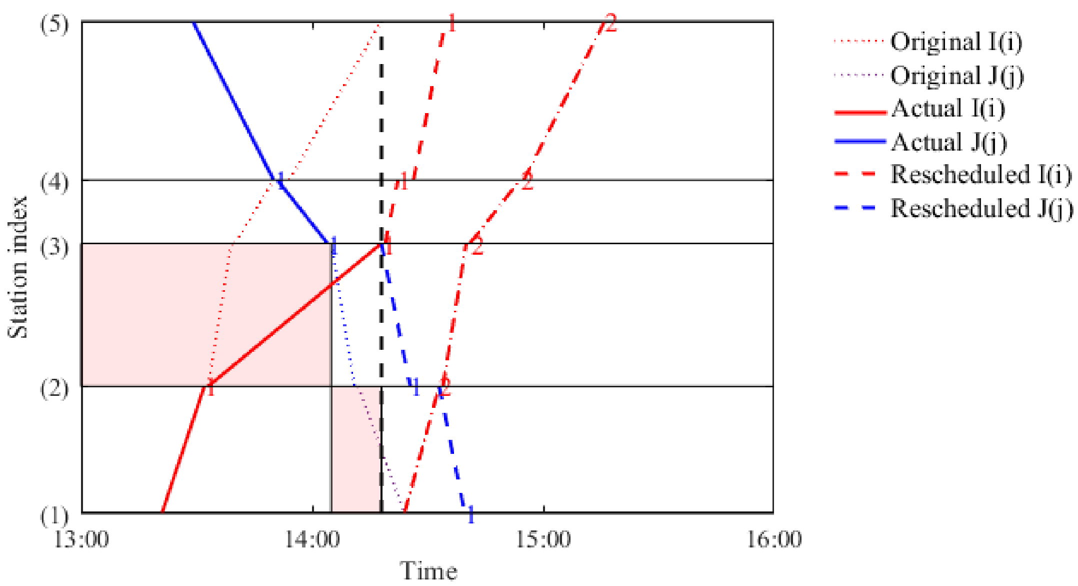
$$t'_{c,k} = \frac{x_c}{v_q} \tag{3-5}$$

When the weather impact is severe, the track blockage will be implied, we consider the track blockage as a special type of speed limitation which is equal to zero. During the impact period, the value of travel time on that link should be infinity. Trains originally scheduled to blocked link can choose either using the alternative lines or wait until the track blockage been cleared.

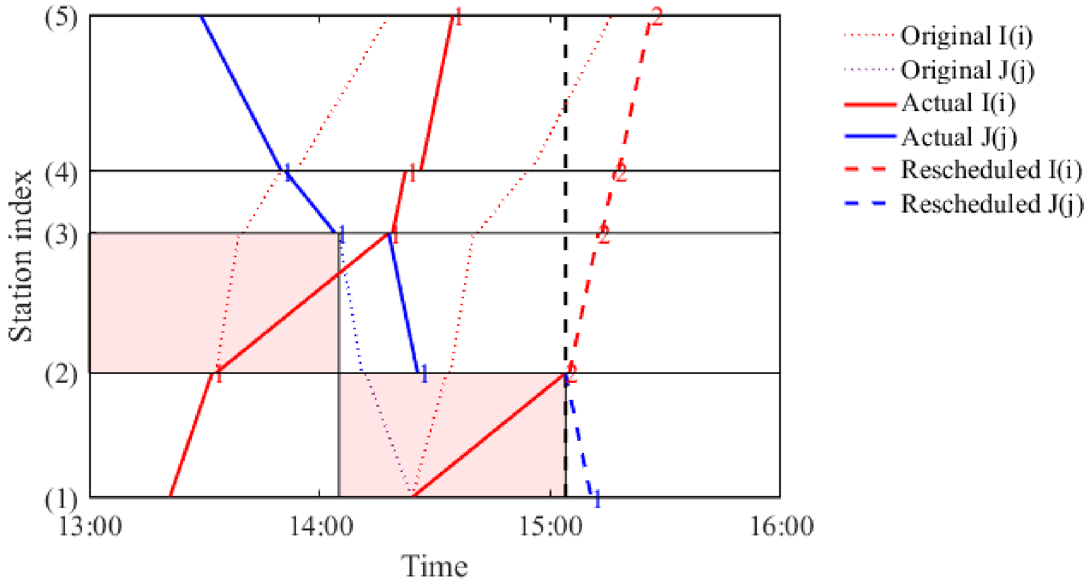
3.4. An illustrative case study for the comparison of UISPC and PCDPC

In this subsection, we further demonstrate the difference in the procedure and performance of UISPC and PCDPC. Figure 3-6 and Figure 3-7 show the rescheduling results by UISPC and PCDPC, respectively. In both figures, the up direction lines are the trajectories of outbound trains (denoted I(i) for outbound train i) and down direction lines inbound trains (denoted J(j) for inbound train j). The numbers alongside the lines are train indexes (red for outbound and blue for inbound). The shadowed zones indicate the weather impacts. The speed restriction in the shadowed zones is 20 km/h, while that in the unaffected zones is 120km/h. In Figure 3-6, the vertical dash lines show the time when the rescheduling is carried out; the train trajectories on the left-hand side of the dash line have already happened, and those on the right-hand side represent the rescheduled result by UISPC.

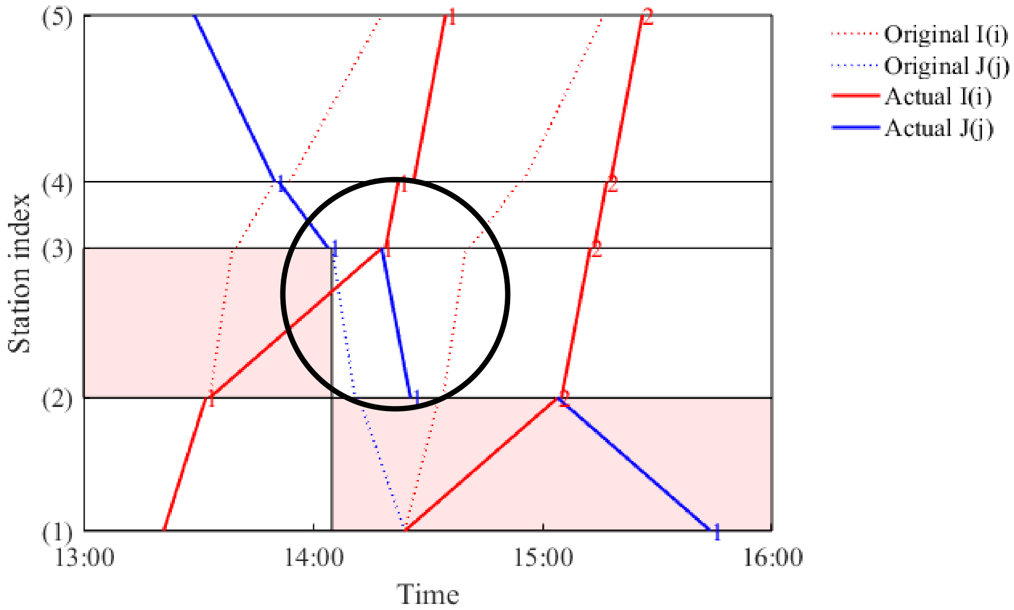
Figure 3-6(a) shows that the first rescheduling is triggered when the delay on train I(1) is detected when arriving at station 3. Without considering any future weather impact, the rescheduling plan is illustrated on the right-hand side of the vertical dash line) suggests that train J(1) is rescheduled to depart from station 3 after train I(1) arrives at station (3) and depart from station 2 after train I(2) arrives at station (2). Under this rescheduling plan, as shown in Figure 3-6(b), due to weather impact, train I(2) has to run slower than expected from station 1 to station 2 and thus delayed, which triggers the second rescheduling. Train J(1) need to wait at station (2) until train I(1) arrives at station (2). Figure 3-6(c) further shows that, train J(1) is also delayed when running from station 2 to station 1 by the weather. The final delays for trains I(1), I(2) and J(1) are 17, 10 and 80 minutes, respectively, and the total system delay amounts to 107 minutes.



(a) Rescheduled result after the occurrence of the first delay



(b) Rescheduled result after the occurrence of the second delay



(c) Final trajectories

Figure 3-6: Rescheduled results from UISPC: (a) first rescheduling when delay detected on train I(1) at station 3; (b) second rescheduling when delay detected on train I(2) at station 2; and (c) the final trajectories.

Figure 3-7 shows the PCDPC conducted before 13:00, by considering all the future weather impacts systematically over the whole planning horizon. As marked by the black circles in Figure 3-6(c) and Figure 3-7, the key difference between

UISPC and PCDPC lie in the decision regarding the priority between trains I(1) and J(1) on traversing the segment between stations 2 and 3. In Figure 3-6, I(1) passed the segment before J(1) and was impacted by the weather. In Figure 3-7, I(1) waits at station 2 and lets J(1) pass the segment first; consequently, both I(1) and J(1) avoid the adverse weather on this segment. The final delays for trains I(1), I(2) and J(1) are 19, 40 and 30 minutes, respectively; the total system delay is 89 minutes, 18 minutes less than that under UISPC.

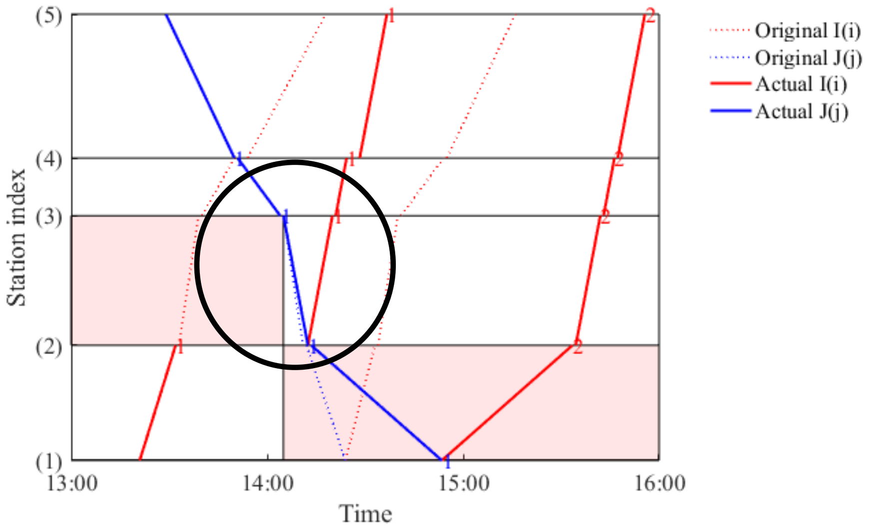


Figure 3-7: Rescheduling results from PCDPC

3.5. Summary

In this chapter, we first introduced the different infrastructure models, i.e. macroscopic, microscopic and mesoscopic and chose the macroscopic model for rescheduling and mesoscopic model for rerouting and rescheduling. Second, we introduce the control actions and the constraints for traffic control, and the solution approaches. Thirdly, how we map the weather data onto the railway network and the difference between our approach and the existing approach in considering weather are analysed, which is also one of the major contributions of this research. At last, we use an illustrative case study to show the effectiveness of our research.

From Chapter 4 to Chapter 6, we will apply the models and approaches described in this chapter and conduct the quantitative analysis for PCDPC and UISPC, the solution quality and computation time for standard solvers and GA, as well as the further application for rerouting under weather impact.

Chapter 4 Single-track rescheduling for weather-induced PCDP with mixed integer programming

4.1. Introduction

Single-track scheduling is one of the most common problems in railway system control; we also choose it to verify our weather-induced control method mentioned in Chapter 3 to start with.

In Section 2.4.2, we introduced that the operational guidance mandates different control actions for different weather severe level. We consider them as two types, one is speed limitation, and another is track blockage. In this Chapter, we consider weather-induced PCDP which only lead to temporary speed restrictions in the network over the next rescheduling horizon. We do not consider track closure nor rerouting of trains. Since the weather forecast is more accurate as it is closer to the time making the forecast, this study focuses on the rescheduling over a short period up to one day or hours ahead, when the weather forecasts are sufficiently reliable. In addition, the conditions are assumed would not change during the planning time horizon.

We will first review the single-track rescheduling models. Second, a MILP model is formulated to solve this PCDP rescheduling (PCDPR) problem to minimise the total delay of all trains at their destinations. In addition to the basic Constraints 1) to 9) in Section 3.2.2, the running time constraints under weather impact in Section 3.3.3 are also considered. Finally, we present examples to illustrate the advantages of considering precise weather forecasting information in the rescheduling process.

4.2. Single-track rescheduling models

In this thesis, we consider only single-track lines. Landex (2009) introduced the differences between single-track and double-track and the way of evaluation single-track capacity. Trains operated over single-track lines can only overtake and cross each other at specific locations such as stations and passing loops with more than one track. Operation restrictions at these locations must be considered in the rescheduling models.

Higgins et al. (1996) proposed a non-linear mixed integer program for a single-track line to minimise the train delays at station and train operating costs. The constraints are train following and overtaking constraints, conflict-avoiding constraints, travel time constraints and departure constraints, where the travel time on a rail segment is bounded from above by the segment length divided by the corresponding upper speed limit.

Li et al. (2014) proposed non-linear mixed integer programming model for single-track rescheduling. The objective function considers the sensitivities factors on train delay, including train types and travelling miles. For the following and overtaking headway constraints, it assumes that at one station, only one departure or arrival event could be arranged at a time, to represent the time constraints of the dispatchers to switch signals for different trains. Therefore, in addition to the safety headway requirements for two consecutive trains on their departure or arrival, they define the headway for every departure or arrival event in a station. Further practical constraints are introduced, such as loading and unloading constraints, stopping/non-stopping constraints, and station capacity constraints.

However, most of the existing rescheduling research are reactive rescheduling which i) model the disturbances to trains instead of tracks; ii) only react when disturbances happened; iii) do not consider future disturbances. Based on Higgins et al. (1996) and Li et al. (2014), this chapter proposes a mixed integer linear programming (MILP) model to solve the rescheduling problem for a single-track line under temporary speed restrictions induced by adverse weather, and new constraints are introduced to represent the weather impact. The temporary speed limits are applied on tracks within only particular time periods, so the trains disturbed are not fixed but determined by the rescheduling decision. Moreover, because the proposed rescheduling method takes all the forecasted disturbances into account in advance, it may help the trains to avoid the initial delays as well as some future delays caused by the future disturbances.

4.3. A mathematical model for rescheduling under adverse weather conditions

As the focus of this chapter is modelling and incorporating the weather impact in train rescheduling, the macroscopic model which omits some details such as station structure, signalling and train length, is sufficient to represent the essential characteristics required. This is a common modelling technique while formulating the train (re)scheduling as mathematical programming problems (Higgins et al.,1996; Schachtebeck and Schöbel, 2010; Li et al., 2014). It makes the modelling possible whilst maintains essential characteristics of the railway system and the major concern of the (re)scheduling process.

In this section, we formulate a mathematical model for the rescheduling of a single-track railway line under abnormal weather conditions. First, we introduce necessitated assumptions based on the characteristics of a single-track line. This is then followed by a detailed description of the novel technique proposed in this thesis for modelling the impact of temporary speed limits induced by adverse weather. Finally, the complete MILP formulation is presented.

4.3.1. Model representation and assumptions of a single-track railway line

Nomenclature

I: the set of outbound trains

J: the set of inbound trains

K: the set of all trains, $K = I \cup J$

L: the set of rail segments in the network, where a segment is the undirected track between two adjacent stations

E: the set of all links, where a link is a directed track from one station to another

E_I: the set of all links available to outbound trains

E_J: the set of all links available to inbound trains

E_w: the set of all weather impacted links

S: the set of all stations on the line

Q: the set of weather events

R_s: the set of all capacity tracks at intermediate station s, $s \in S \setminus \{1, |S|\}$.Where the capacity tracks in one station are defined as the tracks which can be used for trains to dwell at or go through. Each capacity track can be used by only one train at one time.

$|X|$: the number of elements in set X

i, j, k: the train index

s: the station index, numbered in an ascending order along the outbound direction, $s = 1,2, \cdots, |S|$

r: the index of the capacity track in one station

l : the segment index

l^s: the segment between stations s and $s+1$

c_s : the outbound link associated with l^s

$\bar{c}_s$: the inbound link associated with l^s

x_c: the length of link c

$d_{c,k}$: the time of train k entering link c given by the original timetable

$a_{c,k}$: the time of train k leaving link c given by the original timetable

p_q: the impact of weather event $q \in Q$; $p_q = \left[l_q^s, b_q, e_q, v_q\right]$, where l_q^s is the affected segment (between stations s and $s+1$), b_q the start time, e_q the end time, and v_q the temporary speed limit to all trains

$v_{c,k}$: the maximum allowed speed of train k on link c under normal conditions

$t_{c,k}$: the minimum running time of train k on link c under normal condition, $t_{c,k} = x_c / v_{c,k}$

$t_{s,k}^w$: the required minimum dwell time of train k at station s for boarding and alighting

t^h: running headway, the required minimum time headway between two trains which travel on a same segment

t^a: arrival-arrival headway, the required minimum time headway between two opposite trains arriving at the same station

t^d: departure-departure headway, the required minimum time headway between two opposite trains departing from the same station

t^{ad}: arrival-departure headway, the required minimum time headway between two opposite trains arriving at and departing from the same station

Δt_k: the arrival delay of train k at its destination. For outbound trains, $\Delta t_i = (a'_{c_{|S|},i} - a_{c_{|S|},i})_+$; for inbound trains, $\Delta t_j = (a'_{\bar{c}_1,j} - a_{\bar{c}_1,j})_+$, where $Z_+ = \max(Z, 0)$

C: the total arrival delay of all trains at the destination

M: a sufficiently large constant

ε: a sufficiently small constant

We first describe the infrastructure model in this chapter. We consider a two-way single-track railway line, which consists of stations and segments. At the two ends of the line are two terminal stations, and in between them are intermediate stations. As a station is much shorter than a segment, and the travel time in a station is implicitly expressed when using the average travel speed to calculate the arrival time at each station, stations are abstracted as dots with no physical lengths. However, to prevent the conflicts in stations and system deadlocks, the capacity limitations in the stations need to be considered. Each station contains a finite number of tracks, named as capacity tracks, which allow trains to go through or dwell. Each capacity track can be occupied by no more than one train at one time.

Figure 4-1 illustrates a small section of the single-track line. Segment l^s is a directionless track section between two stations s and $s+1$. Trains travelling in opposite directions cannot occupy l^s at the same time. The segment l^s is associated with two directional links: the link from station s to station $s+1$ is denoted as the outbound link c_s, while the link from station $s+1$ to station s is the inbound link $\bar{c}_s$.

The length of link c is denoted x_c. Train routes are fixed; no route change or station skipping is allowed.

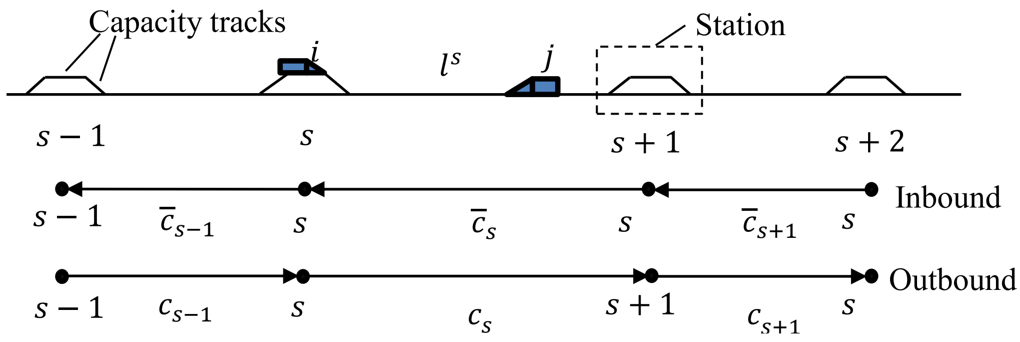


Figure 4-1: A single-track railway line seen from infrastructure level and from train route level

The following assumptions are made for trains.

Assumption 1. The lengths of trains are ignored, as the track segments are much longer than the trains and the time needed for the entire train leaving or entering a segment or a station can be implicitly expressed in running time calculated using average speed.

Assumption 2. Due to different crew or signalling system limitations, some stations can only send or receive one train at the same time but some other stations can send a train and also receive a train at the same time. We set a general rule for minimum headways. Minimum headways should be maintained between two consecutive trains travelling in the same direction, between opposite trains arriving at and departing from the same station, between opposite trains arriving at the same station, and between opposite trains departing from the same station. For any system do not consider any of the above headway, we simply set the value of the headway to zero.

Assumption 3. At the stations, trains need to stop. trains cannot depart from stations before their departure times given by the original schedule to avoid punctual passengers missing the trains. The actual arrival times are free and depending on the train advancing strategy. As passing loops do not have passengers boarding and alighting, the departure time restriction from the original schedule is not needed.

Assumption 4. The running time loss due to acceleration and braking are implicitly considered by using average speed that the train is capable to reach while running freely on a segment.

Assumption 5. Train dwell times at stations should be not less than the required minimum dwell times given by the original schedule to allow passengers loading and unloading. For the stations where a train does not need to stop, the dwell time can be set to zero.

Assumption 6. Train delay at a station is defined as the excess of the actual/rescheduled arrival time over the originally scheduled arrival time.

Assumption 7. In the segment where abnormal weather threshold is predicted to be reached, a lower temporary speed limit will be applied on this segment during the predicted time period. If any part of a train's trajectory falls into the impacted spatiotemporal zone, the train must follow the temporary speed restriction for the whole segment.

4.3.2. MILP formulation for PCDPR

In this chapter, we consider weather-related PCDPs which only lead to temporary speed restrictions in the network, and do not consider track closure or rerouting trains. Since the weather forecast is more accurate as it is closer to the time making the forecast, this study focuses on the rescheduling for a short time period which is hours or at most one day ahead, so that the weather forecasts are sufficiently reliable. A MILP model is formulated to solve this PCDPR problem to minimise the total delay of all trains at their destinations. The new departure and arrival times at all stations are the key decision variables in this model. Variables of train following and overtaking, conflict resolving, weather impact, and capacity track allocating are auxiliary variables.

The train departure can be delayed because:

d1) the arrival at the station is delayed;

d2) to maintain a minimum headway to another train travelling in the same direction which has just departed;

d3) to keep the next track segment clear for a train passing in the opposite direction;

d4) to allow another train behind to overtake; or

d5) to avoid the predicted weather impact zone.

The train arrival can be delayed because:

a1) its departure from previous station is delayed;

a2) it must keep a safety headway to a train in front;

a3) it must wait for the allocated downstream track to be cleared; or

a4) its travel time is increased by the weather-impacted speed limitation

d1) to d4) and a1) to a3) are commonly included in UISPR models as constraints, while d5) and a4) are newly introduced constraints in our PCDPR model.

The PCDPR MILP model includes the following decision variables:

$d'_{c,k}$: the rescheduled time of train k entering link c

$a'_{c,k}$: the rescheduled time of train k leaving link c

and the following auxiliary variables:

$A_{i,j,\mathrm{c}}$: the binary variables, if train *i* traverses link c before train j $A_{i,j,\mathrm{c}} = 1$; otherwise $A_{i,j,\mathrm{c}} = 0$

B_{i,j,l^s}: the binary variables, if train i traverses segment l^s before train j, $B_{i,j,l^s} = 1$; otherwise $B_{i,j,l^s} = 0$

$X_{q,k}$: the binary variables, if the trajectory of train *k* overlaps with weather q, $X_{q,k} = 1$; otherwise $X_{q,k} = 0$

$Y_{q,k}$: the binary variables, if train *k* enters segment before weather *q* starts, $Y_{q,k} = 1$; otherwise $Y_{q,k} = 0$

$\tau^r_{s,k}$: the binary variables, if train k uses track r of station s, $\tau^r_{s,k} = 1$; otherwise $\tau^r_{s,k} = 0$

$t'_{c,k}$: the rescheduled minimum running time of train k on link c

$H_{i,j,s}$: the binary variables, if train *k* arrives station *s* before train *j*, $H_{i,j,s} = 1$; otherwise $H_{i,j,s} = 0$

$G_{i,j,s}$: the binary variables, if train k departs from station s before train j, $G_{i,j,s} = 1$; otherwise $G_{i,j,s} = 0$

The PCDPR model in this chapter is developed based on the UISPR models of Higgins et al. (1996) and Li et al. (2014) with the following modifications as well as additional features. First, when a train does not need to stop at a station, the minimum dwell time is set to be zero; otherwise, it is some predetermined positive value related to passenger boarding and alighting. Second, in order to incorporate the weather related speed restriction, which is the major consideration of PCDPR, we introduce a set of variables to describe whether a train path falls into a weather impact zone. Finally, the travel time constraints are modified so as to capture the effect of travelling through a weather impact zone where a lower speed limit is imposed.

The PCDPR problem is formulated as an optimisation problem to minimise the total arrival delay for all trains at their destinations, i.e.,

$$\min C = \sum_{i \in I} \Delta t_i + \sum_{j \in J} \Delta t_j \tag{4-1}$$

Subject to the following constraints (a)-(i):

(a) **Departure time constraints:**

Constraint (4-2) ensures the rescheduled departure times at all origins and intermediate stations for all trains are not earlier than their original departure times so that punctual passengers will not miss the train.

$$d'_{c,k} \geq d_{c,k} \qquad \forall\, c \in E, k \in K \tag{4-2}$$

(b) Arrival time constraints:

Constraint (4-3) specifies that the rescheduled time of leaving a link (i.e. arrival at a station) should be not earlier than the rescheduled time of entering it plus the minimum running time on it.

$$a'_{c,k} \geq d'_{c,k} + t'_{c,k} \quad \forall\, c \in E, k \in K \tag{4-3}$$

(c) Dwell time constraints:

Constraint (4-4) and (4-5) ensures that the rescheduled dwell time at an intermediate station is not shorter than the required dwell time at this station.

$$d'_{c_s,k} - a'_{c_{s-1},k} \geq t^{w}_{s,k} \quad \forall\, s \in S \setminus \{1, |S|\}\;\; k \in I \tag{4-4}$$

$$d'_{\bar{c}_{s-1},k} - a'_{\bar{c}_s,k} \geq t^w_{s,k} \quad \forall\, s \in S \setminus \{1, |S|\}, k \in J \tag{4-5}$$

(d) Headway constraints for train following and overtaking constraints:

To ensure safety, for two trains i and j travelling on the same link, they have to keep a minimum headway between each other; meanwhile, the headway is needed because normally a station can operate at most one train at one time. Therefore, if train i enters the link c earlier than train j, denoted $A_{i,j,c} = 1$, then the time train j enters (or leaves) this link should be later than the time train i entering (or leaving) this link plus the minimum time headway t^h; in this case, constraints (4-6) and (4-7) will be active. Conversely, if train j enters the link earlier, $A_{i,j,c} = 0$ and constraints (4-8) and (4-9) will be active.

$$t^h + d'_{\mathrm{c},i} \leq d'_{\mathrm{c},j} + M \times (1 - A_{i,j,c}) \tag{4-6}$$

$$t^h + a'_{c,i} \leq a'_{c,j} + M \times (1 - A_{i,j,c}) \tag{4-7}$$

$$t^h + d'_{\mathrm{c},j} \leq d'_{\mathrm{c},i} + M \times A_{i,j,c} \tag{4-8}$$

$$t^h + a'_{c,j} \leq a'_{c,i} + M \times A_{i,j,c} \tag{4-9}$$

where constraints (4-6)- (4-9) apply to all $c \in E_I; i,j \in I$ and all $c \in E_J;\ i,j \in J$.

(e) Headway constraints at stations for trains running in opposite directions:

In some railway systems, a station can normally operate at most one train at one time, and thus a minimum time headway is required between two trains both entering (or leaving) a same station in opposite directions. Considering this situation, the below constraints are introduced. In some other circumstances when trains are allowed to travel opposite directions, the values of the headways can be set zero,

$$t^a_s + a'_{\mathrm{c}_{s-1},i} \leq a'_{\bar{c}_s,j} + M \times (1 - H_{i,j,s}) \tag{4-10}$$

$$t^a_s + a'_{\bar{c}_s,j} \leq a'_{\mathrm{c}_{s-1},i} + M \times H_{i,j,s} \tag{4-11}$$

$$t^d_s + d'_{\mathrm{c}_s,i} \leq d'_{\bar{c}_{s-1},j} + M \times (1 - G_{i,j,s}) \tag{4-12}$$

$$t^d_s + d'_{\bar{c}_{s-1},j} \leq d'_{\mathrm{c}_s,i} + M \times G_{i,j,s} \tag{4-13}$$

where constraints (4-10)-(4-13) apply to all: $\forall s \in S; i \in I;\ j \in J; c_s \in E_I; \bar{c}_{s-1} \in E_j$.

(f) Conflict-avoidance constraints for trains running in opposite directions:

To ensure safety, in the single-track rail system, trains running in opposite directions should not meet on the segment. For outbound train *i* and inbound train *j* using the same segment l^s, if train *i* goes through first, denoted $B_{i,j,l^s} = 1$, the time of train *j* entering segment l^s, i.e. $d'_{\bar{c}_s,j}$ should be not earlier than the time of train *i* leaving segment l^splus the minimum headway t^{ad}, and in this case constraint (4-14) will be active; conversely, if trains *j* uses the segment first, $B_{i,j,l^s} = 0$ and constraint (4-15) will be active.

$$a'_{c_s,i} + t^{ad} \leq d'_{\bar{c}_s,j} + M \times \left(1 - B_{i,j,l^s}\right) \qquad \forall s \in S \setminus \{|S|\}, i \in I, j \in J \tag{4-14}$$

$$a'_{\bar{c}_s,j} + t^{ad} \leq d'_{c_s,i} + M \times B_{i,j,l^s} \qquad \forall s \in S \setminus \{|S|\}, i \in I, j \in J \tag{4-15}$$

(g) Station capacity constraints:

The station capacity constraints are considered to avoid deadlock when the station capacity is limited, and to avoid conflict at stations. In each station *s*, train *k* can only occupy one track at a time, which is specified by Constraint (4-16). Meanwhile, for two trains *i* and *j* allocated to the same capacity track r, i.e. $\tau^r_{s,i} = \tau^r_{s,j} = 1$, $i, j \in K$, regardless of their running directions, if train i occupies r earlier than train j, then the arrival time of train *j* must be later than the departure time of train i plus the minimum headway. Such requirement is specified by Constraints (4-17) to (4-20) where Constraints (4-17) and (4-18) are for trains running in the same direction, and Constraints (4-19) and (4-20) are for opposite directions.

$$\textstyle\sum_{r \in R_s} \tau^r_{s,k} = 1 \qquad \forall\, s \in S \setminus \{1, |S|\}, k \in K \tag{4-16}$$

$$d'_{c,i} + t^h \leq a'_{c,j} + \left(3 - \tau^r_{s,i} - \tau^r_{s,j} - A_{i,j,c}\right) \times M$$

$$\forall\, s \in S \setminus \{1, |S|\}; r \in R_s; c = c_s \text{ and } i, j \in I \text{ or } c = \bar{c}_s \text{ and } i, j \in J \tag{4-17}$$

$$d'_{c,j} + t^h \leq a'_{c,i} + \left(2 - \tau^r_{s,i} - \tau^r_{s,j} + A_{i,j,c}\right) \times M$$

$$\forall\, s \in S \setminus \{1, |S|\}; r \in R_s; c = c_s \text{ and } i, j \in I\ or\ c = \bar{c}_s\ and\ i, j \in J \tag{4-18}$$

$$d'_{c_s,i} + t^{ad} \leq a'_{\bar{c}_s,j} + \left(3 - \tau^r_{s,i} - \tau^r_{s,j} - B_{i,j,l^s}\right) \times M$$

$$\forall\, s \in S \setminus \{1, |S|\}, r \in R_s;\ i \in I;\ j \in J \tag{4-19}$$

$$a'_{\bar{c}_s,j} + t^{ad} \leq d'_{c_s,i} + \left(2 - \tau^r_{s,i} - \tau^r_{s,j} + B_{i,j,l^s}\right) \times M$$

$$\forall\, s \in S \setminus \{1, |S|\}, r \in R_s;\ i \in I; j \in J \tag{4-20}$$

(h) Weather constraints:

The rules in Section 3.3.3 for determining whether a train is impacted by the weather are converted into the weather constraints (4-21)- (4-24) as follows.

Similar to Diego (2016), we introduce two binary variables, $X_{q,k}$ and $Y_{q,k}$. $X_{q,k} = 1$ if train k goes through the weather zone q and thus needs to follow the temporary speed limit, and 0 otherwise. $Y_{q,k}$ specifies, in the case that the train is not impacted by the weather, i.e. $X_{q,k} = 0$, it traverses the segment before or after the weather impact: $Y_{q,k} = 0$ means before and $Y_{q,k} = 1$ means after. The benefit of our model is we define the weather impact variable as weather and train specified, so that it can assign different speed limitation level to the same track during different time intervals in different weather events.

When $X_{q,k} = 1$, constraints (4-21) and (4-22) are active and constraints (4-23) and (4-24) are inactive. Then we have $b_q - t_{c,k} + \varepsilon \le d'_{c,k} \le e_q - \varepsilon$, which is equivalent to Equation (3-4) in Section 3.3.3. When $X_{q,k} = 0$, constraints (4-21) and (4-22) are inactive. If $Y_{q,k} = 0$, constraint (4-24) will be inactive and constraint (4-23) will yield $b_q - t_{c,k} \ge d'_{c,k}$, which is the same to Equation (3-1); otherwise, if $Y_{q,k} = 1$, constraint (4-23) will be inactive and constraint (4-24) will read $e_q \le d'_{c,k}$, which is Equation (3-2) in Section 3.3.3. Therefore, constraints can fully describe whether the train is impacted by the weather impact. Constraint (4-25) is added to reduce the feasible region as $Y_{q,k}$ is valid only when $X_{q,k} = 0$.

$$\left(b_q - t_{c,k}\right) \times X_{q,k} - \left(1 - X_{q,k}\right) \times M + \varepsilon \le d'_{c,k} \tag{4-21}$$

$$e_q \times X_{q,k} + \left(1 - X_{q,k}\right) \times M - \varepsilon \ge d'_{c,k} \tag{4-22}$$

$$\left(b_q - t_{c,k}\right) \times \left(1 - Y_{q,k}\right) + X_{q,k} \times M + Y_{q,k} \times M \ge d'_{c,k} \tag{4-23}$$

$$e_q \times Y_{q,k} - X_{q,k} \times M - \left(1 - Y_{q,k}\right) \times M \le d'_{c,k} \tag{4-24}$$

$$Y_{q,k} \ge X_{q,k} \tag{4-25}$$

Where constraints (4-21)-(4-25) apply to $q \in Q; k \in I \textit{ and } c \in E_w \cap E_I$ or $k \in J \textit{ and } c \in E_w \cap E_J$.

(i) Running time constraints:

Equations (3-3) and in (3-5) Section 3.3.3 formulate the minimum running time under two different scenarios that whether the train goes through a weather impacted zone or not. These two scenarios are combined into Equation (4-26) as follows: when $X_{q,k} = 1$, the train passes through weather impact zone q and follows the reduced speed limit v_q; when $X_{q,k} = 0$, the train is not affected by weather and follows the normal speed limit.

$$t'_{\mathrm{c},k} \geq \frac{x_c}{v_{\mathrm{c},k}} \times \left(1 - X_{q,k}\right) + \frac{x_{\mathrm{c}}}{v_q} \times X_{q,k}$$

$$\forall\, p \in P; k \in I \; and \; c \in E_w \; \cap E_I \text{ or } k \in J \; and \; c \in E_w \; \cap E_J \qquad (4\text{-}26)$$

$$t'_{\mathrm{c},k} \geq \frac{x_c}{v_{\mathrm{c},k}} \qquad \forall\, k \in K \; and \; c \in E \setminus E_w \qquad (4\text{-}27)$$

Constraints (4-1)-(4-27) are then the MILP formulation for solving the train rescheduling problem under adverse weather.

4.4. Case studies

In this section, we will use numerical examples to illustrate the relatively better performance of PCDPR than UISPR.

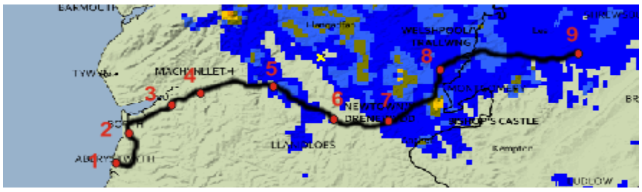
4.4.1. Case study on a real-life railway line

In this subsection, we will apply both the traditional UISPR and the proposed PCDPR to the Cambrian Line, a single-track railway line in the UK, and compare the delays given by the two methods. We will also investigate the influence of information provision on the performance of PCDPR in terms of total delay.

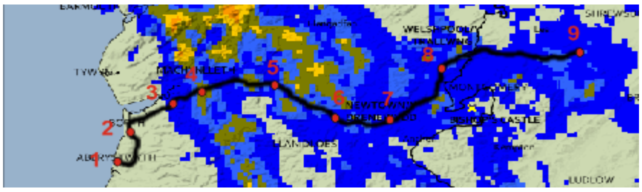
The Cambrian line is an 81.5-mile single-track line in the UK, running from an inland town Shrewsbury, through the Wales mountain range Snowdonia, to the west-coast Aberystwyth. The line goes through nine stations indexed from 1 to 9 for simplicity, from Aberystwyth to Shrewsbury. Consider a test period from 11:00 to 21:00, during which six outbound trains are scheduled to run from station 1 (Aberystwyth) to station 9 (Shrewsbury) and five inbound trains from station 9 to station 1. The real world original departure and arrival times for each train in each station are listed in Table 4-1 and Table 4-2, respectively. The normal speed limit is

100 mph. The capacity of each intermediate station is two. The minimum headway for the same direction is set to 2 minutes; the arrival-arrival, departure-departure and arrival-departure headways are all set to 3 minutes; the required minimum dwell time is the dwell time in the original timetable, which can be worked out from the scheduled departure time in Table 4-1 and the scheduled arrival time in Table 4-2.

Assume that a weather front passed by the line and brings the rainfall. Figure 4-2 shows the gridded rainfall data at 11:00 and 14:00, respectively. The black line and the red dots indicate the railway line and stations, respectively, where the station indexes are marked aside. Assume the bright yellow squares indicate the rainfall amount which triggers the temporary speed restriction of 20mph. The time resolution of applying temporary speed restriction is one hour. Therefore, according to Figure 4-2(a), no temporary speed limit is applied during 11:00 to 12:00; according to Figure 4-2(b), segments between stations 4 and 6 will be under temporary speed restriction during 14:00 to 15:00.



(a)



(b)

Figure 4-2: Gridded rainfall amount at (a) 11:00 and (b) 14:00

The speed restriction is mapped along the line and over the planning time period, as shown in Figure 4-3. Also plotted in Figure 4-3 are the original/planned train schedules, whose detailed departure and arrival time are showed in Table 4-1 and Table 4-2.

Table 4-1: Departure time of each train at each station (hhmm)

Station index	Mileage (mile)	I(1)	I(2)	I(3)	I(4)	I(5)	I(6)	J(1)	J(2)	J(3)	J(4)	J(5)
1	0	1230	1330	1530	1730	1830	1930	-	-	-	-	-
2	8.25	1241	1341	1541	1741	1841	1941	1305	1506	1705	1904	2006
3	16.50	1257	1352	1552	1751	1852	1956	1255	1456	1655	1854	1956
4	20.50	1306	1407	1608	1805	1909	2008	1248	1449	1649	1846	1947
5	30.85	1320	1324	1625	1817	1923	2017	1239	1424	1625	1817	1923
6	42.25	1333	1434	1631	1828	1932	2031	1212	1412	1613	1810	1914
7	47.75	1340	1441	1642	1839	1943	2041	1205	1405	1606	1803	1907
8	61.75	1354	1455	1656	1854	1957	2056	1151	1351	1552	1749	1853
9	81.50	-	-	-	-	-	-	1129	1329	1530	1727	1831

Table 4-2: Arrival time of each train at each station (hhmm)

Station index	Mileage (mile)	I(1)	I(2)	I(3)	I(4)	I(5)	I(6)	J(1)	J(2)	J(3)	J(4)	J(5)
1	0	-	-	-	-	-	-	1316	1517	1719	1916	2019
2	8.25	1240	1340	1540	1740	1840	1940	1304	1505	1704	1903	2005
3	16.50	1256	1351	1551	1750	1851	1955	1254	1455	1654	1853	1955
4	20.50	1304	1358	1607	1759	1859	2003	1243	1442	1643	1840	1944
5	30.85	1320	1324	1625	1817	1923	2017	1239	1424	1625	1817	1923
6	42.25	1332	1433	1630	1827	1931	2030	1211	1411	1612	1809	1913
7	47.75	1339	1440	1641	1838	1942	2040	1204	1404	1605	1802	1906
8	61.75	1353	1454	1655	1853	1956	2055	1150	1350	1551	1748	1852
9	81.50	1418	1516	1719	1915	2021	2117	-	-	-	-	-

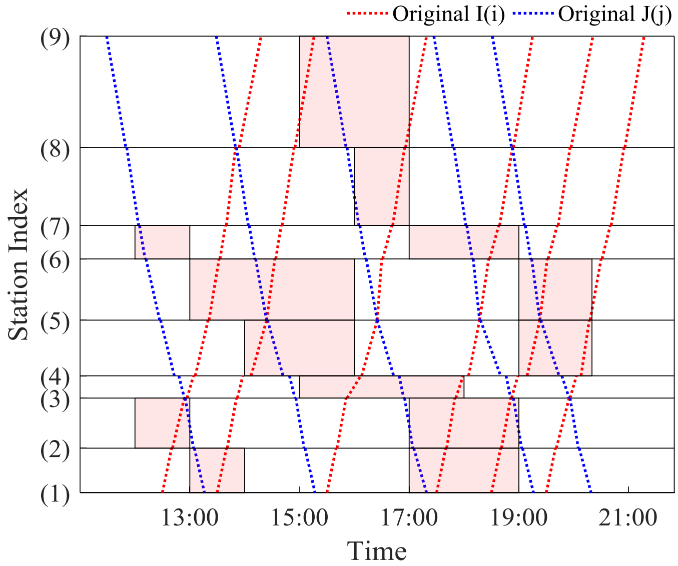


Figure 4-3: The original timetable and the speed restriction zones (indicated by the shadowed area)

Both PCDPR and UISPR MILP models are programmed in MATLAB 2015a and solved by CPLEX 12.6.3 through the interface provided by YALMIP (20150919). The values of M and ε are set to be 10000 and 0.001, respectively. The working computer has an Intel Core i5 3.20 GHz processor and 8.0 GB RAM, and the operating system is Windows 7.

4.4.2. Compare UISPR and PCDPR

We first compared the UISPR method in Experiment 4-1 (E 4-1) and the PCDPR method in Experiment 4-2 (E 4-2). In E 4-1, when a train departs from a station, according to the most recent timetable, it is informed whether the temporary speed limit is effective on the segment ahead: if not, it runs according to the most recent timetable; otherwise, it runs at the scheduled speed or temporary speed limit, whichever is lower. When a train arrives at a station, the arrival time is compared with the most-recently-scheduled arrival time at this station; if a delay is detected, the rescheduling is trigged. In E 4-2, rescheduling is conducted only once at 11:00, by taking account of all the planned temporary speed limits from 11:00 to 21:00. Delays of all trains at their destinations by the two methods are shown in Table 4-3. All trains expect I(1) experience less delay in E 4-2 than in E 4-1. Overall, the total delay resulted from E 4-2 is 163 minutes less than that from E 4-1, which means that

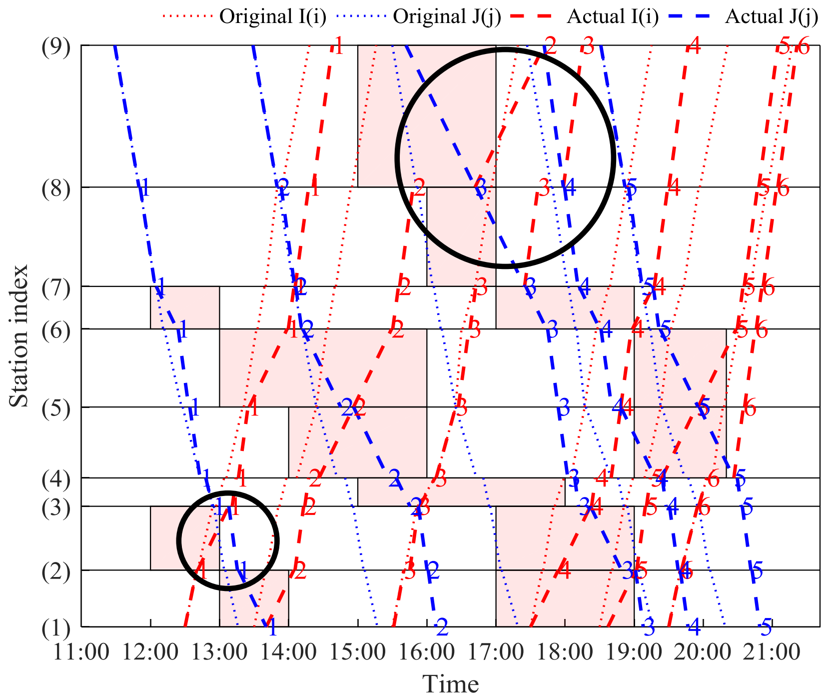
PCDPR considering all future weather information in the rescheduling is better than UISPR which considers no weather information in the rescheduling and deal the delays individually.

Table 4-3: Arrival delays (unit: minute) of all trains at the destinations in both E 4-1 (UISPR) and E 4-2 (PCDPR).

Experiment	I(1)	I(2)	I(3)	I(4)	I(5)	I(6)	J(1)	J(2)	J(3)	J(4)	J(5)	Sum	Computation time (s)
E 4-1	20	146	56	32	44	4	25	51	109	31	30	547	310
E 4-2	31	85	14	20	40	0	17	37	108	7	25	384	19
Difference	-11	61	42	11	4	4	8	14	1	24	5	163	298

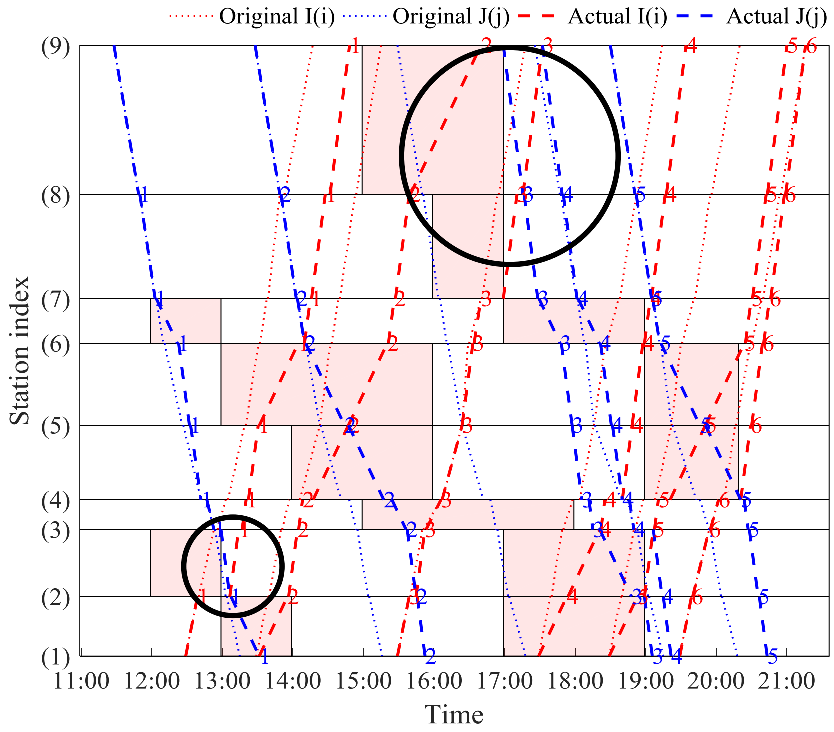
The detailed rescheduled results by UISPR and PCDPR are shown in Figure 4-4 (a) and (b), respectively. The outbound trains are running from station (1) to station (9), while inbound from station (9) to station (1). The originally scheduled train trajectories are represented by dotted lines, and the actual ones by dashed lines. The numbers alongside the trajectories are indexes of trains (red for outbound and blue for inbound).

The circles in Figure 4-4 mark out the two major differences between the experiment results. The smaller circle in Figure 4-4 highlights the different rescheduled trajectories for trains I(1) and J(1) on track segment 2 (between station 2 and station 3). In E 4-1, without prior knowledge of the weather impact, following the original plan, train I(1) went through the segment first while following the temporary speed restriction.



(a)

This then led to a series of knock-on effects, including delayed departure of train J(1) from station 2 and late arrival at its destination (i.e. station 1), and then delayed departure of train I(2) from its origin (station 1), and further the delayed departure of train J(2) at station 5. For these affected trains I(1), I(2), J(1) and J(2), the delays at the destinations are 20, 146, 25 and 51 minutes, respectively.



(b)

Figure 4-4: Planned and actual train trajectories by: (a) E 4-1 using UISPR and (b) E 4-2 using PCDPR.

In E 4-2, while taking account of all the weather events, train J(1) is rescheduled to go through segment 2 before I(1) and depart after the weather impact to avoid the temporary speed restriction. The delay of J(1) at the destination is then 25 minutes, and that of trains I(1), I(2) and J(2) are 31, 85 and 37 minutes, respectively. Compared to E 4-1, E 4-2 leads to 11 minutes more delay on I(1), but 83 minutes less delay in total on trains I(2), J(1) and J(2).

Similarly, as highlighted by the larger circle in Figure 4-4, the two experiments are significantly different in dealing with the conflict between train I(2) and train J(3) on track segment 8 between station 8 and 9. In E 4-1, train I(2) has to wait at station (8) for entering segment 8 until J(3) leaves this segment; while in E 4-2, train I(2) is scheduled to traverse this segment before train J(3). This results in E 4-2 generating in total 91 minutes less delay on trains I(3), I(4), I(5), I(6), J(3), J(4) and J(5), compared with E 4-1.

In E 4-1, the rescheduling is conducted 26 times, using a computation time of 310s in total, which is on average 12s per rescheduling. E 4-2 conducts rescheduling only once using 19s. It is reasonable that PCDPR is more time-consuming than the average of each UISPR rescheduling as more constraints are considered in the former; however, through the whole operation period from 11:00 to 21:00, the PCDPR consumes less computing resource than the overall UISPR process and performs better in reducing the passenger delays.

4.4.3. A rolling PCDPR with partial information

Section 4.4.2 shows that by taking account of forecasted weather disturbances in timetable rescheduling, our proposed PCDPR method can result in less delay than the conventional UISPR method. In terms of the advanced information on disturbance events, PCDPR has full information while UISPR has no information.

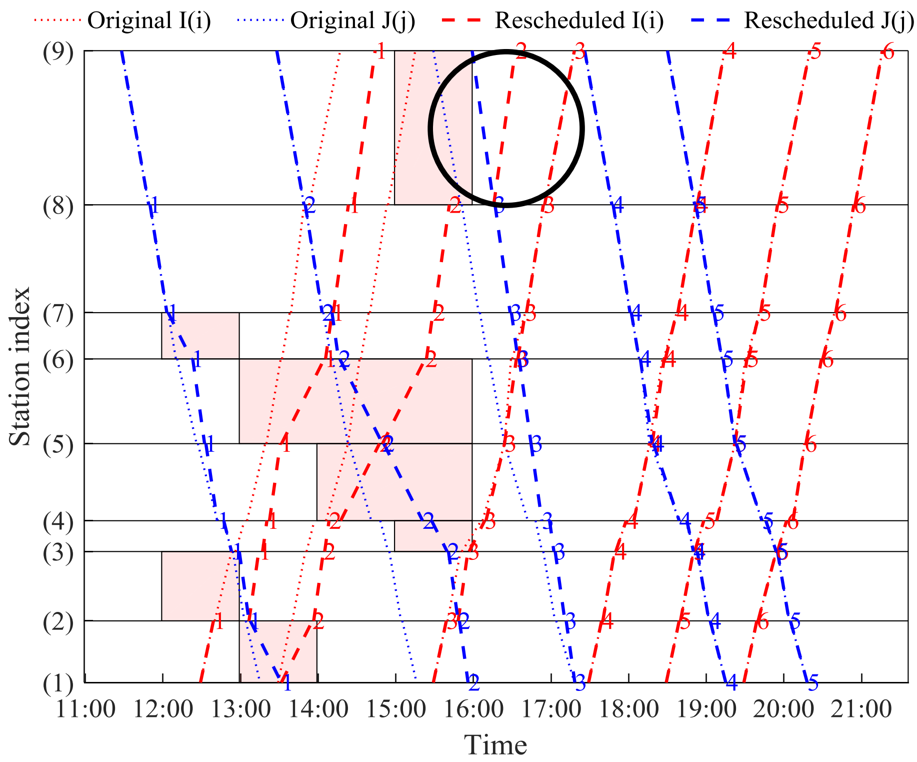
However, as the weather forecast is more accurate when it is closer in time. Therefore, we further consider a rolling PCDPR as when a shorter-term weather forecast is available.

We use the same line as that in Section 4.4.2, and consider the same 10-hour period between 11:00 and 21:00; the difference is that, although the rescheduling is still for the whole 10-hour operation period, now the weather forecast is made (or accurate enough) only for the next five hours. We conduct a new Experiment 3 (E 4-3) where the PCDPR is conducted twice: one at 11:00 when the weather forecast for 11:00-16:00 is available, and the other at 16:00 when the weather forecast from 16:00 to 21:00 is available.

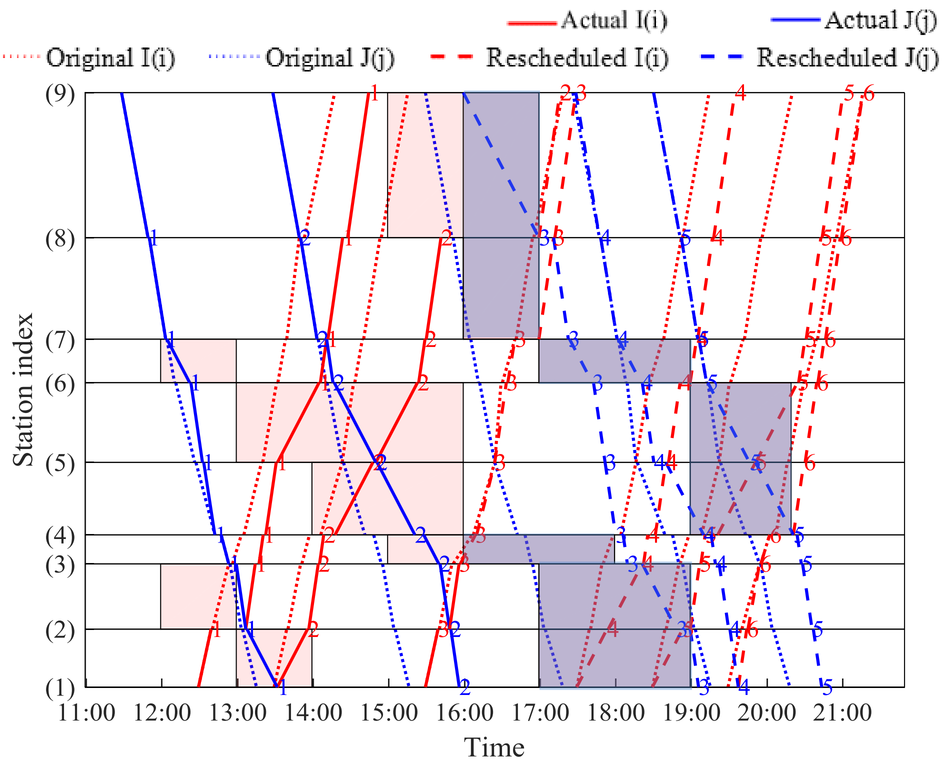
Given the first weather forecast for 11:00 – 16:00, the rescheduling result is shown in Figure 4-5(a).The legends are the same as in Figure 4-4. Notably, as highlighted by the circle, different from the result of E 4-2, train J(3) is scheduled to depart after 16:00 as it is assumed that the weather impact will end at 16:00, and train I(2) to depart from station 8 after train J(3) passed station 8.

As the new weather forecast is revealed at 16:00, the PCDPR is rerun, and the result is shown in Figure 4-5(b). The solid lines in Figure 4-5(b) represent the trajectories which have already commenced before 16:00; the blue shaded rectangular areas mark the weather events forecasted only at time 16:00 and are only considered in the second PCDPR.

In Figure 4-5(b), due to the weather impact on segment 8 from 16:00 to 17:00, train J(3) runs through a weather-impacted zone and arrives at station 8 late. This further leads to departure delay for train I(2) from station 8 and its subsequent arrival delay at station 9.



(a) First rescheduling at 11:00 considering weather data from11:00 to 16:00



(b) Second rescheduling at 16:00 considering weather data from16:00 to 21:00

Figure 4-5: E 4-3 rescheduling results from PCDPR. In (b), the pink shaded areas represented the weather events considered in the first rescheduling period, while the blue areas the weather events considered in the second rescheduling period.

The final delays for individual trains at their terminals in E 4-3 are shown in Table 4-4. Compared with the delays from E 4-2, E 4-3 results in higher overall delays. This is expectable as the full (accurate) information is always more beneficial than the partial information.

Table 4-4: Arrival delay minutes of each train at their destinations by E 4-2 and E 4-3

Experiment	I(1)	I(2)	I(3)	I(4)	I(5)	I(6)	J(1)	J(2)	J(3)	J(4)	J(5)	Sum	Computation time (s)
E 4-3	27	121	10	20	40	0	17	40	108	22	25	430	24
Difference to E 4-2	-4	36	-4	0	0	0	0	3	0	15	0	46	5

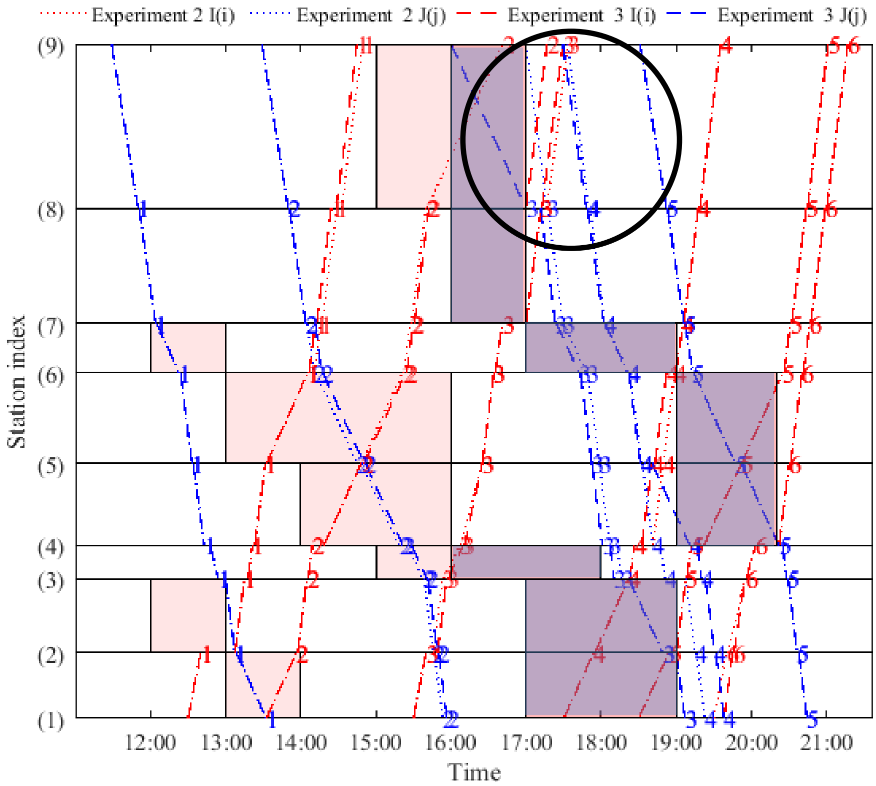


Figure 4-6: Comparing trajectories of E 4-2 and E 4-3

This can be shown in detail in Figure 4-6 where the rescheduled results from E 4-2 and E 4-3 are displayed together. It can be seen that, in dealing with the conflict between train I(2) and train J(3) on segment 8, in E 4-2, train I(2) is scheduled to go first while in E 4-3, J(3) is scheduled to go first, which result in train I(2) getting 36 minutes more delay than E 4-2. This is due to in E 4-3, the order of train I(2) and train J(3) is already decided in the first schedule which did not consider the weather impact from 16:00 and train J(3) is already scheduled to depart after 16:00. When the second rescheduling is conducted at 16:00, time already ‘wasted’ and train J(3) can only depart from 16:00 and operate under weather impact, which result in 36 minutes knock-on delay at the destination.

The computational time for the first and second rescheduling are 12 seconds in E 4-3. The total computational time is 5 seconds more than E 4-2 which is again reasonable, as there are some more calculation in E 4-3, which is for the second half timetable in the first rescheduling.

4.4.4. Sensitivity of the rescheduling results to the spread of adverse weather

In this subsection, we examine the sensitivity of the rescheduling results to the spatiotemporal spread of the adverse weather under different numbers of weather affected zones, and compare the gains of our proposed PCDPR approach over the traditional UISPR approach.

We use the same Cambrian Line and examine 12 different categories indexed 1 to 12 according to the number of weather impacted zones considered. Each category N has 20 randomly-generated test cases, and each test case has N randomly-generated weather impacted zones. The zones in different test cases differ in the spatiotemporal span of the zone as well as the level of temporary speed restriction applied. The test scenarios are randomly generated to represent the dynamic and stochastic nature of the weather disturbances, each tested case in category N is generated by the following two-step approach:

Step 1: randomly choose N segments (a segment can be chosen multiple times) on the line to be the impacted areas, and randomly choose a temporary speed limit from the following two values, 20 mph and 40 mph, to represent the different severity level of the weather disturbance.

Step 2: randomly assign a start time and an end time to each of the above-chosen segments. After such random assignment, it is possible that a same track segment may be impacted by two or more weather events whose durations may overlap; in this case, methods will be deployed to ensure no temporal overlapping among them.

We conduct both the UISPR and PCDPR for each of the 240 test cases generated. Figure 4-7 shows the difference of the total arrival delays between UISPR and PCDPR (i.e. delay of UISPR minus delay of PCDPR) for all of the 240 cases. PCDPR resulted in 184 minutes less (about 21%) delay than UISPR on average. And all the PCDPR gives not more delay than the UISPR. There is a general trend that the more adverse weather events, the bigger gains by adopting the PCDPR approach using forecasted disturbances than reacting to already-happened disturbances as in the UISPR approach. Notably, there are 31 cases where PCDPR and UISPR result in the same total delay. It is because a) the trains do not go through the weather-impacted zones when they follow the original timetables and thus not be disturbed, or b) UISPR and PCDPR make the same decision, or c) they make different decisions but lead to the same total delay.

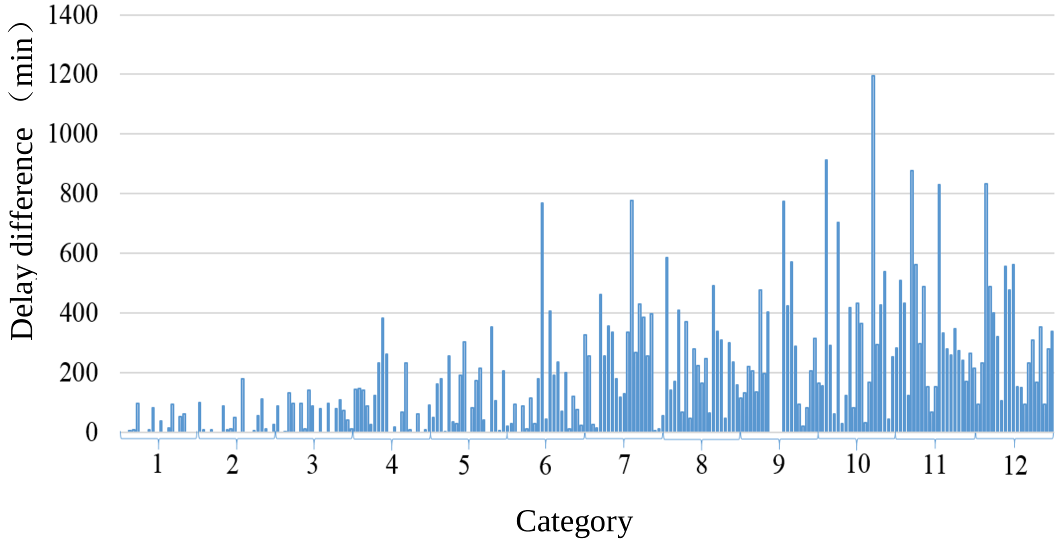


Figure 4-7: The difference in the total arrival delays between UISPR and PCDPR

4.5. Summary

In this chapter, we first introduced the existing single-track rescheduling mixed integer programming models and identified the research gaps. We considered a single-track railway line and formulated PCDPR as a mixed integer linear programming (MILP) problem, which considers the general constraints for train rescheduling (such as departure and arrival times, headway, overtaking, capacity and conflict avoidance), as well as new constraints corresponding to the weather impacts. According to the severity of the forecasted weather events and the speed restriction guidance of the rail industry, the reduced speed limit is applied to each weather-impacted zone, and trains passing through these zones will have to follow the reduced speed limits.

The effectiveness of the proposed PCDPR was demonstrated on the Cambrian Line in the UK. Compared with the traditional UISPR which is conducted after a delay has happened, our proposed PCDPR led to lower overall delays by incorporating the forecasted weather disturbance. And the complete information on weather forecast is better than partial information. The sensitivity analysis is also conducted

The case studies demonstrate that by using the general-purpose MILP solvers, our method is able to solve small-scale rescheduling problems in feasible time and achieve optimal solutions. However, for large-scale problems, the general-purpose solvers may not be efficient enough so specialised solution algorithms may be developed. To take the proposed method from concept to practical operations on larger and more complex networks, more efficient model formulation and solution algorithms are needed.

Chapter 5 Single-track rescheduling for weather-induced PCDP with genetic algorithm

5.1. Introduction

When the computation time is very limited, the optimal solutions might not be achievable and thus feasible (or near optimal) solutions would be accepted. In this chapter, for the train rescheduling problem with weather-induced PCDP, we propose a genetic algorithm (GA) to generate near optimal results within short computation time. GAs are randomised search algorithms inspired by natural genetics and natural selection. They mimic the “survival of the fittest” rule to generate better individuals (solutions) generation by generation. By assigning high choosing probability to individuals with high fitness for survivability, the possibility of reaching the optimal solution is increased.

In the remainder of this chapter, we will first review the relevant rescheduling literature that applied the GA method. This is followed by a customised GA for train rescheduling and the corresponding process applied to PCDPR. Last but not least, we conduct the numerical tests to show the computational efficiency and the solution quality of GA.

5.2. GA for train timetable (re)scheduling

Just like in biology, a computer based GA also has genes, chromosomes and generations. A gene is the smallest element in GA and represents a certain decision in the train timetable rescheduling algorithm. A chromosome is composed of genes, which represents an entire solution. A generation is a group of chromosomes representing many potential solutions for the problem. Genes and chromosomes determine the solution quality and computing efficiency for GA. In this subsection, we will review some of the GAs used in train timetable (re)scheduling and specifically focus on how people encode genes and form the chromosomes.

Salim and Cai (1995) conducted the pioneering works which applied GA to railway traffic control. They used *n* by *m* chromosomes to represent the stopping patterns of *n* stations with *m* trains. In each chromosome, a gene is a binary and represents the state of a train in a station, where 0 means stopping and 1 means passing. Due to the computation capability limitation at that time, a problem of 12 stations and 9 trains needs 1.5 hours to reach a feasible solution.

Chang and Chung (2005) developed a GA which used matrices to represent chromosomes. The row of chromosome indicated trains and the columns indicated the stations. Each of the elements in the matrices consisted of dwelling time, arrival time and the departure time of train *i* at station *j*. It took about 20 min to get a feasible solution on a personal computer with a Pentium III-800 CPU.

Tormos et al. (2008) considered train timetabling problem as a Job-Shop problem and solved it with GA. The solution was encoded as precedence feasible list pairs (t, t_i^t), where *t* represents train *t* and t_i^t represents the i^{th} track section of its journey. Real-world cases considering adding new trains to the line were tested based on Spanish railway lines ranging from 96 km to 401 km. The results showed that GA outperforms the random and parameterised regret biased based random sampling methods in terms of the average deviation to the optimal solution.

Dündar and Şahin (2013) considered conflicts of the number of all the opposite train pairs, as well as potential conflicts from same direction train pairs. Meanwhile, to constitute a conflict free schedule, they used a binary variable to guide the train priority for each conflict. All these binary variables constituted a chromosome.

Instead of representing each binary variable as a gene, Higgins et al. (1997) used variable-length chromosomes to reduce the length of the chromosome. One gene included three pieces of data: the train delayed, the train with priority, and the corresponding track segment. The genes which did not impact the fitness were eliminated. This caused the issue that a produced offspring solution might not represent a fully resolved schedule. To resolve it, new genes would be added when the offspring had more conflicts than the parents, and the infeasible offspring was replaced by one of the parents.

5.3. A genetic algorithm formulation of the PCDPR problem

The generation updates between parents and offspring are dependent on process of evolution, i.e. selection, crossover and mutation. In this section we define the elements and the process of our proposed GA model. We will start by introducing the concept of conflict resolution matrix, and then a genetic representation for scheduling, followed by the operators considered in this thesis. Finally, we introduce the processes of generation alternation, i.e. how the GA solutions are improved.

5.3.1. Train conflict resolution

As introduced in Section 3.2.2 of Chapter 3, in single train scheduling problem, conflicts on the single-track segment are the key issues to solve. In this study, we introduce the concept of Conflict Resolution Matrix $Mtrx^n$ to represent the n^{th} solution for all the conflicts in the system.

In the matrix, the subscripts of rows and columns represent the outbound and inbound trains, respectively. The value of each elements $Mtrx^n(i,j) = g^n{}_{i,j}$ represents the priority between outbound train *i* and inbound *j* when they conflict with each other. If $g^n{}_{i,j} = 1$, train *i* will have the priority to travel the conflicted zone, whilst if $g^n{}_{i,j} = 0$, train *j* will have the priority.

We take a small single-track line as an example to illustrate the concept of Conflict Resolution Matrix. As shown in Figure 5-1, there are two outbound trains and three inbound trains, where each outbound train *i* has a potential conflict $C_{i,j}$ with each inbound train *j*. If there is a conflict between outbound train *i* and inbound train *j*, $C_{i,j} = 1$, otherwise, $C_{i,j} = 0$.

Figure 5-2 shows the conflict free timetable with two different Conflict Resolution Matrices. In Figure 5-2 (a), all the outbound trains have the priority, i.e. $Mtrx^1$=[1,1,1;1,1,1]. As outbound train I(1) and inbound train J(2) meet at station (4) which has enough tracks for meeting, there is no conflict between them, i.e. $C_{1,2} = 0$. The $g^1{}_{1,2} = 1$ is not effective. In Figure 5-2 (b), we change $Mtrx^1$ to

$Mtrx^2$ by letting $g^1{}_{1,3} = 0$, which means J(3) has the priority to use the segment when conflicting with I(1).

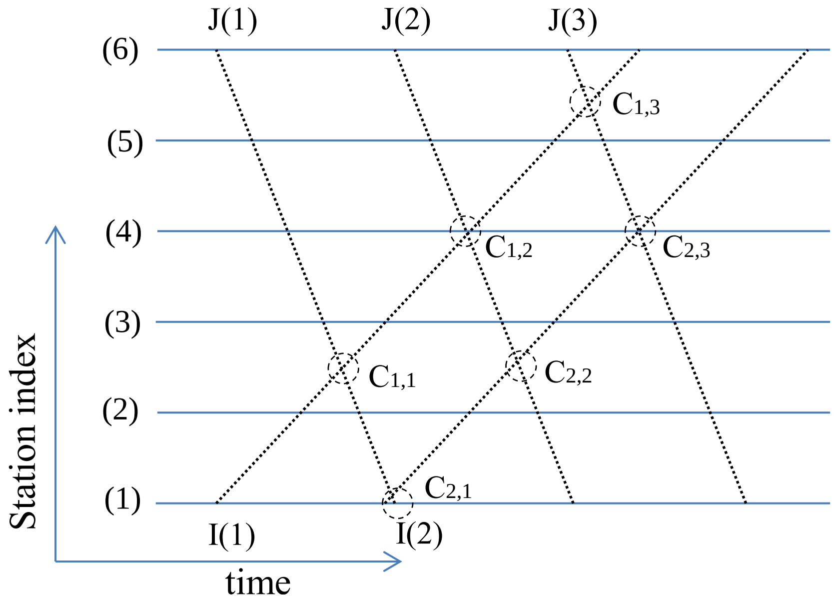
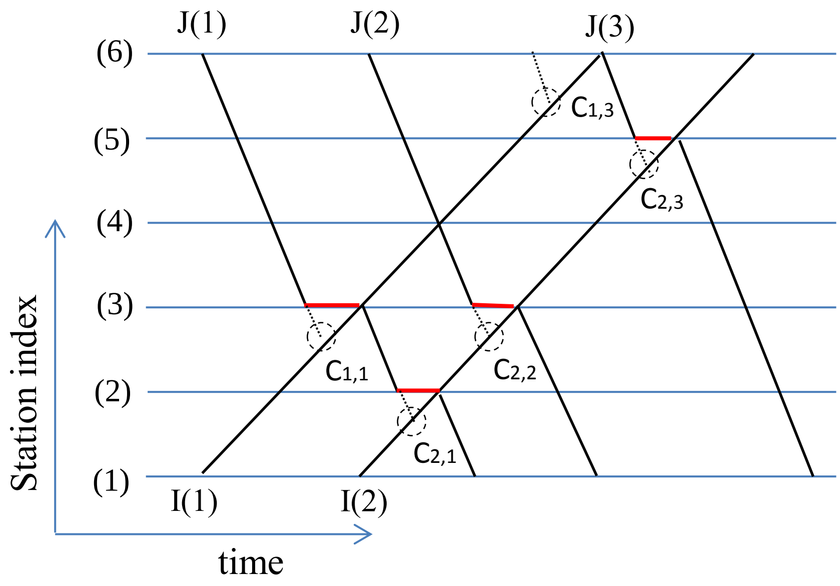
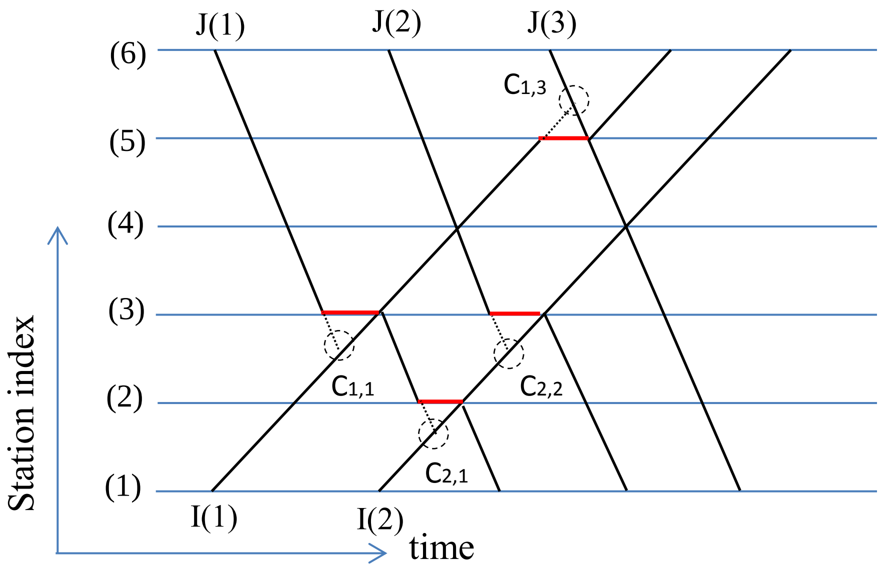


Figure 5-1: All the potential conflicts on a single-track line



(a) $Mtrx^1$=[1,1,1;1,1,1]



(b) $Mtrx^2$=[1,1,0;1,1,1]

Figure 5-2: Conflicts free diagrams with (a) Mtrx^1 =[1,1,1;1,1,1] and (b) Mtrx^2=[1,1,0;1,1,1]

5.3.2. Genetic representation of train schedules

Gene $g_{i,j}$: In the studied single-track railway system, the conflict $C_{i,j}$ will happen when two opposite trains claim the same track. The gene $g_{i,j}$ represents the solution for the conflict. Note that we only study homogeneous trains in this chapter, which means all trains are considered the same type, i.e. no trains have priority and all trains have the same power and braking system. Therefore, trains travelling in the same direction do not need to, nor can they, overtake each other.

Chromosome $Mtrx$: In the rescheduling context, as genes stand for the solution for each individual conflict, chromosome then represents the feasible schedule. We structure the Conflict Resolution Matrix $Mtrx$ as the chromosome. For the solution of each individual conflict, $Mtrx$ *(i,j)*= $g_{i,j}$.

Fitness function $f(Mtrx)$: Fitness f is the indicator for the "health" of the chromosome, which corresponds to the value of the objective function. The healthier a chromosome is, the higher chance the individual will be selected as parent for breeding the offspring. We use the total system delay as the fitness function. The less the delay, the better the solution, the higher chance the individual will be selected.

$$f(Mtrx) = \sum_{i \in I} \Delta t_i + \sum_{j \in J} \Delta t_j \tag{5-1}$$

Generation: In our case, the generation is a set of feasible train conflict-resolving plans. Fitter individuals in each generation will be picked out as parents with higher probability to generate offspring through a group of operation which will be introduced in the next section.

The end condition of the algorithm is normally set as when the rate of convergence reaches a certain level, or a maximum number of iterations have been executed. In this research, we choose both of them as end condition. The algorithm will be stopped as soon as any of the termination criteria is met.

5.3.3. Operators in the GA for train rescheduling

By proper initialisation including the population number (i.e. the total number of chromosomes in a generation) and stopping condition, the first generation of GA can be produced. It will then go through the operations including selection, crossover and mutation to produce the new generation. These operations are set as follows (Dündar and Şahin 2013).

5.3.3.1. Selection

Selection is the operator that individual genomes are chromosomes from a generation for later breeding. We use *roulette wheel selection* method to conduct it. To retain the best genes, we also introduce an elites retaining step. The implementation detailed is as follows:

Step 1: Sort all the *N* individuals in the current generation by descending fitness values, where *N* is the number of population in each generation.

Step 2: Retaining the *Ne* (which is a pre-set number, *Ne*<*N*) best (smallest) individuals for the next generation.

Step 3: Calculate the selection probability $P(Mtrx^n)$ of each individual *n*.

$$P(Mtrx^n)=\frac{1/f(Mtrx^n)}{\sum_{n \in N} 1/f(Mtrx^n)} \tag{5-2}$$

The individual chromosome having the lowest total delay would have the highest probability to be selected.

Step 4: Calculate the accumulated normalised fitness values for each individual chromosome. The accumulated fitness value of an individual is the sum of its own fitness value plus the fitness values of all the previous individuals.

Step 5: Select the rest *N-Ne* individuals. Generate a random number *Rs* between 0 and 1. Select the individual whose accumulated normalised values are greater than or equal to *Rs* to the next generation. If the total selected number of the next generation equals to *N-Ne*, stop selection and go to the crossover operator.

5.3.3.2. Crossover

We choose the single point crossover method. To begin with, the individuals in the new generation are paired randomly. For every pair, a random number *Rc* between 0 and 1 is generated; for both chromosomes, the proportion bigger than *Rc* is exchanged with each other so that new chromosomes are generated.

5.3.3.3. Mutation

A mutation probability R_m is set before the GA starts. For each of the new chromosome, a random number between 0 and 1 is generated to decide whether to conduct the mutation. If the random number is bigger than the mutation probability, this chromosome will be kept to the next generation without any change; otherwise, the mutation will be conducted. In our case, for each of the gene in the chromosome, 1 will be mutated to 0, and 0 will be mutated to 1.

5.3.4. The process in generating chromosome

Based on the concepts and operations introduced above, a complete GA can be constructed to generate improved solutions and output the best feasible solutions.

Step 1: Initialisation: randomly generate *N* chromosomes (feasible schedules);

Step 2: Interpret each of the chromosome and output the fitness value; we will further introduce how to interpret the chromosome in Section 5.4.

Step 3: If the end condition is satisfied, goes to Step 5; otherwise, goes to Step 4.

Step 4: Conduct the Selection, Crossover and Mutation operators in turns to generate the offspring and goes to Step 2.

Step 5: Output the best solutions among the population.

5.4. Rescheduling process with chromosome

Section 5.3 indicates how the chromosome is generated between generations. In this section, we will describe how we interpret a chromosome to work out a feasible timetable mentioned in Step 2 of Section 5.3.4. We will firstly illustrate the model constraints, i.e. how we calculate each train's time point and check whether the constraints mentioned in Section 4.3.2 are satisfied. Then the interpretation steps are illustrated in 5.4.2.

5.4.1. Model constraints

We assume all the trains must satisfy the constraints 1)- 9) in Section 3.2.2. As the rerouting is not involved in this Chapter, we don't consider constraint 10) here. The definition of notification is the same as in Section 4.3.

(a) Arrival time and headway on segments

The arrival time for train k at the downstream station of link *c* is equal to the maximum number between the departure time at the upstream station plus the travel time on link *c* and the arrival time of its the preceding train (if any) plus the headway.

$$a'_{c,k} = \max[d'_{\mathrm{c},k} + t_{c,k}, a'_{c,k-1} + t^h] \qquad \forall\, c \in E, k \in K \tag{5-3}$$

(b) Departure time and headway on segments

The departure time for train k at the upstream station *s* of link *c* is equal to the maximum number between the arrival time at that station *s* plus the required minimum waiting time and the departure time of its preceding train (if any) plus the headway.

$$d'_{c_s,k} = \max[a'_{c_{s-1},k} + t^w_{s,k}, d'_{c_s,k-1} + t^h] \ \ \forall\, c \in E, k \in I\backslash\{1\}, s \in S\backslash\{1\} \tag{5-4}$$

$$d'_{\bar{c}_{s-1},k} = \max\left[a'_{\bar{c}_s,k} + t^w_{s,k}, d'_{\bar{c}_{s-1},k-1} + t^h\right] \quad \forall\, c \in E, k \in J\backslash\{1\}, s \in S\backslash\{1\} \tag{5-5}$$

(c) Minimum departure headway check

As we don't allow overtaking in this model, for train $k \in K \setminus \{|K|\}$, the departure time of train $k+1$ must be no smaller than the departure time of train k plus the headway. If $d'_{c,k+1} < d'_{c,k} + t^h$, then we need to update the departure time of train $k+1$ with the following equation:

$$d'_{c,k+1} = d'_{c,k} + t^d \qquad \forall\, c \in E, k \in K \setminus \{|K|\} \tag{5-6}$$

(d) Minimum arrival headway check

Similar to **Minimum departure headway check,** for train $k \in K \setminus \{|K|\}$, if $a'_{c,k+1} < a'_{c,k} + t^h$, update the arrival time of train $k+1$ with the following equation:

$$a'_{c,k+1} = a'_{c,k} + t^a \qquad \forall\, c \in E, k \in K \setminus \{|K|\} \tag{5-7}$$

(e) Conflict check

For each outbound and inbound train pair, we need to check if they are conflict free. For the same segment, if $a'_{c_s,i} + t^{ad} < d'_{\bar{c}_s,j}$ and $a'_{\bar{c}_s,j} + t^{ad} < d'_{c_s,i}$, then there is a conflict between outbound train i and inbound train j on segment l_s, the conflict resolution will be needed, which will further be explained the conflict resolution approach in Section 5.4.2.

(f) Weather effect check

For train k traveling on segment l_s, if the equation (3-4) is satisfied, $v'_{c,k} = v_p$; otherwise $v'_{c,k} = v_{c,k}$

(g) **Running time**

Running time calculation equation for train k on link c is:

$$t'_{c,k} = x_c / v'_{c,k} \qquad \forall\, c \in E, k \in K \setminus \{|K|\} \tag{5-8}$$

5.4.2. Conflict Resolution with each chromosome

When we get the chromosome, i.e. the Conflict Resolution Matrix, the following process is used to solve all the potential conflicts and output the fitness value of the corresponding chromosome. We also draw a flow chat to illustrate the major steps in Figure 5-3.

For each Conflict Resolution Matrix $Mtrx^n$, the conflict resolution steps are introduced below, where the check conditions and calculate equations in bold letter was described in the 5.4.1.

Step 1: (initialisation) input the original schedule: for all the trains in the system, let $d'_{c,k} = d_{c,k}$ and $a'_{c,k} = a_{c,k}$; input the value of $Mtrx^n$ and weather effect set $Q = \{p_q | p_q = [l_q, b_q, e_q, v_q]\}$, go to step 2.

Step 2: for the first/next expected arrival event in the system, do the **Weather effect check**, if impacted go to step 2.1, otherwise go to step 2.3.

Step 2.1: update its **Running time**, and **Arrival time** considering the weather impact speed limitation. Check if all the trains have arrived at their final destination, if yes, go to Step 5, otherwise, go to step 2.2.

Step 2.2: for all the remaining of downstream stations of current train k, update the **Departure time** and **Arrival time.** If all the affected trains in step 2.1 are updated, go to step 3; otherwise go back to step 2.1.

Step 2.3: check if all the trains have arrived at their final destination, if yes, go to step 5, otherwise, go back to step 2.

Step 3: from the current segment to all the downstream stations, check the headway of train k's following trains, go to step 3.1

Step 3.1: do the **Minimum arrival headway check**, if satisfied, update the **Arrival time** and the **Departure time** of the downstream stations, go to step 3.2.

Step 3.2: do the **Minimum departure headway check**, if satisfied, update **Departure time** and the **Arrival time** of the rest downstream stations, if all the following trains in step 3 are checked, go to step 4; otherwise, go back to step 3.

Step 4: for all the segments and all the trains, do the **Conflict check**, if there is no conflict, go to step 2. otherwise go to step 4.1

Step 4.1: for the earliest (next) $C_{i,j}$, check the corresponding segment l, and the value $Mtrx^n(i,j)$, if $Mtrx^n(i,j) = 1$, then let $d'_{\bar{c},j} = \max[a'_{c,i}, d_{\bar{c},j}, d'_{\bar{c},j-1} + t^h]$; update the **arrival time** and **departure time** of train j in the downstream stations and also its following trains according to minimum arrival/ departure headway check, go to step 4.2, otherwise if $Mtrx^n(i,j) = 0$, then let $d'_{c,i} = \max[a'_{\bar{c},j}, d_{c,i}, d'_{c,i-1} + t^h]$; update the **arrival time** and **departure time** of train i in the downstream stations and its following trains according to minimum arrival/ departure headway check, go to step 4.2.

Step 4.2: if it is the final conflicted train pairs, go to Step 2, otherwise, go to step 4.1.

Step 5: calculate the **Fitness of the chromosome** and output the value.

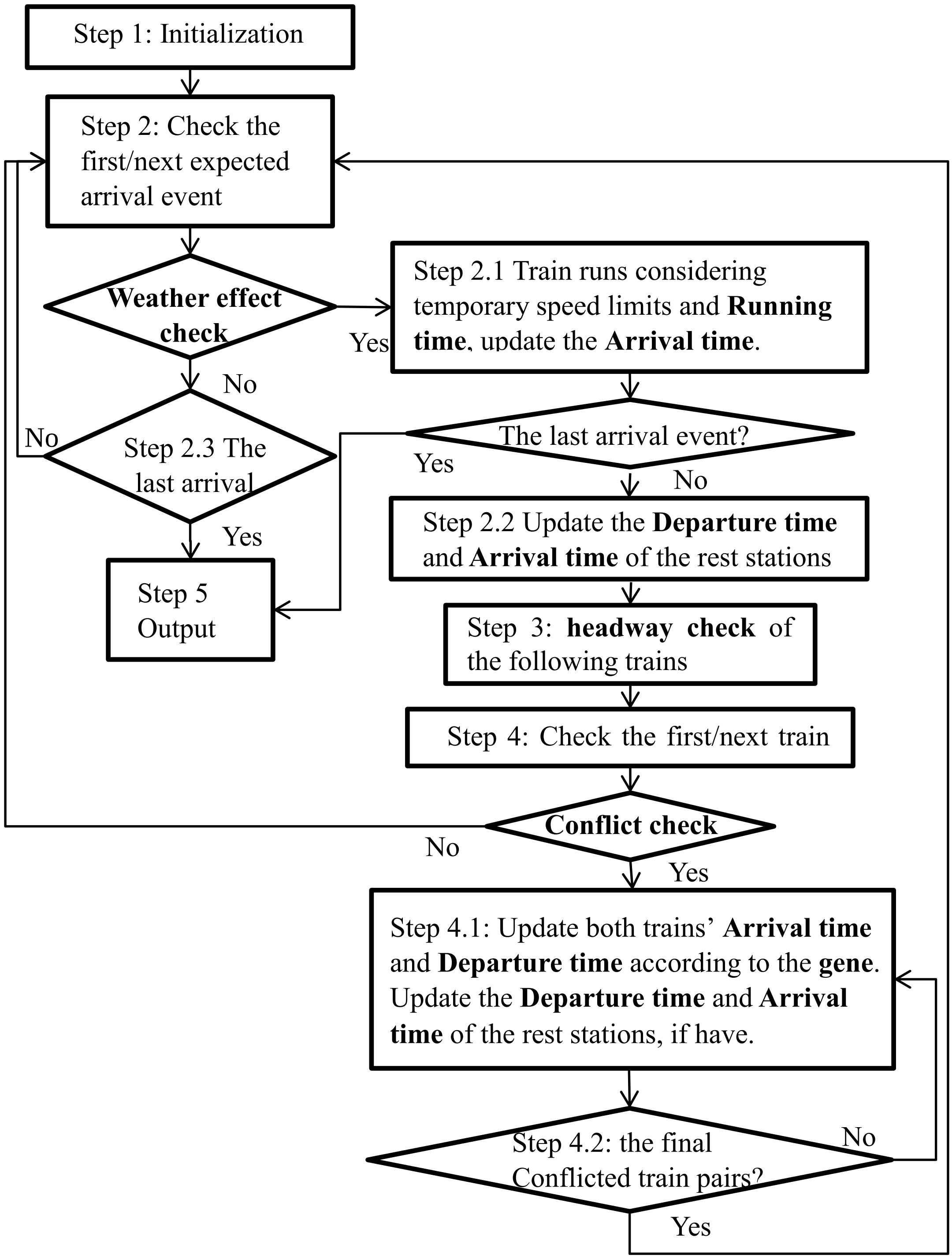


Figure 5-3: Flowchart of rescheduling according to each chromosome

5.5. Experiment and results

We test this GA with the same case described in Section 4.4.1 and solve it using the same computer. We name this case Experiment 5-1 (E 5-1). The population number is 30 in each generation. Two criteria were developed to terminate the algorithm. One is when the ratio of the difference between the highest and the lowest fitness values to the lowest one within a generation is less than 5%. Another is when the maximum generation number reaches 100. The algorithm will be terminated as soon as either of the termination criteria is met (Dündar and Şahin 2013). Since mutation probability in real life is really low, we set the mutation probability as 0.001.

The best fitness value, i.e. the total delays, in each generation is shown in Figure 5-4. In the 26th generation, the first terminate condition is satisfied and the fitness value equals to 496 minutes. The computation time is 2 seconds. The arrival delay of each train and its difference to delay in E 4-2 are shown in Table 5-1.

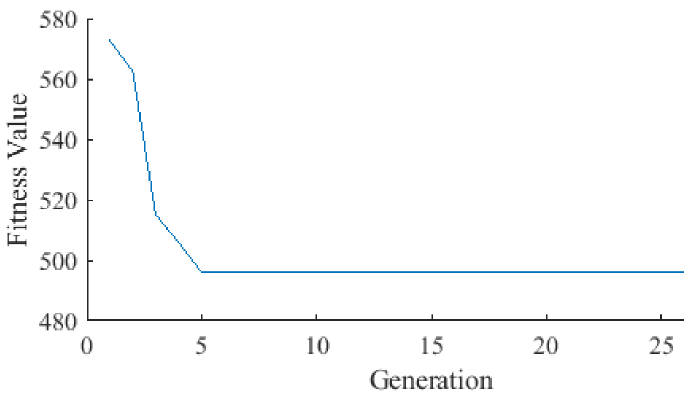


Figure 5-4: Performance of GA

Table 5-1: Arrival delay minutes of each train at their destinations in E 5-1

Experiment	I(1)	I(2)	I(3)	I(4)	I(5)	I(6)	J(1)	J(2)	J(3)	J(4)	J(5)	Sum	Computation time (s)
E 5-1	20	83	48	24	56	3	64	34	117	24	24	496	2
Difference to E 4-1	0	-64	-8	-8	12	-1	39	-17	8	-7	-6	-51	-308
Difference to E 4-2	-11	-3	34	4	16	3	47	-3	9	17	-1	112	-17

The E 5-1 (GA) generates 112 minutes (about 29%) more delay than E 4-2 (PCDPR) but uses 17 seconds (about 89%) less computation time. It is better than E 4-1 (UISPR) in both delay and computation time, i.e. 51 minutes (about 9%) less delay and 308 seconds (99%) less computation time. This means that although GA does not generate the optimal result, the delay is still less than UISPR, and computation time is much less than UISPR and PCDPR.

The detailed rescheduled timetable is shown as in Figure 5-5. The major differences to E 4-2 (the optimal result) are marked out by two black circles. The smaller circle highlights the different rescheduled trajectories for trains J(1) and I(2) on track segment 1 (between station 1 and station 2). In E 4-2 in Section 4.4.2, train J(1) goes through segment 1 before train I(2), while in E 5-1 the order is reversed. This results in train J(1) generating 47 minutes more delay in E 5-1 than in E 4-2.

The bigger circle highlights the different rescheduled trajectories for trains I(3) and J(4) on track segment 8 (between station 8 and station 9). In E 4-2, in Section 4.4.2, train I(3) goes through segment 8 before train J(4), while in E 5-1 the order is reversed. This then leads to I(3) getting 34 minutes more delay in E 5-1 than in E 4-2.

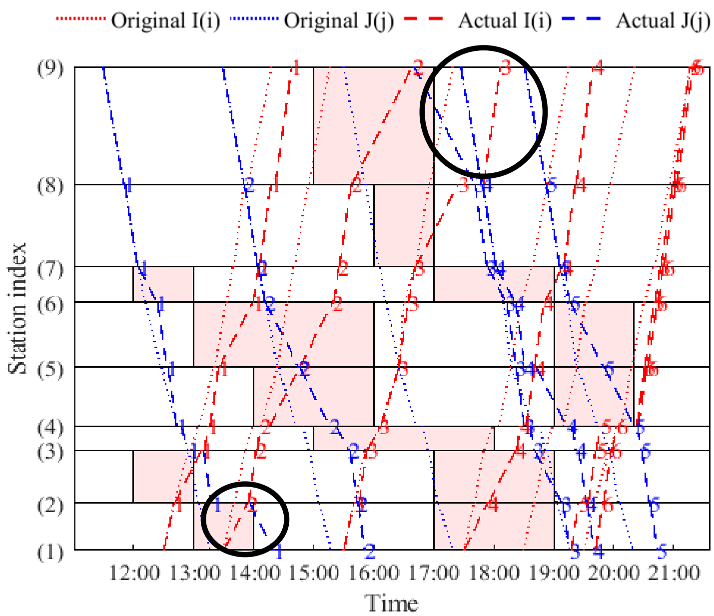


Figure 5-5: Actual operational trajectories under the proposed algorithm

We further run another set of GA experiments to explore if it could generate better results. In the new experiments, we let the GA run 1000 generations. The results of several runs indicate that the algorithm started to converge after a few dozen generations and the result is the same as in E 5-1 (496 minutes). This means for such a problem scale which is a 11-train and 9-station single track line, a GA with elite remaining process can produce near optimal solutions within much shorter computation time than solving the MILP in chapter 4 (using CPLEX).

5.6. Summary

In this Chapter, we first reviewed the GA for train timetable (re)scheduling. To get feasible solutions quicker than MILP solver, we designed a GA in Section 5.3. The binary Conflict Resolution Matrix is introduced and set as chromosome in GA. The rows and columns in the matrix correspond to the outbound and inbound trains, respectively. The values of the elements indicate the conflict resolution decisions for train pairs. The GA operations, i.e. selection, crossover and mutation, are conducted between generations to generate a better solution.

The experiment results indicate that the PCDPR-GA in E 5-1 generates 496 minutes delay, which is about 29% more delay than the optimal solution of PCDPR-MILP in E 4-2. However, the computation time is only 2 seconds, which is 89% less than that of E 4-2. Although PCDPR-GA does not generate the optimal result, it is better than UISPR in E 4-1; more specifically, it produces about 9% less delay and uses 99% less computation time. The advantage of GA in computation time will be more pronounced on large networks, as it will not exponentially increase, but MILP will.

Chapter 6 Rerouting and rescheduling under disruption

6.1. Introduction

Chapter 4 presented the advantages of PCDPR approach for train timetable rescheduling. However, it dealt with only adverse weather conditions which only cause speed reduction to the railway system. With more severe weather, the impacts on the railways may lead to not only speed reduction, but also track blockage according to the operation guidance introduced in Section 2.4.2. Under this situation, services may be cancelled, rerouted, detoured, and/or rescheduled. The dispatchers need a tool to help them to make specific decisions for each different weather situation so that the total loss could be minimised.

In this chapter, we consider the operational responses to severe weather conditions which cause speed limitation as well as partial and temporal track blockage to the network. We also adapt the weather modelling methods mentioned in Chapter 3. For the track blockage, we set the speed limit to zero so trains cannot move on the blocked sections. As the blockage caused by weather might be just for a short period, e.g. an hour or two, we do not consider returning back as a control option, but only rescheduling and rerouting. To reach globally optimal, we optimise rerouting and rescheduling simultaneously instead of sequentially.

6.2. A train rerouting and rescheduling problem

Though existing rerouting literature rarely consider disruptions caused by weather, we can still find many caused by maintenance, infrastructure failure, or deviant on the track, etc. A train route is a sequence of links. When any one of the links is blocked, the route is not passable. To continue the service, dispatchers need to identify alternative passage, or instruct trains to wait until the blockages are cleaned. Either way, rescheduling will also be required as trains can no longer follow their timetable. Literature dealing with the combination rerouting and rescheduling problems can be classified into two approaches: sequential or simultaneous optimisation.

In the sequential approach, Lee and Chen (2009) proposed an heuristic optimisation model and solved real-world instances with it. It firstly generated a simple initial solution and then iteratively improved the solution with a four-step process: (1) order trains on inter-station blocks; (2) assign trains to tracks in stations; (3) order trains on intra-station tracks; and (4) solve for the schedule. Between iterations, a threshold accepting rule was used to decide either accepting or rejecting the solutions

Pellegrini, et al. (2014) proposed a routing and scheduling mixed-integer linear programming formulation to tackle real-time traffic management when perturbation happened. The optimisation objectives are to minimise either individual train's maximum secondary delay or the total system secondary delay. In their system, routes did not include any intermediate stops. A two-step cycle was used to speed up the solution process. The first step rescheduling conducted the optimisation without considering the route changes. Based on the solution obtained, the rescheduling optimisation was performed with all possible train routes.

Among the simultaneous approaches, Rodriguez (2007) built a constraint programming model to solve rerouting and reordering problem at a junction. They firstly defined assignment constraints and sequence constraints, then they connected these two constraints together by using a third constraint. To reduce computing time, they relaxed the acceleration constraints to assume trains can reach any speed at no time. The disturbances they considered are initial delays of trains at certain stations.

Fang et al. (2017) studied a routing and scheduling problem with a time window for transporting hazmat. A mixed integer model considering risk threshold constraints was firstly built and a heuristic lower bounding scheme was proposed to solve the problem. The numerical tests showed that medium to large instances could be solved in several minutes.

Meng and Zhou (2014) developed a simultaneous rerouting and rescheduling model for the multiple-track train dispatching problems. The decomposition mechanism was applied through modelling track capacities as side constraints by reformulating them as a vector of cumulative flow variables. Then track capacities are further dualised through a proposed Lagrangian relaxation solution framework. Compared to the common sequential train rerouting and rescheduling approaches, the numerical experiments demonstrated the benefits of simultaneous train rerouting and rescheduling. However, their study did not consider the case of speed restriction, platform requirement for trains loading and unloading, or the occupation of bi-directional tracks by two opposite trains.

In Diego (2016), a microscopic approach was proposed to adjust the timetable for the planned maintenance activities. There are two differences between this maintenance model and our research. First, the disturbances due to maintenance and weather are modelled differently. In maintenance, there is only one directly impacted main section and one or two affected sections which run in parallel to the main one. One location is only impacted during one maintenance period. While weather-related disruptions, especially with a moving weather front, the impacts can be felt at different locations during different time periods. Second, the way of transferring maintenance to rescheduling model is different. In Diego (2016), all possible train routes are required input to the model; this requires significant effort in model initialisation. While in the PCDP modelling, only the forecasted weather locations and time periods are the required input into the model, which significantly simplifies the data requirement, and thus making it practically feasible to conduct a rescheduling response to abnormal weather effects in a relatively short time horizon.

6.3. Problem formulation

6.3.1. Problem description

When the weather impact reaches a severe level, according to the operation requirements, the impacted services need to be suspended in the impacted area. We abstract these suspensions as track blockages applied to the impacted period. When this happens, alternative control strategy may involve rerouting and rescheduling for the affected services.

We consider both local and global rerouting. The local rerouting means changing tracks locally, e.g. changing platforms at the same station or changing to parallel tracks if they are available at the same location, while global rerouting means trains changing to an entirely different route which may involve in passing different stations to the original ones.

The operation constraints to the model are the same as those described in Section 3.2.2. We set the constraints 1) to 8) as the constraints that must be satisfied while constraints 9) and 10) as those allowed to be broken with a certain penalty.

An important element of modelling rerouting is the topological structure of the modelled network. In this study, we use a mesoscopic representation of the rail network, which has features sufficient to reflect the track connections and track choices in the network, but not a microscopic model as we don't consider the gradient, signalling system, etc.

Model features

(A1) Network: The railway network is represented as nodes, edges and virtual links as shown in Figure 6-1. Nodes represent rail switches, the network entry and exit points, and connection points on platforms or passing loops. An edge is a track between two nodes and is nondirectional. A virtual link is a directional arc indicating possible travel direction on the edge; where an edge is associated with two opposite links. The stations and passing loops are represented as several points and edges as illustrated in Figure 6-1. We build a mesoscopic model in which the track lengths in stations are reflected and travel time in stations are not omitted either. Trains with loading and unloading tasks must stop at the edges aside platforms.

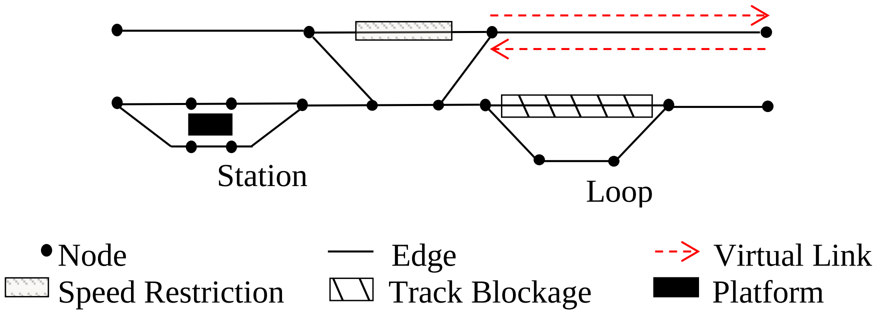


Figure 6-1: Illustration of a mesoscopic representation of a rail network.

(A2) Time: Trains are abstracted as dots with no length. A train’s arrival time at a link is the moment when it reaches the start node of the link and the departure time from a link is the moment when it leaves the end node of the link. Train's acceleration and deceleration are infinite, which means it will need no time to accelerate or decelerate to the operation speed. Trains cannot depart before their planned departure time from station tracks. Dwell time should be no shorter than the required minimum dwell time. Two trains can simultaneously travel on the same link in the same direction with a minimum headway time, while two opposite trains cannot travel at the same edge at the same time.

(A3) Weather: On the edge where adverse weather threshold is breached, a corresponding temporary speed restriction will be applied on both the associated links during the impacted time period. The weather-impacted speed restriction is mapped onto the whole section of a track between two nodes. If any part of a train's trajectory falls into the impacted zone, the train will have to operate under the speed limit while travelling on the impacted edges. On the edge where severe weather threshold is breached, the entire edge will be blocked during the predicted impact time. No trains would be allowed to pass through any link associated with this blocked edge. Impacted trains can either wait until the blocked track clear or reroute to other links. If rerouted, a penalty cost may be added to represent the penalty for the potential safety issue.

Under normal condition, each track can be traversed in only one predetermined direction. Under adverse weather condition when some tracks are blocked, trains will be allowed to change tracks and the unblocked tracks parallel to the blocked ones will be allowed to serve trains from both directions. As shown in Figure 6-2, this small network consists of eight nodes and eight edges. Blue Route 1 (node sequences 4-3-2-1) and red Route 2 (node sequences 5-6-7-8) are original routes for inbound and outbound trains, respectively. When track b between node 6 and node 7 is blocked, the outbound trains can travel on the track between node 2 and node 3, which is originally used by inbound trains only. The new route for outbound trains is Route 3 (node sequences 5-6-2-3-7-8). The timetable for outbound and inbound trains need be adjusted accordingly due to 1) the possible longer travel time between node 6 and node 7 for outbound trains, and 2) the capacity reduction for inbound trains on the edge between node 2 and node 3, due to the temporary use by outbound trains.

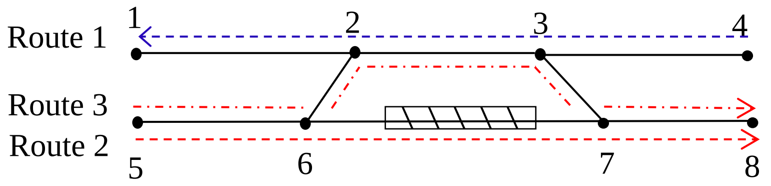


Figure 6-2: Illustration of a possible rerouting along small rail network.

Based on the abstraction above, we formulate the rerouting and rescheduling problem as a mixed integer liner programming (MILP) model which aims to minimise the total cost under abnormal weather impact. In building our model, we adopt the P1 formulation structures in Meng and Zhou (2014), and add-on new weather events constraints and edge borrow constraints in forming our proposed model.

6.3.2. Variables

We first introduce some new notations used in this model.

α, γ, θ: the node index, $\alpha, \gamma, \theta \in N$, where N is the set of nodes

$c_{\alpha,\gamma}$: the edge index between α and γ, $c_{\alpha,\gamma} \in C$, where C is the set of all the edges

e: the link index, denoted by $(\alpha, \gamma), e \in E, E$ is the set of links

$x_{\alpha,\gamma}$: the length of edge $c_{\alpha,\gamma}$ (km)

E_r: the set of links under adverse weather impact with speed restriction, $E_r \subset E$

E_k: the set of possible links train k may use, $E_k \subset E$

E'_k : the set of opposite links of train k, $E'_k \subset E$

E_k^a : the set of backup links for train k, $E_k^a \subset E$

E_s: the set of links in station s, $E_s \subset E$

F_s: the set of trains that must stop in station s, $F_s \subset F$

$E^o(i)$: the set of links starting from node i

$E^d(i)$: the set of links ending at node i

AT_k: the planned arrival time of train k at its destination (min)

$DT_{k,s}$: the planned departure time of train k from station s (min)

$t_k(\alpha,\gamma)$:the planed running time for train k drive through link (α,γ) (min)

$t_k^w(\alpha,\gamma)$: the minimum dwell time for train k on link (α,γ) (min)

$t^a(\alpha,\gamma)$: arrival-arrival headway, the required minimum time headway of two opposite trains arriving at link (α,γ) (min)

$t^d(\alpha,\gamma)$: departure-departure headway, the required minimum time headway of two opposite trains departing from link (α,γ) (min)

$t^{da}(\alpha,\gamma)$: arrival-departure headway, the required minimum time headway of two trains departing from and arriving at link (α,γ) (min)

O_k: the original node of train k

S_k: the destination node of train k

$v_k(\alpha,\gamma)$: the normal speed limitation of train k on link (α,γ) (km/min)

β_1: the cost coefficient of delay

β_2: the cost coefficient of the times using the opposite links (including both links on running lines and sidings)

β_3: the cost coefficient of the times using the alternative links

The decision variables are as follows.

$Z_k(\alpha,\gamma)$*:* the binary train routing variables. If train k selects link *(α,γ)* on the network, $Z_k(\alpha,\gamma) = 1$; otherwise $Z_k(\alpha,\gamma) = 0$

$a_{fk}(\alpha,\gamma)$: the arrival time of train k at link (α,γ)

$d_k(\alpha,\gamma)$: the departure time of train k from link (α,γ)

$\theta_{k,k'}(\alpha,\gamma)$: the binary train routing variables, if train $k$arrives at edge $c_{\alpha,\gamma}$ before train k', $\theta_{k,k'}(\alpha,\gamma) = 1$; otherwise $\theta_{k,k'}(\alpha,\gamma) = 0$

$TT_k(\alpha,\gamma)$: the occupation time of train k on link (α,γ)

6.3.3. Objective function and constraints

The objective is to minimise a weighted combination of all trains' total arrival delay at their destinations, the penalty cost of the times borrowing links used by opposite direction trains and the penalty cost using backup links. The proper value of the weight can be decided according to the industry's assessments; we do not discuss this problem in this chapter. Regarding the constraints of the model Constraints (6-2)-(6-11) are referenced from Meng & Zhou (2014). The rest of constraints are modelled according to the feature of this chapter by the authors.

$$\min cost = \sum_{k\in K} \beta_1 \cdot \Delta t_k + \sum_{k\in K}\sum_{\alpha,\gamma:(\alpha,\gamma)\in E_k'} \beta_2 \cdot x_k(\alpha,\gamma) + \sum_{k\in K}\sum_{\alpha,\gamma:(\alpha,\gamma)\in E_k^a} \beta_3 \cdot x_k(\alpha,\gamma) \qquad (6\text{-}1)$$

Where β_1, β_2 and β_3 are the cost coefficients for delay, times for using opposite links and times for using backup links. $\sum_{k\in K} \beta_1 \cdot \Delta t_k$ is the cost for the total arrival delay of all trains at their destinations. $\sum_{k\in K}\sum_{\alpha,\gamma:(\alpha,\gamma)\in E_k'} \beta_2 \cdot x_k(\alpha,\gamma)$ is the sum of all trains' penalty cost of the times borrowing links used by trains traveling in the opposite direction. $\sum_{k\in K}\sum_{\alpha,\gamma:(\alpha,\gamma)\in E_k^a} \beta_3 \cdot x_k(\alpha,\gamma)$ is the sum of all trains' penalty cost of the times using backup links.

Subject to the following groups of constraints:

(a) Flow balance constraints

This group of constraints is similar to the Group I constraints of Meng and Zhou (2014). Constraints (6-2) to (6-4) ensure that all trains can commence their journey and go through the network from the origin node to destination node. Constraint (6-2) is flow balance constraints at the origin nodes, it ensures a train will go out from the original node and only use one of the links which are joint together by that node. Constraint (6-3) is flow balance constraints at the intermediate nodes. It ensures the numbers of links chosen by a train are equal when arriving and leaving the same intermediate node. Constraint (6-4) is flow balance constraints at the destination nodes, which enforces a train will only use one link when reaching its destination node.

$$\sum_{\alpha,\gamma:(\alpha,\gamma)\in E^o(O_k)\cap E_k} Z_k(\alpha,\gamma) = 1 \qquad \forall k \qquad (6\text{-}2)$$

$$\sum_{\alpha:(\alpha,\gamma)\in E^d(\gamma)\cap E_k} Z_k(\alpha,\gamma) = \sum_{k:(\gamma,\theta)\in E^o(\gamma)\cap E_k} Z_k(\gamma,\theta)$$

$$\forall k, j \in N - O_k - S_k \tag{6-3}$$

$$\sum_{\alpha,\gamma:(\alpha,\gamma)\in E^d(S_k)\cap E_k} Z_k(\alpha,\gamma) = 1 \qquad \forall k \tag{6-4}$$

(b) Time-space network constraints

This group is similar to constraints Group II in Meng and Zhou (2014). Constraint (6-5) is link to link transition constraint which guarantees departure time and arrival time of a train at two connected links are equal. Constraint (6-6) and (6-7) are mapping constraints between the time-space network and physical network. They make sure when a link (α,γ) is not selected by train k, i.e. $Z_k(\alpha,\gamma) = 0$, its departure time and arrival time should be 0 as well.

$$\sum_{\alpha,\gamma:(\alpha,\gamma)\in E_k} d_k(i,j) = \sum_{\gamma,k:(\gamma,\theta)\in E_k} a_k(j,k) \qquad \forall k, j \in N - O_k - S_k \tag{6-5}$$

$$Z_k(\alpha,\gamma) - 1 \le a_k(\alpha,\gamma) \le Z_k(\alpha,\gamma)\cdot M \qquad \forall k,\ (\alpha,\gamma)\in E_k \tag{6-6}$$

$$Z_k(\alpha,\gamma) - 1 \le d_k(\alpha,\gamma) \le Z_k(\alpha,\gamma)\cdot M \qquad \forall k,\ (\alpha,\gamma)\in E_k \tag{6-7}$$

(c) Occupation constraints

Occupation time of train k on the link (α,γ) is calculated by constraint (6-8), i.e. the departure time minus the arrival time. We consider that in order to meet the loading and unloading task, trains have to stop at their designated stopping stations. Constraint (6-9) is improved based on P1 Group III of Meng and Zhou (2014). We introduce (6-9) as station stop constraint to ensure that specified stations will not be missed by certain trains and make sure trains must choose one of the platform links in the station, while they did not require trains to pass certain important stations while escaping some unimportant trains in making rerouting decisions. Constraint (6-10) ensures if a link (α,γ) is selected by train k, i.e. $Z_k(\alpha,\gamma) = 1$, the occupation time must be not shorter than the planned running time plus the required dwell time. When a train does not need to stop on a station link, the dwell time is set to zero. Constraint (6-11) ensures a train cannot depart earlier than the planned departure time from station to allow all the punctual passengers boarding the train.

$$TT_k(\alpha,\gamma) = d_k(\alpha,\gamma) - a_k(\alpha,\gamma) \qquad \forall k,\ (\alpha,\gamma)\in E_k \tag{6-8}$$

$$\sum_{\alpha,\gamma:(\alpha,\gamma)\in E_k\cap E_s} Z_k(\alpha,\gamma) = 1 \qquad \forall s, k \in F_s \tag{6-9}$$

$$TT_k(\alpha,\gamma) + (1 - Z_k(\alpha,\gamma)) \cdot M \geq t_k(\alpha,\gamma) + t_k^w(\alpha,\gamma) \ \forall k, (\alpha,\gamma) \in E_k \quad (6\text{-}10)$$

$$\sum_{\alpha,\gamma:(\alpha,\gamma)\in E_s \cap E_k} d_k(\alpha,\gamma) \geq DT_{k,s} \qquad \forall s, k \in F_s \quad (6\text{-}11)$$

(d) Mapping constraints between train order and usage on the same track:

P1 Group IV in Meng and Zhou (2014) claimed to be "mapping constraints between train orders and cell usage on the same track"; however, the formulations were only suitable for trains travelling in the same directions, as they did not constrained trains travelling from the opposite direction using the same track, which means they considered only unidirectional tracks. This thesis introduces $\theta_{k,k'}(\alpha,\gamma)$ as precedence variables on the line $c_{\alpha,\gamma}$ instead of on the link (α,γ), and introducing occupation variables for both directions in the constraints, making the constraints groups suitable for both unidirectional and bi-directional tracks.

Constraint (6-12) makes sure if two trains, travelling in opposite directions or the same direction are to use the same track (edge), one train will have priority over another to go through the edge to avoid conflict. Constraints (6-13) to (6-15) are auxiliary constraints to mandate only one train getting the priority when two trains are applying for the same track. Constraint (6-13) ensures two different trains will not have the priority at the same time on the same link; constraint (6-14) ensure if train f is not taking link (α,γ) or (γ,α), i.e. $Z_k(\alpha,\gamma) = Z_k(\gamma,\alpha) = 0$, then $\theta_{k,k'}(\alpha,\gamma) = \theta_{k',k}(\alpha,\gamma) = 0$; constraint (6-15) makes sure if a train chooses one direction of an edge, then it will not use the other direction any more.

$$Z_k(\alpha,\gamma) + Z_{k\prime}(\alpha,\gamma) + Z_k(\gamma,\alpha) + Z_{k'}(\gamma,\alpha) - 1 \leq \theta_{k,k'}(\alpha,\gamma) + \theta_{k',k}(\alpha,\gamma)$$

$$\leq 3 - Z_k(\alpha,\gamma) - Z_{k'}(\alpha,\gamma) - Z_k(\gamma,\alpha) - Z_{k'}(\gamma,\alpha)$$

$$\forall k, k', k \neq k', (\alpha,\gamma) \in E_k \cap E_{k'}, (\gamma,\alpha) \in E_k \cap E_{k\prime} \quad (6\text{-}12)$$

$$\theta_{k,k'}(\alpha,\gamma) + \theta_{k',k}(\alpha,\gamma) \leq 1 \qquad \forall k, k', k \neq k', (\alpha,\gamma) \in E_k \cap E_{k\prime} \quad (6\text{-}13)$$

$$\theta_{k,k'}(\alpha,\gamma) + \theta_{k',k}(\alpha,\gamma) \leq Z_k(\alpha,\gamma) + Z_k(\gamma,\alpha)$$

$$\forall k, k', k \neq k', (\alpha,\gamma) \in E_k \cap E_{k\prime}, (\gamma.\alpha) \in E_k \cap E_{k\prime} \quad (6\text{-}14)$$

$$Z_k(\alpha,\gamma) + Z_k(\gamma,\alpha) \qquad \forall k, (\alpha,\gamma) \in E_k \cap E_{k\prime}, (\gamma,\alpha) \in E_k \cap E_{k\prime} \quad (6\text{-}15)$$

(e) Capacity constraints on the same track

Following the introducing of $\theta_{k,k'}(\alpha,\gamma)$, constraints (6-16) and (6-17) ensure that if two trains travelling in the opposite directions are using the same edge, one train can enter the edge only after the other train has left. For example, if trains k and k' travelling in the opposite direction are both to use the same edge, i.e. $Z_k(\alpha,\gamma) = Z_{k'}(\gamma,\alpha) = 1$ and train k has priority over train k', $\theta_{k,k'}(\alpha,\gamma) = 1$, then constraint (6-16) guarantees that train k' will only enter link (γ,α) $t^{da}(\alpha,\gamma)$ minutes after train k has departed from link (α,γ); Likewise, if train k' has priority over train k, $\theta_{k,k'}(\alpha,\gamma) = 1$, constraint (6-17) guarantees that train k will enter link (α,γ) at least $t^{da}(\alpha,\gamma)$ minutes after train k' has departed from link (γ,α). If the two trains running on the same direction, constraint (6-18) ensures that one train will go first and the other will follow it with a headway time of at least $t^{a}(\alpha,\gamma)$. Constraint (6-19) makes sure that the overtaking will not happen on the same edge.

$$a_{k'}(\gamma,\alpha) + \left(3 - Z_k(\alpha,\gamma) - Z_{k'}(\gamma,\alpha) - \theta_{k,k'}(\alpha,\gamma)\right) \cdot M$$

$$\geq \mathrm{d_k}(\alpha,\gamma) + t^{da}(\alpha,\gamma) \qquad \forall \mathrm{k} \in \mathrm{I}, \mathrm{k}' \in \mathrm{F}_\alpha, (\alpha,\gamma) \in \mathrm{E_k}, (\gamma,\alpha) \in \mathrm{E}_{k'} \qquad (6\text{-}16)$$

$$a_k(\alpha,\gamma) + \left(3 - Z_k(\alpha,\gamma) - Z_{k'}(\gamma,\alpha) - \theta_{k',k}(\alpha,\gamma)\right) \cdot M$$

$$\geq d_{k'}(\gamma,\alpha) + t^{da}(\alpha,\gamma) \qquad \forall k \in F_o, k' \in F_i, (\alpha,\gamma) \in E_k, (\gamma,\alpha) \in E_{k'} \quad (6\text{-}17)$$

$$a_{k'}(\alpha,\gamma) + \left(3 - Z_k(\alpha,\gamma) - Z_{k'}(\gamma,\alpha) - \theta_{k,k'}(\alpha,\gamma)\right) \cdot M \geq a_k(\alpha,\gamma) + t^{a}(\alpha,\gamma)$$

$$\forall k, k' \in F, \quad k \neq k', \;\; a(\alpha,\gamma) \in E_k \cap E_{k'} \qquad (6\text{-}18)$$

$$d_{k'}(\alpha,\gamma) + \left(3 - Z_k(\alpha,\gamma) - Z_{k'}(\gamma,\alpha) - \theta_{k,k'}(\alpha,\gamma)\right) \cdot M \geq d_k(\alpha,\gamma) + t^{d}(\alpha,\gamma)$$

$$\forall k, k' \in K, \quad k \neq k', \;\; (\alpha,\gamma) \in E_k \cap E_{k'} \qquad (6\text{-}19)$$

(F) Weather impact constraints

This group is newly introduced in this thesis to map the weather information into the model. Constraints (6-20) and (6-21) will decide whether a train will be affected by the q^{th} weather impact. When the train is not affected by the weather effect, i.e. $X_{q,k} = 0$, constraints (6-22) and (6-23) will active and figure out that the train is going before or after the *rth* weather impact. Constraint (6-24) ensures that if $X_{q,k} = 1$, the train will travel under the speed restriction v_q. If weather impact is very severe, the track will be blocked, i.e. the speed restriction $v_q = 0$, and then according to constraints (6-25), the travel time on that impacted link will be M, a sufficiently large constant which will further result in that impacted track will not be chosen. Constraints (6-26) and (6-27) are auxiliary constraints. Constraints (6-26) ensures when $X_{q,k} = 1$, $Y_{q,k}$=1; constraint (6-27) ensures when train k is not using (α,γ) train will not impact by weather on link (α,γ), i.e. $Z_k(\alpha,\gamma) = 0$, $X_{q,k} = 0$, where $p_q = [l_q, b_q, e_q, v_q]$, $(\alpha,\gamma) = l_q$.

$$\left(b_q - t_k(\alpha,\gamma)\right) \cdot X_{q,k} - \left(1 - X_{q,k}\right) \cdot M - \left(1 - Z_k(\alpha,\gamma)\right) \cdot M \le a_k(\alpha,\gamma)$$

$$\forall k,\ (\alpha,\gamma) = e_r, q \in Q \tag{6-20}$$

$$\mathrm{e_q} \cdot \mathrm{X_{q,k}} + \left(1 - \mathrm{X_{q,k}}\right) \cdot \mathrm{M} + \left(1 - \mathrm{Z_k}(\alpha,\gamma)\right) \cdot \mathrm{M} \ge a_k(\alpha,\gamma) + \varepsilon$$

$$\forall k,\ (\alpha,\gamma) \in E_r, q \in Q \tag{6-21}$$

$$\left(b_q - t_k(\alpha,\gamma)\right) \cdot \left(1 - Y_{q,k}\right) + X_{q,k} \cdot M + Y_{q,k} \cdot M + \left(1 - Z_k(\alpha,\gamma)\right) \cdot M$$

$$\ge a_f(\alpha,\gamma) + \varepsilon \qquad \forall f,\ (\alpha,\gamma) \in E_r, q \in Q \tag{6-22}$$

$$e_q \cdot Y_{q,k} - X_{q,k} \cdot M - \left(1 - Y_{q,k}\right) \cdot M - \left(1 - Z_k(\alpha,\gamma)\right) \cdot M \le a_k(\alpha,\gamma) + \varepsilon$$

$$\forall k,\ (\alpha,\gamma) = l_q, q \in Q \tag{6-23}$$

$$TT_k(\alpha,\gamma) + \left(1 - Z_k(\alpha,\gamma)\right) \cdot M \ge X_{q,k} \cdot \frac{x_{i,j}}{v_q} + \left(1 - X_{q,k}\right) \cdot \frac{x_{i,j}}{v_k(\alpha,\gamma)} + t_k^w(\alpha,\gamma)$$

$$\forall k,\ (\alpha,\gamma) = l_q,\ q \in Q\ , v_q{}' \ne 0 \tag{6-24}$$

$$TT_k(\alpha,\gamma) + \left(1 - Z_k(\alpha,\gamma)\right) \cdot M \ge X_{q,k} \cdot M + \left(1 - X_{q,k}\right) \cdot \frac{x_{i,j}}{v_k(\alpha,\gamma)} + t_k^w(\alpha,\gamma)$$

$$\forall k,\ (\alpha,\gamma) = l_q,\ q \in Q, v_q{}' = 0 \tag{6-25}$$

$$Y_{q,k} \ge X_{q,k} \qquad \forall k,\ (\alpha,\gamma) = l_q, q \in Q \tag{6-26}$$

$$Z_k(\alpha,\gamma) \ge X_{q,k} \qquad \forall k,\ (\alpha,\gamma) = l_q, q \in Q \tag{6-27}$$

6.4. Experiments and results

In this section, we will use numerical examples to analyse the performance of this model. All the cases are running using the same computer as described in Section 4.4.

6.4.1. Case study 1: East Coast Main Line

We will firstly choose a simple double track network to start with.

6.4.1.1. Case description

We conduct a case study based on a 63 km section of the East Coast Main Line in the UK, as shown in Figure 6-3. There are two parallel main lines, i.e. Main line 1 (Main1) and Main line 2 (Main2) and several parallel siding lines aside the main ones at some parts. For the modelling perspective, two dummy nodes, i.e. Start and End, are added at the two ends (or use ‘north and south ends’) of the section. As shown by the arrows in the figure, the route of Main1 is End-Newark North Gate-Grantham-Start, and the route of Main2 is Start-Grantham- Newark North Gate-End. Grantham and Newark North Gate are two stations where trains need to dwell for at least two minutes to allow loading and unloading. The entire network is displayed in Figure 6-3. Under severe weather impact, we assume that all the tracks become bi-directional, which means that if one track is blocked, the train is allowed to pass through the opposite track.

Four outbound trains are going from Start to End and four inbound trains are going from End to Start during the study period between 12:00 – 14:30. Trains’ original timetable is shown in Figure 6-4. For better visualisation, in Figure 6-4, trains travelling on Main1 are marked as blue and trains travelling on Main2 are marked as red, and when trains are travelling on siding tracks, the trajectories will be marked as black. For simplicity, we use horizontal lines to mark the positions of the stations and loops in the timetable diagram, and we do not show all the sidings and junctions. Dotted lines stand for loops while solid lines stand for stations. All the required minimum headways between two trains are two minutes.

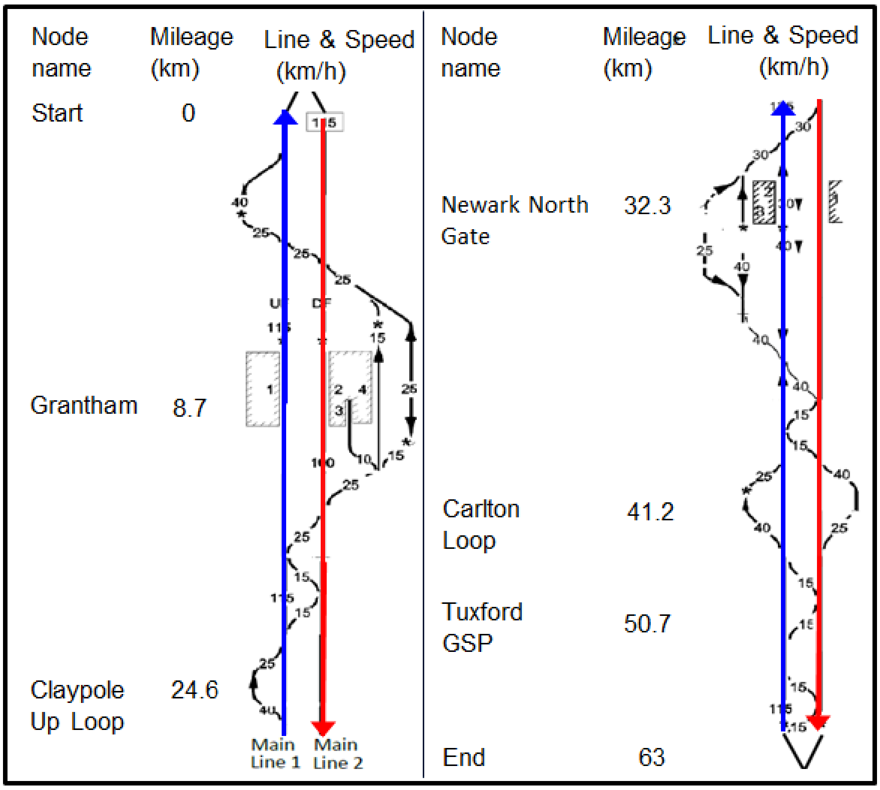


Figure 6-3: The modelled East Coast Main Line section.

Assume there is a thunderstorm front moving from Highdyke to Newark North Gate, from 12:10 to 14:00. The effect of the moving weather front on the operations of the network are described below in terms of the speed restriction, track blockage, locations and time period:

a) Speed restriction of 60 km/hr, between Highdyke to Grantham station, from 12:10 to 12:40, 60 km/hr;
b) Speed restriction of 30 km/h from Grantham station to Claypole Up Loop from 12:40 to 13:20 on both two main lines.
c) Track blockage on the Main Line 1 from Claypole Up Loop to Newark North Gate, the train cannot get through this section during 12:00 to 14:00.

The detailed impact distributions are shown as the shadow squares in Figure 6-4.

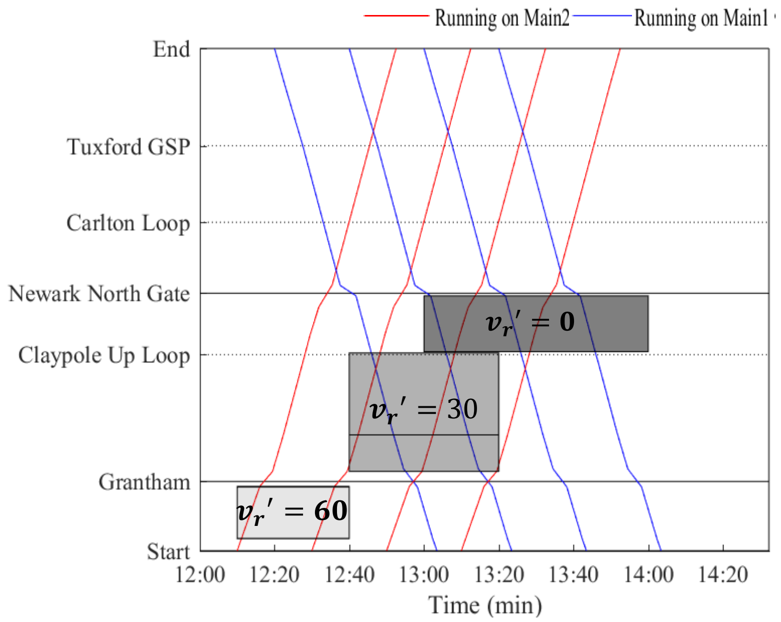


Figure 6-4: The original timetable and the forecasted weather impact.

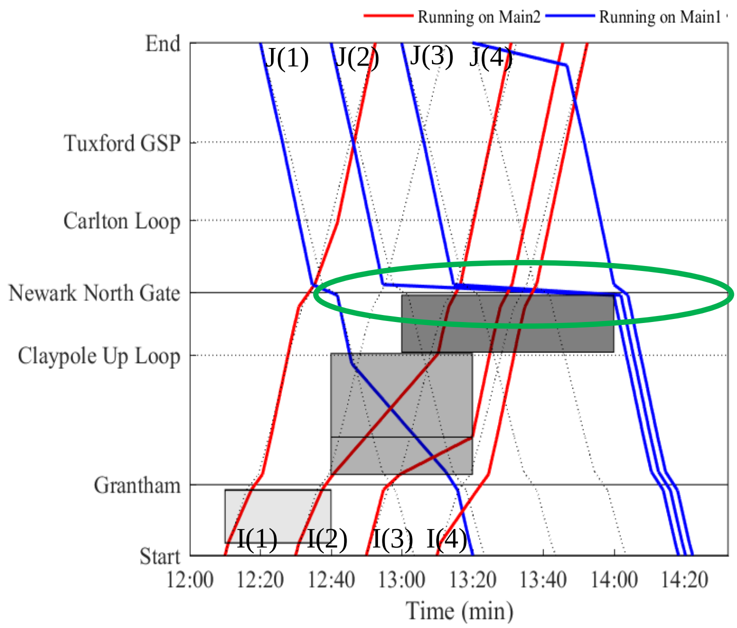
6.4.1.2. Results of the rerouting and rescheduling

There are no back-up lines in this network so the value of β_3 is set to 0. For simplicity, we set $\beta_1 = 1$ and test different value of $\beta_2 = 1,2,3, \ldots$. When the penalty of changing to opposite tracks (β_2) reaches a certain value, no trains will change tracks, i.e. the value of the second part of the objective function: $\sum_{k \in K} \sum_{\alpha,\gamma:(\alpha,\gamma) \in E_k'} \beta_2 \cdot x_k(\alpha, \gamma)$ will always be zero. By then, the total cost value will remain the delay cost of trains waiting until the blockages disappear. For this reason, we don't test the cases after β_2 reaches that level.

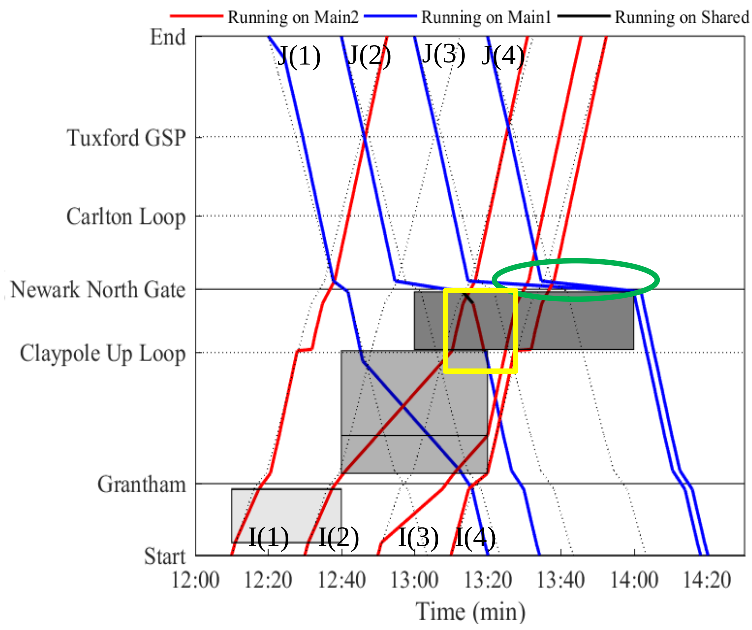
The statistics of the test results are shown in Table 6-1. The adjusted timetables are shown in Figure 6-5. To save space, we present the results in groups, which are aggregated by total delay minutes. In each group, although β_2 values are different, all trains' route choices and the total delay minutes, i.e. the adjusted timetables, are the same. As for the total cost in each group, since the timetables are the same, the total cost will increase while β_2 increases. For example, in the second group, where the total delay is 111 minutes and total penalty is three times, when $\beta_2 = 10$, the total cost is 111+3×10=141; when β_2 increases to 16, the total cost is 111+3×16=159. For the computation time, when β_2 is bigger, the computation time is less, as with bigger penalty cost, feasible domain is small, which will use less time. The detailed analysis for the four figures in Figure 6-5 is presented subsequently.

Table 6-1: Test result for different β_2 value

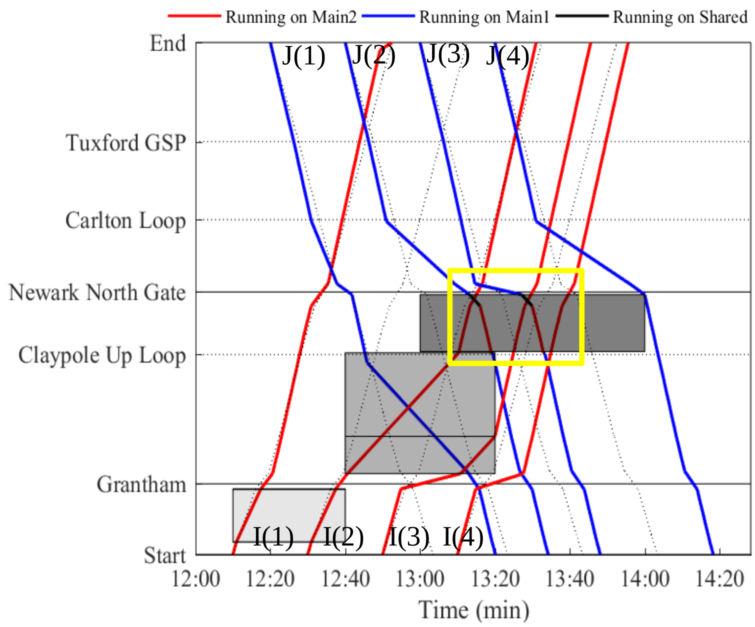
β_2	Total cost	Total delay (min)	Total penalty times	Computation time (s)
≥17	159	159	0	≤94
≥10 and ≤16	141-159	111	3	≥141 and ≤210
≥5 and ≤9	111-135	81	6	≥232 and ≤617
≤4	67-103	67	9	≤1295



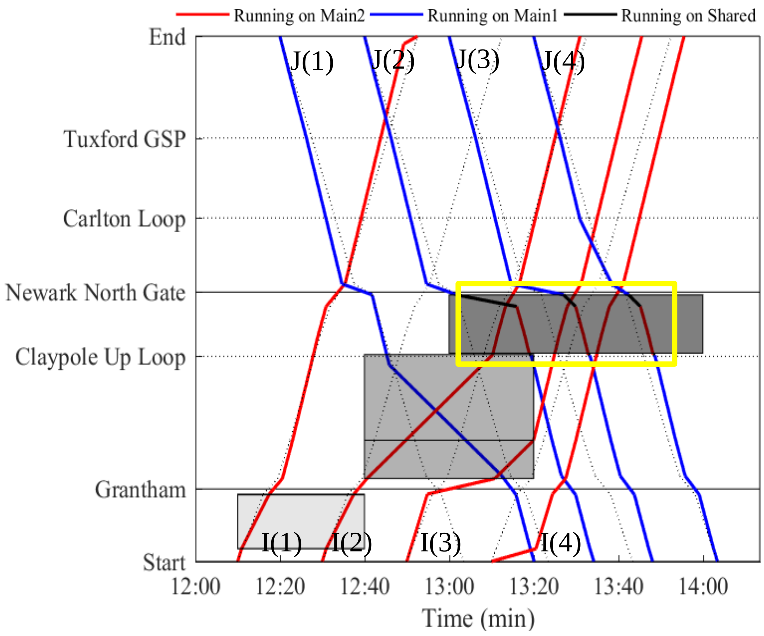
(a)


Running on Main2
Running on Main1
Running on Shared
End
Tuxford GSP
Carlton Loop
Newark North Gate
Claypole Up Loop
Grantham
Start
J(1)
J(2)
J(3)
J(4)
I(1)
I(2)
I(3)
I(4)
12:00
12:20
12:40
13:00
13:20
13:40
14:00
14:20
Time (min)

(b)


Running on Main2
Running on Main1
Running on Shared
End
Tuxford GSP
Carlton Loop
Newark North Gate
Claypole Up Loop
Grantham
Start
J(1)
J(2)
J(3)
J(4)
I(1)
I(2)
I(3)
I(4)
12:00
12:20
12:40
13:00
13:20
13:40
14:00
14:20
Time (min)

(c)



(d)

Figure 6-5: Rescheduled timetable when (a) $\beta_2 \geq 17$; (b) $10 \leq \beta_2 \leq 16$; (c) $5 \leq \beta_2 \leq 9$ and (d) $\beta_2 \leq 4$

All the four rescheduled timetables in Figure 6-5 indicate trains going through the first and second weather periods (given by the earliest two grey square zones) are slowed down due to the speed restrictions. No inbound trains go through the section of Newark North Gate to Claypole Up Loop by line Main1 from 13:00 to 14:00 due to the track blockage on Main1. One train's trajectory is plotted as several two-end lines which connect the arrival time and departure time pairs on links. As we only control the departure time at stations and aim to reduce the total delays and penalty cost, train trajectories on intermediate links look quite random but have no an impact on the objective.

When line Main1 is blocked from 13:00 to 14:00, the model will deliver different results by different penalty coefficients (β_2). When $\beta_2 \geq 17$, as shown in the green ellipse in Figure 6-5 (a), no impacted inbound trains change to line Main2 (train will be marked as red lines if using Main2), instead, all the trains depart Newark North Gate until the blockage is cleared.

When $10 \leq \beta_2 \leq 16$, as shown in the yellow square in Figure 6-5 (b), only the first impacted inbound train J(2) shifts to line Main2 and shifts back to Main1 using side tracks (marked as solid black lines) in Figure 6-5 (b). Constraints Group 5 ensure when two opposite direction trains using the same track, a priority will be issued to one of the trains to avoid the potential conflict. As shown in the near green ellipse, the remaining two trains will not depart from Newark North Gate until the blockage on Main1 is cleared. This indicates that when using the siding track, the total cost of the delay and penalty is smaller than the total delay when waiting until the impact is cleared for the first impacted inbound train, and the other way round for the last two impacted trains when $10 \leq \beta_2 \leq 16$.

When $5 \leq \beta_2 \leq 9$ and $\beta_2 \leq 4$, as shown in the yellow squares in Figure 6-5 (c) and (d), the first two impacted trains J(2) and J (3) and all the three impacted trains J(2), J(3) and J(4) shift to the opposite tracks from Newark North Gate to Claypole Up Loop, respectively. This indicates, with the penalty cost reducing, more trains are allowed to change to the opposite tracks.

6.4.2. Case study 2: a larger network

6.4.2.1. Case description

We used the same network in

to test our method in a larger network. We neglect some irrelevant nodes (signalling) between points. Instead, we only keep nodes which reflect the network structure. The network is shown as Figure 6-6. It consists of 36 nodes and 50 edges, with a total track length of 287.7 km. We use the same setting as in Meng and Zhou (2014) for safety headways which is 3 minutes.

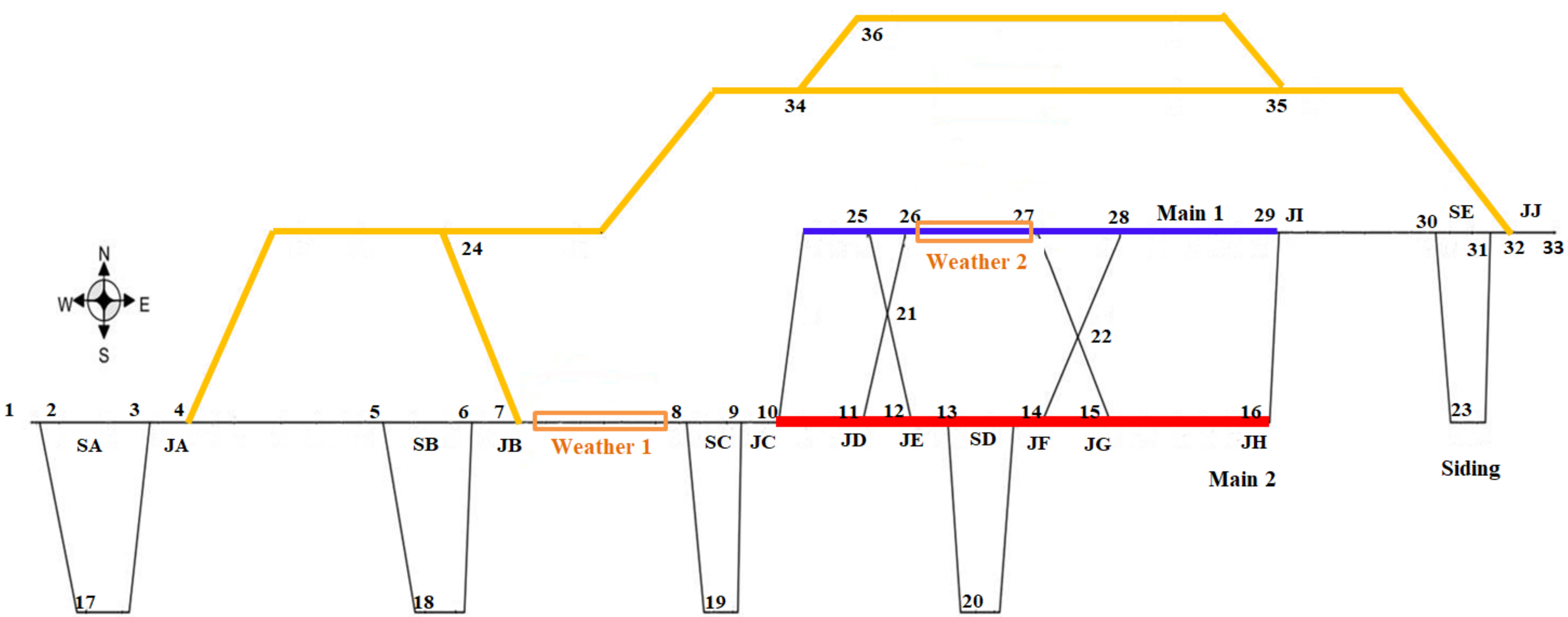


Figure 6-6: A bigger network of Case 2 (source: Meng and Zhou , 2014)

The original timetable is shown as Figure 6-7. The black tracks can be used by any trains, the tracks in red (Main2) are allocated to outbound trains only and tracks in blue (Main1) are allocated to inbound trains only under normal conditions.

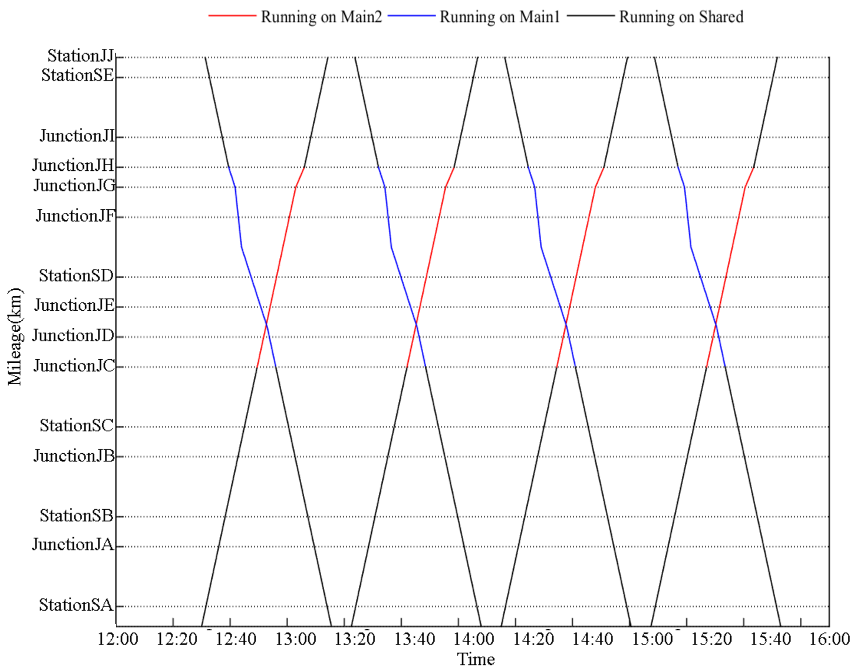


Figure 6-7: The original timetable of Case Study 2

We use the same objective function and constraints described in Section 6.3.3. The delay coefficient is set to $\beta_1 = 100$. Under sever disruptions, the following conditions are applied:

- Main2 can be used by inbound trains and Main1 can be used by outbound trains with an opposite track penalty cost rate β_2
- The alternative tracks in yellow can be used with an alternative track penalty cost rate β_3, which is caused due to missing their scheduled stations;

We choose a combination of two different locations (i.e. marked as Weather 1 from node 7 to node 8, and Weather 2 from node 26 to node 27 as shown in Figure 6-6) and different weather types to test the weather impact to the network. Weather 1 represents the impact to a shared track segment for both outbound and inbound trains. It is used to test the possible usage of the alternative tracks for trains from both directions. Weather 2 represents the impact to a directed track segment. It is used to test the possible usage of its opposite directed tracks and the alternative tracks. We consider speed limitation of 20km/h which might be caused by the heavy rain and the track blockage due to the strong wind.

To control the variables, we set all the disturbance time to be from 13:40 to 14:40. This period of time covers the trains from both outbound and inbound. The perturbation combinations are as follows:

Perturbation (1): Weather 1, one-hour speed reduced to 20km/h

Perturbation (2): Weather 2, one-hour speed reduced to 20km/h

Perturbation (3): Weather 1, one-hour blockage

Perturbation (4): Weather 2, one-hour blockage

Perturbation (5): Weather 1 and Weather 2, one hour speed reduced to 20km/h

Perturbation (6): Weather 1 and Weather 2, one hour blockage

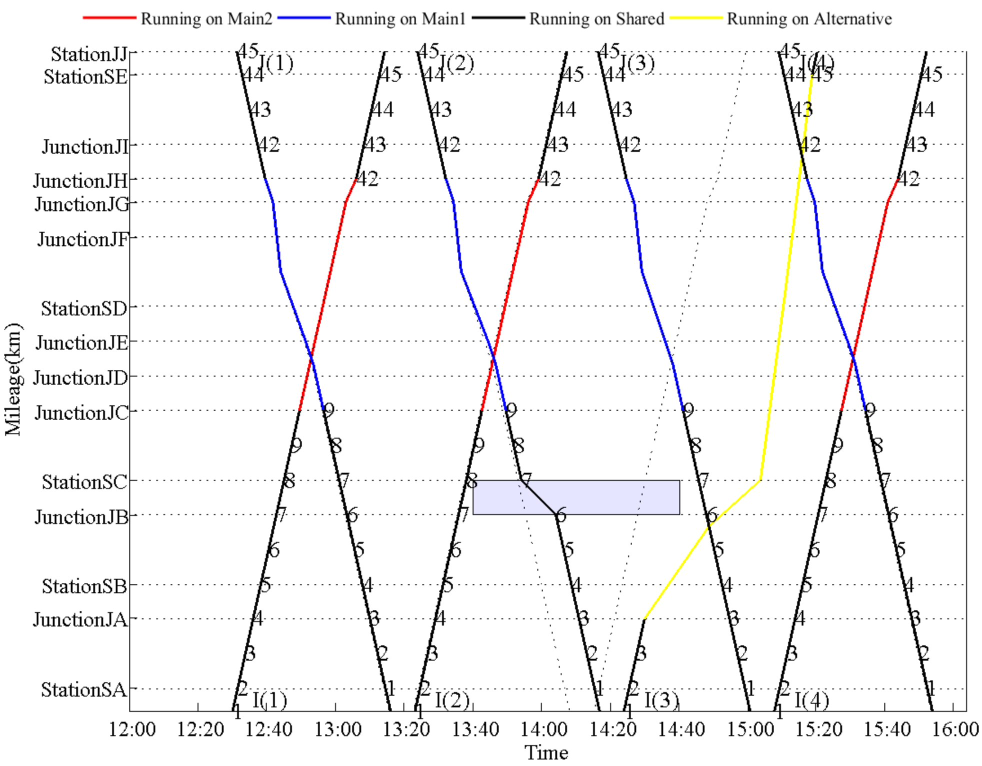
Weather 1 happens on the single-track, which is the original route of both inbound and outbound trains. If perturbations happened, trains from both directions have two choices, either to follow their original routes with some potential delay, or to choose the alternative tracks with some alternative track penalty costs. Weather 2 happens on routes for inbound trains. If perturbations happened, the affected inbounded trains three choices: first, trains can follow their original routes with some potential delay; second, trains can choose the alternative tracks with some alternative track penalty costs; and third, trains can choose the opposite tracks with some opposite track penalty costs. As the objective is to minimise the total weighted cost, the model will compare the costs under choosing different options and made the best decisions.

6.4.2.2. Result

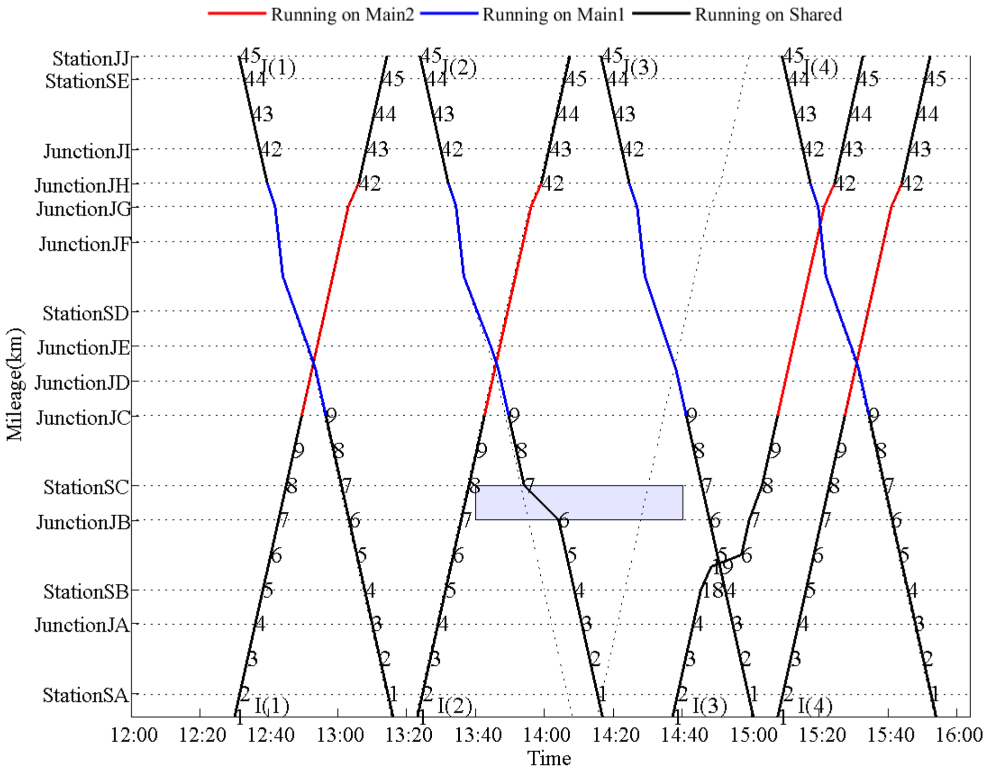
The statistics of the test results are as shown in Table 6-2. The adjusted timetables are shown in Figure 6-8. To save space, we also present the results in groups, which are aggregated by total delay minutes. The group criteria are the same as those described in Section 6.4.1.1. The detailed analysis of the figures in Figure 6-8 is shown in the following paragraphs. The number under each diagram corresponds to the case No. indicated in the first column of Table 6-2.

Table 6-2: Test result for different perturbations and different β_2 and β_3

No.	Perturbation	β_2	Times using opposite tracks	β_3	Times using alternative tracks	Total cost	Delay (minutes)	Average Computation time (s)
1	(1)	-	-	0-305	4	3100-4320	31	52
2	(1)	-	-	≥ 306	0	4325	43.25	47
3	(2)	0-732	3	-	-	3625-5821	36.25	52
4	(2)	≥ 733	0	-	-	5825	58.25	49
5	(3)	-	-	0-224	8	5350 3750-	37.5	46
6	(3)	-	-	225-987	4	4650-8598	46.5	46
7	(3)	-	-	≥ 988	0	8600	86	45
8	(4)	0-2449	3	-	-	3625-10972	36.25	48
9	(4)	≥2450	0	-	-	11000	110	54
10	(5)	-	-	0-474	4	5725-7621	57.25	52
11	(5)	-	-	≥ 475	0	7625	76.25	56
12	(6)	-	-	0-1862	4	5725-13173	57.25	52
13	(6)	-	-	≥1863		13175	131.75	48

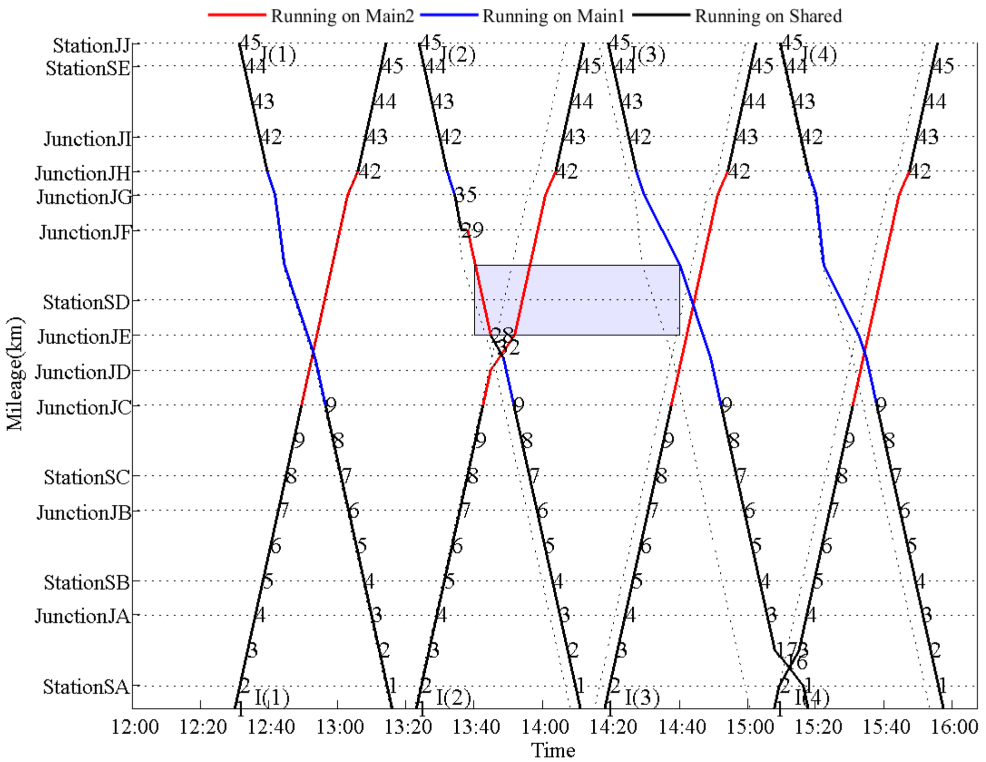


No. 1



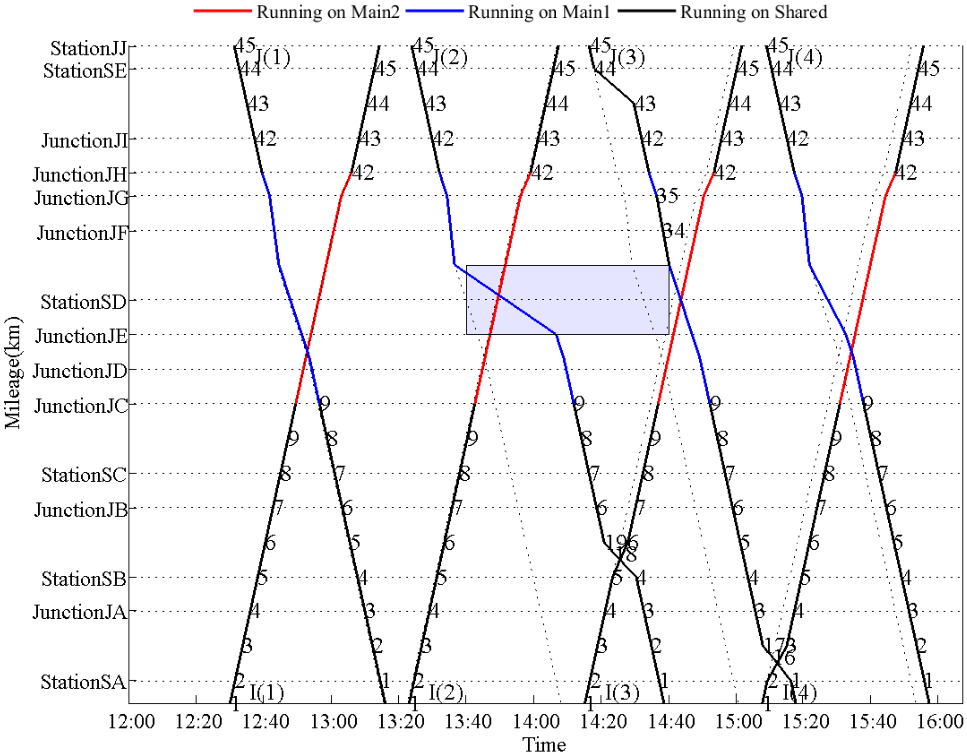
Running on Main2
Running on Main1
Running on Shared
StationJJ
StationSE
JunctionJI
JunctionJH
JunctionJG
JunctionJF
StationSD
JunctionJE
JunctionJD
JunctionJC
StationSC
JunctionJB
StationSB
JunctionJA
StationSA
Mileage(km)
12:00
12:20
12:40
13:00
13:20
13:40
14:00
14:20
14:40
15:00
15:20
15:40
16:00
Time
I(1)
I(2)
I(3)
I(4)

No. 2

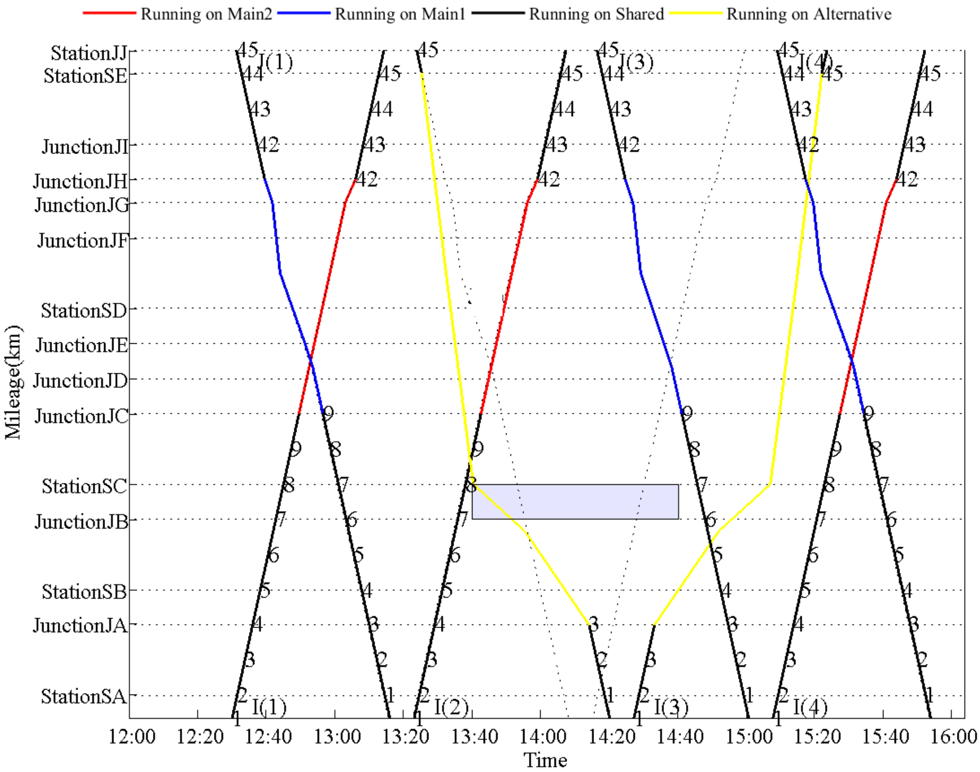


Running on Main2
Running on Main1
Running on Shared
StationJJ
StationSE
JunctionJI
JunctionJH
JunctionJG
JunctionJF
StationSD
JunctionJE
JunctionJD
JunctionJC
StationSC
JunctionJB
StationSB
JunctionJA
StationSA
Mileage(km)
12:00
12:20
12:40
13:00
13:20
13:40
14:00
14:20
14:40
15:00
15:20
15:40
16:00
Time
I(1)
I(2)
I(3)
I(4)

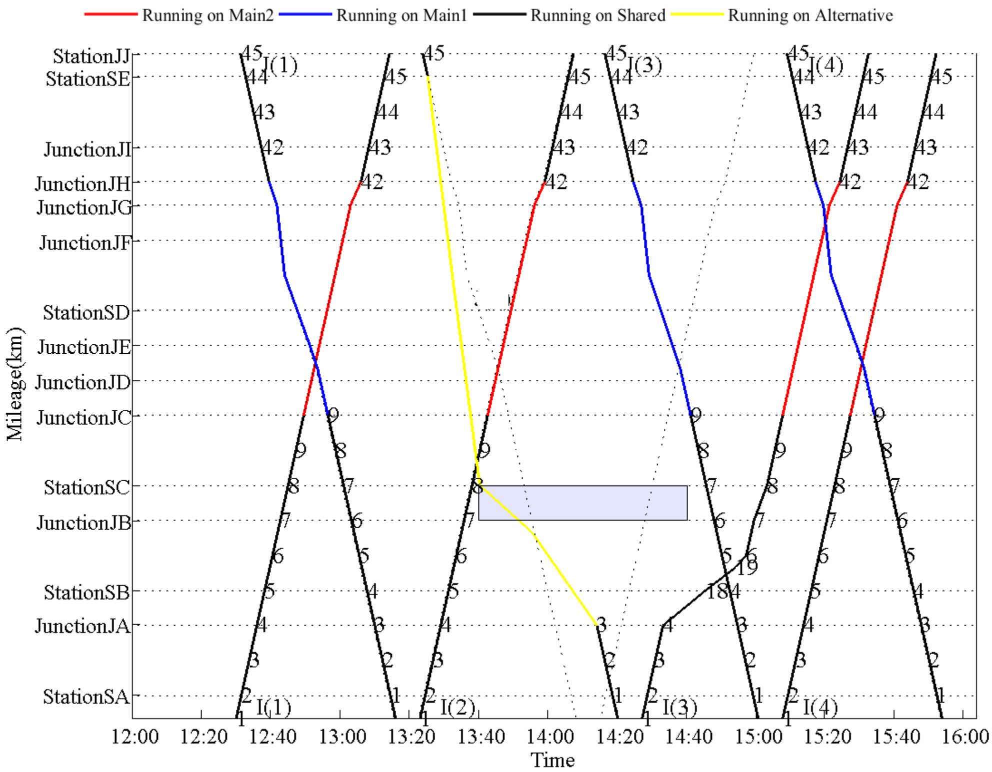
No. 3


Running on Main2
Running on Main1
Running on Shared
StationJJ
StationSE
JunctionJI
JunctionJH
JunctionJG
JunctionJF
StationSD
JunctionJE
JunctionJD
JunctionJC
StationSC
JunctionJB
StationSB
JunctionJA
StationSA
Mileage(km)
12:00
12:20
12:40
13:00
13:20
13:40
14:00
14:20
14:40
15:00
15:20
15:40
16:00
Time
I(1)
I(2)
I(3)
I(4)

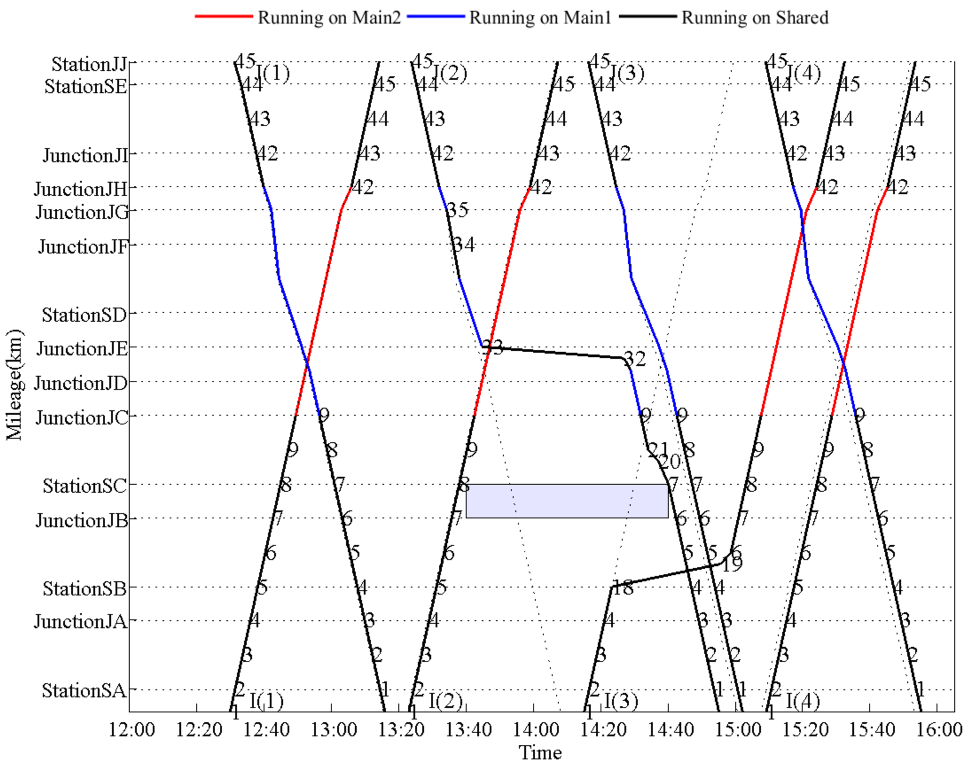
No.4


Running on Main2
Running on Main1
Running on Shared
Running on Alternative
StationJJ
StationSE
JunctionJI
JunctionJH
JunctionJG
JunctionJF
StationSD
JunctionJE
JunctionJD
JunctionJC
StationSC
JunctionJB
StationSB
JunctionJA
StationSA
Mileage(km)
12:00
12:20
12:40
13:00
13:20
13:40
14:00
14:20
14:40
15:00
15:20
15:40
16:00
Time
I(1)
I(2)
I(3)
I(4)

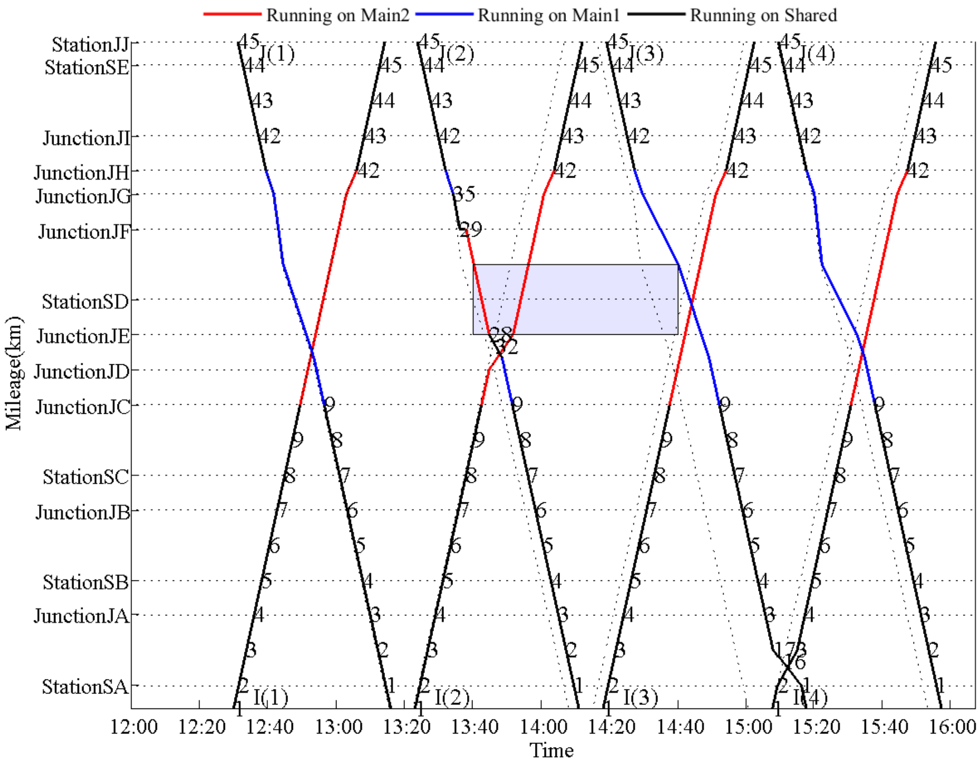
No. 5


Running on Main2
Running on Main1
Running on Shared
Running on Alternative
StationJJ
StationSE
JunctionJI
JunctionJH
JunctionJG
JunctionJF
StationSD
JunctionJE
JunctionJD
JunctionJC
StationSC
JunctionJB
StationSB
JunctionJA
StationSA
Mileage(km)
12:00
12:20
12:40
13:00
13:20
13:40
14:00
14:20
14:40
15:00
15:20
15:40
16:00
Time

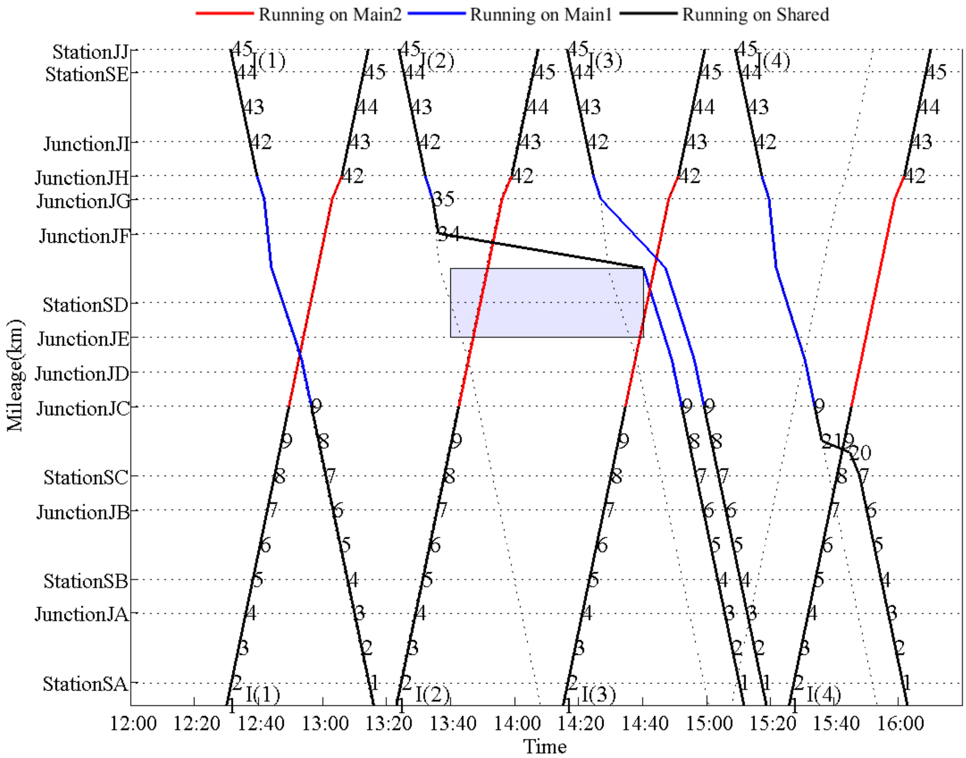
No. 6


Running on Main2
Running on Main1
Running on Shared
StationJJ
StationSE
JunctionJI
JunctionJH
JunctionJG
JunctionJF
StationSD
JunctionJE
JunctionJD
JunctionJC
StationSC
JunctionJB
StationSB
JunctionJA
StationSA
Mileage(km)
12:00
12:20
12:40
13:00
13:20
13:40
14:00
14:20
14:40
15:00
15:20
15:40
16:00
Time

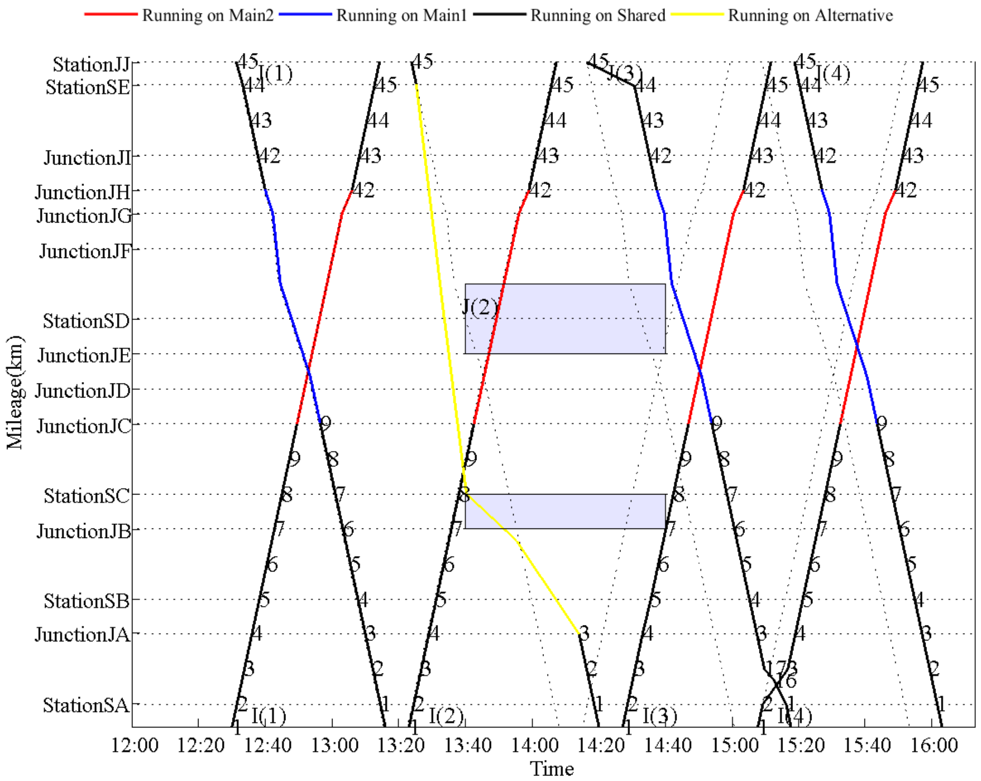
No. 7


Running on Main2
Running on Main1
Running on Shared
StationJJ
StationSE
JunctionJI
JunctionJH
JunctionJG
JunctionJF
StationSD
JunctionJE
JunctionJD
JunctionJC
StationSC
JunctionJB
StationSB
JunctionJA
StationSA
Mileage(km)
12:00
12:20
12:40
13:00
13:20
13:40
14:00
14:20
14:40
15:00
15:20
15:40
16:00
Time
I(1)
I(2)
I(3)
I(4)

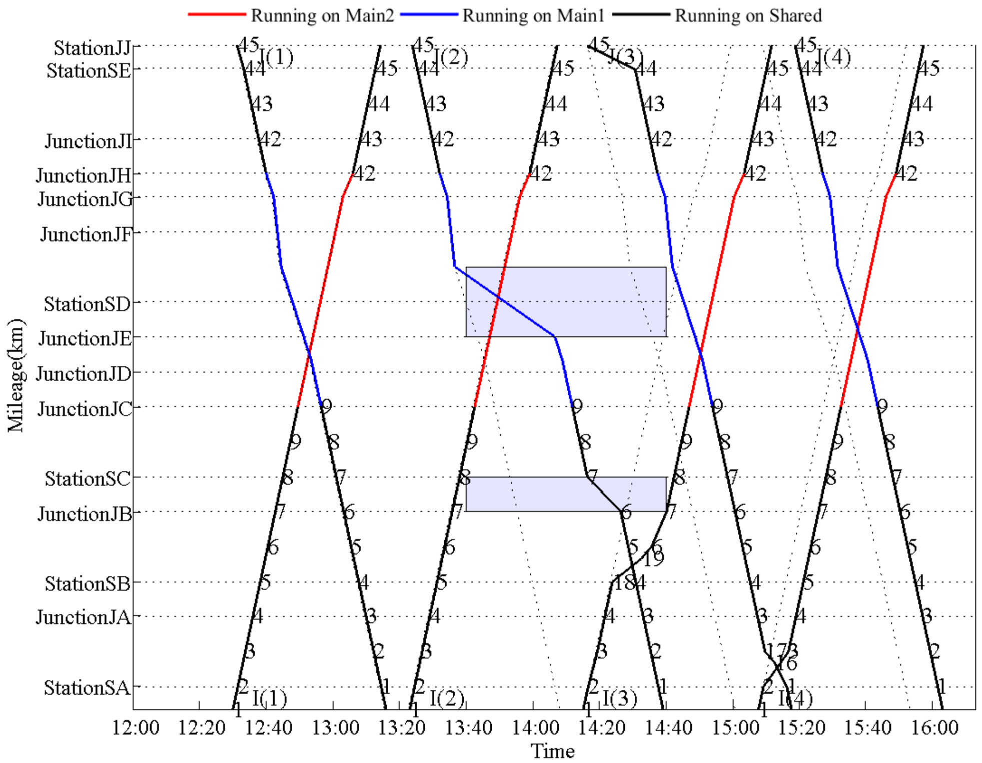
No. 8


Running on Main2
Running on Main1
Running on Shared
StationJJ
StationSE
JunctionJI
JunctionJH
JunctionJG
JunctionJF
StationSD
JunctionJE
JunctionJD
JunctionJC
StationSC
JunctionJB
StationSB
JunctionJA
StationSA
Mileage(km)
12:00
12:20
12:40
13:00
13:20
13:40
14:00
14:20
14:40
15:00
15:20
15:40
16:00
Time
I(1)
I(2)
I(3)
I(4)

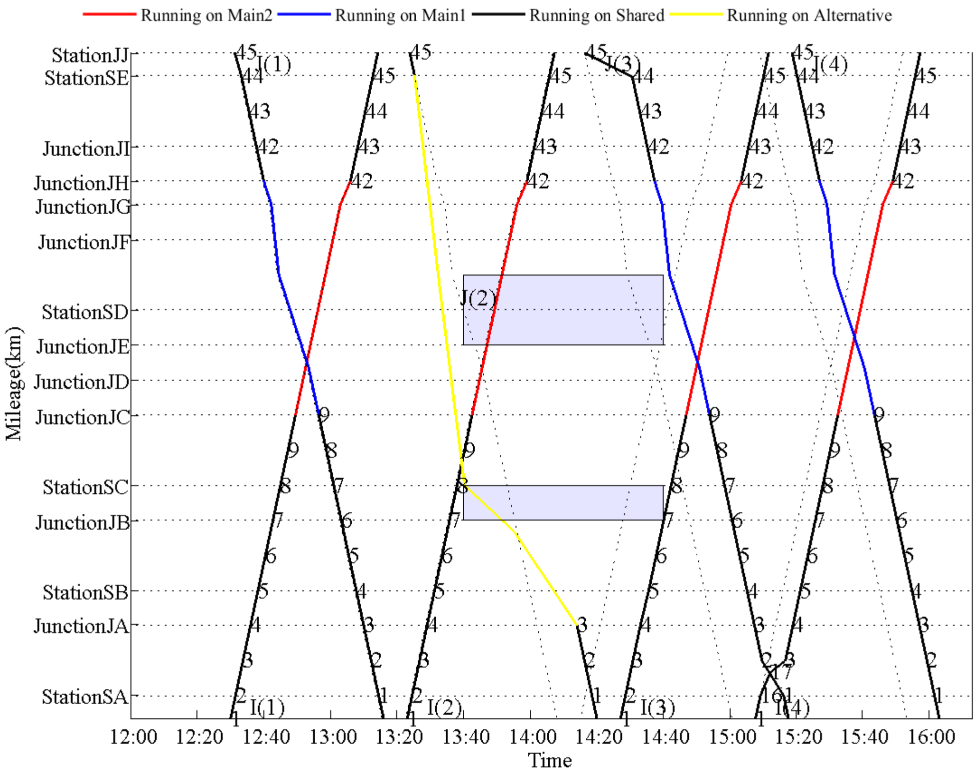
No. 9


Running on Main2
Running on Main1
Running on Shared
Running on Alternative
StationJJ
StationSE
JunctionJI
JunctionJH
JunctionJG
JunctionJF
StationSD
JunctionJE
JunctionJD
JunctionJC
StationSC
JunctionJB
StationSB
JunctionJA
StationSA
Mileage(km)
12:00
12:20
12:40
13:00
13:20
13:40
14:00
14:20
14:40
15:00
15:20
15:40
16:00
Time
I(1)
I(2)
I(3)
I(4)
J(1)
J(2)
J(3)
J(4)

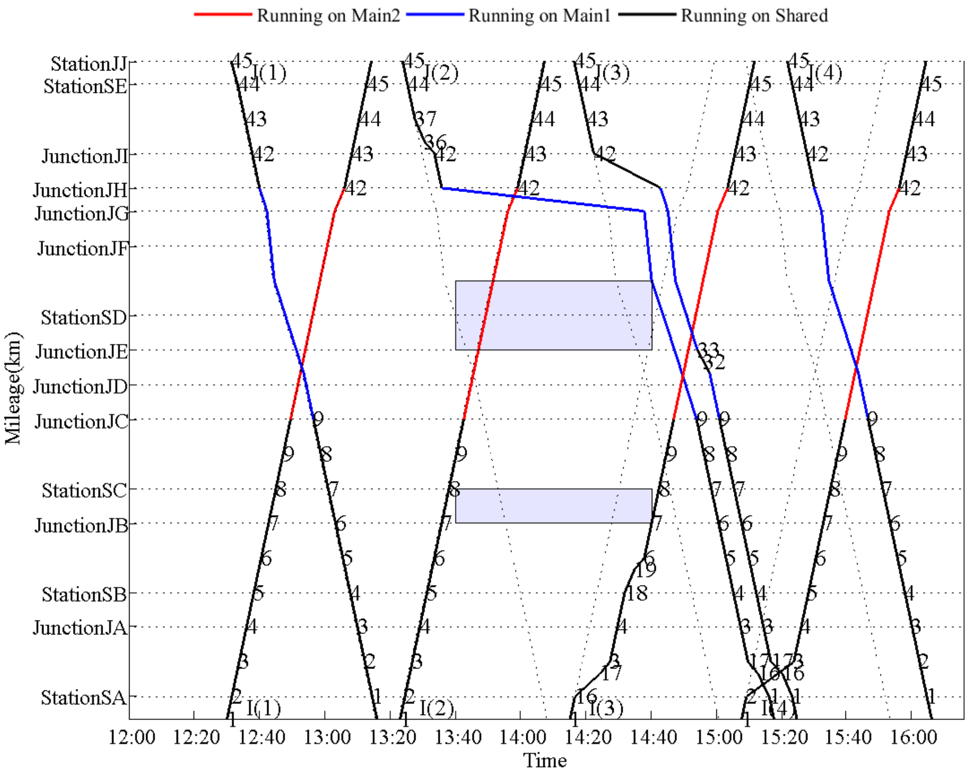
No. 10


Running on Main2
Running on Main1
Running on Shared
StationJJ
StationSE
JunctionJI
JunctionJH
JunctionJG
JunctionJF
StationSD
JunctionJE
JunctionJD
JunctionJC
StationSC
JunctionJB
StationSB
JunctionJA
StationSA
Mileage(km)
12:00
12:20
12:40
13:00
13:20
13:40
14:00
14:20
14:40
15:00
15:20
15:40
16:00
Time
I(1)
I(2)
I(3)
I(4)
J(1)
J(2)
J(3)
J(4)

No. 11



No. 12



No. 13

Figure 6-8: The rescheduling and rerouting results

By comparing the results shown in Figure 6-8 impacted by the same disturbance but with different penalty costs, we could draw the following conclusions:

1. In diagrams No. 1 and No. 2, the weather impact is as described in Perturbation (1) in Section 6.4.2.1. Train J(2) goes through the one-hour speed reduction and is delayed by about 10 minutes at its destination, which further impacts the departure time of train I(3) from its origin station. In diagram No. 1, when $\beta_3 \leq 305$, train I(3) will change to the alternative tracks. In diagram No. 2, when $\beta_3 \geq 306$, train I(3) will change to use the side track on the single-track part to avoid conflict with train J(3).

   In diagrams No. 3 and No. 4, the weather impact is as described in Perturbation (2) in Section 6.4.2.1. Train J(2) and J(3) originally going through the one-hour speed reduction zone is directly impacted. In diagram No. 3, when $\beta_2 \leq 732$, train J(2) will change to the opposite tracks which is originally only be used by outbound trains. In diagram No. 4, when $\beta_2 \geq 733$, train J(2) will keep travelling on its original track under a lower speed limitation given by the weather condition. On its downstream trip, to avoid conflict with train I(3), it changes to the side track which leads to some further delay.

   By comparing No. 1 and No. 2, and No. 3 and No. 4, we can conclude that when speed restriction happened, if the weighted sum of the delay cost and the penalty cost of using the alternative tracks or the opposite tracks is smaller than the delay cost using the speed restriction tracks or siding tracks, the alternative tracks or opposite tracks will be used.

2. In diagrams No. 5 to No. 7, the weather impact is as described in Perturbation (3) in Section 6.4.2.1. One hour track blockage happens to the single-track section between JunctionJB and StationSC. In diagram No. 5, when $\beta_3 \leq 224$, train J(2) and train I(3) change to the alternative tracks. In diagram No.6, when $225 \leq \beta_3 \leq 987$, only train J(2) changes to the alternative track, while train I(3) is delayed when travelling on the side track to avoid the track blockage period and also the conflict with train J(3). In diagram No. 7, when $\beta_3 \geq 988$, no train will change to the alternative track, both impacted trains travell on side track to avoid the blockage period as well as the conflict with other trains.

   By comparing No. 5 to No. 7, we can conclude that when track blockage happened on shared tracks, and potentially impacted two trains, the model will specific suggestions to each train. With the penalty increase, less alternative tracks will be used.

3. In diagrams No. 8 to No. 9, the weather impact is as described in Perturbation (4) in Section 6.4.2.1. One hour track blockage happens to the inbound track section. Train J(2) and train J(3) originally go through that period are directly impacted. In diagram No.8, when $\beta_2 \leq 2449$, train J(2) will change to the opposite tracks which originally belongs to outbound trains. In diagram No. 9, when $\beta_2 \geq 2450$, the opposite track will not be borrowed instead, train J(2) will keep travelling to its side track and wait until the blockage disappears.

   By comparing No. 8 to No. 9, we can conclude that, if track blockage happened to inbound tracks, the opposite track tends to be used, unless the penalty is very high (in the test case, more than 20 times than the delay minutes weight). It can be also noted that No. 3 are the same with No. 8, though the impact in No. 3 is speed reduction, and the impact in No. 8 is track blockage. When a penalty is relatively low, trains tend to change to the opposite tracks as long as the total delay is less comparing travel through the impacted area or wait until the impact end.

4. In diagrams No. 10 and No. 11, the weather impact is as described in Perturbation (5) in Section 6.4.2.1. Train J(2), train J(3) and train I(3) will travel under a speed limitation if following their original timetable. By comparing train J(2) in both diagrams, we can have the similar conclusion as described in point 1.

5. In diagrams No. 12 and No. 13, the weather impact is as described in Perturbation (6) in Section 6.4.2.1. Train J(2), train J(3) and train I(3) cannot travel under their original timetable due to the blockage. By comparing the trains in both diagrams, we can have the similar conclusion as described in point 3. That is when the alternative track penalty is not very high, trains tend to use alternative tracks instead of wait blockage period end.

   By comparing No.10 and No. 12, we notice, though No. 10 has speed reductions while No. 12 has track blockages at the same place, the delay minutes and the rescheduled timetable are the same, i.e. both leading to using alternative longer tracks. We can conclude that when the alternative track penalty cost is relatively low, the disturbance severity does not have much difference in impacting the delay minutes.

In this section, we show that this rerouting and rescheduling model can be used in deciding when and which alternative/opposite tracks can be used under the presence of temporary speed restrictions and track blockages considering the weights of different components of total costs. When we have the weather forecasting data, we can use this model to test which option is optimal and achieve the least cost.

6.5. Conclusions

In this Chapter, we first reviewed the existing rescheduling and rerouting researches and proposed a new method which models the adverse and severe weather impact as compounded speed restrictions and track blockages. Second, we proposed a simultaneous rerouting and rescheduling MILP model which maps the forecasted weather impact and aims to minimise trains' total delay and times of using alternative/opposite tracks. We assumed when a track was blocked, the impacted trains were allowed to shift to the opposite normal condition tracks with certain penalty costs to reduce the delay at their destinations. With the modified capacity constraints on tracks, the potential conflicts were avoided between two opposite trains on the same track when trains were borrowing opposite tracks.

The effectiveness of the proposed model was demonstrated in a real-life 63-km long corridor of the East Coast Mainline in the UK, and one larger network taken from the literature. From the case studies, we could see that optimised new routing and timetable plans were generated by commercial solvers in feasible computing time. Under the more conservative situation, i.e. with bigger changing tracks penalty cost, fewer trains would shift to opposite tracks. This resulted in larger system delay minutes and larger total cost. The model could be used to help controllers dispatching trains under bad weather impact. With the advanced weather forecasting information and a given penalty cost, this model could advise an optimised timetable.

Chapter 7 Conclusion

7.1. Summary

The abnormal weather such as strong wind, high temperature and flood causes a massive amount of financial loss for the railway industry and passengers. For examples, in the Great Britain, the weather related service delay amounts to over two million minutes each year. This cause the infrastructure manager Network Rail to pay tens of millions of pounds to train operating companies for weather related delay and cancellation compensations (as shown in Figure 1-1 in Chapter 1).

We further investigated the connection between abnormal weather on train delays and cancellations. The weather hazards may result in railway buckles, point failure, structure damage, etc. To ensure safety, the railway authorities mandate detailed mitigations to deal with different hazards, such as temporal speed limitation and service suspension. This further leads to train delays and cancellations. Though some countries like the Netherlands prepared backup train timetables for extreme weather. They may not always work well as the temporal and spatial characteristics of weather are different each time.

Though the weather forecasting technology is improving and can have high accuracy in short term forecasting, existing research in railway traffic control still treat weather as unpredictable, independent and static perturbations (UISP), which react after the weather impact has happened and a certain amount of delay been observed. This results in that controllers have to adjust the timetables several times during the impact period and the result may be spacial optimal in each individual adjustment but not be globally optimal in temporal dimension.

To fill this gap, this research focuses on designing control methods incorporating weather impact by mapping the weather data into train control model. The weather related initial delay could be eliminated and the future weather impact will be considered so that the solutions are globally optimal from both temporal and spatial dimensions.

Chapter 3 analysed the difference between the existing way of considering weather impact and the proposed method. We introduced a new concept of the predictable, compound and dynamic perturbation (PCDP) as a representation of possible abnormal weather impacts on railway operations.

According to the structure of the railway line, the gridded weather data was mapped onto the time-space diagram to identify the impact of each abnormal weather event in terms of duration and impacted segments. We considered adverse weather conditions which lead to reduced speed limits, as well as the severe weather conditions which lead to the track blockages. According to the severity of the forecasted weather events and the speed restriction guidance of the rail industry, the reduced speed limit was applied to each weather-impacted zone, and trains passing through these zones would have to follow the reduced speed limits or chose alternative routes under the situation of track blockages. In this way, the weather forecasting data could be included in the traffic control models.

Chapter 4 introduced a MILP formulation of timetable rescheduling under PCDP, named PCDPR, to minimise the total arrival delays of all trains at their destinations. In the PCDPR, we studied train traffic on a single-track railway line and formulated the PCDPR as a mixed integer liner programming (MILP) problem, which considers the general constraints for train rescheduling (such as departure and arrival times, minimum headway, overtaking, capacity and avoidance of potential train conflicts), as well as new constraints corresponding to the weather impacts.

The effectiveness of the proposed PCDPR was demonstrated on the Cambrian Line in the UK. Compared with the traditional UISPR which is conducted after trains had been delayed, our proposed PCDPR led to 163 minutes (about 30%) less overall delay by incorporating the forecasted weather disturbance. We also quantified how much the complete information on weather forecast enabled better quality train schedules than partial information. Conducting one rescheduling with the next 10 hours' weather forcasting data resulted in 43 minutes (about 11%) less delay than conducting two rescheduling and each with the next five hours forecasting data.

We also tested 240 randomly generated cases for sensitivity analysis on the same railway line, in which a mix of two different types of speed limitations were grouped into 12 weather categories which corresponded to the different number of weather impacted zones. PCDPR resulted in 184 minutes less delay than UISPR on average and PCDPR gave not more delay than the UISPR. A general trend was also observed:

the more adverse weather events, the bigger gains by adopting the PCDPR approach instead of the UISPR approach.

Chapter 5 designed a GA method to solve the PCDPR problem in Chapter 4 to improve the computation efficiency. We introduced the concept of conflict resolution matrix (chromosome) in which each element (gene) represents the solution for the potential conflict of each train pair. GA generated 112 minutes (about 29%) more delay with 17 seconds (about 89%) less computation time than PCDPR MILP in CPLEX, and 51 minutes (about 9%) less delay with 308 seconds (99%) less computation time UISPR MILP in commercial solver. This indicated GA could generate feasible results with far less computation time than PCDPR MILP , and is absolutely better than UISPR MILP solved by commercial solver in both rescheduled delay and computation time.

Chapter 6 considered a MILP formulation of simultaneous rerouting and rescheduling under PCDP (PCDPRR) with not only speed limitation but also track blockages. The PCDPRR was designed to help dispatchers make specific better routing and train timetabling decisions for each different weather situation so that the total loss could be minimised.

In the PCDPRR, we assumed when a track was blocked, the impacted trains were allowed to shift to the opposite normal condition tracks or back up lines with certain penalty costs to reduce the delay at their destinations. With the modified track occupation constraints on tracks, the rerouting model is suitable for a bidirectional-track network other than just unidirectional-track networks. The potential conflicts were avoided between two opposite trains on the same track when trains were borrowing opposite tracks.

In the numerical examples, we first studied a double track railway line which has only two main routes, one for inbound trains and the other for outbound trains, respectively. The optimised new routes and timetables were generated by commercial solvers in feasible computing time. Under a more conservative situation with bigger changing track penalty cost, less trains would shift to opposite tracks, which resulted in larger system delay minutes and larger total cost.

Then we studied a more complicated network which had an alternative route in addition to two main lines. The case study showed that with different values of penalty cost for delay, using backup lines and using opposite tracks, the model could generate different optimised suggestions under different weather impact situations. When the penalties of using other tracks was sufficiently large, impacted trains would wait until the blockage disappeared rather than switch to other lines.

7.2. Conclusions

The main objective of this thesis is to incorporate weather impact to railway traffic control so that the weather related delay and cost could be minimised. We designed the way to map weather data into railway line and further transfer it to RTC MILP models, i.e. PCDPR and PCDPRR. The experiments showed our PCDPR model can generate 21% less delay on average compared to the existing rescheduling model. The PCDPRR model can help to produce cost effective route and timetable decisions in severe conditions.

As far as we are aware, we are the first to point out the weather impact can be treated as PCDP instead of UISP, so that the future weather impact could be considered and minimised spatially and temporally. We firstly designed a method to map the weather forecast data to the railway line, so that the weather condition on railway can be described precisely in fine resolution. According to the railway industry weather management standards, we then interpreted the abnormal weather as restrictions, i.e. speed limitations and track blockages.

The PCDPR (train rescheduling) model can be used in the adverse weather conditions in which controllers want to optimise the system delay by rescheduling when speed limitations are applied on tracks and trains are running late. In case the computation time is very limited, the designed heuristic GA could be used in generating feasible solutions. The advantage of GA can be especially highlighted in large networks, as GA is polynomial time while the MILP model is exponential time and moreover the parallel computing could be applied in GA.

We also modified the track occupation constraints based on previous research so that the rerouting model can be applied to bidirectional-track networks rather than just unidirectional-track networks. When situation goes worse and the rerouting is needed, the PCDPRR method can help the controllers find the most cost effective route considering the potential risks in using the opposite tracks as well as the cost in missing some stations in the journey when using other lines.

7.3. Perspectives

We have showed the benefits of incorporating the weather impact in railway traffic control. However, there is still a big gap between the theoretical model and their application in industry. Further research should be conducted to make the models more practical in modelling infrastructure and weather, handling deviations between computer models and real world human operations and machinery, and managing the expectation of users such as train operation companies (TOCs), controllers and passengers, etc.

To simplify the model formulation, the proposed models adopted a rather simplified representation of the railway systems and did not explicitly consider railway signalling and safety systems on the interlocking of inbound and outbound routes at stations. Moreover, the proposed models used an average speed limitation for the entire segment and used an average speed in calculating the running time. These simplifications might lead to deviation between the computed timetable and the real world, or even make the model impractical. Further studies are necessary to examine the impact of ignoring these realistic features of the railway system and make the models more precise before they could be used in real world.

There are also scopes to improve the precisionof the PCDPC models regarding the actual weather information provided. (1) In this research, temporary speed limit and track blockage are uniform over the entire track segment between two adjacent stations. However, as in the UK practice, the weather forecast data is routinely mapped onto 2.2km-by-2.2km grids, or even more precisely 1.5km-by-1.5km grids. One possible improvement on practical significance is to adopt finer speed limit regulations which vary along the inter-station segment according to the detailed weather mapping. (2) On another front, as the weather forecast is not absolutely accurate due to the dynamic and stochastic nature of the weather, and some other stochastic disturbances could also happen, a robust train control model considering the uncertainty in weather forecasts may be worthy of investigation. (3) Meanwhile, to ensure efficiency, an automatic mapping program which transfers the weather forecast data to speed restrictions and blockages for the railway control program is needed. (4) An efficient open loop amending progress is also needed in case of large weather forecasting error in real time operation.

To handle deviations between computer models and real world human operations and machinery, real-life analysis and experiments are needed. Speed limitations and track blockages in the algorithms are interpreted from the mitigation requirements. However, in practice, deviations might accrue when drivers implement the requirements. Empirical analysis to quantify the effects of different weather types on actual speed limitations, track blockages and train delays would help to identify the gap between the drivers' accrual operations and the railway industry requirements under weather impact.

In addition to making the model more robust to the uncertainties and deviations, the railway industry will also need to consider the satisfaction or expectation from the user side. New compensation agreements between infrastructure managers and TOCs, and between TOCs and passengers, are needed with the application of PCDPC methods. A step by step information system for passengers regarding the potential delays corresponding to the weather uncertainly is also needed to help passengers make better travel plans while avoiding promising too much on timetables which might mislead them.

References

Akin, D., Sisiopiku, V.P. and Skabardonis, A. 2011. Impacts of Weather on Traffic Flow Characteristics of Urban Freeways in Istanbul. *Procedia - Social and Behavioral Sciences*. **16**, pp.89–99.

Akin, D., Sisiopiku, V.P. and Skabardonis, A. 2011. Impacts of Weather on Traffic Flow Characteristics of Urban Freeways in Istanbul. *Procedia - Social and Behavioral Sciences*. **16**, pp.89–99.

Albrecht, A.R., Panton, D.M. and Lee, D.H. 2013. Rescheduling rail networks with maintenance disruptions using Problem Space Search. *Computers and Operations Research*. **40**(3), pp.703–712.

Andersson, E. V 2014. *Assessment of Robustness in Railway Traffic Timetables*. PhD thesis, Linköping University.

Ariano, A.D., Pranzo, M., Hansen, I.A., D’Ariano, A., Pranzo, M. and Hansen, I.A. 2007. Conflict resolution and train speed coordination for solving real-time timetable perturbations. *IEEE Transactions on Intelligent Transportation Systems*. **8**(2), pp.208–222.

Balakrishnan, H. 2007. Techniques for Reallocating Airport Resources in Adverse Weather *In*: *Proceedings of the 46th IEEE Conference on Decision and Control*. New Orleans, LA, pp.2949–2956.

Balakrishnan, H. and Chandran, B.G. 2014. Optimal Large-Scale Air Traffic Flow Management. Unpublished.

Boccia, M., Mannino, C. and Vasilyev, I. 2013. The dispatching problem on multitrack territories: heuristic approaches based on mixed integer linear programming. *Networks*. **62**(4), pp.315–326.

Brazil, W., White, A., Nogal, M., Caulfield, B., O’Connor, A., Morton, C., Caul, B., Connor, A.O. and Morton, C. 2017. Weather and rail delays: Analysis of metropolitan rail in Dublin. *Journal of Transport Geography*. **59**, pp.69–76.

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

Cai, X. and Goh, C.J.J. 1994. A fast heuristic for the train scheduling problem. *Computers & Operations Research*. **21**(5), pp.499–510.

Caimi, G., Burkolter, D. and Herrmann, T. 2004. Finding delay-tolerant train routings through stations *In*: *Operations Research Proceedings 2004*. Zurich, Switzerland, pp.136–143.

Caimi, G., Fuchsberger, M., Laumanns, M. and Lüthi, M. 2012. A model predictive control approach for discrete-time rescheduling in complex central railway station areas. *Computers and Operations Research*. **39**(11), pp.2578–2593.

Campbell, S.E. and Delaura, R.A. 2011. *Convective Weather Avoidance Modeling for Low-Altitude Routes* [Online]. Lexington, Massachusetts. [Accessed 18 January 2019]. Available from: https://pdfs.semanticscholar.org/e237/b06f23b8e5986ac9fb69b398dfb463b4c30f.pdf.

Carey, M. 1994. A model and strategy for train pathing with choice of lines, platforms, and routes. *Transportation Research Part B*. **28**(5), pp.333–353.

Chan, W.N., Refai, M. and Delaura, R. 2007. An Approach to Verify a Model for Translating Convective Weather Information to Air Traffic Management Impact *7th AIAA Aviation Technology, Integration and Operations Conference* [Online]. Belfast, Northern Ireland, pp.1–13.

Chang, S.C. and Chung, Y.C. 2005. From timetabling to train regulation - a new train operation model. *Information and Software Technology*. **47**(9), pp.575–585.

Chen, L., Schmid, F., Dasigi, M., Ning, B., Roberts, C. and Tang, T. 2010. Real-time train rescheduling in junction areas. *Proceedings of the Institution of Mechanical Engineers Part F-Journal of Rail and Rapid Transit*. **224**(F6), pp.547–557.

Cheung, B.S.N., Chow, K.P., Hui, L.C.K. and Yong, A.M.K. 1999. Railway track possession assignment using constraint satisfaction. *Engineering Applications of Artificial Intelligence*. **12**, pp.599–611.

Chigusa, K., Sato, K. and Koseki, T. 2012. Passenger-oriented optimization for train rescheduling on the basis of mixed integer programming. *IEEJ Transactions on Industry Applications*. **132**(2), pp.170–177.

Corman, F., D’Ariano, A., Hansen, I.A., D’Ariano, A. and Hansen, I.A. 2014. Evaluating disturbance robustness of railway schedules. *Journal of Intelligent Transportation Systems*. **18**(1), pp.106–120.

Corman, F., D’Ariano, A., Pacciarelli, D. and Pranzo, M. 2010. A tabu search algorithm for rerouting trains during rail operations. *Transportation Research Part B*. **44**(1), pp.175–192.

Corman, F., D’Ariano, A., Pacciarelli, D. and Pranzo, M. 2012. Bi-objective conflict detection and resolution in railway traffic management. *Transportation Research Part C*. **20**(1), pp.79–94.

Corman, F. and Meng, L. 2015. A review of online dynamic models and algorithms for railway traffic control. *IEEE Transactions on Intelligent Transportation Systems*. **16**(3), pp.1274–1284.

Corman, F. and Quaglietta, E. 2015. Closing the loop in real-time railway control: Framework design and impacts on operations. *Transportation Research Part C*. **54**, pp.15–39.

D’Ariano, A. 2008. *Improving Real-Time Train Dispatching: Models , Algorithms and Applications*. PhD thesis, Delft University of Technology.

D’Ariano, A., Pranzo, M. and Hansen, I.A. 2007. Conflict resolution and train speed coordination for solving real-time timetable perturbations. *IEEE Transactions on Intelligent Transportation Systems*. **8**(2), pp.208–222.

Dey, K.C., Mishra, A. and Chowdhury, M. 2015. Potential of intelligent transportation systems in mitigating adverse weather impacts on road mobility: A review. *IEEE Transactions on Intelligent Transportation Systems*. **16**(3), pp.1107–1119.

Diego, L. 2016. *Contributions on microscopic approaches to solve the train timetabling problem and its integration to the performance of infrastructure maintenance activities*. PhD thesis, University of Valenciennes and Hainaut Cambresis.

Dollevoet, T., Huisman, D., Kroon, L.G., Veelenturf, L.P. and Wagenaar, J.C. 2017. Application of an iterative framework for real-time railway rescheduling. *Computers & Operations Research*. **78**, pp.203–217.

Dorfman, M.J. and Medanic, J. 2004. Scheduling trains on a railway network using a discrete event model of railway traffic. *Transportation Research Part B*. **38**(1), pp.81–98.

Dündar, S. and Şahin, I. 2013. Train re-scheduling with genetic algorithms and artificial neural networks for single-track railways. *Transportation Research Part C*. **27**, pp.1–15.

Enayatollahi, F. and Atashgah, M.A.A. 2018. Wind effect analysis on air traffic congestion in terminal area via cellular automata. *Aviation*. **22**(3), pp.102–114.

Fang, K., Ke, G.Y. and Verma, M. 2017. A routing and scheduling approach to rail transportation of hazardous materials with demand due dates. *European Journal of Operational Research*. **261**(1), pp.154–168.

Fang, W., Yang, S., Yao, X., Member, S. and Yao, X. 2015. A survey on problem models and solution approaches to rescheduling in railway networks. *IEEE Transactions on Intelligent Transportation Systems*. **16**(6), pp.2997–3016.

Federal Aviation Administration. 2017. FAQ: Weather Delay. [Accessed 11 January 2019]. Available from: https://www.faa.gov/nextgen/programs/weather/faq/.

Flathers, B., Fronzak, M., Huberdeau, M., Mcknight, C., Wang, M. and Wilhelm, G. 2013. *A Framework for the Development of the ATM-Weather Integration Concept*. [Accessed 21 January 2019]. Available from: https://www.mitre.org/sites/default/files/pdf/13_0903.pdf.

Forsgren, M., Aronsson, M. and Gestrelius, S. 2013. Maintaining tracks and traffic flow at the same time. *Journal of Rail Transport Planning & Management*. **3**(3), pp.111–123.

Gatto, M., Glaus, B., Jacob, R., Peeters, L. and Widmayer, P. 2004. Railway delay management: exploring its algorithmic complexity. *Lecture Notes in Computer Science*. **3111**, pp.199–211.

Golding, B.W., Ballard, S.P., Mylne, K., Roberts, N., Saulter, A., Wilson, C., Agnew, P., Davis, L.S., Trice, J., Jones, C., Simonin, D., Li, Z., Pierce, C., Bennett, A., Weeks, M., Moseley, S., Golding, B.W., Ballard, S.P., Mylne, K., Roberts, N., Saulter, A., Wilson, C., Agnew, P., Davis, L.S., Trice, J., Jones, C., Simonin, D., Li, Z., Pierce, C., Bennett, A., Weeks, M. and Moseley, S. 2014. Forecasting Capabilities for the London 2012 Olympics. *Bulletin of the American Meteorological Society*. **95**(6), pp.883–896.

Goodwin, L.C. 2012. *Best practices for road weather management*. Washington, DC.

Hansen, I.A. and Pachl, J. 2014. *Railway Timetabling & Operation* (I. A. Hansen & P. Jorn, eds.).

Higgins, A. 1998. Scheduling of railway track maintenance activities and crews. *The Journal of the Operational Research Society*. **49**(10), pp.1026–1033.

Higgins, A., Kozan, E. and Ferreira, L. 1996. Optimal scheduling of trains on a single line track. *Transportation Research Part B*. **30**(2), pp.147–158.

Higgins, a, Kozan, E. and Ferreira, L. 1997. Heuristic techniques for single line train scheduling. *Journal of Heuristics*. **3**(1), pp.43–62.

Hooghiemstra, J.S., Kroon, L.G., Odijk, M.A., Salomon, M. and Zwaneveld, P.J. 1999. Decision support systems support the search for win-win solutions in railway network design. *Interfaces*. **29**(2), pp.15–32.

Jacobs, J. 2004. Reducing delays by means of computer-aided 'on-the-spot' rescheduling *In*: *Computers in Railway Six*., pp.603–612.

Jardine, A.K.S., Lin, D. and Banjevic, D. 2006. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*. **20**, pp.1483–1510.

Jaroszweski, D., Quinn, A., Baker, C., Hooper, E., Kochsiek, J., Schultz, S. and Silla, A. 2014. *Guidebook for Enhancing Resilience of European Rail Transport in Extreme Weather Events*.

Lake, M., Ferreira, L. and Murra, M. 2010. Minimising costs in scheduling railway track maintenance *In*: *Computers in Railways VII 1*., pp.865–902.

Landex, A. 2009. Evaluation of railway networks with single track operation using the UIC 406 capacity method. *Networks and Spatial Economics*. **9**(1), pp.7–23.

Larsen, R., Pranzo, M., D'Ariano, A., Corman, F., Pacciarelli, D., D'Ariano, A., Corman, F. and Pacciarelli, D. 2014. Susceptibility of optimal train schedules to stochastic disturbances of process times. *Flexible Services and Manufacturing*

Journal. **26**(4), pp.466–489.

Lee, Y. and Chen, C.Y. 2009. A heuristic for the train pathing and timetabling problem. *Transportation Research Part B*.

Li, F., Gao, Z., Li, K. and Yang, L. 2008. Efficient scheduling of railway traffic based on global information of train. *Transportation Research Part B*. **42**(10), pp.1008–1030.

Li, F., Sheu, J. and Gao, Z. 2014. Deadlock analysis, prevention and train optimal travel mechanism in single-track railway system. *Transportation Research Part B*. **68**, pp.385–414.

Lidén, T. and Joborn, M. 2017. An optimization model for integrated planning of railway traffic and network maintenance. *Transportation Research Part C*. **74**, pp.327–347.

Lim, W.X. and Zhong, Z.W. 2018. Re-Planning of Flight Routes Avoiding Convective Weather and the 'Three Areas'. *IEEE Transactions on Intelligent Transportation Systems*. **19**(3), pp.868–877.

Lin, J., Yu, W., Yang, X., Yang, Q., Fu, X. and Zhao, W. 2015. A Novel Dynamic En-Route Decision Real-Time Route Guidance Scheme in Intelligent Transportation Systems. *Proceedings - International Conference on Distributed Computing Systems*. **2015**–**July**, pp.61–72.

Meng, L. and Zhou, X. 2014. Simultaneous train rerouting and rescheduling on an N-track network: A model reformulation with network-based cumulative flow variables. *Transportation Research Part B*. **67**, pp.208–234.

Met Office 2017. Met Office Numerical Weather Prediction models. [Accessed 11 August 2017]. Available from: http://www.metoffice.gov.uk/research/modelling-systems/unified-model/weather-forecasting.

MetOffice 2016. How our forecasts measure up. [Accessed 20 August 2017]. Available from: https://blog.metoffice.gov.uk/2016/07/10/how-our-forecasts-measure-up/.

Mladenovic, S., Veskovic, S., Branovic, I., Jankovic, S. and Acimovic, S. 2016.

Heuristic based real-time train rescheduling system. *Networks*. **67**(1), pp.32–48.

Mu, S. and Dessouky, M. 2011. Scheduling freight trains traveling on complex networks. *Transportation Research Part B*. **45**(7), pp.1103–1123.

Narayanaswami, S. and Rangaraj, N. 2013. Modelling disruptions and resolving conflicts optimally in a railway schedule. *Computers and Industrial Engineering*. **64**(1), pp.469–481.

Network Rail 2013. *Managing track in hot weather (NR/L2/TRK/001/mod14)*. London.

Network Rail 2014. *Network Rail Weather Analysis Report*. London.

Network Rail 2011. *Weather - Managing the Operation Risks (NR/L2/OCS/021)*. London.

Nielsen, L.K., Kroon, L. and Maróti, G. 2012. A rolling horizon approach for disruption management of railway rolling stock. *European Journal of Operational Research*. **220**(2), pp.496–509.

Oliveira, E.S. De 2001. *Solving Single-Track Railway Scheduling Problem Using Constraint Programming*. PhD thesis, University of Leeds.

Pellegrini, P., Marliere, G., Pesenti, R. and 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**(5), pp.2609–2619.

Pellegrini, P., Marliere, G. and Rodriguez, J. 2014. Optimal train routing and scheduling for managing traffic perturbations in complex junctions. *Transportation Research Part B*. **59**, pp.58–80.

Pender, B., Currie, G., Delbosc, A. and Shiwakoti, N. 2013. Disruption recovery in passenger railways. *Transportation Research Record: Journal of the Transportation Research Board*. **2353**(1), pp.22–32.

Peng, F., Kang, S., Li, X. and Ouyang, Y. 2011. A heuristic approach to the railroad track maintenance scheduling problem. *Computer-Aided Civil and Infrastructure Engineering*. **26**, pp.129–145.

Rodriguez, J. 2007. A constraint programming model for real-time train scheduling at junctions. *Transportation Research Part B*. **41**(2), pp.231–245.

Salim, V. and Cai, X. 1995. Scheduling cargo trains using genetic algorithms. *Proceedings of 1995 IEEE International Conference on Evolutionary Computation*. **1**, pp.224–227.

Santos, R., Fonseca, P. and Pais, A. 2015. Planning and scheduling efficient heavy rail track maintenance through a Decision Rules Model. *Research in Transportation Economics*. **54**, pp.20–32.

Sato, K., Tamura, K. and Tomii, N. 2013. A MIP-based timetable rescheduling formulation and algorithm minimizing further inconvenience to passengers. *Journal of Rail Transport Planning and Management*. **3**(3), pp.38–53.

Schachtebeck, M. and Schöbel, A. 2010. To wait or not to wait and who goes first? delay management with priority decisions. *Transportation Science*. **44**(3), pp.307–321.

Snelder, M. and Calvert, S. 2016. Quantifying the impact of adverse weather conditions on road network performance. *European Journal of Transport and Infrastructure Research*. **16**(1), pp.128–149.

The Secretary Delay Attribution Board 2015. Delay Attribution Guide. London

Tormos, P., Lova, A., Barber, F., Ingolotti, L., Abril, M. and Salido, M.A. 2008. A genetic algorithm for railway scheduling problems. *Metaheuristics for Scheduling in Industrial and Manufacturing Applications*. **128**, pp.255–276.

Törnquist, J. 2012. Design of an effective algorithm for fast response to the re-scheduling of railway traffic during disturbances. *Transportation Research Part C*. **20**(1), pp.62–78.

Tornquist, J. and Persson, J.A. 2007. N-tracked railway traffic re-scheduling during disturbances. *Transportation Research Part B*. **41**(3), pp.342–362.

Veelenturf, L., MP, K., V, C., LG, K. and P, T. 2016. A railway timetable rescheduling approach for handling large-scale disruptions. *Transportaion Science*. **50(3)**, pp.841–62.

Veelenturf, L.P., Potthoff, D., Huisman, D. and Kroon, L.G. 2012. Railway crew rescheduling with retiming. *Transportation Research Part C*. **20**(1), pp.95–110.

Virgin Trains 2015. *Managing Service Disruption*. London.

Xia, Y., Van Ommeren, J.N., Rietveld, P. and Verhagen, W. 2013. Railway infrastructure disturbances and train operator performance: The role of weather. *Transportation Research Part D*. **18**, pp.97–102.

Xu, W., Zhao, P. and Ning, L. 2018. Last train delay management in urban rail transit network: Bi-objective MIP model and genetic algorithm. *KSCE Journal of Civil Engineering*. **22**(4), pp.1436–1445.

Yang, L., Li, F., Gao, Z. and Li, K. 2010. Discrete-time movement model of a group of trains on a rail line with stochastic disturbance. *Chinese Physics B*. **19**(10), p.100510.

Yang, Y. 2018. Practical Method for 4-Dimentional Strategic Air Traffic Management Problem With. *IEEE Transactions on Intelligent Transportation Systems*. **19**(6), pp.1697–1708.

Yap, M.D. 2014. *Robust public transport from a passenger perspective: A study to evaluate and improve the robustness of multi-level public transport networks*. Master thesis, Delft University of Technology.

Zhou, Y. and Mi, C. 2013. Modeling and simulation of train movements under scheduling and control for a fixed-block railway network using cellular automata. *Simulation*. **89**(6), pp.771–783.

Zografos, K.G., Madas, M.A. and Androutsopoulos, K.N. 2017. Increasing airport capacity utilisation through optimum slot scheduling : review of current developments and identification of future needs. *Journal of Scheduling*. **20**(1), pp.3–24.

Appendix

Comparison of models in key literature and Wang (2019)

	Rescheduling - Timing model			Rerouting & rescheduling - Network flow model	
Constraints	Higgins et al. (1996)	Li et al. (2014)	**Wang (2019) PCDPR**	Meng and Zhou (2014) P1 model	**Wang (2019) PCDPRR**
Following and overtaking	✓	✓	(4-6)- (4-9)	implied in *	implied in **
Conflict avoidance	✓	✓	(4-14), (4-15)	implied in *	implied in **
Travel time	✓	✓	(4-3)	✓	(6-8)
Departure time	✓	✓	(4-2)	Only origin stations	Only key stations
Opposite train headway	-	✓	(4-10)- (4-13)	-	implied in **
Capacity at stations	-	✓	(4-16)- (4-20)	implied in *	implied in **
Dwell time	-	✓	(4-4) , (4-5)	✓	(6-10)
Stopping stations	-	✓	-	-	(6-9), (6-11)
Weather related	-	-	**(4-21)-(4-27)**	-	**(6-20)-(6-27)**
Flow balance	-	-	-	✓	(6-2)-(6-4)
Time-space	-	-	-	✓	(6-5)-(6-7)
* One way track occupation	-	-	-	✓	implied in **
** **Two way track occupation**	-	-	-	-	**(6-12)-(6-19)**
Objective	Cost of delay + Train operating costs	Cost relating to train types and travel mileages	Total delay	Total delay	Total delay
Linear	No	No	Yes	Yes	Yes