Evaluation of Railway Performance through Quality of Service

by

Menglei Lu

A thesis submitted to

The University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

School of Electronic, Electrical and Systems Engineering

College of Engineering and Physical Sciences

University of Birmingham

March 2016

Abstract

Railways worldwide have experienced unprecedented growth over the last decade. The infrastructure of many European railways is becoming increasingly saturated, while the operators face strong demands to increase services and carry more passengers and freight. There has been a high demand for improving the performance of current railway networks, either by upgrading the infrastructure, or developing better management strategies, to provide better service to customers. This problem is drawing increasing attention from many parties within and beyond the railway industry.

Even though performance is viewed as a very significant problem by the railway industry, different stakeholders (e.g. infrastructure managers, operators, funding agencies) have varied perceptions and requirements towards performance.

In this research, the author first reviewed and analysed the existing performance measures used in industries, including railway transportation. A new generic framework for the measurement and improvement of railway network performance has been proposed, based on the concept of quality of service. The key factors affecting quality of service are identified and analysed.

Secondly, the quantification of performance has been identified based on the Quality of Service framework. Multi criteria decision making has been applied to determine the weights of each Key Performance Indicator in the framework.

Finally, factors within the railway system have been analysed for their impact on performance outputs, to support the development of performance improvement plans. A case study has been conveyed to show the influence of the system properties on the performance measured by quality of service.

Acknowledgements

I am extremely grateful to my supervisors, Prof. Clive Roberts and Prof. Felix Schmid, for their patient, continuous guidance during my study. I much appreciate their suggestions and support they provided in both academic research and daily life. Their extended knowledge, plentiful experience and enthusiastic styles have encouraged me through many of the difficult times.

I would like to thank Dr Gemma Nicholson for her kind support and sparkling inspirations for this thesis and the publications, as well as the pleasant time that we had together doing research and working on projects.

I would also like to extend my gratitude to Prof. Chris Barkan, and all my friends in the UIUC RailTEC group, who have offered me generous help during my academic visit and afterwards.

Many thanks to all members of the Birmingham Centre for Railway Research and Education for their kind help.

As well, I would like to thank all of my friends for their support.

Finally, I would like to thank my husband, my parents and my parents-in-law. It is their love, and understanding that have made my study and life in UK a wonderful time.

Table of Contents

List of Figures

List of Tables

Abbreviations

AC	Accommodation
AHP	Analysis of Variance
ANOVA	Analytical Hierarchy Process
API	Application Program Interface
ARS	Automatic Route Setting
ATC	Automatic Train Control
BRaVE	Birmingham Railway Virtual Environment
CaSL	Cancellation and Severe Lateness
CN	Connectivity
CR	Consistency ratio
DF	Degree of Freedom
DfT	Department for Transport
DOE	Design of Experiment
ECML	East Coast Main Line
ELECTRE	ELimination and Choice Expressing REality
EN	Energy
FPM	Freight Performance Measure
GSM-R	Global System for Mobile communication – Railway
IFs	Influencing Factors
IM	Infrastructure Manager
JT	Journey Time
KPIs	Key Performance Indicators
LTPP	Long Term Planning Process
MAUT	Multi Criteria Decision Making
MCDM	Multi-Attribute Utility Theory
OD	One-Variable-At-a-Time
OVAT	Origin-destination
PC	Passenger Comfort
PIS	Passenger Information System
PPM	Public Performance Measure
PT	Punctuality
QoS	Quality of Service
RS	Resilience
RU	Resource Usage
SIMONE	Simulation Model for Networks
S/N ratio	Signal-to-Noise ratio
TOCs	Train Operating Companies

Chapter 1: Introduction

1.1 The role of railway transportation

Railway has long been one of the most important modes of public transportation. It serves as an important means for transporting both passengers and freight. Ever since its first appearance in the 1800s, railways worldwide have prospered with a lot of innovations and engineering breakthroughs.

Apart from being a vital part of the public transport around the world, railways also help support economic growth, manage road congestion, and combat climate change. With its connectivity and universal access, railways serve not only as a means of transportation, but also as a modern national identity for customers (Network Rail, 2012a).

In recent years, railways worldwide are experiencing unprecedented growth in both market demand and capacity supply. As one of the busiest and fastest growing railways in the world over the last decade, the British rail network has expanded greatly ever since privatisation, conveying a large percent of British passenger and freight movements. There had been more than 1.7 billion passenger journeys in the year 2015, and 12% of freight goods have been moved by rail (Department for Transport, 2016b).

These new developments have brought challenges as well as opportunities. Despite the rapid development, many mainline railways have been facing challenges to their capacity and operations due to the fast growing demand. The fast developments in road and air transportation are also bringing pressure to railways.

In recent years, with the fast growing international businesses, the demand for an integrated transport system which provides a smart mix of rail, road and air traffic has drawn a lot of attention. A robust railway network is vital in realising this integrated transport system.

1.2 Railway operations and performance

With rapid economic growth and social development, there has been a high demand for more train paths and services (Office for Official Publications of the European Communities, 2008). The ever growing demand for a smarter and more competitive railway system is also urging the current rail networks to introduce more effective strategies to improve their performance (Wardman, 2006, Paulley et al., 2006). At the same time many of the existing mainline railways are becoming more and more congested, with little room left for service improvement (Watson et al., 2003, Network Rail, 2008b, Goverde, 2007). Because of the complexity of railway operations, a lot of capacity and resources are not used efficiently (Woodland, 2004, Kontaxi and Riccia, 2012, Krueger, 1999). Passengers, governments, infrastructure managers and train operators all require that the current operations are improved to produce better performance.

Performance measurement and improvement is a key component in the strategic management process (Poister, 2008). In developing strategies for railway performance improvement, the first step is to adopt an effective approach for the evaluation of performance. There have been a number of methods applied for performance evaluation in the railway industry, such as the measurement of effectiveness and efficiency, customer satisfaction and service quality measures (Lan and Lin, 2005, Lan and Lin, 2006, Azadeh et al., 2008, Yu and Lin, 2008, Yu, 2008, Lundberg et al., 2009, Chiou et al., 2010). However, none of these measures can satisfy the performance requirements of all stakeholders when used individually (dell'Olio et al., 2011, Cirillo et al., 2011, Román et al., 2014, Stathopoulos and Marcucci, 2014, De Ona et al., 2015). For sustainable development of the performance of a railway network, a comprehensive framework for performance evaluation and improvement is required.

1.3 Research hypothesis

To meet the stakeholders' expectations towards performance, an effective and efficient measure for performance evaluation and improvement of a railway system is required, as described above, which is mainly due to:

- The varied opinions on factors affecting performance;
- The requirements for the evaluation of technical and operational factors in the railway system;
- The need for improvement in performance;
- The need for efficient utilisation of current performance;
- The shortcomings of conventional performance evaluation measures.

The focus of this research and the hypothesis is:

- The performance of a railway system can be evaluated with a number of technical and operational parameters.

To demonstrate that the statement is true, the hypothesis is split into the following elements:

- A wide range of parameters are involved in the evaluation of railway performance;

- The performance of a railway system cannot be evaluated with only technical or only operational parameters;
- Current performance measures can only satisfy the performance requirements of some stakeholders;
- An evaluation framework based on the concept of quality of service can be developed to cover most parameters for performance evaluation;
- The quality of service framework can be used to compare performance on different routes, compare different future options for the same route, and to develop more efficient and effective strategies for the improvement of performance.

All of these statements will be investigated in the research presented in the thesis.

1.4 Methodology

In this thesis, a railway performance evaluation framework has been constructed, using the techniques of multi-criteria decision analysis. The criteria weights are generated from a focus group study, and then applied in the calculation of the performance value. This value is analysed to reveal the impact of system elements on the overall performance. A performance improvement plan can be developed with the results from the analysis.

1.5 Document Structure

This thesis is structured into eight main chapters, listed as below:

- Chapter 1 gives an introduction of the thesis;
- Chapter 2 provides an introduction to the railway planning and operations process;

- Chapter 3 reviews related literature of railway performance measurement;
- Chapter 4 illustrates the performance evaluation framework, based on the Quality of Service concept;
- Chapter 5 specifies the quantification of Quality of Service, from the determination of criteria weights to the calculation of the performance value;
- Chapter 6 discusses the application of the Quality of Service framework;
- Chapter 7 explains the procedures of using the framework for performance evaluation and improvement with a case study based on a real world scenario;
- Chapter 8 presents the conclusions and discussions about future work.

Chapter 2: Fundamentals of railway planning and operations

This chapter gives an overview of the railway operational process, from the planning phase to daily operations. A review of the British railway architecture is presented, including timetable planning, traffic management and train control.

2.1 The railway planning and operation process

2.1.1 The hierarchical railway planning and operation process

British railway companies face the challenge of accommodating the expected growth of transport demand while improving train punctuality. One of the approaches to satisfy the increased demand is to construct more railway networks. In recent years, there have been a number of large rail investment projects taking place, such as Crossrail and HS2. However, it is not always possible to put in new railway lines in many areas, due to time constraints, capital, environmental and even political and social causes. For many existing networks, a more efficient use of the available resources is necessary.

Railways are complex systems consisting of interconnections and interactions of several subsystems (e.g. track, rolling stock, operation). The composition of each subsystem and their behaviour makes it even harder to predict the overall performance of a railway network. To ensure that the railway delivers the desirable results, all system components need to be carefully examined, including infrastructure, timetable, rolling stock, crew, etc. Typical outputs expected from the railway include (Network Rail et al., 2014):

- Safety;
- Capacity;

- Train performance;
- Availability;
- Quality of service;
- Carbon impact;
- Cost efficiency.

The usual way railways manage their performance is through carefully designed plans of operations, followed by real-time policies to manage disturbances. Railway planning involves a number of steps from prediction of traffic demand to the control of daily operations. In this process, trade-offs are required to balance between traffic demand and provided services, to achieve both high customer satisfaction and smooth operations.

With regard to timescale, the railway planning and operation process can be represented by the following hierarchical process (Figure 2-1):

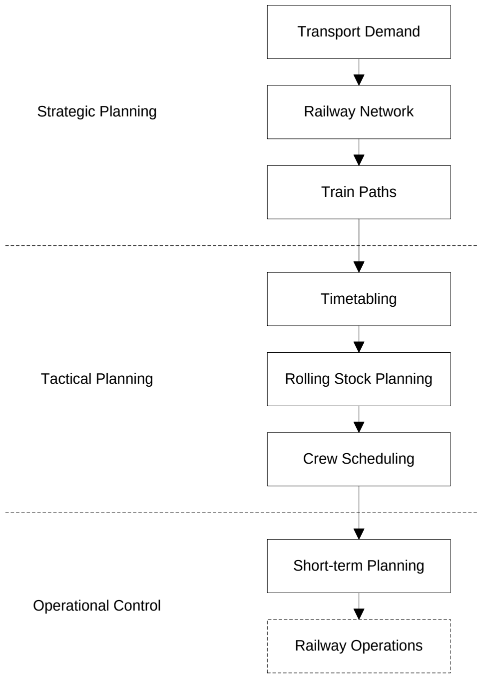


Figure 2-1: The railway planning and operation process

In a typical railway planning process, three levels of planning horizons are usually distinguished: strategic planning, tactical planning and operational management (Huisman et al., 2005).

Strategic planning is concerned with strategic, fundamental policy decisions regarding the capability of the railway network to meet future demand. In this phase major investments and long-term planning are engaged， such as construction of new lines, hiring and training new staff and amending operational rules. The time horizon is often several years. Through management of capacity and resources, strategic planning aims at delivering the sufficient structure of train services to meet market requirements.

Tactical planning deals with intermediate phase management which translates strategic decisions into specific plans relevant to each area of operation. With tactical planning, schedules for rolling stock and crew are planned alongside timetables.

Operational control is concerned with short-term management of traffic and operations. The strategies are performed to form weekly timetable plans, and to deal with rescheduling during disruptions and disturbances on a day-to day basis.

A detailed description of these three horizons will be demonstrated in the following sections.

2.1.2 British Railways planning and operation process

As the infrastructure manager, timetable planner and system operator of the British railway, Network Rail is responsible for the planning and operation of the network. They aim for safe, reliable and efficient operation of the network, with trains timetabled in a way that can make the best use of the network (Network Rail, 2017). They achieve these goals by following a standard process (Figure 2-2):

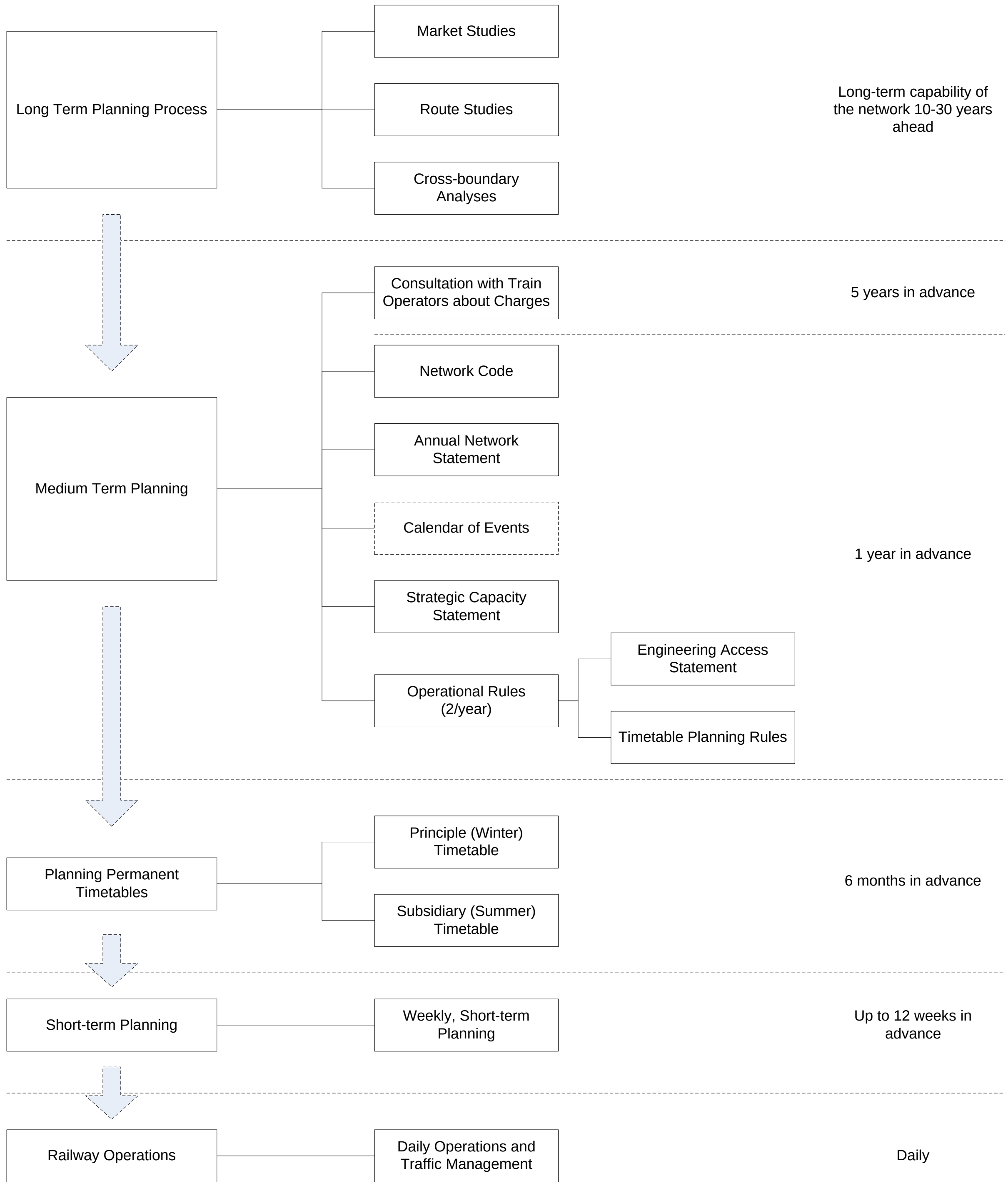


Figure 2-2: Network Rail planning and operation process (summarised from (Network Rail, 2012b))

This process corresponds to the three planning horizons as described in Figure 2-1, with strategic planning covering the long term planning process and part of the medium term planning; tactical planning being the medium term planning of capacity, events, rules and permanent timetables; and operational control being the short-term planning and operations.

The process starts with the development of strategies to meet future demands on a whole network basis using the Long Term Planning Process (LTPP). This is when Network Rail works with its stakeholders to identify strategic goals and forecast traffic demand. Major investments are assessed, and the priorities for capacity use are decided. The LTPP delivers the future railway capability in 10-30 years time.

In medium term planning, detailed plans are made for the railway network, capacity and performance, operational rules, and timetable events. The British railway is planned on a five year time span defined by Network Rail, and the current control period (CP5) runs from 2014 to 2019. The planning process for CP6 will start with train operator consultations on changes from CP5, and runs until one year before operations. This process provides guidelines for the planning of permanent timetables.

Two permanent timetables are produced each year, a principle (winter) timetable and a subsidiary (summer) timetable, about six months before operations. Both working timetables and passenger timetables are published with each release. Rolling stock movements are regulated within the working timetables, and crew schedules are developed accordingly.

Short term adjustments to the timetable can be made at least one week before operations. The weekly timetables are published 12 weeks before the commencement date of the permanent timetable (Network Rail, 2015a), and last minute path requests can be processed until the day before operations.

Traffic management and train control are performed in daily operations to deal with disruptions, disturbances and accidents that occur on the railway network.

2.2 Strategic planning

2.2.1 Transport demand

The first step of strategic planning is to estimate transport demand. Factors considered in demand forecasting include not only attributes within the railway, but also a number of external elements such as social-economic factors and competition from other transport modes (Wardman, 2006, Paulley et al., 2006). The railway industry in Great Britain generally uses the principal demand forecasting parameters as defined in the Passenger Demand Forecasting Handbook (PDFH). This is the industry standard document which defines the forecasting framework and relevant parameters, and provides recommended values for parameters (ATOC, 2016). Demand forecasts based on the PDFH provide guidance for investment planning, pricing, timetabling and operating decisions, and business planning and budgeting.

2.2.2 Railway network

Railway network planning aims at the long term development of railway infrastructure based on the railway transport and traffic demand. It involves the planning of stations, tracks, signalling, electrification and safety facilities. The construction of new railway infrastructure requires a huge amount of investment and time, and has a great environmental, economical and social impact. The planning process can take decades, with thorough strategic studies and political discussions.

For existing railway infrastructure, upgrades, maintenance and renewal can be carried out to accommodate future traffic. Great Britain has the longest history of railway transport in the world, and the current infrastructure is the result of multiple historical strategic developments. Figure 2-3 gives an overview of the British national railway network. The undergoing projects for new lines and the upgrading of existing lines are defined in the Route Studies published by Network Rail on an occasional basis (Network Rail, 2014a).

Figure 2-3: National Railway Map Great Britain (ATOC, 2013)

2.2.3 Train paths

A train path is a railway connection between an origin station and a destination station, characterised by a certain train service. The train paths form the base of a working timetable. Network Rail defines the Strategic Train Paths in their annual Strategic Capacity Statements, and the train operating companies need to reserve train slots in the working timetable to ensure their access rights to train paths (Network Rail, 2015b).

2.3 Tactical planning

2.3.1 Timetabling

Timetabling is the process of planning a feasible schedule for each train path based on the available infrastructure, with consideration of track, junctions, platforms, and the signalling system. A railway timetable should perform the following functions (Pachl et al., 2008):

- Plan the train paths for optimum use of infrastructure;
- Ensure sufficient train separations and avoid train traffic conflicts;
- Provide traffic information to passengers;
- Support traffic control, locomotive and rolling stock usage and crew scheduling.

In Great Britain, the Department for Transport (DfT) regulates the delivery of rail services via franchising, and thus it plays a role in the timetabling process (Department for Transport, 2016a). Timetable generation begins with the train operators submitting their service bids and track access requirements to the infrastructure manager, i.e. Network Rail. With the collected train service specifications, Network Rail produces a conflict-free draft timetable. After that the train operators negotiate with Network Rail to resolve the operational differences and make modifications to the draft timetable. The final timetables are published 6 months before implementation (Ho et al., 2012).

The published passenger timetable is based on a detailed working timetable for railway personnel and traffic management systems. It shows all train movements on the railway network, including freight trains, empty trains and trains going in and out from depots (Network Rail, 2016). For each individual train service, the working timetable specifies:

- dates of operation and service type;
- origin and destination station;
- train stops at stations and intermediate locations;
- running tracks, including tracks through junctions/sidings, stations and platforms;
- scheduled departure and arrival times at stops;
- margins and allowances for performance and operational requirements.

Rolling stock movements and the allocation of crew members can also be planned based on the working timetables.

2.3.2 Rolling stock planning

The efficient circulation of rolling stock is an important issue for train operators, as the investment is huge, and the life cycle of rolling stock is usually decades (Peeters and Kroon, 2008). In tactical planning, the train operators need to determine the type of rolling stock and the number of units to be allocated to train services. The decisions are made to meet the capacity goals and speed requirements, to accommodate the passenger flow on each line.

Customer preferences need to be taken into consideration as well. For instance, if it is an intercity line, passengers would value the comfort and cleanliness of train carriages; on the other hand, for regional services, journey time can be the main preference (Abbink et al., 2004). The total length of trains should also match the lengths of platforms along the route, if Selective Door Opening (SDO) is not available.

For operational efficiency, it is preferred to use only a limited number of rolling stock types, and the ones that are easy to be coupled or decoupled. This would improve the robustness of rolling stock circulation. Energy and maintenance cost are also main concerns of train operators (Huisman et al., 2005).

2.3.3 Crew scheduling

Railway crew scheduling is the process of matching the number of crews available to required duties on train runs. It is a key element to ensure efficient and reliable operations. The work activity of all crew members, such as drivers, guards, and catering staff, are planned on a daily basis.

For each train service, the crew start from a certain crew depot at a designated start time, perform a number of duties (sequences of transport tasks), then return to the destination depot at the required end time, according to the timetable (Mingozzi et al., 1999).

To plan a crew schedule, the transport tasks are grouped into feasible duties, with constraints from operational and contractual requirements, then assigned to the crews (Hanafi and Kozan, 2014). Crew rests, activities and spare coverage need to be implemented in the plan.

Train operators seek to reduce the number of duties as much as possible. In Great Britain, train crews account for about 20 – 25% of the total operating cost (Kwan, 2011). A good crew schedule would have the duties planned in a way to minimise costs, with consideration of the characteristics of different crew, such as qualifications, pre-assigned tasks, and individual requests (Freling et al., 2004). It is also necessary to have a number of available crews that can await orders. Thus trade-offs are necessary between provision of spare crews and staffing costs. Furthermore, with efficient crew schedules, there would be spare room for crew movements, and the robustness of timetables can be improved.

2.4 Operational control

2.4.1 Short-term planning

Short term modifications can be made to the working timetables to accommodate ad hoc demand, especially those additional train paths required by freight operators. Timetable variations are planned on a week by week basis, and the weekly timetables are published 12 weeks in advance. After that, necessary restrictions of use can be issued by Network Rail to produce a finalised timetable. All timetable participants need to be notified as soon as practicable.

2.4.2 Daily operations and traffic management

2.4.2.1 Railway signalling

The signalling system plays a vital part in maintaining safety of the railway network. It controls the movements of trains along the tracks, and keeps trains from colliding into each other by retaining a safe separation between them. The signalling system is also responsible for controlling train movements at junctions and regulating train movements according to service and speed requirements.

In modern railways, most networks are equipped with fixed block signalling, where tracks are split into a series of sections, known as block sections. The safety distance is maintained by allowing only one train at a time to occupy each block section. The position of each train is detected and reported by the block section(s) it occupies. A block section is protected by signals which give instructions to drivers. If the block is empty, the driver would be given permission to proceed, or if another train has entered the block, the driver must stop at the signal. With multiple-aspect signalling, the minimum distance between two successive trains can be reduced significantly while maintaining a safe separation.

In some advanced networks, moving block signalling systems are adopted. The positions of trains are reported continuously, thus permitting more flexible and accurate control of train movements. The safety distance is kept by ensuring the minimum headway between trains (Takeuchi et al., 2003). This requires an efficient communication system between trains and control centres.

2.4.2.2 Train control

The European standard has required member countries to implement the same train control system to ensure rail transport competitiveness. The European Train Control System (ETCS) is presented with four levels (UIC, 2008):

- Level 0: ETCS-compliant trains operate on the existing non-ETCS-compliant network;
- Level 1: ETCS equipment is installed both on board and lineside, and train data is transmitted on spot with Eurobalises;
- Level 2: Train location and running data is transmitted continuously with Global System for Mobile communication – Railway (GSM-R), and signalling information is communicated to the driver in-cab;

- Level 3: No trackside equipment is used. Train data is transmitted continuously with GSM-R, and moving block signalling is applied.

In Great Britain, in areas where ETCS is not yet available, a variety of train control and protection systems are implemented, such as Train Protection and Warning System, Automatic Train Protection, Automatic Train Control (consisting of Automatic Train Protection and Automatic Train Operation) and Automatic Route Setting.

Train Protection and Warning System (TPWS) and Automatic Train Protection (ATP) have been installed on most of the British railway network to reduce signals passed at danger (SPAD). These two systems consist of both track and train equipment, and can initiate a brake demand automatically (Office of Rail and and Road, 2017).

Automatic Train Control (ATC) is conducted with a communication network based on Eurobalise for intermittent communication, or Global System for Mobile communication – Railway (GSM-R) for continuous communication, to report train positions, check the driver's response to signalling commands, give instructions on driving and intervene when necessary. With ATC, the safety level of railways is enhanced, and the train running performance in terms of smoothness, journey time and energy consumption can be improved.

Automatic Route Setting (ARS) is an electronic or relay based system which sets train routes automatically based on information from trains, tracks, interlocking and timetables. A command is generated to set the route to the next signal when a train reaches a certain signal. When exiting, ARS checks if the succeeding route is available. If not, a message is displayed to the driver about the actions to be taken (Kuhn, 1998). An ARS system helps the control of train movements through junctions.

2.4.2.3 Rescheduling

In railway operations, train services are designed to run to a pre-defined conflict free timetable. However, in daily practice, trains are not always operated according to the timetable. When disturbances occur, one or more trains may be delayed, disrupting the traffic. In this scenario, a signaller must produce a new conflict-free train schedule that can help reduce the influence of delays and recover from the timetable disturbances.

In Great Britain, train control and rescheduling is implemented by a number of Network Rail control centres, which are responsible for control regions. When a disturbance happen, new start and end times would be computed for the affected trains, along with the sections they will traverse (Törnquist, 2007). The sequence of trains passing junctions would be adjusted accordingly. With successful rescheduling strategies, the total and accumulated delays would be reduced, and operations would return to schedule within a short time.

2.5 Summary

This chapter provides an overview of the planning and operation process of passenger railways in Britain. Railway traffic relies on a wide range of technical elements such as infrastructure, rolling stock and signalling systems, as well as a number of human factors. All these would bring a lot of unpredictability in operations. There might be a lot of fluctuation in the performance of a railway network. In the next chapter, the measurement of railway performance will be discussed. It plays a vital role in supporting the optimisation of railway planning and operations.

Chapter 3: Evaluation of railway performance

3.1 The importance of performance measurement in the railway industry

Performance is a comprehensive concept which covers a number of aspects, such as availability, reliability and punctuality of fulfilling a given claim. Industry and businesses measure performance to evaluate and monitor their behaviour and to make improvement plans accordingly. A variety of performance measures have been applied in many management processes, such as strategic and operational management, human resources, organisational management, information systems, and marketing (Franco-Santos and Bourne, 2005, Neely, 2005). With an effective performance management system, the manager can maintain control over the organisation, and relevant parties can monitor and make sure the organisation produces the desired results.

Performance measurement aims to produce objective, quantitative information of the performance of organisations that would support management, decision making and performance improvement (Poister, 2008). People measure performance for a variety of reasons. A summary of the common purposes of performance measurement has been given by Behn (2003) as to (1) evaluate how well the organisation works; (2) control the behaviour of subordinates; (3) manage budget; (4) motivate staff, collaborators and stakeholders; (5) promote the organisation; (6) celebrate the accomplishments; (7) learn the strength and weaknesses; and (8) improve performance. In general, performance management involves monitoring and in-depth examination of performance towards pre-established goals.

In the railway industry, performance evaluation has long been a problem as well. In the rail context, performance shows the capability of a railway system to fulfil their claim of transporting passengers and freight. European railway companies face the challenge of accommodating the expected growth of transport demand while improving train punctuality. Because of the complexity of railway systems, a lot of capacity and resources are not used efficiently (Woodland, 2004, Kontaxi and Riccia, 2012, Krueger, 1999). At the same time, many of the existing mainline railways are congested (Watson et al., 2003, Network Rail, 2008b, Goverde, 2007). There is a high demand for more train paths and services. Passengers, the government, infrastructure managers, timetable planners and train operation companies all require that the current operations be improved. Competition from other transportation modes is also urging the railways to adopt more effective strategies to improve their performance (Lan and Lin, 2005, Lan and Lin, 2006, Azadeh et al., 2008, Yu and Lin, 2008, Yu, 2008, Lundberg et al., 2009, Chiou et al., 2010). A more efficient use of existing infrastructure is necessary. It is expected that better performance can be achieved with improved performance management strategies (Poister, 2008). Measuring and improving the performance of railway networks has thus become a main theme in railway planning and management.

In developing strategies for railway performance improvement, the first step is to adopt an effective approach for the evaluation of performance. There have been a number of methods applied for performance evaluation in the railway industry, such as the measurement of effectiveness and efficiency, customer satisfaction and service quality measures (Lan and Lin, 2005, Lan and Lin, 2006, Azadeh et al., 2008, Yu and Lin, 2008, Yu, 2008, Lundberg et al., 2009, Chiou et al., 2010). However, none of these measures can satisfy the performance requirements of all stakeholders when used individually (dell'Olio et al., 2011, Cirillo et al., 2011, Román et al., 2014, Stathopoulos and Marcucci, 2014, De Ona et al., 2015). For sustainable development of the performance of a railway network, a comprehensive framework for performance evaluation and improvement is required.

3.2 Stakeholders' requirements towards performance

Railways are huge, complex systems involving a number of stakeholders, and each individual stakeholder has got its own requirements and expectations towards performance. Thus when it comes to performance management, a wide range of civil, electrical, mechanical and environmental engineering elements within the railway system are considered. Table 3-1 gives a glimpse of the differences in stakeholders' views towards performance from the aspect of capacity. The market, infrastructure planners, timetable planners and operators each have their own understanding of and requirements for capacity. When evaluating performance, more elements would need to be taken into consideration, such as resource management, information needs and environmental concerns.

There is a vast variety of stakeholders in the railway system; the main ones include passengers, government, infrastructure managers (IM), timetable planners and train operation companies (TOCs). In many countries, including Great Britain, there is not an independent timetable planner role, as the infrastructure manager is also responsible for producing timetables.

Market (customer needs)	**Infrastructure planning**	**Timetable planning**	**Operations**
Expected number of train paths (peak) Expected mix of traffic and speed (peak) Infrastructure quality need Journey times as short as possible Translation of all short- and long-term market-induced demands to reach optimised load	Expected number of train paths (average) Expected mix of traffic and speed (average) Expected conditions of infrastructure Time supplements for expected disruptions Maintenance strategies	Requested number of train paths Requested mix of traffic and speed Existing conditions of infrastructure Time supplements for expected disruptions Time supplements for maintenance Connecting services in stations Requests out of regular interval timetables (system times, train stops, etc.)	Actual number of trains Actual mix of traffic speed Actual conditions of infrastructure Delays caused by operational disruptions Delays caused by track works Delays caused by missed connections Additional capacity by time supplements not needed

Table 3-1: Different views of capacity (UIC, 2004)

In the European Union, a number of authorities such as governments and independent regulators, have proposed their own performance goals and performance indicators to assist their strategic planning. For Great Britain, the Department for Transport (DfT) has published a set of performance indicators as part of their business plan (Department for Transport, 2012):

- Rail subsidy per passenger mile: the amount of subsidy paid to train operating companies, both directly by the DfT and indirectly via Network Rail;
- Cost of running the rail network: the operating cost of running the railways, including the costs incurred by DfT franchised train operators and the grant paid to Network Rail as part of the government subsidising the railways;
- Proportion of trains running on time, which is measured by the Public Performance Measure (PPM);

- Rail passenger miles: the total number of miles covered by passengers on rail journeys over a set time period.

From these indicators, it can be seen that the government is concerned about the cost efficiency of TOCs and the infrastructure managers (IM), the available train paths and passenger seats provided, and the punctuality and reliability of train operations.

For IMs the main concerns are stability of the network and efficient investment. The timetable planner, on the other hand, aims to provide resilient and stable timetables that can offer competitive travel times as well as withstand perturbations. As the IM and timetable planner, Network Rail measures their performance in terms of (Office of Rail and Road, 2015, Network Rail, 2014b):

- Safety of passengers and staff;
- Asset management;
- Network operations: the punctuality and reliability of operations, measured by the Public Performance Measure (PPM), Freight Performance Measure (FPM), Cancellation and Severe Lateness (CaSL);
- Customer service: passenger satisfaction and customer satisfaction;
- Infrastructure development;
- Financial performance;
- Data quality.

All these indicators apply to the performance measurement of the railway network, apart from safety, financial performance and data quality. Financial performance and data quality are performance monitors for the behaviour of the company. As to safety, it is a factor taken as given by most customers (as shown in safety surveys), and therefore not used for this analysis.

The train operating companies view railway performance as the punctuality and reliability of their services, using assessments such as PPM[1] and other self-defined punctuality measures (London Midland, 2016, Virgin Trains East Coast, 2016). Because their aim is to attract customers and make profit, TOCs also value customer satisfaction and efficiency of investment.

From the passengers' perspective, a high level performance is the main reason to use railway transportation. Their requirements towards performance are shown in the National Passenger Surveys, categorised as satisfaction over the whole journey, at stations and on trains (Transport Focus, 2016). The elements affecting satisfaction fall into a wide range, from comfort, passenger information, station and train facilities to available seats, length and mix of journeys, punctuality, delay management, etc.

To make changes to the existing railway network for performance improvements, the needs of all stakeholders are to be taken into consideration. It is possible that different stakeholders have varied opinions on the same matter. It is necessary to choose a performance measure that can accommodate the requirements from all stakeholders while allowing trade-offs between conflicting opinions.

[1] The Public Performance Measure (PPM) is a measure of punctuality and reliability of passenger services in Great Britain. It compares train runs against the planned timetable, and shows the percentage of trains arriving at their terminating stations 'on time' (within 5 minutes for London and South East or regional services, or 10 minutes for long distance services).

3.3 The approaches for performance measurement

3.3.1 The effectiveness and efficiency measure

Many studies have dealt with railway performance evaluation. On a system level, a number of researchers have focused on the evaluation of railway performance by assessing effectiveness and efficiency in the industry (Lan and Lin, 2005, Lan and Lin, 2006, Azadeh et al., 2008, Yu and Lin, 2008, Yu, 2008, Lundberg et al., 2009, Chiou et al., 2010). The effectiveness and efficiency measure is one of the most popular performance measurement strategies. Effectiveness is the system's ability to meet certain goals, while efficiency measures the relationship of the produced outputs to the resource inputs.

A number of effectiveness and efficiency definitions are used in the literature to assess performance. Among these, technical efficiency and technical effectiveness are the most recognised (Oum and Yu, 1994, Kerstens, 1996, Chapin and Schmidt, 1999, Karlaftis, 2004). Technical efficiency measures the transformation of inputs (lines, vehicle, labour, etc) into outputs (passenger train kms, freight train kms), and technical effectiveness measures the relationship between inputs and consumed services (passenger kms, freight kms). Other definitions include cost efficiency, service effectiveness, and cost effectiveness. Fielding *et al.* (1985) defined cost efficiency as the ratio of service outputs (vehicle hours, vehicle miles, capacity miles, service reliability) to service inputs (labour, capital, fuel), cost effectiveness as the ratio of service consumption (passengers, passenger miles, operating revenue) to inputs, and service effectiveness as the ratio of service consumption to service outputs.

To summarise, in effectiveness and efficiency measures, the main entities include:

- inputs: lines, vehicles, labour, energy, capital;

- outputs: passenger train kms, freight train kms, vehicle journeys, service reliability;
- consumptions: passenger kms, freight kms, passengers, revenue.

The entities have covered a great part of railway performance, yet there are still a lot of elements within the system that have not been considered. As discussed in the previous section, stakeholders often have high expectations towards many other elements, requiring less delays, more resilient timetables, connectivity, passenger comfort, etc. In particular, none of these effectiveness and efficiency measures has put enough attention on the requirements and demands from the market. This is due to the fact that performance measures are oriented from the position of infrastructure managers or government agencies. A method that takes in the demands of service providers and customers would become helpful to implement these performance measures.

3.3.2 Service quality as a performance measure

Quality of Service is an important concept widely used in many domains. It is a measure of how well the service level delivered matches customer expectations (Lewis and Booms, 1983). The aim of using the Service Quality concept is to take satisfaction in consideration when managing performance, to achieve higher customer loyalty, higher profitability and lower cost (Cavana et al., 2007).

In modern public transport systems, the quality of passenger services is one of the major concerns of transportation providers. Their main objective is to achieve customer satisfaction and loyalty, to maintain their competitiveness in passenger transportation. The achieved service quality is highly dependent on the service provider's decisions on where, when and how they'll provide the service within their financial plan. Better service quality would mean higher customer satisfaction as well as potentially higher costs. The management of service quality is the process of finding the balance between customer expectations and service provider's capability, with a certain base level of service (Transportation Research Board, 2013).

Customers' perceptions towards service quality, however, vary significantly from service provider's views. Train operating companies (TOCs) tend to use a bottom up approach to monitor service quality, focusing more on the central attributes of service quality. In the bottom up approach, the TOCs advocate changes and strategies in the railway network, and then work to meet market requirements. The quality criteria involved are outlined in major European research initiatives (EQUIP, 2000, European Committee for Standardization, 2002), which include: availability, accessibility, information, time, customer care, comfort, security, environmental impact, reliability, asset utilisation, technical performance, customer satisfaction, safety and security, and external influences on the operator. TOCs usually select from these criteria to support their measurement of service quality.

Conversely, customers prefer a top down approach, which emphasises the service delivered to them, and then goes down into the detailed specifications of daily operations. Research on customers' perceptions has shown that service characteristics such as reliability, punctuality, travel time, frequency, comfort and cleanliness, information, fare level, safety, customer service and integration with other modes of transport are found to be important in the evaluation of service quality (dell'Olio et al., 2011, Cirillo et al., 2011, Román et al., 2014, Stathopoulos and Marcucci, 2014, De Ona et al., 2015)

To generate a comprehensive model for service quality measurement, views of service providers and customers both need to be taken into consideration. The factors involved need to be categorised to allow systematic assessment of performance.

3.4 The need for a new performance measurement framework

The usual way in which railways manage their performance is through carefully designed plans of operations, followed by real time policies to manage disturbances. From previous discussions, the use of effectiveness and efficiency or service quality individually cannot satisfy the need of performance evaluation for all stakeholders. If infrastructure managers, train operators, government and passengers are all to be involved in the performance measurement, a novel framework is required to present the demands from all these stakeholders in a clear and organised way. Both technical characteristics and operational functions must be integrated, with consideration to passenger satisfaction and human factors.

Based on conventional performance measures discussed in this chapter, the new performance measurement framework should cover a number of performance indicators, including passenger or freight kms, vehicle journeys, timetable related factors (e.g. infrastructure occupation, timetable stability, feasibility, robustness, and resilience), connections, comfort, energy, and staff, etc. When assessing the performance of a railway network, customer satisfaction survey and passenger flow model are not always feasible or accurate. Thus some indicators would not be utilised in the framework, such as cleanliness, information, security, fare level and customer service. Capital and revenue are also not considered, as the evaluation would involve a number of social and economic parameters outside the railway system. These indicators would become helpful with a traffic demand analysis, which is beyond the scope of this research.

Safety is not taken into account in the measurement either. The railway is widely recognised as one of the safest modes of transport in the EU. In the rail industry, safety is measured with Safety Performance Indicators (Office of Rail and Road, 2015, Network Rail, 2014b). Furthermore, the current safety level in a railway network can be reflected by safety surveys. With this information available, most customers take safety as a given, and thus it is not used in this research for quality of service analysis. The construction of a comprehensive performance evaluation framework based on the service quality concept will be illustrated in the next chapter. This framework would be capable to assist the development of more efficient and effective strategies for the improvement of railway performance.

3.5 Summary

In this chapter, conventional measures for performance evaluation in railway transportation are discussed. The performance of a railway system can be evaluated with a number of technical and operational parameters, as described in the effectiveness and efficiency measures, and the service quality measures. However, the use of either class of measures individually cannot cover the performance requirements of all stakeholders, such as infrastructure managers, train operators, government and passengers. A comprehensive framework which takes in the opinions of all stakeholders, based on the existing performance measures, particularly service quality, would be more effective to convey performance evaluation.

Chapter 4: The railway performance evaluation framework

As summarised in Chapter 3, when evaluating railway performance, it is difficult to address the key requirements of all stakeholders in the system. In order to include more aspects in the assessment, a novel framework, conceptually based on 'quality of service', is proposed for the evaluation of railway performance. In this chapter, a detailed description of the Quality of Service framework is demonstrated. The structure, the inputs and the outputs of the framework are illustrated. With this framework, railway performance can be evaluated with a number of key criteria which reflect the behaviour of the network. The content of this chapter has been based on a previous published paper written by the author (Lu et al., 2013).

4.1 The Quality of Service framework

The Quality of Service framework is focused on providing a comprehensive framework that covers the main attributes of the railway system, which supports performance evaluation and improvement. In the proposed framework, the concept of "quality of service" is used to represent the expectations towards the performance of railway systems, covering the factors affecting the network performance from the perspectives of all stakeholders, such as operators, timetable planners, infrastructure managers, passengers and the government. The framework combines the traditional performance measure of effectiveness and efficiency with the service quality assessment from the train operator and customer's perspective. Railway network performance in terms of quality of service is achieved by a combination of strategic factors and tactical factors that influence the way in which a railway system or network supports commercial objectives. With the Quality of Service (QoS) framework, it is possible to reveal the relationship between the system elements with the overall performance. A detailed overview of the QoS framework is shown in Figure 4-1. The full framework with all the key performance indicators, key measures and influencing factors is listed.

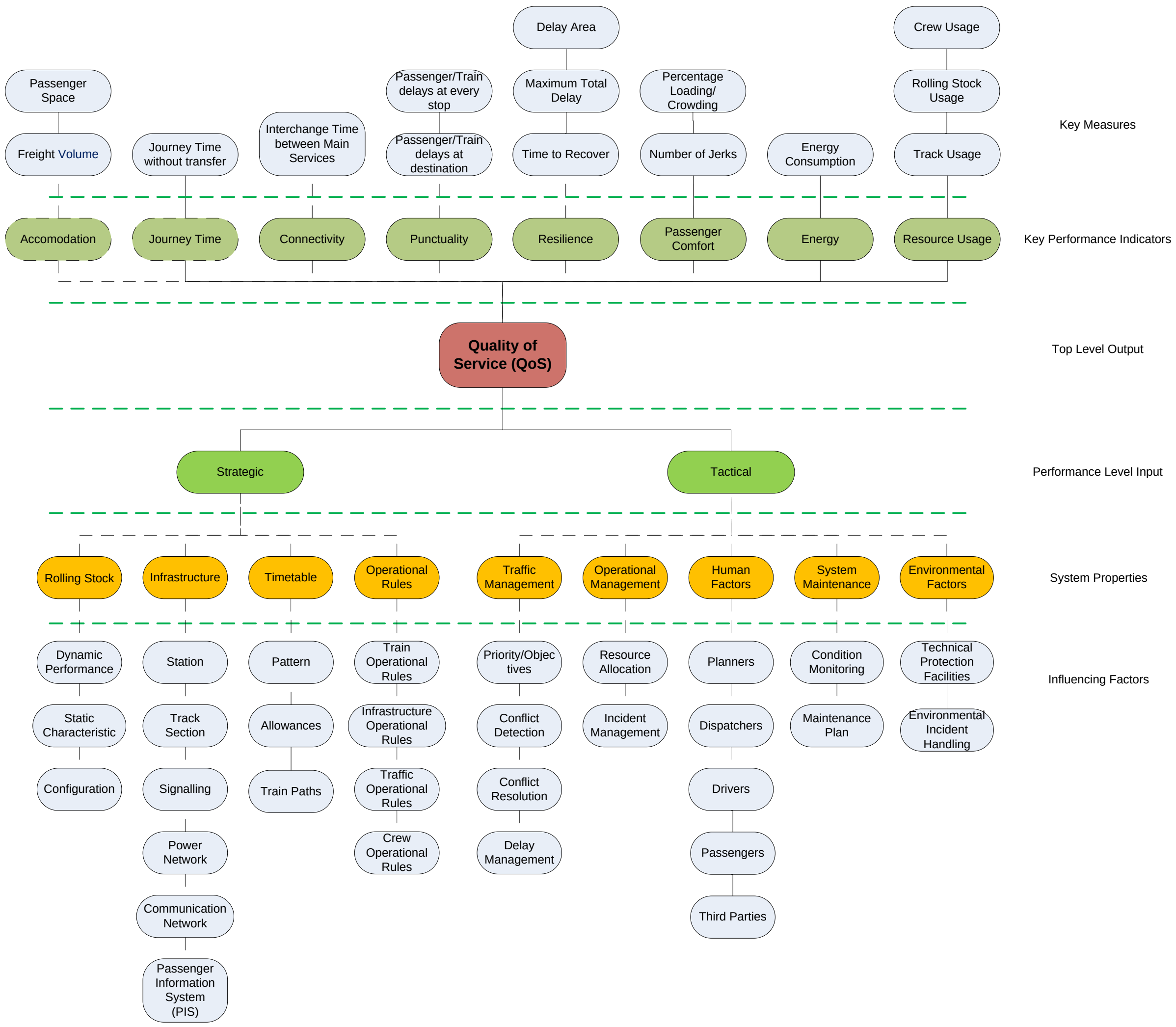


Figure 4-1: The quality of service framework

In the QoS framework, to measure the performance of a system, a number of measures called key performance indicators (KPIs) are considered. The KPIs cover parameters as diverse as accommodation, journey time, connectivity, punctuality, resilience, passenger comfort, energy, and resource usage. These KPIs focus on the aspects that are most critical to the current and future success of the railway network. They reflect the performance on a day-to-day basis. Because the QoS framework is aimed at performance improvement as well as evaluation, many of the existing indicators used to measure railways are not adopted directly, for example, customer satisfaction, capital expenditure and profit, as discussed in Section 3.4. The existing indicators (sometimes referred to as result indicators) generally attempt to quantify the overall performance of the railway over a specific period of time, but struggle to provide support for future performance improvement (Parmenter, 2015). The direct outputs of the framework are managed through the KPIs (shown in the top boxes) that are monitored by means of the measures derived from a railway's operations. The bottom boxes in Figure 4-1 show the influencing factors (IFs) within the system, which are the inputs of this framework. At the performance level, these factors are categorised as 'strategic' factors on the left and 'tactical' factors on the right. The former are factors that are expensive and time consuming to change while the latter are relatively easily changed over a shorter timescale. The KPI values are delivered based on different combinations of IFs. Different combinations of IF values produce different KPI and QoS outputs. By modifying the IFs, the values of the KPIs of the framework will change, thus indicating that the performance of the system has changed.

The QoS framework can be applied to compare and evaluate the impact of changes to subsystem properties of a railway network, such as its rolling stock, infrastructure, timetable or operational and traffic management methods.

This research has been influenced and developed by a group of infrastructure managers, operators, suppliers and academics working within the EC FP7 ON-TIME project (ON-TIME Consortium, 2012). The members of this project have contributed in the development of a capacity analysis network, which formed the basis of the performance evaluation approach presented in this thesis. During the research in each work package, a number of KPIs, such as Accommodation, Journey Time, Connectivity, Punctuality, Resilience and Resource Usage were applied for benchmarking purposes. Responses and suggestions from other researchers in the project have helped the selection and assessment of KPIs and IFs in the development of the Quality of Service approach.

4.2 Selection of Key Performance Indicators

Railway performance in terms of Quality of Service is measured with the eight KPIs in the framework: accommodation, journey time, connectivity, punctuality, resilience, passenger comfort, energy, and resource usage. Each KPI is evaluated with one or more Key Measures derived from operations. The selection of KPIs and Key Measures are produced based on expert opinions about the structure of railway operations and stakeholder needs, as collected from the ON-TIME project (ON-TIME, 2012a). Detailed definitions and calculation of the KPIs and Key Measures are given in this section.

4.2.1 Accommodation (AC)

This is the rolling stock's maximum capacity to carry passengers or freight, in terms of available passenger seats and standing room, and available freight container volume subject to a permitted maximum tonnage. This KPI shows the capability of the services to provide train paths to customers. The Key Measures for this KPI are:

- Passenger Space Kilometre: this is the total number of seats and standing space in the defined area. For each individual service, its passenger space kilometre is calculated as the total passenger space multiplied by the distance travelled;
- Freight Volume: this is the available freight container volume that can be moved in the defined area. For each individual service, its freight volume is calculated as the available freight container volume multiplied by the distance travelled.

Accommodation is closely linked to the passenger km and tonne km which are used in the traditional effectiveness and efficiency measure. The passenger km and tonne km measures only consider the design capacity of each vehicle, whereas accommodation also considers the extra space on each carriage that can be used in practice.

4.2.2 Journey Time (JT)

Journey time is considered as the total practical consumed time for trains to complete their trips, without connections with other services. It is the actual time trains consume rather than the planned time in the timetable. With different traffic and driving conditions, the journey time of individual trains may vary a lot from the timetable.

When assessing journey time, the total journey time of all services is not always helpful. It is worthwhile counting the sectional journey time (i.e. the total journey time of services when passing a defined section) or the journey time of specific services (e.g. fast trains). For example, when assessing the journey time of trains running on a large network, some busy sections can be picked out to be examined. As in Figure 4-2, an origin-destination (OD) pair has been chosen, with trains running from Welwyn Garden City to King's Cross. The differences in journey times reflect the mixture of fast and slow train services.

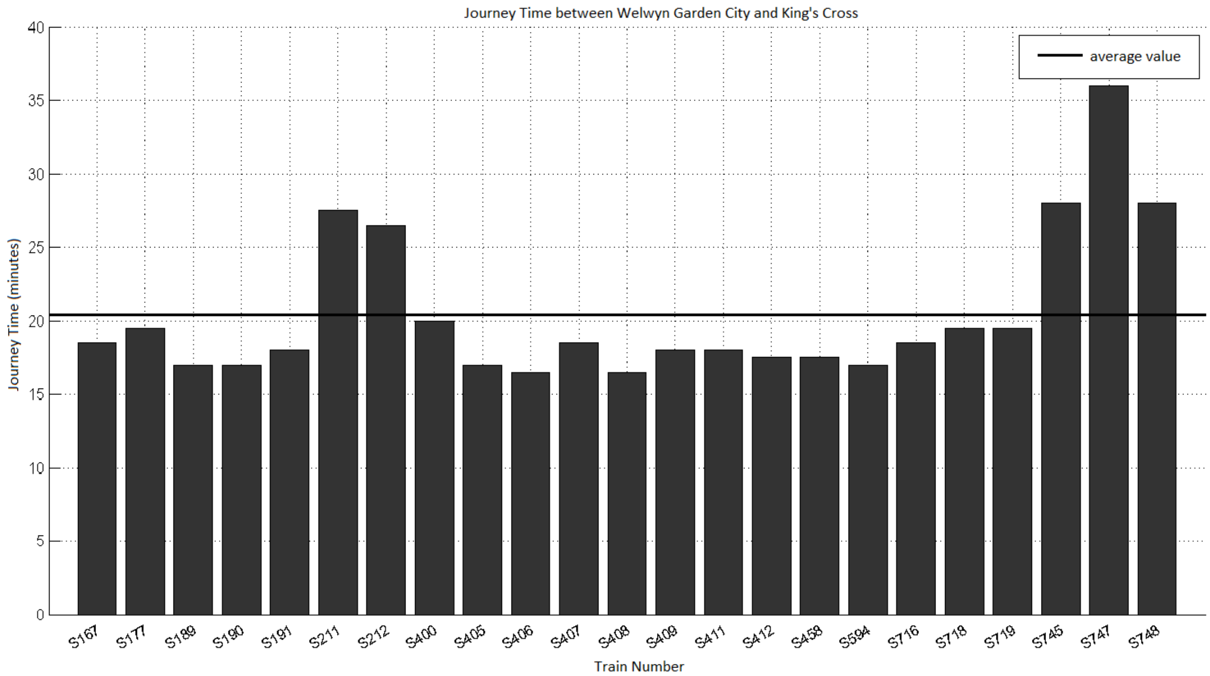
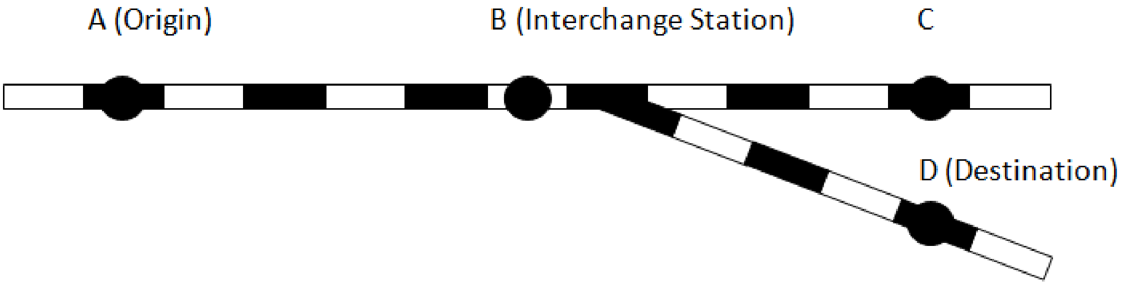


Figure 4-2: Assessment of Journey Time

4.2.3 Connectivity (CN)

Connectivity shows the passenger or cargo interchange time between any two services at a given interchange (Nuzzolo et al., 2001). It is related to both the number of possible connections and connection times at the interchange station. Connection time is the timetabled time in seconds between a passenger/cargo arriving on the first service and departing on the second service. In practice, a minimum interchange time is defined to make the connection feasible for passenger/ cargo movement. An example of connectivity calculation is given in Figure 4-3, with the minimum interchange time of 5 minutes. In the timetable, two different types of train service operate: Services S1 and S2 from A-B-C and Services S3, S4 and S5 from B-D. Service S2 cannot be connected with S4, because the interchange time is only 1 minute. Thus in total there are two possible connections, and the average connection time is 6 minutes.



Timetable

	S1	S2
A	7:00	7:10
B	7:15	7:25
C	7:20	7:30

	S3	S4	S5
B	7:21	7:26	7:31
D	7:31	7:36	7:41

Connections

No	Departure Time at A	Arrival Time at D	Service from Origin	Connecting Service	Transfer Time (minutes)
1	7:00	7:31	S1	S3	6
2	7:10	7:41	S2	S5	6

Figure 4-3: Example of connectivity assessment

Figure 4-4 shows the results of a connectivity assessment using MATLAB. A high level of connectivity would mean a larger number of available connections, and a low average connection time. These two measures are used to evaluate the connectivity KPI.

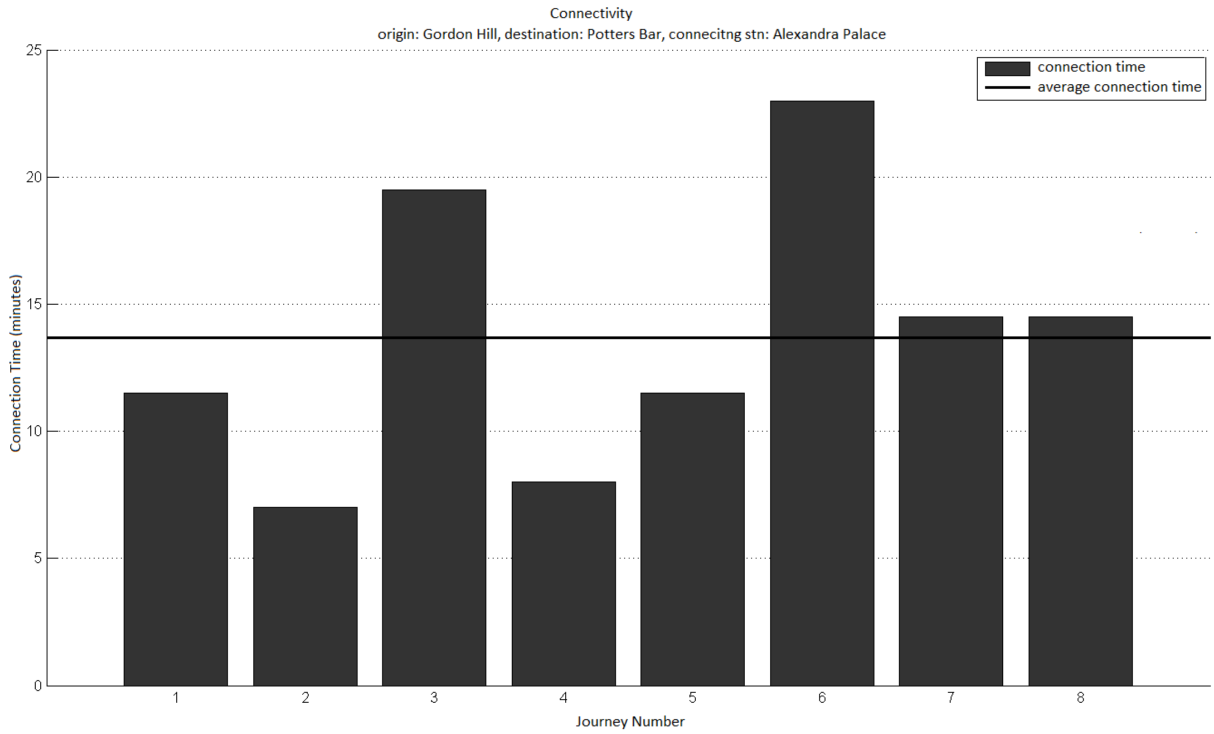


Figure 4-4: Connectivity assessment result from MATLAB

4.2.4 Punctuality (PT)

Punctuality is "the characteristic of being able to complete a required task or fulfil an obligation before or at a previously designated time" (Roberts et al., 2012b). A train is defined as punctual if it arrives within a time range before or after the timetabled arrival time. Railway traffic is usually managed based on an off-line timetable, which is designed before daily operations begin; generally there is some consideration of possible conflicts. However, in real-time operations, train runs can get perturbed, causing delays and congestions. The punctuality KPI monitors the total delays of trains running in the area, and at designated stations. The selection of stations should include those with high passenger/ cargo flow, and places where delays/ perturbations happen frequently. In certain research problems, some other stations may also be included for the specific scenario (e.g. stations with old facilities, stations of a certain size). Figure 4-5 shows an example of punctuality assessment using MATLAB, focussing on stations where delays are most likely to occur. The punctuality data is collected for selected stations on ECML on a weekday, from 7am to 10am. In this case, as the driving style is set to be running to line speed, there is in fact a lot of earliness in the network.

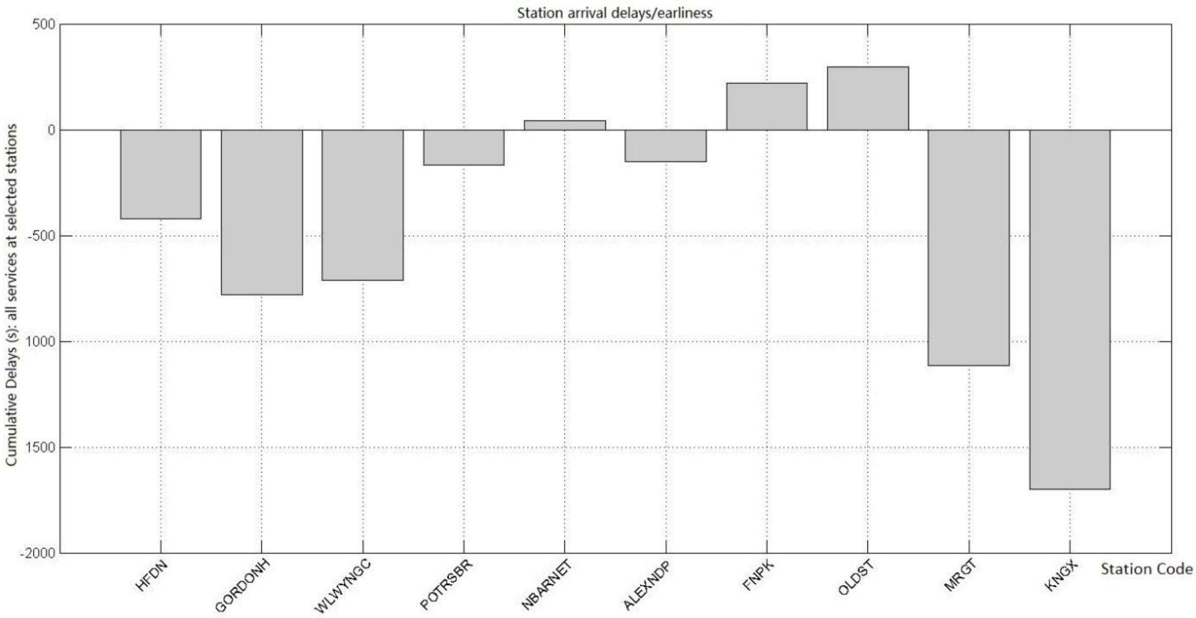


Figure 4-5: Punctuality assessment result from MATLAB

4.2.5 Resilience (RS)

Resilience is defined as the ability of a system to withstand stresses, pressures, perturbations, unpredictable changes or variations in its operating environment without loss of functionality. The concept is based on the traditional timetable stability and robustness definitions (Goverde, 2005). Three levels of resilience are defined:

- Stability: the ability to recover without active train rescheduling;
- Robustness: the ability to recover with active train rescheduling/ordering;
- Recoverability: the ability to recover with operational management measures such as train cancellation, rolling stock re-allocation etc.

In most scenarios, the resilience KPI is evaluated by considering the delays to all trains that travel within a given area, during a given time period. The delay of each train is calculated with reference to scheduled timings at stations (both for arrivals and departures) and at selected signals or similar timing points. These delays to each train are then plotted against time. Additionally, the system delay is calculated as the sum of the delay values of each train at its most recent timing point (see Figure 4-6). This is a continuous, event-driven function. The key measures of resilience are based on this system delay curve; they are: i) the system delay maximum value, (the maximum total delay [seconds] during the time period under consideration); ii) the time to recover [seconds], the time taken for the system delay curve to return below a threshold value after a delay was introduced and iii) the delay area [$seconds^2$] calculated as the area under the system delay versus time curve.

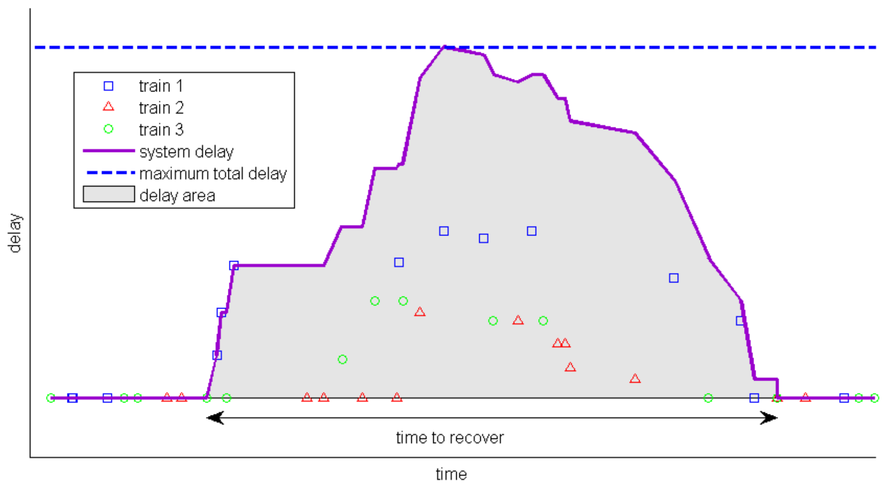


Figure 4-6: Example key measures used to evaluate the resilience KPI using an example with three delayed trains (Lu et al., 2013)

There exists some scenarios where trains run faster than planned, and arrive at stations long before the timetabled times. The trains have to wait at the stations until the timetabled arrival times, blocking platforms so that other trains cannot arrive. In this case, it is worthy to monitor the early arrivals as well as delays. The key measures of resilience would be based on the deviations rather than delays, measuring: i) the maximum total deviations [seconds] during the time period under consideration; ii) the time to recover [seconds], the time taken for the system deviation curve to return below a threshold value after a deviation was introduced and iii) the deviation area [seconds^2] calculated as the area under the system deviation versus time curve.

In Figure 4-7, delays and deviations are both evaluated, and the trains with the most significant delay/ earliness are listed.

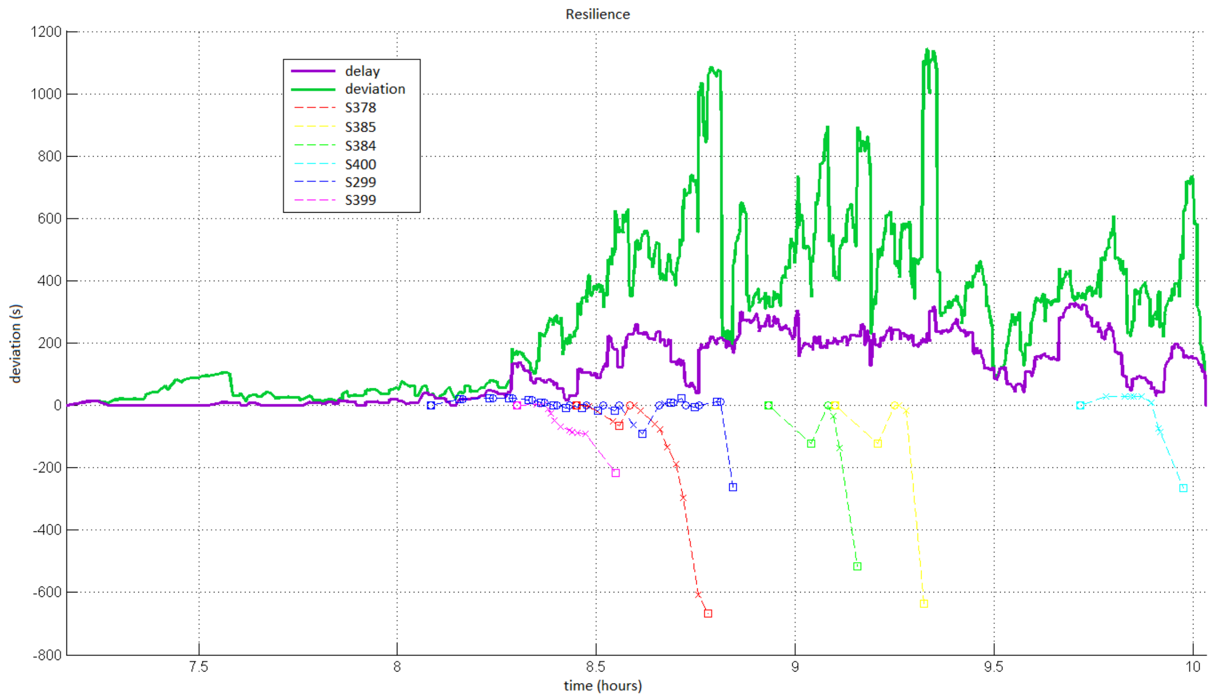


Figure 4-7: Resilience assessment result from MATLAB

4.2.6 Passenger Comfort (PC)

In railway transport, to guarantee a good level of comfort for passengers on their journeys, many factors need to be considered. Among these, the smoothness of the train driving performance has a significant impact. To improve passenger comfort, the number of jerks over a particularly limit (typically 0.75 m/s^3) needs to be reduced (Pearson et al., 1998). Jerk, J(t), is the rate of change of acceleration with respect to time,

$$J(t) = \frac{d\boldsymbol{a}}{dt} = \frac{d^2\boldsymbol{v}}{dt^2} = \frac{d^3\boldsymbol{x}}{dt^3}, \quad (4\text{-}1)$$

where a is acceleration, v is velocity, x is position and t is time. This key measure is often determined by the driving style of trains.

Another key measure for passenger comfort is the percentage loading/ crowding of trains. Passengers expect to have enough room on trains, while train operators want to fill the trains to improve their operational efficiency. To measure the percentage loading, the passenger flow data needs to be counted at each station and across the whole network.

To improve passenger comfort, both the number of jerks and crowding need to be reduced. However, these two key measures are independent of each other, and need not be combined. Passenger surveys show that crowding has more impact on comfort (Transport Focus, 2017). Thus when measuring passenger comfort, crowding is the first thing to be considered. It has a higher priority over jerks. If the percentage loading is high in the railway network, then the improvement of passenger comfort shall be focused on improving the crowding conditions. However, if there is moderate percentage loading, the number of jerks during the journey should be assessed.

4.2.7 Energy (EN)

Generally, energy consumption in the railway system includes energy consumed both by running rolling stock and infrastructure such as stations, signalling systems, etc (Gunselmann, 2005). The most important part of energy used is that needed to move the trains (González-Gil et al., 2014). In practice, it is not easy to quantify all the energy used. Energy consumption of trains is often used as a representation. By optimising the train speed profile, the energy consumption can be minimised (Miyatake and Ko, 2010).

To calculate the tractive effort applied at the wheels, the required vehicle and line data include:

- M mass of the train (kg);
- λ rotary allowance;
- α gradient angle (rad);
- g acceleration due to gravity (m/s2);
- F_R resistance of motion (N);
- K curvature resistance coefficient (Nm);
- r radius of curvature of the track (m).

Then the tractive effort *F* is calculated as a function of velocity *v* and time *t*:

$$M(1+\lambda)\frac{dv}{dt} = F - Mgsin(\alpha) - F\_R - K/r. \quad (4\text{-}2)$$

The energy consumed by trains is calculated as the total traction energy of all trains in the chosen area. In a time period 0 to *T*, if there are *N* trains, the energy consumption is:

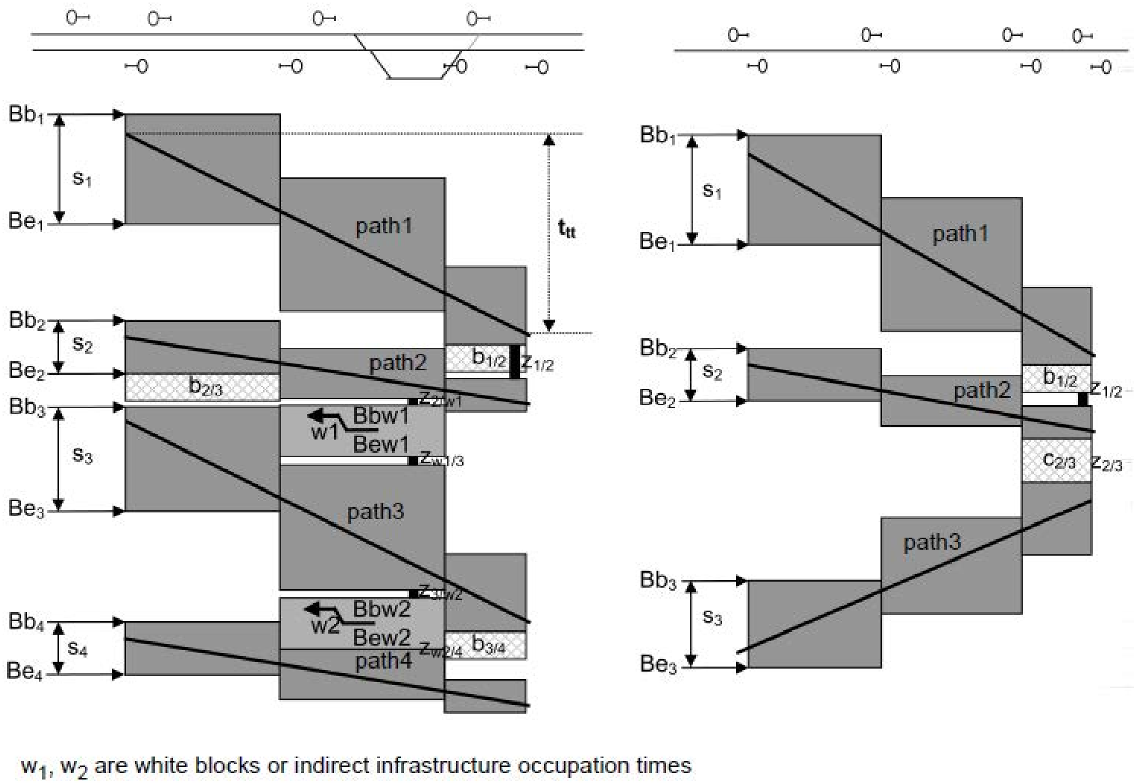
$$E = \sum_{1}^{N} \int_{0}^{T} F \times v(t)dt. \quad (4\text{-}3)$$

4.2.8 Resource Usage (RU)

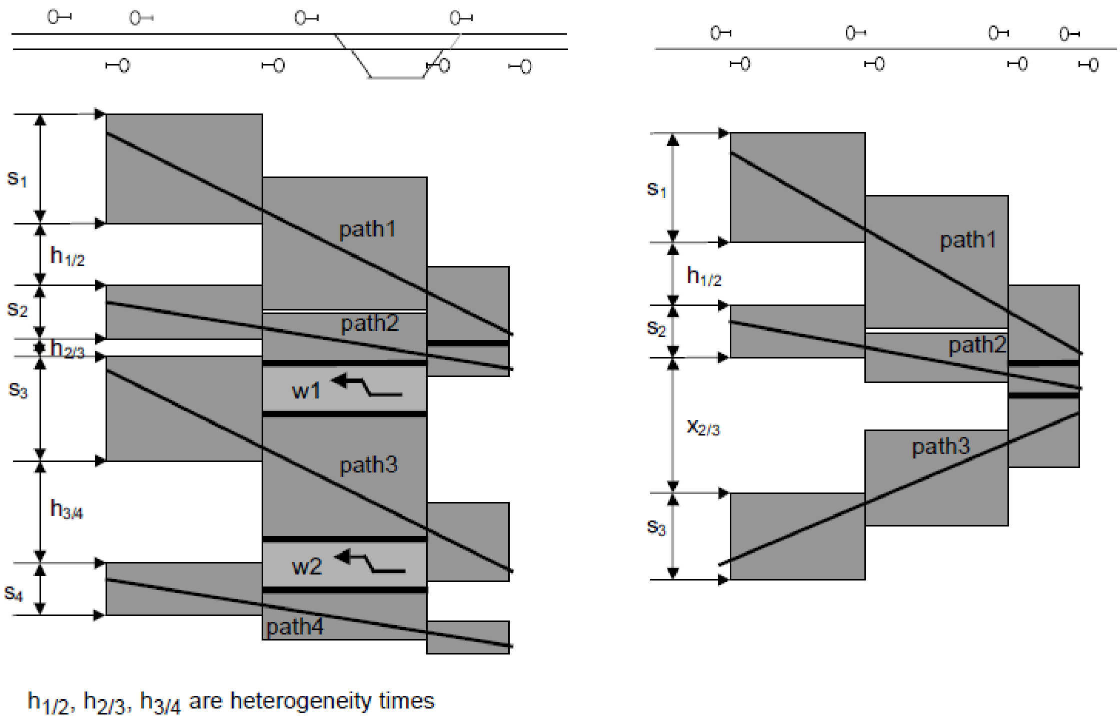
The resources used in the railway system include three main aspects: track usage, rolling stock usage and crew usage.

The track usage percentage is a reflection of the infrastructure occupation over a given time window (in minutes). The simplest way to measure track usage is by counting the number of trains passing certain track sections in the time period. A more precise measurement can be based upon the measurement method of capacity consumption outlined in UIC406, if buffer time and headway data are available.

The UIC 406 (UIC, 2004) method is widely used in the European Union to evaluate railway infrastructure capacity consumption, using a timetable compression method as shown in Figure 4-8. For compression purposes, all single train paths are pushed together up to the minimum theoretical headway, without any buffer time. This compression can be done by undertaking a graphical analysis, using suitable tools, or through analytical calculations. The total infrastructure occupation time is the total time required by the timetable after compression.



Timetable shares within a timetable



Timetable shares after compression

Figure 4-8: UIC 406 timetable compression (UIC, 2004)

The rolling stock usage percentage is defined as the average percentage of the rolling stock in use over a given time window, relative to the maximum amount of rolling stock available.

The crew utilisation is a measure of the number of paid man-hours worked by the crew over a given time period.

4.3 Selection of Influencing Factors

The inputs of the QoS framework are categorised as two groups of IFs: strategic and tactical. The strategic factors are defined as the ones that take a lot of time and resources to improve, and are often modified in the strategic planning phase. The tactical factors, on the other hand, are much easier to be changed in tactical planning and short term operations. A detailed description of all IFs is given in this section.

4.3.1 Strategic Influencing Factors

4.3.1.1 Rolling stock

Rolling stock comprises all the vehicles running on the railway network. They are the main dynamic element in the railway system. The performance of rolling stock is one of the major concerns of all stakeholders, and attracts a lot of attention from passengers and freight customers (Huisman et al., 2005). The capability of rolling stock is usually shown with the following characteristics:

- Dynamic Performance: this is closely linked with the vehicle’s running dynamics. When trains run on railway lines, a number of dynamic characteristics would affect their running curve, such as braking, acceleration, resistance, and traction force, etc;
- Static Performance: some of the vehicle’s technical details may not affect train runs directly, yet would influence the behaviour of trains at certain positions such as

locations with curvature or gradients. These elements include vehicle length, mass, maximum design speed, vehicle capacity, etc. Here both dynamic and static performances apply to individual vehicles;

- Configuration: this is the way rolling stock is formed. Trains in service are usually composed of one or more locomotives and several attached coaches, or they may be made up of multiple units. Different combinations of locomotives and carriages would result in the different total length, mass, seats/ freight tonnage, and running dynamics. These characteristics would not be a simple addition of the behaviour of all the individual vehicles.

Conventionally, a vehicle's dynamic and static performances are fixed since it is manufactured, with only vehicle capacity adjustable to TOCs. The configuration, however, is the feature that TOCs arrange and change a lot in practice.

4.3.1.2 Infrastructure

Infrastructure is a vital component of the railway system. It has great influence on the train run, energy utilisation and potentially passenger comfort.

The elements of railway infrastructure that contribute to performance changes include:

- Stations: the location, size and layout of stations and platforms have a great effect on passenger flows and vehicle movements, especially under perturbations (Edwards, 2013). Passenger and vehicle waiting locations need to be provided;

- Track sections: the track provides a dependable surface for vehicles to move through the railway network. The gradient, curvature, and radius of the track affect train running greatly. The location and structure of switches and crossings also have a great impact on train operations;

- Signalling: signalling is set on all railway lines to ensure traffic safety by keeping an adequate separation of trains (Pachl, 2002). The drivers need to follow the signals so that the timetabled operation and regularity of services is kept. The type of signalling (block signalling, cab signalling, etc) and the structure of interlocking equipped in the network affects train runs, operations and capacity greatly;

- Power network: the power network supports the full spectrum of electrified railway infrastructure, from traction supplies, track equipments, station and platform services to maintenance (UK Power Networks Services, 2015). The traction power, signalling, communication, lighting and passenger handling and safety equipment are all dependant on the power network;

- Communication network: the railway communication system allows the secure and speedy contact between drivers and signallers, and improves punctuality, reliability and safety of operations. Great Britain railway communication network uses the Global System for Mobile Communications - Railway (GSM-R) technology (Network Rail, 2015c), which involves trackside radio stations, in-cab and fixed equipment and central switching equipment;

- Passenger Information System (PIS): real-time travel information is provided to the passengers in a variety of ways. Passengers can plan journeys, book tickets, access timetables and get service updates from online, mobile services as well as the ticket office. Live departure board displays and audio announcements are presented in station halls and at platforms. In-car displays and audio announcements are also installed to keep passengers updated of the service status and to give safety instructions while on the train.

4.3.1.3 Timetable

A railway timetable is a detailed plan of trains departing and arriving at stations. In European railways, working timetables are planned in the long-term and mid-term planning phase at least 6 months in advance. Minor adjustments might be applied up until one week before operation. In daily operations, temporary path requests are addressed by dispatchers with traffic and operational management. Timetable parameters include:

- Pattern: this is the time slot arrangement pattern for all the trains in the nominal timetable, e.g. the mixture of patterns for fast trains and slow trains. It affects the resilience of the nominal timetables;

- Allowances: the allowance time, buffer time and recovery time are time slots inserted into the timetable, in addition to the minimum travel time, to help improve punctuality. It is not usually possible to achieve the minimum travel time due to "the unavoidable variability of physical characteristics, driver behaviours, passengers boarding and alighting variations and other potential influencing factors to train operations in real life conditions" (Roberts et al., 2012a). Thus allowance times are added in the nominal timetable to compensate for the additional travel times. Recovery times are reserved for trains to be recovered from initial/ primary delays, and buffer times are added to reduce or avoid knock-on delays;

- Train paths: a train path shows the details of the whole journey of a service running from its origin station to the destination. The route of the train with all the stopping stations and passing stations is defined, as well as the arrival, departure and passing times at these stations.

4.3.1.4 Operational rules

Operational Rules provides direct instructions to the railway staff. The main aspects are:

- Train operational rules: long term regulations for train operations;
- Infrastructure operational rules: long term regulations for infrastructure operations;
- Traffic operational rules: conflict/ delay management plan, priority, train mix etc;
- Crew operational rules: the allocation plan of crew.

These rules regulate the actions and procedures to be followed in daily operations and during incidents.

4.3.2 Tactical Influencing Factors

4.3.2.1 Traffic management

Traffic Management controls the movement of rolling stock during its operation. It deals with the following aspects:

- Priority: different classes of train services are usually given a different priority. Passenger and intercity services are often considered of greater importance, and when approaching stations and passing junctions these trains are often allowed to emerge first. In real-time operations, the priority of services can be changed to reduce the effect of delays and perturbations;

- Conflict detection and resolution: the dispatchers face the daily task of recovering train services when timetables are disrupted by disturbances. The dispatchers use the process of conflict detection and resolution to reduce delays by making short-term adjustments to train schedules and cancelling trains of severe delays (D'Ariano et al., 2007);

- Delay management: various dispatching strategies can be applied to reduce delays, such as rule-based strategies and optimisation-based strategies (Kliewer and Suhl, 2011).

4.3.2.2 Operational management

The operational management deals with short-term operations plans to support traffic management. The main contents are:

- Resource allocation: making rolling stock, crew and other resource plans to support new train schedules;

- Incident management: the movement of resources when faults and break-downs occur in the operation.

4.3.2.3 Human factors

The design and operations of a railway system is supported by people. It is essential to consider the capability and the demands of the individuals involved in the system. There are a number of human factors in the system, such as planners, dispatchers, drivers, passengers and third parties. Their behaviour introduces uncertainties into the operations.

4.3.2.4 System maintenance

The equipment and facilities are essential foundations of the railway system. Their condition should be observed and maintained regularly. The practices are:

- Condition Monitoring: monitors and reports on the state and quality of the railway hardware;
- Maintenance Plan: a regular maintenance plan is also vital in keeping the "health" of the system.

4.3.2.5 Environmental factors

Environmental factors such as wind, rain, snow and lightning are the source of many railway accidents and perturbations (Thornes and Davis, 2002). Technical protection facilities and emergency plans are applied to protect against the influences from these environmental factors:

- Technical Protection Facilities: these are the facilities equipped to the vulnerable parts, including wind shields, rain shields, lighting conductors, etc;
- Environmental Incident Handling: plans are made to deal with emergencies caused by environmental factors.

4.4 Summary

In this chapter a novel framework has been proposed for the evaluation and improvement of railway performance. This Quality of Service framework assesses performance with eight Key Performance Indicators: accommodation, journey time, connectivity, punctuality, resilience, passenger comfort, energy, and resource usage. The concepts and calculations of these KPIs are discussed. The influencing factors within the railway system are also defined, which would affect the KPI values and the overall performance. Based on the framework, the performance of the network can be evaluated, and the impact of each influencing factor can be assessed. An illustration of the approaches would be given in Chapter 5 and Chapter 6.

Chapter 5: The quantification of Quality of Service

In this chapter, the steps for quantifying QoS are discussed. The quantification of QoS is a complex problem involving a large number of elements. The concept of Multi Criteria Decision Making (MCDM) has been applied in this framework, which requires the determination of criteria weights for quantitative measurement. Methods and approaches for the calculation of criteria weights are also discussed. A focus group approach has been applied to derive the weights. A case study is carried out to explain the focus group methodology and process, based on a real world railway operational scenario. Analysis and discussion of the results is also included in this chapter.

5.1 Quality of Service as a Multi Criteria Decision Making framework

Railway performance in terms of QoS is evaluated with eight KPIs: AC, JT, CN, PT, RS, PT, EN, and RU. With the large number of KPIs to be considered, it is hard to tell the importance of each KPI as to the network. In this case, Multi Criteria Decision Making (MCDM) can be applied to combine the KPIs together and analyse the overall performance.

MCDM is a class of decision making methods which has been widely used in the railway industry. In operational decision making situations, when high level planning and management is needed, quantitative analysis becomes complicated as defining a single decision making goal becomes very difficult (Stewart, 1992). This problem exists in many railway decision making procedures. The railway system contains complex interactions between engineering, social and economic elements. When making decisions at a system level, a number of factors will be involved, such as timetable stability, transport volume, resource usage and crew management. A single goal is usually not sufficient for an overall evaluation of a wide range of performances. For complex decision making problems that involve a number of factors, one of the most recognised solutions is to use MCDM, which attempts to elaborate the qualitative goal with several individual, relatively precise criteria (Stewart, 1992, Figueira et al., 2005).

The most popular MCDM methods include weighted sum, goal programming, AHP, MAUT and ELECTRE (Wallenius et al., 2008, Triantaphyllou, 2013).

Weighted sum is the most popular and simplest MCDM method. If the decision making problem is made up of n criteria and each criterion follows the rule "the higher value the better", then for each alternative, its importance is:

$$A = \sum_{i=1}^{n} w_i a_i,\ i = 1,2,3 \ldots n. \quad (5\text{-}1)$$

Here, w_i is the weight of the i_{th} criteria, and a_i is the performance value of the alternative when tested with the i_{th} criteria.

In the weighted sum method, all the criteria data need to be presented with the same metric to allow them to be added together.

Goal programming aims to produce one or more solutions that meet or approach the defined goals (targets), subject to the resource constraints. When dealing with a problem using goal programming, a target is pre-defined for the objective function, and the solutions are found by minimising the deviations from the target. For a target *t*, an objective function *f(x)* containing all the criteria information is defined. Some constraints are applied to one or more criteria, and two positive deviations, *n* and *p* are defined. The relationship between *f(x)* and *t* is:

$$\mathrm{f(x)} - p + n = t, \quad (5\text{-}2)$$

where p represents the overachievement of the goal, and n is the underachievement. An adequate solution would provide a combination of decision vectors which minimise n and p.

Goal programming is widely applied in computer science and economics. The limitation of this method is that when the number of decision criteria gets larger, the complexity of resolving the problem is increased enormously.

The Analytical Hierarchy Process (AHP) is one of the most popular MCDM methods for complicated problems that contain a lot of criteria. It is also applicable when both qualitative and quantitative data are involved. AHP was introduced by Saaty in the 1970s (Saaty and Bennett, 1977), which arranges factors in a hierarchic structure, and compares criteria and alternatives with an approach named pairwise comparison to generate a ranking of the alternatives. Suppose a problem can be divided into several hierarchies, then in each hierarchy the criteria are given weights by comparing them in pairs. If n alternatives are provided, then under each individual criterion, the alternatives are also compared in pairs to assess their priorities. These priorities are then multiplied with the criteria weights to get the composite impact of each alternative. An example of the AHP method is provided in Section 5.3 to explain the process in detail.

Another two conventional MCDM methods are MAUT (Multi-Attribute Utility Theory) and ELECTRE (ELimination and Choice Expressing REality). MAUT is a class of MCDM methods that are suitable for solving problems involving trade-offs and risky choices (Figueira et al., 2005). Since its first introduction in the 1970s, it has been popular in decision making for public policy (Ananda and Herath, 2005), health care (Torrance et al., 1982) and power plant planning (Voropai and Ivanova, 2002). This method uses a systematic approach to quantify preferences. A scale of 0 – 1 is applied to turn preferences into numerical values for comparison. The result is a rank of the alternatives reflecting the decision maker's preferences. ELECTRE is a class of MCDM methods that can deal with a large number of criteria (more than 3 criteria can be treated, and the number can go up to 12 or 13). It is also suitable if there is heterogeneity among criteria. Unlike the methods mentioned before, ELECTRE is an outranking approach to model preference, which helps to discard some alternatives. Thus ELECTRE is often used to limit the number of alternatives before applying other MCDM methods for further analysis. Both the ELECTRE and MAUT methods try to take uncertainty and risks into consideration, which are not discussed in most MCDM methods.

Many applications of MCDM have been made in various areas of railway research, such as site selection (Mohajeri and Amin, 2010, Mateus et al., 2008), system management (Nystrom and Soderholm, 2010, Azadeh et al., 2008, An et al., 2011) and evaluation (Sivilevicius and Maskeliunaite, 2010, James, 1991, Maskeliunaite and Sivilevicius, 2012, Bureika, 2011).

Figure 5-1 shows a common decision making process used by MCDM methods.

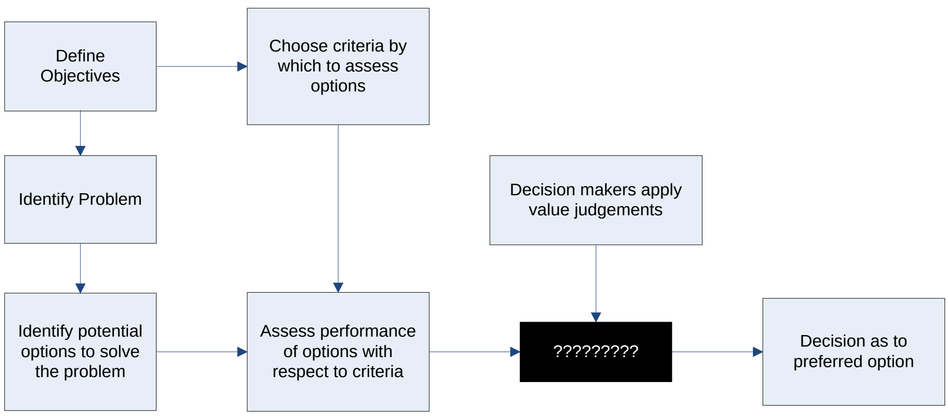


Figure 5-1: A simplified representation of the decision making process (Sayers et al., 2003)

The first step in solving any MCDM problem is problem definition and data collection (Rosanty et al., 2012). The problem domain, the possible alternatives and required criteria all need to be precisely defined. The alternatives are then assessed with respect to each of the criteria to obtain criterion specific scores.

The black box (Figure 5-1) represents one of the most important problems in MCDM: how to combine criteria specific information together to precisely evaluate the importance of alternatives? In this procedure, criteria weights are applied for the aggregation of scores into overall preference values (Triantaphyllou, 2013). To get an accurate quantitative analysis, precise weights are needed.

When treating the QoS framework as a multi-criteria problem, eight decision criteria (KPIs) are involved. The objective function for the QoS calculation is the “black box” in this problem. To get a quantitative description of the QoS, the first step is to address weights for the KPIs. For different networks, the relationship between KPIs may vary a lot. Thus the criteria weights need to be determined subject to each specific railway network.

5.2 Weighting methodology

MCDM is widely used when selecting alternatives with regard to a number of well-defined, usually independent criteria (Roy, 2005). MCDM methods that generate preference scores for each alternative require decision makers to provide information on the relative importance (weightings) of criteria.

Many different methods have been proposed for generating criteria weights (Choo et al., 1999). A good comparison of some weighting methods is given by Hobbs (1980), Barron and Barrett (1996), and Bell et al. (2001). Methods to derive cardinal weights include point allocation, hierarchical weighting (Guinto, 2008), trade off and pricing-out methods (Keeney and Raiffa, 1994), ratio method and swing weight (Winterfeldt and Edwards, 1986), conjoint procedures (Green and Srinivasan, 1990), Analytical Hierarchy Process (Saaty, 1990) and multi objective linear programming (Costa and Clímaco, 1999). In this section, some of these weight selection techniques are discussed. Weighting methods based on outranking have been discussed in methods such as ELECTRE and PROMETHEE (Figueira et al., 2005).

Point allocation and hierarchical weighting are methods that use a direct rating approach. Decision criteria are categorised, and the decision maker rates each criteria according to their importance. These ratings are normalised to get the final weights. These approaches are very straightforward and easy to carry out; however, the ratings are based strictly upon the decision maker's subjective judgements, thus theoretically valid results cannot be guaranteed (Guinto, 2008).

The ratio method and the swing method both yield weightings through a ranking and rating process. In ratio weighting, the decision maker would first directly rank the criteria, and then give each criterion a numerical score to indicate its level of importance. The swing method forces the decision maker to consider the range of values in each criterion. The decision maker then considers the possibility to "swing up" these values from the worst to the best, resulting in a rank of criteria. The most preferred criterion is given a rating of 100, others rated with scores proportional to it. The final weightings are obtained by normalising these ratings.

In trade off weighting, pairs of alternatives are compared by each pair of criteria, assuming that both alternatives have the same performance in all the other criteria. By choosing the preferred alternative (solution), the decision maker decides on the relative importance of the two criteria. Then the decision maker needs to yield the indifference values by trading off one criterion for another. With numerous trade-off processes the final weights can be derived.

Another frequently used method is the analytical hierarchy process (AHP). Decision criteria are compared pairwise, based on a subjective scale of 1 – 9. A reciprocal pairwise comparison matrix is constructed through this systematic process. Criteria weights are obtained with an Eigenvector approach.

Other weighting techniques cited in the literature are, e.g. approximate weighting methods such as equal weighting, rank sum and rank-order centroid (Jia et al., 1998); ratio questioning, the Churchman-Ackoff procedure, and Metfessel allocation (Hobbs, 1980). These methods can yield good quality decisions if time is limited. However, the accuracy is not as good as quantitative ratio weights, such as swing weights, trade off weights and AHP.

5.3 Deriving criteria weights in the Quality of Service framework

In the QoS framework, the KPIs are relatively independent, and there isn't a set 'target' for the QoS level. Getting a single QoS value is more meaningful for comparison purposes. Among all the methods discussed in Section 5.1, weighted sum and AHP would meet the needs for quantitative analysis of quality of service. In practice, there exist a large number of potential alternatives, and it would become very subjective to do pairwise comparison of these alternatives. Thus the AHP approach is not adopted. However, the pairwise comparison approach for determining criteria weights can be a good method for weighting decisions. In the Quality of Service approach, the weighted sum model is chosen. This method is simple and straight forward to carry out. The KPIs need to be normalised before going into the model.

To calculate the QoS value, weightings are required for each KPI. Not all KPIs have the same influence on the overall performance. As the number of decision criteria is large, the approach to derive weights needs to be carefully selected (Nijkamp and Delft, 1977).

As discussed in Section 5.2, a lot of methods have been proposed for the determination of criteria weights. In the framework to be assessed, methods related to direct ranking and rating of criteria become very imprecise. This is due to the complexity of making judgements with a relatively large number of parameters. In this case, methods that compare criteria in pairs are preferred. In this research a pairwise comparison method is adopted.

Pairwise comparison was introduced in conjunction with the AHP method by Saaty in the 1970s. The aim is to improve the accuracy of judgements by concentrating on comparing two elements on a single property each time (Saaty, 2008). The relationship between the two criteria compared is indicated with the ratio of one criterion over the other. In this way the problem of having different measurement scales is avoided. The ratios are mapped in a reciprocal matrix, and the final weightings are derived by calculating the eigenvector of the matrix.

In the quality of service framework, the KPIs are relatively independent. It is hard to determine the relative importance of the KPIs. Some KPIs are author defined, and previous research has only addressed some of the KPIs considered, such as accommodation, journey time and punctuality (Hansen et al., 2013, Schittenhelm and Landex, 2013). It is not sufficiently rigorous to determine the criteria weights based on the author's own understanding of the KPIs. These have made it essential to gather stakeholders' opinions about the KPIs and their relative importance. A Focus Group approach was thus designed to determine the weights in this particular MCDM problem.

5.3.1 The Focus Group process

The focus group technique is a qualitative investigation method to obtain data towards a specific subject. It is based on workshop-style meetings using a group of individuals specially chosen for the purpose. The target of the focus group method is to achieve a group consensus on the relative importance of the different criteria via discussion and negotiation between group members (Tsiporkova and Boeva, 2006, Herrera et al., 1996). This method succeeds in reflecting both the opinions of individuals and the "group opinion" which is formed through the interactions between group members. Successful applications of the focus group technique have been made in many areas, such as social science and health care (Hydén and Bülow, 2003). In transportation, the focus group method has been applied to gather opinions about passenger behaviour and their opinions towards service quality (Golob and Hensher, 2007, Iki et al., 2012, Simons et al., 2014, dell'Olio et al., 2011).

In a focus group meeting, a group size of 6 to 12 people is preferred (Massey, 2011), with a moderator to direct the discussion. The group members are selected to represent the people related to the problem as closely as possible. The length of the meeting is usually around two hours. During the meeting, questions and discussions are presented, and interview and observation techniques are used by the moderator to elicit the group members' attitudes and opinions. The moderator is also responsible to keep the discussions on track, and give directions to help the group members to understand the problem. The group members may modify their opinions during the discussions based on the information they acquire during the process.

For the QoS research, the focus group meeting is organised loosely in order that group members are encouraged to express their own perceptions towards the problem. The moderator tries to "keep distance" to give group members more freedom to resolve the problem together. Because four types of stakeholders (IM, TOC, government and passengers) are involved in the problem under consideration, it is necessary to include a panel of several representatives of each stakeholder group in the focus group. In each group, the background of the members shall be varied (e.g. recruit timetable planners, operators, maintenance staff in the IM group), to ensure the quality of information collected in the research. The group members, passengers in particular, would also need to have adequate knowledge of railway transportation as to understand the basic terms in the QoS framework. These people would be given information about the QoS framework, the scenario to be assessed, and the focus group process. The procedure of the focus group discussion is as follows:

1. The moderator introduces the purpose and the process of the discussion;

2. The participants introduce their background, expertise and interests in railway transportation;

3. The moderator introduces the QoS framework and the scenario;

4. The participants are divided into groups where they play the role of stakeholders to decide initial weights with pairwise comparison;

5. The stakeholder groups present their initial weights for discussion and negotiation, and derive the revised weights;

6. The participants make revision and adjustment to get the final weights of the KPIs;

7. Summary and discussion.

5.3.2 The numerical evaluation scale

In step 4 of the focus group discussion, when doing pairwise comparison, participants are asked to provide their opinions on the importance of the KPIs. Individuals are given the linguistic phrases in Table 5-1 to make judgements towards pairs of attributes. They would decide for each pair (1) which attribute is more important, and (2) how much more important it is to the other attribute. All focus group members work together to generate a table containing the relative importance of the KPIs. Their linguistic judgements are translated using the following scale:

Intensity	Definition	Explanation
1	Equal Importance	Two elements contribute equally to the objective
3	Moderate Importance	Experience and judgement slightly favours one element
5	Strong Importance	Experience and judgement strongly favours one element
7	Very Strong Importance	One element is favoured very strongly over the other
9	Extreme Importance	One element is favoured extremely over the other
2,4,6,8 can be used to express intermediate values; Reciprocals are used for inverse comparisons.		

Table 5-1: 9-point scale for attitude measurement

This 9-point scale is defined by Saaty as a standard ratio scale used in AHP (Saaty, 1990). It assumes that people can distinguish clearly the extent of difference between two attributes based on the evenly distributed numerical values. There have been a lot of debates about the accuracy of this scale to quantify people's opinions (Ji and Jiang, 2003, Finan and Hurley, 1999, Dong et al., 2008). It is argued that this scale is often used as it is simple to use rather than psychologically appropriate. Studies on people's perceptions suggest that the difference between two neighbouring levels gets larger when the levels are higher (Schoner and Wedley, 1989, Triantaphyllou et al., 1994). That is, when people feel that A gets much more important than B, the difference between these two attributes are actually higher than expected.

To improve the consistency of results, the numerical values of the intensity are transferred into power scale values in calculation:

Intensity	Numerical Value	Explanation
1	$(\sqrt[8]{9})^0 = 1$	A is equally important to B
2	$(\sqrt[8]{9})^1 \approx 1.32$	A is weakly or slightly more important than B
3	$(\sqrt[8]{9})^2 \approx 1.73$	A is moderately more important than B
4	$(\sqrt[8]{9})^3 \approx 2.28$	A is moderately plus more important than B
5	$(\sqrt[8]{9})^4 = 3$	A is strongly more important than B
6	$(\sqrt[8]{9})^5 \approx 3.95$	A is strongly plus more important than B
7	$(\sqrt[8]{9})^6 \approx 5.20$	A is very strongly more important than B
8	$(\sqrt[8]{9})^7 \approx 6.84$	A is very, very strongly more important than B
9	$(\sqrt[8]{9})^8 = 9$	A is extremely more important than B

Table 5-2: Numerical values of the power scale (Elliott, 2010)

5.3.3 Weights calculation with pairwise comparison

From the focus group meeting, a pairwise comparison table containing the relative importance of the KPIs is generated. The criteria weights are calculated with these data. To demonstrate the calculation process of criteria weights through pairwise comparison, a simple example is provided below.

Suppose a parent is picking a school for a child, and there are 3 choices: School A, School B and School C. The parent considers four criteria: learning, friends, school life and convenience, as in Figure 5-2.

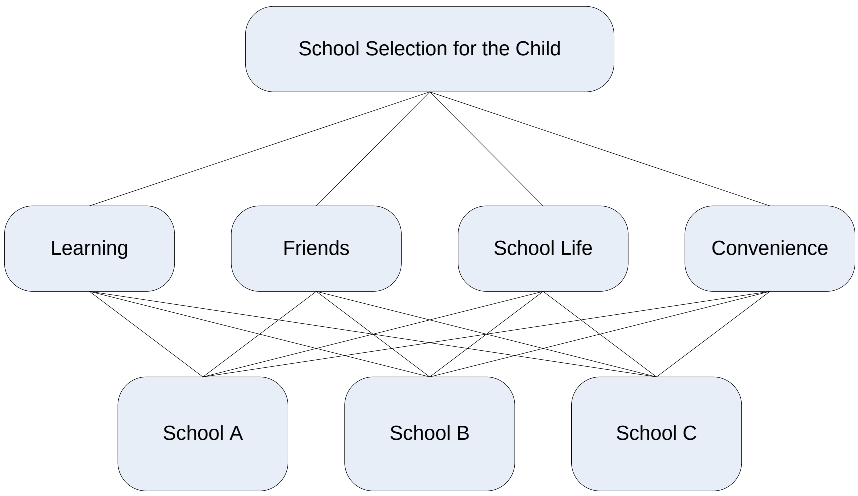


Figure 5-2: Decomposition of the problem into a hierarchy

Using Saaty's 9-point scale, the parent rates the relative importance of each criteria, and the corresponding pairwise comparison matrix is:

	Learning	Friends	School Life	Convenience
Learning	1	4	3	7
Friends	1/4	1	2	1/5
School Life	1/3	1/2	1	1/6
Convenience	1/7	5	6	1

Table 5-3: The pairwise comparison matrix

The priority of each criterion is calculated as the Eigenvector of the matrix (Table 5-4).

	Weight
Learning	0.51
Friends	0.11
School Life	0.09
Convenience	0.29

Table 5-4: The criteria weights for school selection

If a weight sum or weight product model is applied, the criteria weights can be applied right away. If the selected MCDM method is AHP, then pairwise comparison of alternatives are required. In this example, it is assumed that the parent continues to compare the schools in pairs, and gets the following results:

Learning				
	A	B	C	**Priorities**
A	1	1/3	1/2	0.16
B	3	1	3	0.59
C	2	1/3	1	0.25

Friends				
	A	B	C	**Priorities**
A	1	1/2	1	0.25
B	2	1	2	0.5
C	1	1/2	1	0.25

School Life				
	A	B	C	**Priorities**
A	1	9	7	0.75
B	1/9	1	1/5	0.06
C	1/7	5	1	0.19

Convenience				
	A	B	C	**Priorities**
A	1	6	4	0.69
B	1/6	1	1/3	0.09
C	1/4	3	1	0.22

Table 5-5: The priorities of alternatives under each criterion

For each alternative, the composite impact is the sum of weighted priorities. This produces the composite impact of schools as:

	Learning	Friends	School Life	Convenience	**Impact**
	0.51	0.11	0.09	0.29	
A	0.16	0.25	0.75	0.69	**0.38**
B	0.59	0.5	0.06	0.09	**0.39**
C	0.25	0.25	0.19	0.22	**0.24**

Table 5-6: The composite impact of schools

5.4 Case Study

5.4.1 Scenario setting

The scenario that the focus group discussion is based on is part of the East Coast Main Line (ECML) in Great Britain. This section runs between Welwyn Garden City, Hertford North and London King's Cross. A mixture of long distance and local traffic are presented in this network. A detailed description can be found in Chapter 7: Case Study.

In this discussion, participants were asked to think about the current situation of peak time (7am to 10am) passenger traffic on the ECML. The main stakeholders in this network are infrastructure manager, train operator, timetable planner, government and passengers.

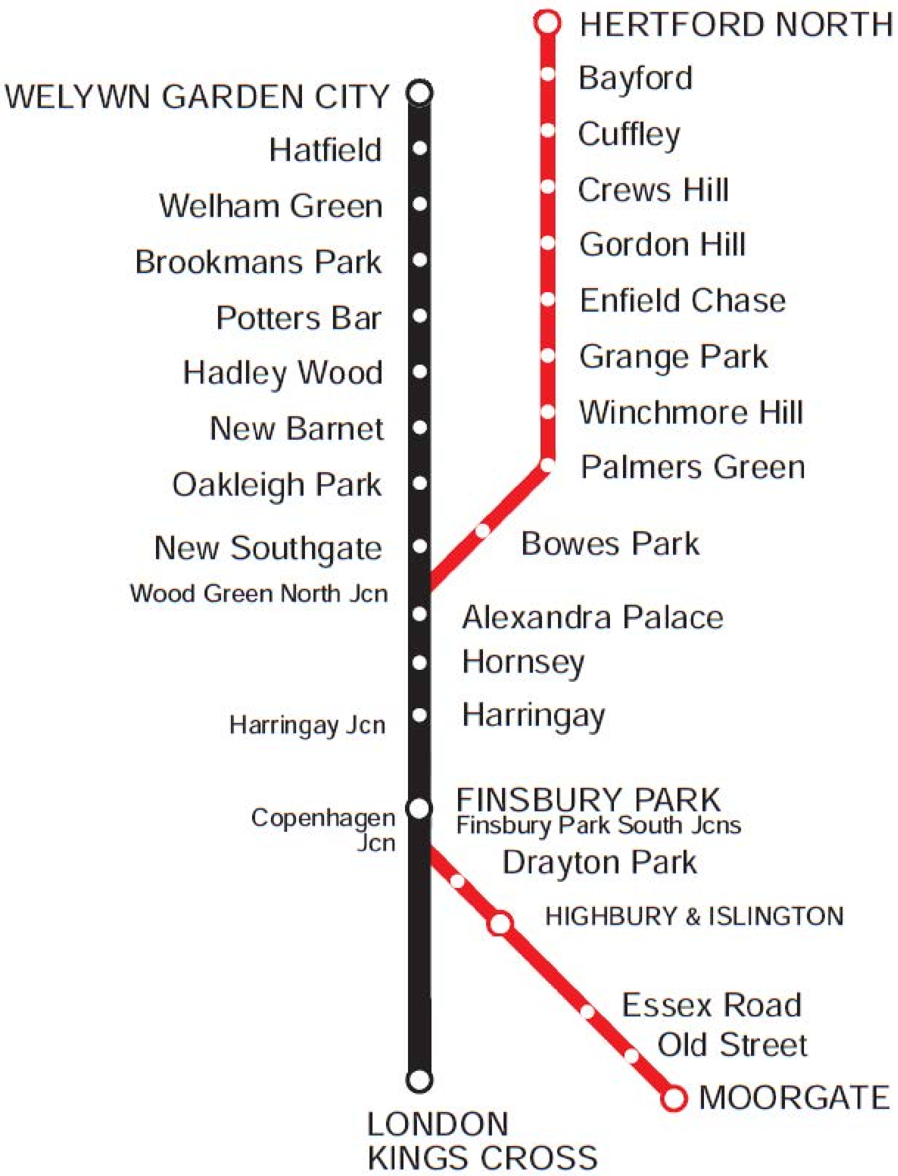


Figure 5-3: Map of the ECML

5.4.2 Selection of group members

In this case study, the focus group is set in an academic environment. To get the KPI weightings, only people who have worked in the railway industry or research institutions are selected, as some terms are very unfamiliar to ordinary rail users. The focus group consists of 1 moderator and 8 participants, all with a rail research background, and some have work experience in the railway industry. Each of the participants has exclusive knowledge about the railway industry and its operation, and is able to identify the requirements of different stakeholders. As all participants are “well informed” candidates, only one focus group meeting needed to be held. In pilot work, the participants had all been informed of the QoS research. They had thought about the rankings of KPIs, and had performed pairwise comparison of KPIs individually. It was discovered that all participants have different preferences towards the KPIs. For example, accommodation was favoured by all participants, while views on energy, resilience and connectivity varied a lot.

5.4.3 The decision making process

The members were divided into 4 groups: infrastructure manager (IM), train operating companies (TOCs), government and passengers. In the British system, the infrastructure manager (Network Rail) is also responsible for producing timetables, thus there wasn’t a timetable planner group. Each group was made up of two participants.

The four groups were first asked to each fill in a pairwise comparison form as shown in Table 5-7. Within each group the members need to reach agreement on the ratios from the viewpoints of their role. The results were four initial pairwise comparison matrices.

Then each group talked about their opinions on the relationship of KPIs. The four groups negotiated and made trade-offs based on their initial answers. Four revised matrices were derived at this stage.

The final weights were obtained with a discussion between all participants. All groups expressed their preferences and considerations. A finalised pairwise comparison matrix was produced. The participants had also discussed their concerns and some thoughts about the focus group study.

Which characteristic do you think is more important and to what extent?

A	B	More important? A or B	Scale (1-9)
Accommodation (AC)	Journey Time (JT)		
	Connectivity (CN)		
	Punctuality (PT)		
	Resilience (RS)		
	Passenger Comfort (PC)		
	Energy (EN)		
	Resource Usage (RU)		
Journey Time (JT)	Connectivity (CN)		
	Punctuality (PT)		
	Resilience (RS)		
	Passenger Comfort (PC)		
	Energy (EN)		
	Resource Usage (RU)		
Connectivity (CN)	Punctuality (PT)		
	Resilience (RS)		
	Passenger Comfort (PC)		
	Energy (EN)		
	Resource Usage (RU)		
Punctuality (PT)	Resilience (RS)		
	Passenger Comfort (PC)		
	Energy (EN)		
	Resource Usage (RU)		
Resilience (RS)	Passenger Comfort (PC)		
	Energy (EN)		
	Resource Usage (RU)		
Passenger Comfort (PC)	Energy (EN)		
	Resource Usage (RU)		
Energy (EN)	Resource Usage (RU)		

Table 5-7: The pairwise comparison table

5.5 Analysis of results

5.5.1 Calculation of weights from matrices

The weights are calculated from the pairwise comparison matrices from the discussions. Table 5-8 shows the initial answers of the IM that are used to generate the initial weights.

	AC	JT	CN	PT	RS	PC	EN	RU
AC	1	7	8	1/7	1	6	4	1/5
JT	1/7	1	1	1/5	1/7	4	1	1/6
CN	1/8	1	1	1/6	1/5	4	1/4	1/7
PT	7	5	6	1	6	8	5	6
RS	1	7	5	1/6	1	1	5	3
PC	1/6	1/4	1/4	1/8	1	1	1/3	1/6
EN	1/4	1	4	1/5	1/5	3	1	1/2
RU	5	6	7	1/6	1/3	6	2	1

Table 5-8: The pairwise comparison matrix

In this case, when comparing AC and JT, AC is considered very strongly important (intensity 7). Thus the corresponding cells are filled with 7 and 1/7.

From the pairwise comparison matrices, the criteria weights can be derived. By looking up the numerical values in Table 5-2, the result of each pairwise comparison can be translated into a score, as in Table 5-9.

	AC	JT	CN	PT	RS	PC	EN	RU
AC	1.00	5.20	6.84	0.19	1.00	3.95	2.28	0.33
JT	0.19	1.00	1.00	0.33	0.19	2.28	1.00	0.25
CN	0.15	1.00	1.00	0.25	0.33	2.28	0.44	0.19
PT	5.20	3.00	3.95	1.00	1.00	6.84	3.00	3.95
RS	1.00	5.20	3.00	1.00	1.00	1.00	3.00	1.73
PC	0.25	0.44	0.44	0.15	1.00	1.00	0.58	0.25
EN	0.44	1.00	2.28	0.33	0.33	1.73	1.00	0.76
RU	3.00	3.95	5.20	0.25	0.58	3.95	1.32	1.00

Table 5-9: The pairwise comparison matrix translated

The criteria weights are gained by calculating the eigenvector of this matrix. The result is:

AC	0.15
JT	0.06
CN	0.05
PT	0.33
RS	0.15
PC	0.04
EN	0.07
RU	0.15

Table 5-10: The eigenvector of the comparison matrix: criteria weights

All data used for the weights calculation can be found in Appendix A.

5.5.2 Initial weights

After independent group discussions, four groups of initial weights were generated (Table 5-11). All the stakeholders have favoured accommodation and punctuality, and the IM, TOC and the government have also valued punctuality and resource usage, while showing little concern about connectivity and passenger comfort. The passenger group, however, had almost completely opposite opinions on these KPIs. Opinions on other KPIs were even more varied. This phenomenon is closely linked to the interests and benefits of different stakeholders.

	IM	TOC	GOV	PASSENGER
AC	0.15	0.27	0.19	0.22
JT	0.06	0.11	0.04	0.12
CN	0.05	0.05	0.09	0.14
PT	0.33	0.11	0.12	0.23
RS	0.15	0.09	0.27	0.09
PC	0.04	0.05	0.05	0.16
EN	0.07	0.13	0.06	0.02
RU	0.15	0.19	0.18	0.02

Table 5-11: Initial weights

5.5.3 Revised weights

After negotiation across stakeholders, several changes had been made to the pairwise comparison ratios. The infrastructure manager was persuaded to change PT/RU from 6 to 1, as timetable planning and infrastructure utilisation are both important themes for the IM. The ratio was modified after the IM group had gained a better understanding of Resource Usage.

The Government group had also made some modifications. As the current connectivity and passenger comfort in ECML is good, the government group lowered the importance of these KPIs. The ratios changed were: JT/CN from 1/3 to 2, CN/PT from 1/3 to 1/5, CN/EN from 3 to 1/2, PC/EN from 3 to 1/3.

The modifications had led to a small change in the weights and rankings. The revised weights are as follows:

	IM	TOC	GOV	PASSENGER	AVERAGE
AC	0.15	0.27	0.18	0.22	0.21
JT	0.06	0.11	0.04	0.12	0.08
CN	0.05	0.05	0.08	0.14	0.08
PT	0.28	0.11	0.13	0.23	0.19
RS	0.15	0.09	0.27	0.09	0.15
PC	0.04	0.05	0.04	0.16	0.07
EN	0.07	0.13	0.08	0.02	0.08
RU	0.19	0.19	0.18	0.02	0.15

Table 5-12: Revised weights

5.5.4 Final weights

After deriving the revised weights, all the participants discussed the final weights together. Compared to taking an average across all stakeholders, deciding weights through negotiation takes a lot more time, yet it has produced more accurate results. For some KPIs, there can be a lot of variation in stakeholders' opinions about their importance. When taking an average, there might be a misjudgement on the importance of the KPIs that have been valued differently by stakeholders. Encouraging effective communication improves the understanding of the scenario and the KPIs, and it also helps to discover problems and possible improvement plans in the studied railway network.

In the discussion of final weights, some ratios had been agreed on by all stakeholders, e.g. AC/JT, JT/CN, and CN/PT. On ratios regarding Energy and Resource Usage, the IM, TOC and the government had similar views. The passengers, however, had always given these two KPIs the lowest importance. They claimed that these KPIs don't affect ticket prices a lot, and their impacts are not obvious in everyday journeys. The final decision between all participants was to combine the opinions of the IM, TOC and government on these ratios, as passengers' opinions on the two KPIs were very vague. Some significantly different viewpoints had been placed on other ratios, mainly because of the stakeholders' different expectations and concerns (Table 5-13).

	IM	TOC	GOVERNMENT	PASSENGER
Main concerns	• Maximise *Resource Usage*; • Produce a resilient timetable (*Resilience*) and reduce delays (*Punctuality*); • Provide as many train paths (*Accommodation*), while insuring a certain level of *Resilience*.	• Accommodate more passengers and freight (*Accommodation*); • Lower *Resource Usage* and *Energy* costs; • Improve *Journey Time* and *Punctuality*.	• Ensure smooth running and disruption free operations (*Resource Usage* and *Resilience*); • Provide sufficient *Accommodation*; • Reduce delays (*Punctuality*).	• Have good *Punctuality* and less delays; • Long distance journeys requires good *Passenger Comfort*; • Get connections (*Connectivity*) and tickets (*Accommodation*).
KPIs that are given lower priorities	• *Journey time* and *Connectivity* can't be changed much on the ECML; • As long as *Resource Usage* is good, *Journey Time* and *Connect*ivity are not valued much; • *Energy* costs are paid by TOCs; • *Passenger Comfort* is not bad in ECML.	• *Resilience* is decided by timetable, which is planned by IM; • *Connectivity* and *Passenger Comfort* are not bad in ECML; • Good *Passenger Comfort* attracts people to use rail services, yet it affects only part of passengers' perception.	• Reducing *Energy* costs may help with government debts; • *Connectivity* needs to be improved within a RU limit; • *Journey Time* and *Passenger Comfort* are less important if people can get seats.	• *Journey Time* is less important if people can get on trains; • Passengers care about the train they're on more than other services in the network; they experience delays rather than *Resilience*; • *Energy* doesn't affect ticket prices a lot; • Bad *Resource Usage* can hardly be experienced by passengers.

Table 5-13: Stakeholders' concerns and preferences

When the stakeholders cannot persuade each other to make trade-offs, it was agreed that an average of ratios was to be taken as the final answer. With the Eigenvector approach, the final weights and rankings were gained:

	Weights	Ranking
AC	0.20	1
JT	0.08	6
CN	0.07	7
PT	0.17	3
RS	0.14	4
PC	0.06	8
EN	0.09	5
RU	0.20	2

Table 5-14: Final weights and ranking

Compared to the average weights in Table 5-12, there are a few variations, and a significant difference in the weight of Resource Usage. It is a reflection that with the negotiation method, the accuracy of the final weights has been improved.

To testify that all participants have given reasonable responses, the consistency ratios (CRs) of the pairwise comparison matrices are calculated (Triantaphyllou, 2000).

For the matrices used in the calculation of initial and revised weights, the CR values are varied among stakeholders. TOC and the government have delivered matrices of significantly small CR values (about 10% or less), while the outputs from the IM and passenger shows high CRs. The inconsistency of IM and passenger's views may come from the fact that they have very strong preferences towards some of the KPIs, thus affecting their judgement towards weightings. The large number of decision criteria also contributed to the difficulty of determining the relative importance of factors (Saaty, 1990). The CR value of the final pairwise comparison matrix is less than 10% (5.49%). This suggests that the final criteria weights are theoretically valid, and the negotiations between stakeholders have improved the consistency of the result.

The results reflected some of the common themes across all stakeholders: better punctuality, and improved accommodation (more seats). Resilience and Resource Usage are also drawing attention as they are the foundation of a smooth running and profitable railway network. Energy is also becoming a concern. Journey Time is valued a lot with the participants, yet it has got a low score as the current journey times on the ECML are satisfactory. The Connectivity and Passenger Comfort KPIs may need more attention. The results show that there is not a significantly underestimated KPI. Even the KPI with the lowest weight (*Passenger Comfort*) shows a significance of more than 5%.

In its current state, the ECML is facing the problem of demand exceeding capacity. A number of improvements are expected all over the network, in particular more seats, fewer delays and perturbations, better utilisation of the current services and infrastructure, and more efficient energy plans (Network Rail, 2008a, Network Rail, 2009, Network Rail, 2010). Overall, the ranking gained from this study is in line with the current situation on the ECML. The weightings may help to identify which elements to enhance in the system.

5.5.5 Discussion

The main problem that all participants of the focus group had raised is whether to take an average of opinions. In this scenario, passengers' views on Energy and Resource Usage are "ignored". It is a question whether to rely more on one or several stakeholders' viewpoints. In some cases, some stakeholders may not be involved. In future practice, the number of stakeholder groups and group members need to be adjusted with each scenario.

The participants also mentioned providing more detailed information about the scenario. For instance, when the IM and TOCs consider some KPIs (e.g. punctuality, energy), they may relate these closely to cost. Thus charges for capacity, delays and electricity all need to be available. It is worth communicating with the participants before the meeting to give them a better understanding of the scenario and the research.

Another thing to improve is the selection of participants. In this study the participants are all railway researchers. They are selected because they have all had some knowledge about the scenario and the KPIs. By including some participants with other backgrounds, the discussion may become broader, and hence more valid results can be produced.

Apart from the KPI weightings, the focus group meeting has also helped in gathering thoughts on how to improve the overall QoS. This helps in defining options for QoS improvement in the scenario. It is to be noted that this study is scenario sensitive. When another railway network is studied, a different focus group meeting must be held to derive scenario specific weights for KPIs.

5.6 Calculation of QoS

The QoS values are calculated with a weighted sum model, with weights generated from the focus group process and pairwise comparisons. It is not feasible or accurate to apply AHP to compare alternatives, as potential alternatives are not defined. In practice, a large number of alternatives may arise, and the pairwise comparison of alternatives would become very subjective with the complexity of the railway network.

In the QoS framework, the eight KPIs are measured separately with their Key Measures. To apply the weighted sum model, the KPIs need to be normalised. This is because not all of the

KPIs improve in the same direction. For accommodation and resource usage, a larger number shows a higher level, yet for other KPIs the smaller the better. The KPIs need to be of the same metric to be considered in this problem. The normalisation of KPIs is done by defining a maximum or best value for each KPI, and compare all the data to these numbers. This would produce criteria values within the (0,1) range.

After the determination of KPI weights, the QoS of a specific railway network can be calculated with the formula:

$$QoS = \sum wn \times KPIn. \tag{5-3}$$

The QoS results are then used for further evaluation and analysis.

5.7 Summary

In this chapter the quantification of Quality of Service is discussed. As a complex multi criteria framework, it is not easy to assign weights to the KPIs for QoS evaluation. Different techniques for weight determination have been discussed. Saaty's pairwise comparison method has been chosen to derive the weights.

A focus group approach has been developed to gather stakeholders' opinions towards the KPI weights. Through the structured focus group discussion, data has been collected about people's expectations and concerns towards the system, and the relative importance of KPIs when compared two at a time. The criteria weights are then derived from the pairwise comparison matrices. A case study based on a part of the East Coast Main Line has been carried out based on the decision making process.

The overall QoS is calculated as the weighted sum of all KPI values (normalised). To assess the cause of KPI/QoS changes, i.e. the factors that affect performance in the network, it is necessary to execute an analysis of the impact of Influencing Factors on KPIs and QoS. The process will be demonstrated in the next chapter.

Chapter 6: Application of the Quality of Service framework

6.1 Conventional measures for performance enhancement

The performance of a railway network, in terms of the quantity and quality of services that it offers to its users, is affected by many factors, such as traffic characteristics, infrastructure and operational methods. Therefore, there is a wide variety of potential solutions for improving the performance and capacity of a railway line or part of a network (UIC, 1996, Roberts et al., 2010, Strategic Rail Authority, 2003), among which improved approaches towards scheduling is an important concern (Burdett and Kozan, 2010, Burdett and Kozan, 2009). Depending on how easily such solutions can be implemented, they can be classified as 'strategic' or 'tactical'.

The 'strategic' solutions involve modifying the more fixed parameters of the network or system by improving components and features that require a long time or significant investment, if a change is to be achieved. Several aspects are involved, listed here in ascending order of difficulty of intervention:

- Timetable parameters: traffic pattern, buffer times, recovery times, etc;
- Rolling stock parameters: quantity of passenger and freight accommodation, train performance in terms of acceleration and braking rates;
- Infrastructure parameters: upgrades to the signalling system, track alignment, junction layouts, station tracks and new lines, etc.

Solutions that address these aspects of network performance may lead to a significant increase in the quantity and quality of network capability.

The 'tactical' solutions are those that are relatively easy to implement in terms of time and monetary investment. The factors changed are included in the category of tactical factors in Figure 4-1. These are the short to medium term solutions.

In the short term, measures can be taken to improve the quantity and quality of existing traffic capacity and to influence the market demand. These measures include:

- Tariff changes;
- Minimal timetable modifications, such as amending dwell times and turnaround times;
- Better staff management.

The short term solutions can lead to a temporary improvement in network performance, particularly during peaks caused by certain events, such as sporting events and exhibitions. However, dissatisfaction may be caused amongst passengers, staff and trade unions. Generally, the measures are also not compatible with the need to accommodate an ever-increasing traffic volume.

In the medium term, solutions can involve, again in increasing order of difficulty:

- Adjustment of allowances and minor timetable changes such as small train path amendments;
- Changing operating parameters: delay management, etc;
- Changing traffic parameters: priorities, train mix, etc.

The medium term solutions are often considered the best approaches for network performance improvement, especially with a limited budget. With little change to the characteristics of the current railway networks, these solutions can maximise the quantity and quality of the existing network. Yet, on already overloaded networks, the effects may be minor.

6.2 Performance improvement based on the QoS framework

6.2.1 The selection of KPIs and key measures

In the QoS framework, a large number of KPIs and key measures are applied to quantify performance. For different performance goals, the importance of these KPIs and key measures may vary according to the scenario. With different characteristics of each railway network, it is necessary to check the use of these entities.

In practice, the criteria weights for the KPIs vary a lot from scenario to scenario, and sometimes a number of KPIs may even be neglected in the evaluation. The choice of KPI weights is relevant to the purpose of the project and the stakeholders involved. For example, if the infrastructure manager/timetable planner wants to investigate the influence of a timetable change, they might not be interested in passenger comfort or energy usage. On the other hand, if the project is aimed at attracting more passengers to use the network, passenger comfort would be valued a lot, while energy and resource usage might be given a very low priority. At the beginning of the performance evaluation and improvement process, the relative importance of the KPIs in the specific scenario must be determined.

The key measures used to quantify each KPI may also be modified to suit the scenario. There is a lot of flexibility in the key measures used, e.g. the time window and area for assessment, the key stations or timing points in the network, the relationship between the key measures under the same KPI. These issues also need to be defined at the early stages in the evaluation.

6.2.2 Performance assessment and analysis

With the Quality of Service framework, the parameters affecting railway performance are broken down into several categories of influencing factors (IFs). With a defined combination of IFs, the performance of the railway network can be quantified as KPIs and QoS values ready for comparison and further analysis purposes.

When different combinations of IFs are introduced on the same railway network, the corresponding performance may vary to a large extent. To make a plan for performance improvement, it is necessary to analyse the impact of the variations of the framework inputs, i.e. IFs, on the KPI and QoS outputs.

For performance evaluation purposes, a railway network can be considered using a complex multi-criteria model which can be quantified with a mathematical expression. This mathematical equation links the IFs with the KPIs and QoS, and can be retrieved through experimentation. The process is to input a good variety of IFs into the system, and analyse the performance outputs. In this study, as the number of inputs is large, the regression method is adopted to establish the link between inputs and outputs. From the results of the analysis, the impact of the IFs can be assessed with their percentage contributions on the outputs. The decision maker is able to identify the strengths and weaknesses in the system, and also predict performance with given IF values.

6.2.3 Developing a performance improvement plan

From the QoS performance analysis, the decision maker can extract the percentage contribution of each input factor as to the outputs. The results can be used to assist the development of performance enhancement plans.

Based on the QoS framework, the IFs are also divided into strategic and tactical factors, according to the time scale for implementation:

- Strategic factors:
    - Rolling Stock;
    - Infrastructure;
    - Timetable;
    - Operational Rules;
- Tactical factors:
    - Traffic Management;
    - Operational Management;
    - Human Factors;
    - System Maintenance;
    - Environmental Factors.

According to the predefined performance goals, predictions can be made on the best combination of IF values to produce the optimal performance. It is also possible to deduce the corresponding performance with specific IF values. When applying the QoS analysis in practice, the more cost effective factors shall be considered first, followed by those that are difficult to be modified. With the rank of the IFs by impact from the performance analysis results, the performance improvement plan can be made under the following rules:

- The tactical factors that have high ranks shall be tested first, by varying their values while keeping all other factors unchanged;
- If some tactical factors have low ranks, they can still be tested before the strategic factors, to check if their impact is satisfactory for the performance goals;
- If the strategic factors are to be assessed, operational rules and timetable factors may be analysed first, followed by rolling stock, then infrastructure factors;
- When considering rolling stock and infrastructure factors, those that have less influence to the timetable can be given higher priority. Some alternatives such as adding a new line may result in a significant change in the timetable, and require a reworking of operational planning. For existing railway networks, these changes would be the least financially friendly choices, thus need to be evaluated closely.

With these rules, several alternative combinations of IF values can be generated for further evaluation. These alternatives will be tested in simulation, and the alternative with the best performance output would be adopted to deliver a final version of the performance improvement plan.

6.3 Using the QoS framework to support strategic management

Strategic management is concerned with implementing strategies and measuring performance as well as monitoring trends and identifying emerging issues that might require strategic responses (Poister, 2008). In the railway strategic management process, performance evaluation and improvement plays a central part, as shown in Figure 6-1.

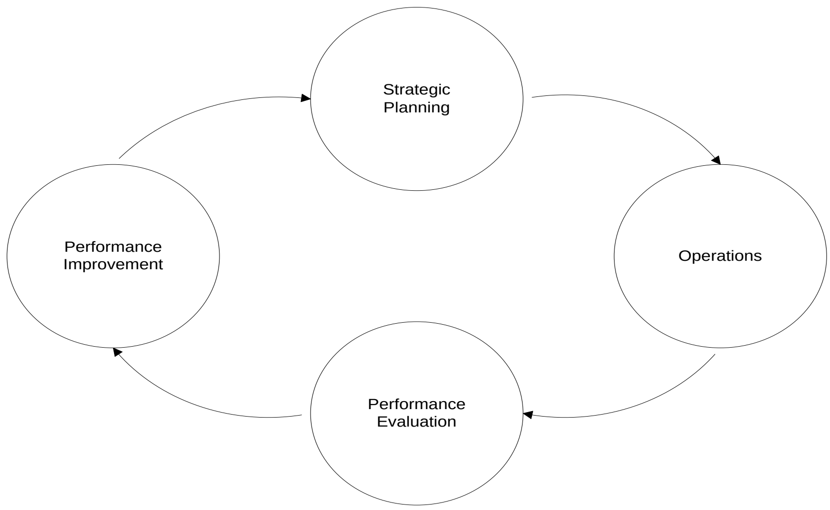


Figure 6-1: The strategic management process

In a railway network, railway operations are managed under a strategic plan constructed in advance. Performance is analysed and evaluated under the current conditions, and then used to make plans for performance improvement. This improvement plan would then be implemented into the strategic planning process to support the system upgrade to deliver better performance.

The QoS framework is a practical tool to conduct railway performance evaluation and improvement. With the QoS framework, performance can be assessed and enhanced in a structured way to produce meaningful and effective outputs to support future development of a railway system. The feasibility of the predefined performance goals from the stakeholders can be checked, and suggestions can be given on the selection of system elements that should be given more consideration.

6.4 Summary

In this chapter, the application of the QoS framework to support the construction of performance improvement plans is discussed. For each individual railway network, the KPIs and their weights need to be determined in the early stages of the performance improvement process. The information then goes into the QoS framework to calculate the current performance value. To generate a performance improvement plan, the influencing factors are tested and analysed to assess their impact on KPIs and QoS. From the analysis results, suggestions can be given on performance enhancement based on the proposed performance goals from stakeholders.

Chapter 7: Case Study

7.1 Introduction to the simulation software

Railway simulation provides an emulation of real world operations in a virtual environment. Compared with on line experiments, simulation saves money and allows for easier performance analysis, evaluation of interactions and timetable modification. (Albrecht et al., 2008)

A number of software tools have been developed for railway simulation purposes. They can be categorised into Macroscopic and Microscopic models.

Macroscopic models describe the network with nodes as stations, connected with directed links (Albrecht et al., 2008). Typical applications include SIMONE (Simulation Model for Networks) and VIRIATO, which are aimed at assessing timetable robustness and supporting strategic timetable planning (Barber, 2007).

Microscopic models replicate the railway operations with detailed infrastructure and timetable information (Nash and Huerlimann, 2004). There are two types of Microscopic simulation models: synchronous and asynchronous. Synchronous models are event-driven, and deal with all trains, routes and signals at the same time. Asynchronous models calculate trains according to their priority (Nash and Huerlimann, 2004). The most commonly used rail simulators are often microscopic, synchronous ones, such as RAILSYS, OPENTRACK, RTC, etc.

For the case study presented, there is a lot of detailed information about infrastructure, rolling stock and operations. All the relative information should be available and accessible in each test run. It is necessary to choose a microscopic, synchronous simulation tool. As the simulation area is large, the amount of time needed for the construction of the model, as well as the difficulty of making modifications in the model, must be taken into consideration.

In this case study, the simulation software chosen is BRaVE (Birmingham Railway Virtual Environment). It is a synchronous, microscopic railway network simulation tool, written in the JAVA development environment, at the University of Birmingham. The simulator forms the core of the Birmingham Centre for Railway Research and Education virtual railway laboratory (Kirkwood, 2014).

The BRaVE simulator is able to process conventional files from other simulators. The file containing the infrastructure information of the simulation area has been taken from the ON-TIME project, and imported into the BRaVE simulator.

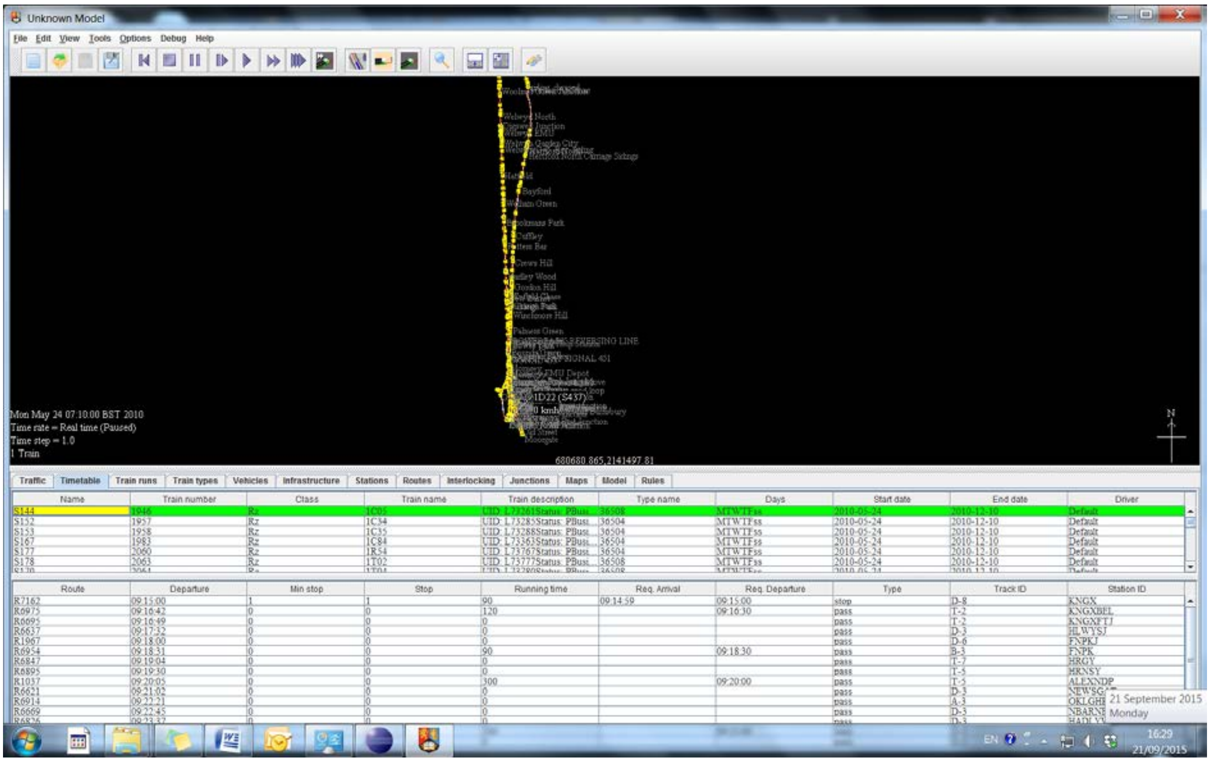


Figure 7-1: The BRAVE simulator

The BRaVE simulator is composed of a graphical panel and an information panel, as shown in Figure 7-1. The graphical panel is a demonstration of the simulation model in a graphical view, showing all elements of the network model, including infrastructure, signals and trains, etc. The network elements are shown in different colours to indicate track occupation and reservation.

The information panel shows the current state of all the entities in the simulation database. This information is grouped in the traffic, timetable, train run, train types, vehicle, infrastructure, routes, interlocking, junction and maps panels. It is possible to make modifications to the timetable, infrastructure and vehicles. Predefined patterns of traffic management, driver behaviour and dispatcher strategies can also be selected in the information panel.

In BRaVE it is possible for users to write their own APIs to control dispatcher, driver, and signaller behaviour. An API (Application Program Interface) is a software component defined by the user to interact with the main program. Incidents such as vehicle breakdown and signal failure can also be inserted through APIs.

BRAVE is capable of providing a variety of simulation outputs. Figure 7-2 shows the train running graphs. The data shown in the graphs can also be exported as .csv files for each individual train.

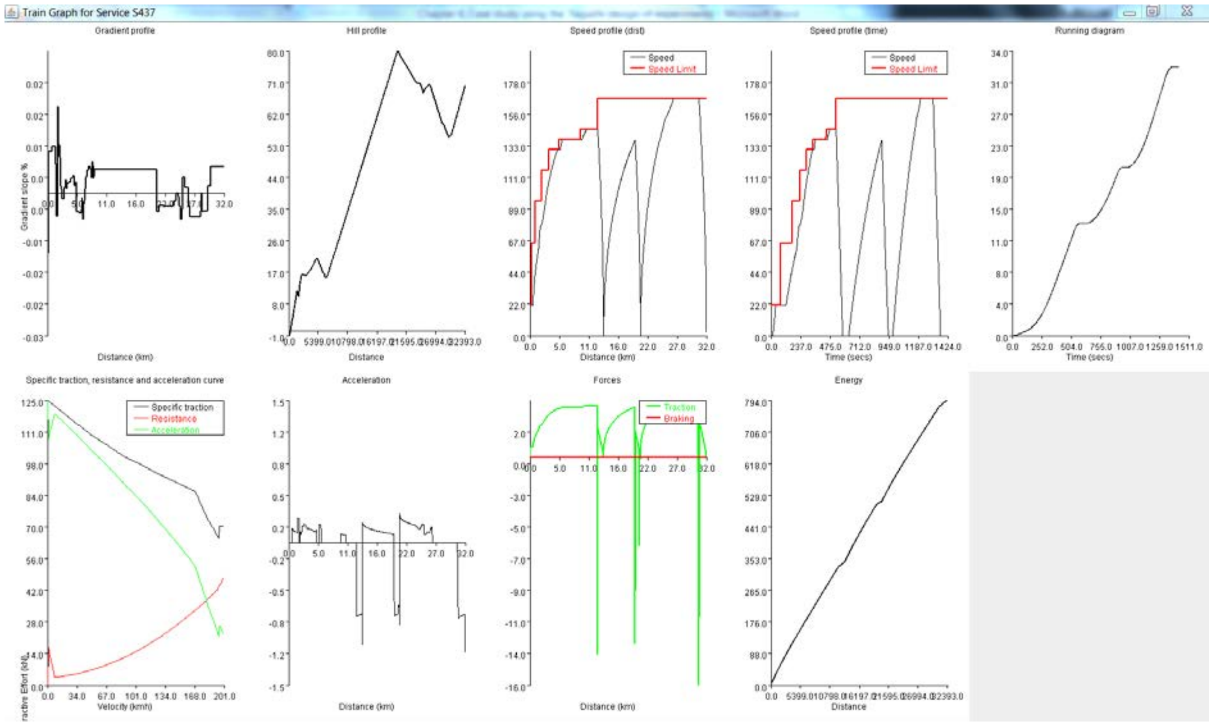


Figure 7-2: Train running graphs

A log file of train passing times at timing points, train positions and energy usage can be produced using the embedded log file tool. Figure 7-3 shows the log file dialog. Users may define the start and end times that the log file should apply, and select signals and stations to be logged.

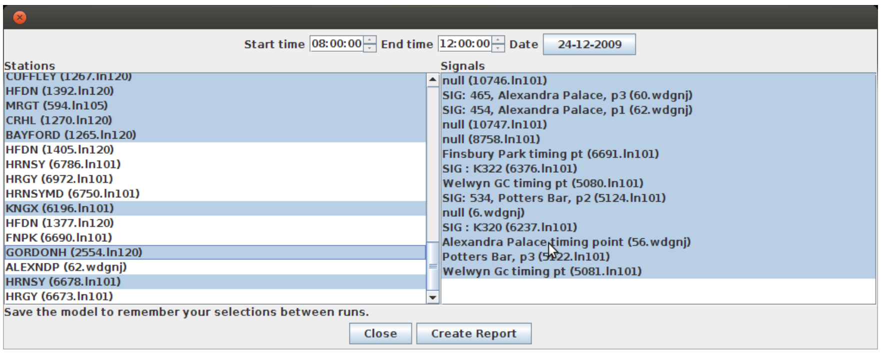


Figure 7-3: Log file dialog

7.2 Scenario setting

A network layout chosen from part of the East Coast Main Line (ECML) in Great Britain is considered in the research.

The East Coast Main Line (ECML) is a 393 mile high-speed railway line linking London and the south east to major cities. It is electrified (with 25 kV OHL) along its whole length. It intersects with a number of other routes at a number of locations, most notably with the North London Line at Copenhagen Junction and Harringay Junction, and the Moorgate branch at Finsbury Park South Junction. The route handles key commuter flows to the north of London, as well as some long distance high speed flows. It also carries regional commuter services, local passenger services and heavy tonnages of freight traffic. It is therefore important to the economic health of a number of areas in Great Britain.

The ECML is a very busy line, with much of the route currently operating at or just below capacity for most of the day. The capacity constraints are mainly due to service mix and stopping patterns. Key routes with saturated capacity include South of Peterborough (capacity usage calculated by the Capacity Utilisation Index: almost 100% during peak hours, and around 70% at off-peak times), Peterborough to Doncaster (70 – 80% peak and off-peak), etc (Network Rail, 2009).

The selected section for the case study (Figure 7-4) runs are between Welwyn Garden City, Hertford North, London King's Cross and Moorgate. In total there are 29 stations. The section comprises four track for most of its length, but widens to six tracks between Alexandra Palace and Finsbury Park, with the Finsbury Park-Moorgate and Alexandra Palace-Hertford North branches being double track. This route section is one of the most vital railway links in the UK, and a good representative of modern European railway networks. It has a good mixture of train speeds and service patterns, and it also runs through two busy junctions and several heavily populated locations with a lot of passenger flow. There is a vast range of variables involved in the operations of this network, and much concern about the performance of this route section. These make the section an ideal scenario for timetable analysis and performance assessment. It is also one of the test areas in the ON-TIME project (ON-TIME, 2012a).

Within the simulation area, in terms of station usage (measured by entries and exits at stations), the most significant stations are King's Cross (31.3 million), Moorgate (9.4 million), Finsbury Park (6.3 million) and Welwyn Garden City (2.7 million) (Steer Davies Gleave, 2015). Figure 7-4 shows the map of the simulation area with relative station usage levels. In each direction, during the busiest part of the day, there are currently 8 long distance high speed trains and 25 suburban trains per hour on the route. Of the suburban trains, more than half are into King's Cross. The highest frequency of services applies to the section between Finsbury Park and Alexandra Palace. Network Rail has made plans to increase the frequency of trains in the near future.

Signalling is using track circuit blocks at present; ERTMS is due to be installed on the route in 2019.

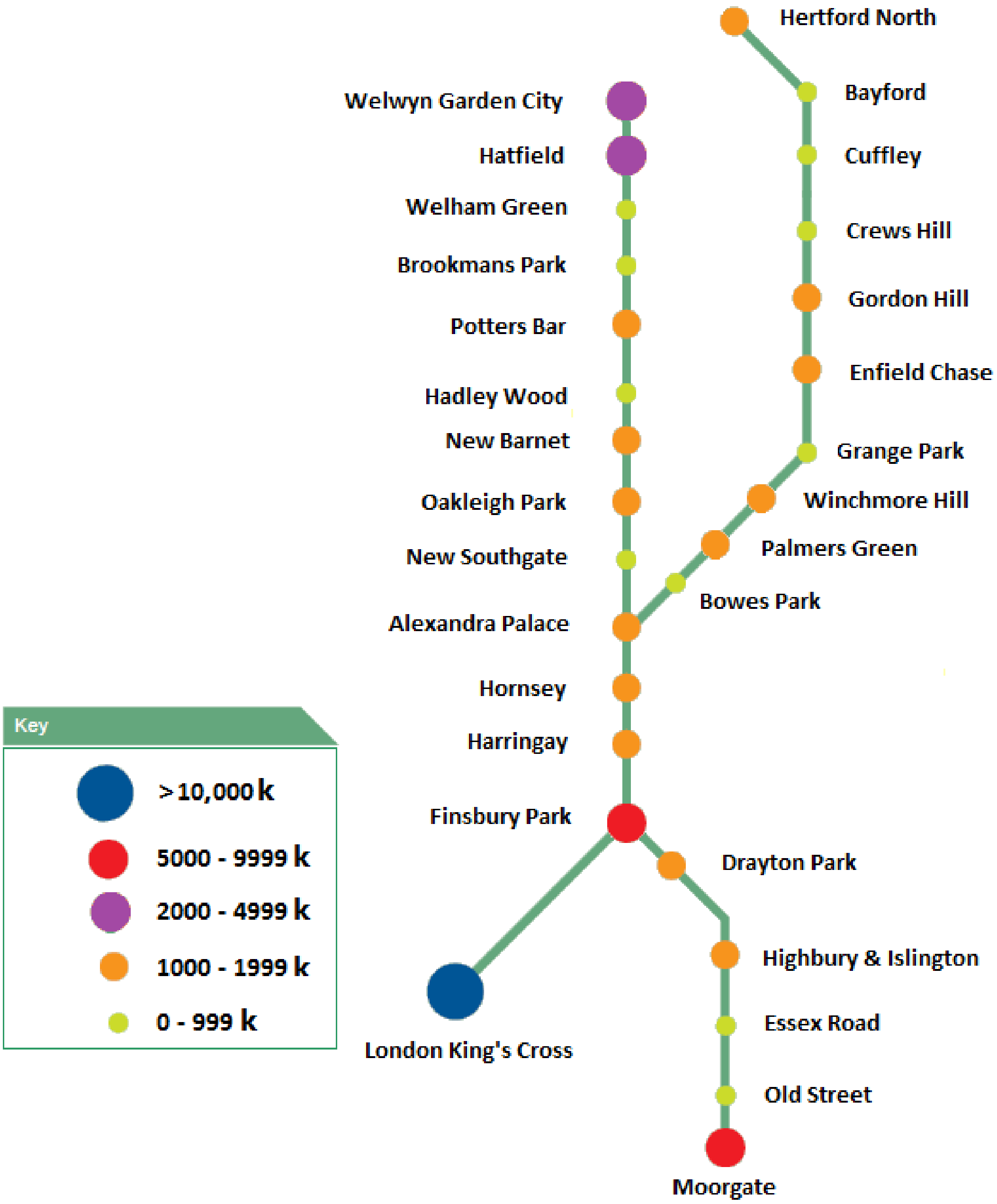


Figure 7-4: Station map of the simulation area (Network Rail, 2009)

According to the delay data collected between 1st April 2005 and 31st March 2007 (Network Rail, 2009), most delays occur at the following stations, as shown in Figure 7-5:

Hertford North, Gordon Hill, Welwyn Garden City, Potters Bar, New Barnet, Alexandra Palace, Finsbury Park, Old Street, Moorgate, Belle Isle, London King's Cross.

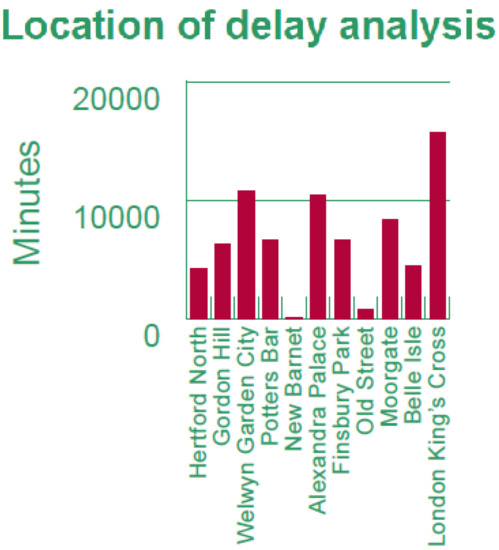


Figure 7-5: Location of delays in the simulation area

7.3 Experiment design

The case study is carried out to test the effectiveness of the Quality of Service application to assess the overall performance. An analysis of IFs and KPIs/ QoS is conveyed to determine the influence of individual factors on the overall performance.

There are a number of influencing factors in the system, as listed in the QoS framework. As the number of factors is large, traditional experiment design procedures would be too complicated, costly and difficult to use. In this case, the Taguchi design of experiments is applied, where orthogonal arrays are used to ensure the quality of experimental data, and also limit the total number of experiments required at the same time.

7.3.1 Design of Experiment (DOE) methods

Design of Experiments was first introduced by Sir Ronald Fisher in the 1920s, when he worked at the Rothamsted Agricultural Field Research Station in England. His experiments were tailored to study the effect of various fertilisers on different plots of land to improve the yield of crops. Fisher applied DOE to differentiate the effect of fertilisers from other relevant factors such as underlying soil condition, moisture, gradient, etc. In this process, he defined the basic principles of factorial design and the data analysis method known as ANOVA (analysis of variance) (Antony, 2014).

Ever since the first applications in agricultural and biological fields, DOE methods have been accepted and applied in many areas of industrial and scientific research, including physics, chemistry, medicine, electronics, computer science, civil engineering, etc. (Chen et al., 2015, Coronado et al., 2015, d'Ambrosio and Ferrari, 2015, Hudovornik and Vrečer, 2015, Lemonakis et al., 2016, Maheshwari et al., 2015, Slanzi et al., 2015, Wu et al., 2015)

All DOE methods share the common goal of extracting the most information possible from a limited set of experiments. For experiments with a large number of factors, the traditional One-Variable-At-a-Time (OVAT) approach is often unreliable, inefficient, time consuming, and may yield false optimum values (Antony, 2014). In a designed experiment, changes are made to input variables in an organised way so that changes in output responses can be observed. DOE is extremely helpful in assessing the influence of key variables on the quality characteristics of interest.

There are many types of experimental design, usually grouped as screening designs (e.g. full factorial design, fractional factorial design), response surface designs (e.g. central composite design, Box-Behnken design), and mixture designs (e.g. simplex lattice design, simplex centroid design) (Anthony et al., 2003, Dejaegher and Vander Heyden, 2011). The latter two design types are often used to refine models after a full factorial analysis of important factors. Figure 7-6 shows examples of screening and response surface designs with three two-level factors. Each point represents a combination of factors to be tested in experiments.

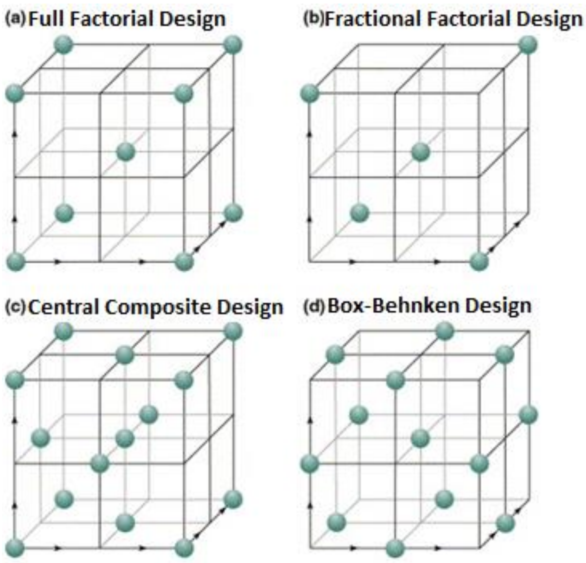


Figure 7-6: Experiment designs (Tye, 2004)

The choice of DOE is a compromise between the amount of information required and the number of experiments affordable. As in this case study, the size of network is quite large (29 stations), as is the number of train services (75 services). Due to limitations of the simulation software, making changes to some factors, such as timetable allowances, would become very time consuming. It is most feasible to apply a fractional factorial design to reduce the number of experiments. Thus the Taguchi design of experiments is applied.

7.3.2 The Taguchi design of experiments

The Taguchi method was developed by the Japanese scientist Dr. Genechi Taguchi in the late 1940s. Since its introduction in the USA in the 1980s, the method has been well developed and widely applied in both industry and research. The main applications include optimising manufacturing processes and designing high quality systems. It has become one of the most effective quality control tools for scientists, engineers and researchers.

In Taguchi designs, orthogonal arrays are used to test factors in pairs instead of testing all combinations of factors. The advantages are that experiment time is saved, costs are reduced, and the significance of factors is quickly detected. This method can also aid a robust design to find the conditions that would lead to the optimum performance.

The process of Taguchi DOE is listed in steps as follows:

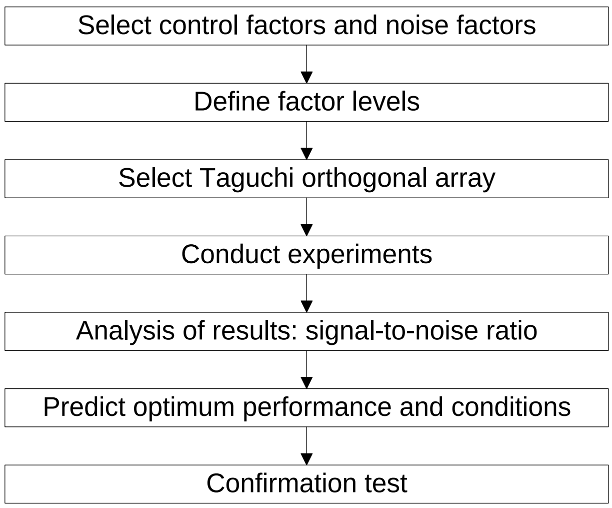


Figure 7-7: The Taguchi process

The first step is factor and level selection. Control factors are the factors that are feasible to control in experiments. Noise factors are those that cause variability in performance, yet cannot be controlled accurately in real life scenarios or in experiments.

The aim of the Taguchi design is to minimise the influence of noise factors by manipulating the S/N ratio. If the aim is to maximise the response, the Signal-to-Noise ratio (S/N ratio) is expected to be "Larger is better". To minimise the response, "Smaller is better" applies. To keep the response to a target, "Nominal is best" is to be selected.

Factors are arranged in an orthogonal array to be tested. The inner array contains all the control factors, and the outer array deals with noise factors. Each combination of control factors are tested under all noise conditions specified in the outer array. The outputs are assessed by their performance values and S/N ratios.

The experiment outputs are then sent to the ANOVA (analysis of variance) analysis. ANOVA is a statistical tool to assess the relative influence of factors and their interactions (Ross, 1996). It detects the difference in the means of factor groups, and identifies the source of the performance variations.

The S/N ratios can also be used to predict the optimum performance. A confirmation test is carried out to test the optimum conditions.

7.3.3 Experiment parameters

According to the Quality of Service framework, several categories of factors may contribute to performance changes (Figure 7-8). Among these, some factors are more controllable in experiments, including Rolling Stock, Timetable and Operational Rules. The tactical factors are considered to be less controllable, or may be treated as noise factors (e.g. Environmental Factors), as they may bring delays and incidents into the system.

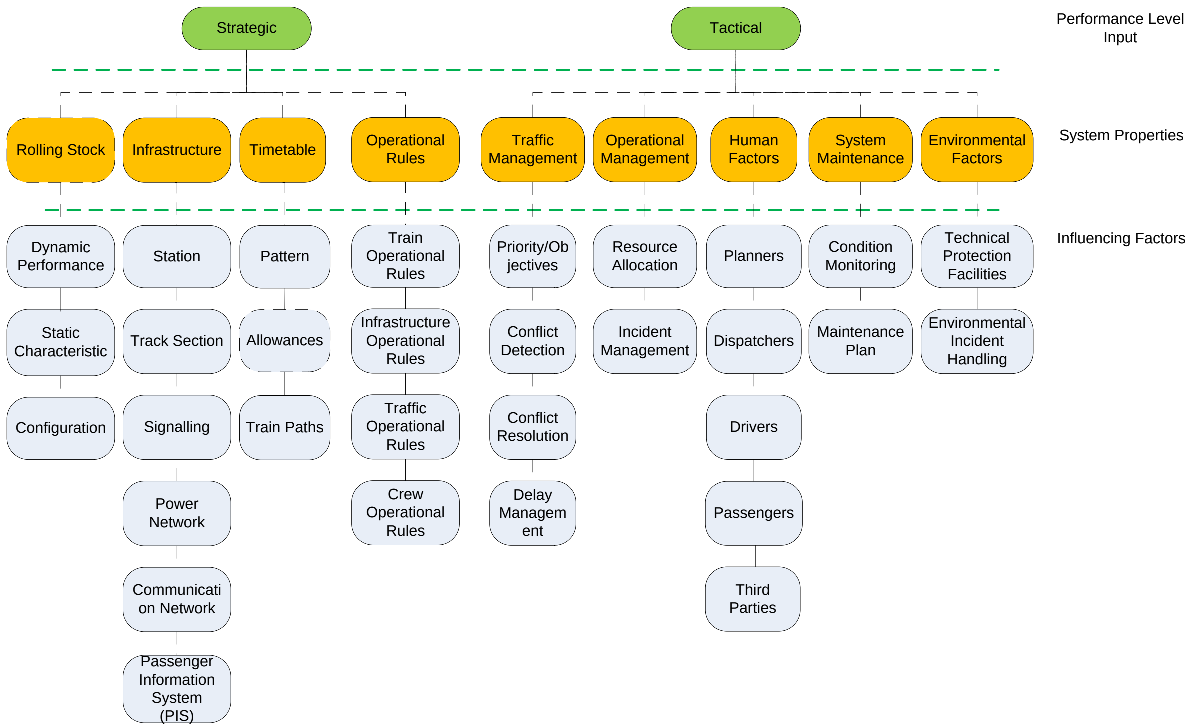


Figure 7-8: Influencing factors in the railway system

Due to limitations of the simulation software, only some of the factors can be tested in BRaVE. These factors are:

Rolling Stock:

- Dynamic performance and configuration: The parameter Train Types in BRaVE covers configuration, acceleration and deceleration;
- Static performance: max speed, number of carriages, length, weight, and number of seats. These factors are defined under Vehicles in BRaVE.

Infrastructure:

- Stations;
- Track sections;

- Signalling.

Timetable:

- Pattern;
- Allowances: buffer times and recovery times;
- Train paths: train routes, train stops, dwell times and running times.

Operational Rules:

- Infrastructure operational rules: line speed limits on routes.

Traffic Management:

- Conflict resolution;
- Delay management.

Human Factors:

- Driver: Driving style is considered in the simulation.

Among these, infrastructure factors are the most difficult to modify in simulations, as infrastructure upgrades, timetables and operations also need to be changed accordingly. It is also unusual to change some rolling stock characteristics, such as vehicle length and weight.

In this case study, 10 control factors and 2 noise factors are considered. These factors are relatively independent, thus correlation between different factors is not considered. Each control factor is tested with 3 levels of values, which are summarised in Table 7-1. Level 1 represents the current condition, and the other two levels with deviations. The value of each level is selected with consideration of its feasibility under the current conditions of the simulated network.

	Seating capacity	Configuration	No. of stops	Scheduled dwell times	Sectional Running Time
Code	A	B	C	D	E
Level 1	100%	Train type 31306: 2 units	original	100%	100%
Level 2	120%	Train type 31306: 3 units	increased	110%	110%
Level 3	80%	Train type 31306: 1 unit	reduced	90%	90%

	Level of Heterogeneity	No. of services	Speed limits	Driving style	Minimum dwell times
Code	F	G	H	J	K
Level 1	original	75	100%	Default	100%
Level 2	increased	85	110%	Early Coasting	110%
Level 3	reduced	63	90%	Target Speed	90%

Table 7-1: Control factor levels

The control factors are coded as A, B, C, D, E, F, G, H, J, and K, in the sequence as in Table 7-1.

The original timetable consists of 75 passenger trains. It runs from 7am to 10am on a weekday, which is the morning peak in the network. A mixture of fast and slow trains is selected. There are three types of trains, grouped by their maximum speed: 201km/h, 161km/h, and 121km/h. Train services are inserted or reduced with consideration of these speed types.

Seating capacity refers to the total passenger seats and standing spaces for each train. Configuration applies to train type 31306, the main type of rolling stock in the timetable. The number of stops, number of services and level of heterogeneity are all changed according to train speed types.

When adding the number of stops, fast trains are made to stop at two busy stations (Potters Bar and Oakleigh Park). When reducing the number of stops, stops made by slow trains at Old Street and Gordon Hill are taken out. These two stations host a moderate level of traffic.

The level of heterogeneity deals with the differences in running times per track section.

Scheduled dwell time, minimum dwell time, sectional running time and driving style applies to all services. Scheduled dwell time is the timetabled time for trains to dwell at a station. Minimum dwell time is the shortest time trains can stay at stations, which is used when trains are running late. Sectional running time is the train running time between two stations. All trains are timed at pre-defined timing points, of which most are around stations.

Speed limit and driving style changes apply to the whole network.

Driving styles are changed by inserting a driver plugin in the BRaVE API. Three kinds of driving styles are defined. For the Default Style, trains run to line speed limits. For Style 2, drivers carry out early coasting and coast if more than 60 seconds early. For Style 3 drivers with a target speed run to 90% of the speed limit, with acceleration and braking when they are 60 seconds early or 30s late.

A detailed explanation of the factor levels for each style can be found in Appendix B.

Two noise factors are considered: conflict resolution and dispatcher delays (Table 7-2). These factors are closely related to traffic management, and have a significant impact on operations. However, they cannot be controlled accurately in practice. This is due to the randomness in the occurrence of delays and disturbances during operation. Conflict resolution strategies are applied when there is a conflict in the network and trains get stuck, blocking the routes of other services. When rerouting, trains are transferred to another available route. Cancellation means cancelling trains before they run to positions where conflicts may occur. Random delays caused by dispatchers can be inserted at each dispatching station. The range is 0 to 15 s.

	Conflict Resolution	Dispatcher delays
Level 1	Rerouting	None
Level 2	Cancellation	Random delays

Table 7-2: Noise factor levels

This results in a Taguchi design of a 10-factor inner array (L27) and a 2-factor outer array (L4), as in Table 7-3. For experiments with 3-level factors, the L27 array is chosen if the number of factors is between 5 and 13. In this study, there are ten control factors, thus the last three columns of the L27 array are not used. Each combination of factors would be tested with two 2-level noise factors, as listed in the outer array.

The inner array shows all the combinations of control factors to be tested in the experiment. In total ten factors (A, B, C, D, E, F, G, H, J, and K) are chosen, each tested with three levels (1, 2, and 3). As stated above, the L27 array is chosen. Thus for each noise situation, there would be 27 tests to be carried out.

The outer array lists the noise situations in the case study. Two noise factors are defined, each tested with two levels. This leads to four types of noise that need to be studied. All combinations of factor levels defined in the inner array would be tested under these four noise situations.

The total number of tests in the case study is calculated as 27×4 (108).

	Inner Array									
	A	B	C	D	E	F	G	H	J	K
T1	1	1	1	1	1	1	1	1	1	1
T2	1	1	1	1	2	2	2	2	2	2
T3	1	1	1	1	3	3	3	3	3	3
T4	1	2	2	2	1	1	1	2	2	2
T5	1	2	2	2	2	2	2	3	3	3
T6	1	2	2	2	3	3	3	1	1	1
T7	1	3	3	3	1	1	1	3	3	3
T8	1	3	3	3	2	2	2	1	1	1
T9	1	3	3	3	3	3	3	2	2	2
T10	2	1	2	3	1	2	3	1	2	3
T11	2	1	2	3	2	3	1	2	3	1
T12	2	1	2	3	3	1	2	3	1	2
T13	2	2	3	1	1	2	3	2	3	1
T14	2	2	3	1	2	3	1	3	1	2
T15	2	2	3	1	3	1	2	1	2	3
T16	2	3	1	2	1	2	3	3	1	2
T17	2	3	1	2	2	3	1	1	2	3
T18	2	3	1	2	3	1	2	2	3	1
T19	3	1	3	2	1	3	2	1	3	2
T20	3	1	3	2	2	1	3	2	1	3
T21	3	1	3	2	3	2	1	3	2	1
T22	3	2	1	3	1	3	2	2	1	3
T23	3	2	1	3	2	1	3	3	2	1
T24	3	2	1	3	3	2	1	1	3	2
T25	3	3	2	1	1	3	2	3	2	1
T26	3	3	2	1	2	1	3	1	3	2
T27	3	3	2	1	3	2	1	2	1	3

	Outer Array	
	N1	N2
T1	1	1
T2	1	2
T3	2	1
T4	2	2

Table 7-3: Taguchi orthogonal arrays

7.4 Data processing

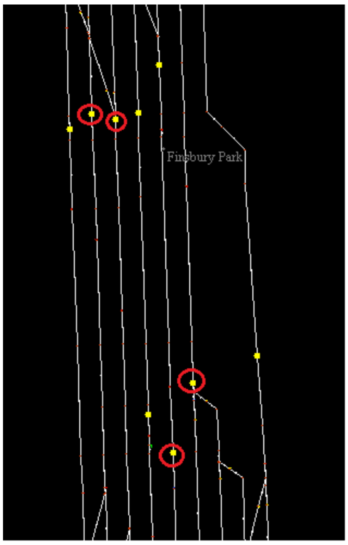
From the BRaVE simulations, train runs are recoded as log file outputs. This information is imported into a MATLAB program to calculate the corresponding KPI values: AC, JT, CN, PT, RS, EN and RU. The number of seats and station distances are recorded in .mat files to be processed in MATLAB.

The KPIs are calculated as:

- Accommodation: total seat km of all services in the network;
- Journey Time: the total journey time of long distance services;
- Connectivity: average connection time by services which make a stop at the interchange station;
- Punctuality: total delay of services at designated stations;
- Resilience: the total delay area;
- Passenger Comfort: frequency of jerks (the rate of changes of acceleration);
- Energy Usage: total energy consumed by all trains;
- Resource Usage: the number of trains used and track occupation.

Long distance services are those running from border to border, i.e. trains whose origin and destination stations are among Welwyn Garden City, Hertford North, London King’s Cross, and Moorgate. Punctuality is measured at stations and timing points which are frequently used by Network Rail for monitoring delays, as shown in Figure 7-5, i.e. Hertford North, Gordon

Hill, Welwyn Garden City, Potters Bar, New Barnet, Alexandra Palace, Finsbury Park, Old Street, Moorgate, Belle Isle, and London King's Cross. Connectivity measures interchanges at Alexandra Palace, for the trip from Gordon Hill to Oakleigh Park. Passenger Comfort is determined by the driving style in this case. If the driving style would lead to more jerks, it is considered to provide worse Passenger Comfort. The default driving style is marked as 0.8, Early Coasting as 0.9, and Target Speed as 0.6. Track occupation is measured by the total passing times of trains at timing points. In the studied area, the busiest section is between Alexandra Palace and Finsbury Park. As no trains in the original timetable start or terminate their journeys within this section, all trains entering from Alexandra Palace or Finsbury Park would run through the whole section. Thus the track occupation in this section can be assessed with a few timing points on each track of the route. In this case study a number of timing points around Finsbury Park are chosen (Figure 7-9). These timing points are denoted by the red circles. In the experiment timetable, only four of the tracks have been used by services, and the timing points are picked on these tracks.



Left to Right:
N14714
N14715
N14434
N14693

Figure 7-9: Selected timing points to measure track occupation

Each KPI is normalised as a percentage to be applied in the QoS calculation. A standard maximum value is set for each KPI. In this case these are the maximum from all experiments. All KPI values are normalised by dividing them by the maximum values.

The KPIs are multiplied by their weights from the Focus Group study Table 7-4.

	Weights	Ranking
AC	0.1972	1
JT	0.0763	6
CN	0.0663	7
PT	0.1696	3
RS	0.1371	4
PC	0.0646	8
EN	0.0933	5
RU	0.1955	2

Table 7-4: KPI weights

The QoS result of each experiment is calculated as the weighted sum of the normalised KPI values. For all 108 tests, the QoS values are listed in Table 7-5.

	Quality of Service				Taguchi Analysis
	N1	N2	N3	N4	S/N (dB)
1	0.7239	0.7149	0.7239	0.7149	-2.8606
2	0.6962	0.7386	0.6962	0.7386	-2.8966
3	0.4741	0.4754	0.4741	0.4754	-6.4707
4	0.6538	0.6552	0.6538	0.6552	-3.6820
5	0.6828	0.7026	0.6790	0.6996	-3.2132
6	0.5211	0.5118	0.5211	0.5118	-5.7406
7	0.5579	0.5566	0.5579	0.5566	-5.0792
8	0.6454	0.6902	0.6474	0.6931	-3.5063
9	0.4027	0.4105	0.4027	0.4105	-7.8182
10	0.5903	0.6014	0.5903	0.6014	-4.4980
11	0.6484	0.6867	0.6484	0.6867	-3.5210
12	0.6970	0.6863	0.6970	0.6863	-3.2029
13	0.7328	0.7170	0.7328	0.7170	-2.7959
14	0.7272	0.7643	0.7243	0.7643	-2.5655
15	0.6543	0.6563	0.6543	0.6563	-3.6712
16	0.6505	0.6061	0.6505	0.6061	-4.0533
17	0.6332	0.6826	0.6332	0.6826	-3.6551
18	0.6364	0.6349	0.6364	0.6349	-3.9355
19	0.7133	0.6869	0.7133	0.6869	-3.1018
20	0.5270	0.5735	0.5270	0.5735	-5.2121
21	0.5192	0.5247	0.5155	0.5203	-5.6815
22	0.7559	0.7683	0.7559	0.7683	-2.3607
23	0.6673	0.7087	0.6673	0.7087	-3.2603
24	0.5510	0.5685	0.5510	0.5635	-5.0618
25	0.6195	0.6150	0.6195	0.6150	-4.1912
26	0.4848	0.5220	0.4848	0.5220	-5.9801
27	0.5407	0.5411	0.5407	0.5411	-5.3376

Table 7-5: Experiment results: Quality of Service and Signal to Noise ratios

7.5 Interpretation of experiment data:

7.5.1 S/N ratio

By importing the QoS values into MINITAB, a Taguchi analysis is executed. The aim is to maximise the QoS value (the response of the experiment), thus the Signal-to-Noise (S/N) ratio is expected to be "Larger is better".

$$S/N = -10 \times \log\left(\sum\left(\frac{1}{Y^2}\right)/n\right)$$

The S/N results are shown in Table 7-5. In all experiments, Test 9 and Test 22 denote the smallest and largest S/N ratio.

Table 7-6 lists the responses for Signal to Noise ratios for different factor levels. The data is visualised in a Main Effects Plot (Figure 7-10), showing the mean of S/N ratios of each factor at Level 1, Level 2 and Level 3. The factors ranked by S/N ratio (high to low) are: No. of services (G), Sectional Running Time (E), Configuration (B), Seating Capacity (A), No. of stops (C), Driving style (J), Minimum dwell times (K), Level of Heterogeneity (F), Scheduled dwell times (D), and Speed limits (H). To produce the optimal overall S/N ratio, all factors need to be set at their best response level, i.e. the level at which each factor produces the largest S/N ratio. From the Main Effects Plot, the factor levels that would yield the optimal response are A2B2C1D1E1F1G2H2J1K1.

Level	A	B	C	D	E	F	G	H	J	K
1	-4.585	-4.161	-3.839	-4.085	-3.625	-4.098	-4.16	-4.231	-3.871	-3.944
2	-3.544	-3.595	-4.374	-4.253	-3.757	-4.116	-3.342	-4.173	-4.373	-4.262
3	-4.465	-4.84	-4.381	-4.256	-5.213	-4.381	-5.092	-4.191	-4.351	-4.389
Delta	1.041	1.245	0.542	0.171	1.589	0.282	1.75	0.057	0.502	0.445
Rank	4	3	5	9	2	8	1	10	6	7

Table 7-6: Response table for Signal to Noise ratios (Larger is better)

Figure 7-10: Main effects plot for S/N ratios

7.5.2 ANOVA analysis

The purpose of the ANOVA analysis is to determine which factor affects the quality characteristics significantly (Taguchi, 2005). First, the total sum of squared deviations from means is calculated as:

$$SS_T = \sum_{i=1}^{n} (Y_i - \bar{Y})^2, \quad (7\text{-}1)$$

where n is the number of experiments in the orthogonal array, Y_i is the mean S/N ratio for the i_{th} experiment, and $\bar{Y}$ is the total mean S/N ratio. SS_T is decomposed of $Seq\ SS$, the sum of squared deviations due to each factor, and sum of error $SS_{e.}$

The percentage contribution (P%) can be calculated as:

$$P\% = Seq\ SS/SS_T. \quad (7\text{-}2)$$

The degree of freedom (DF) implies the amount of information in the data. For each factor, DF = number of levels -1; the total DF = number of experiments -1.

$$Adj\ MS = Seq\ SS/DF. \tag{7-3}$$

The data are checked with an F test to assess the model variance to error variance. If the F value is smaller than 1, the corresponding factor may have less significance on the response.

Source		DF	Seq SS	Adj MS	F	P%
A	Seating capacity	2	5.8389	2.9194	8.89	12.33
B	Configuration	2	6.9945	3.4973	10.65	14.77
C	No. of stops	2	1.7386	0.8693	2.65	3.67
D	Scheduled dwell times	2	0.1717	0.0859	0.26	0.36
E	Sectional Running Time	2	13.9886	6.9943	21.30	29.54
F	Level of Heterogeneity	2	0.45	0.225	0.69	0.95
G	No. of services	2	13.8006	6.9003	21.02	29.14
H	Speed limits	2	0.0155	0.0078	0.02	0.03
J	Driving style	2	1.4473	0.7236	2.20	3.06
K	Minimum dwell times	2	0.9467	0.4734	1.44	2.00
Residual Error		6	1.97	0.3283		4.16
Total		26	47.3625			100

Table 7-7: Analysis of variance for S/N ratios

From the ANOVA results (Table 7-7), Sectional Running Time (E) and No. of services (G) contribute most to the response. Rolling stock characteristics Configuration (B) and Seating capacity (A) have a medium level of influence. Changing No. of stops (C), Driving style (J) and Minimum dwell times (K) would also make a difference to the overall performance.

From the percentage contribution, the least significant factors are F (Level of Heterogeneity), D (Scheduled dwell times), and H (Speed limits). They also have low F values, which means the factors with the defined levels do not show much influence on performance.

There is a slight difference in the factor ranks from the response table and the ANOVA analysis. Sectional Running Time (E) and No. of services (G) are ranked differently in these two analyses. This suggests that both factors have a similar significance on the response.

For individual KPIs the percentage contributions of factors are shown in Figure 7-11. Details can be found in Table 7-8.

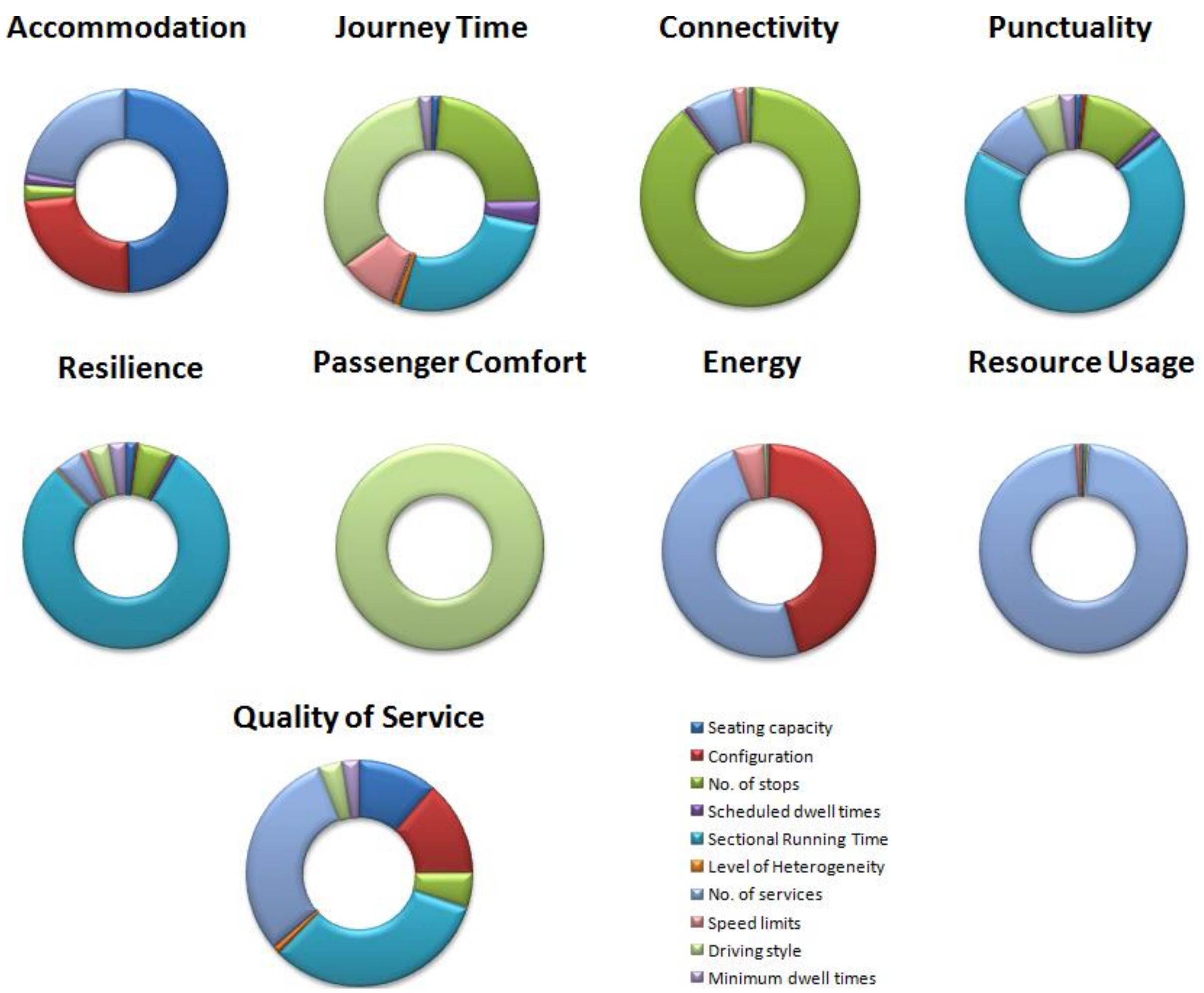


Figure 7-11: Percentage contributions of factors on KPIs and Quality of Service

	AC	JT	CN	PT	RS	PC	EN	RU
A	49.513%	0.948%	0.307%	0.983%	1.625%	0%	0.135%	0.264%
B	23.828%	0.160%	0.119%	0.631%	0.353%	0%	45.196%	0.089%
C	2.604%	18.603%	85.362%	10.515%	5.162%	0%	0.088%	0.331%
D	1.795%	2.987%	0.592%	1.532%	0.806%	0%	0.035%	0.039%
E	0.011%	21.124%	0.063%	64.529%	76.449%	0%	0.052%	0.043%
F	0.002%	0.852%	0.234%	0.108%	0.443%	0%	0.026%	0.036%
G	22.033%	0.149%	6.762%	8.622%	4.150%	0%	48.863%	97.640%
H	0.051%	7.344%	1.773%	0.006%	1.193%	0%	4.486%	0.824%
J	0.041%	26.374%	0.565%	5.262%	3.132%	100%	0.550%	0.102%
K	0.017%	1.564%	0.051%	2.052%	2.489%	0%	0.298%	0.246%

Table 7-8: Percentage contributions of factors against KPIs

It can be noted that each factor would have a different influence on individual KPIs. For instance, Sectional Running Time (E) has great significance on Journey Time, Punctuality and Resilience, yet contributes only a little to several other KPIs. Passenger Comfort here is purely determined by Driving Style (J), as it is the only factor considered in measurement in this study. For each KPI, different factors also take different portions. Factors with great importance on one or more KPIs (e.g. Driving Style) can sometimes have a small influence to the overall QoS. Factors D (Scheduled dwell times), F (Level of Heterogeneity), H (Speed limits) and Minimum dwell times (K) have little influence on all KPIs, which also makes them less influential for the overall performance.

7.5.3 Confirmation test

Once the optimal combination of factors is selected, a confirmation test is requested to verify the improvements in quality characteristics based on the prediction. According to the response table, to get the response with the optimal S/N ratio, factor levels should be set at A2B2C1D1E1F1G2H2J1K1 (refer to Table 7-1). Compared to the current condition (Experiment 1), the factors improved are Seating Capacity, Configuration, Heterogeneity, Number of Services and Speed Limits. The test conditions are listed in Table 7-9:

Seating capacity	Configuration	No. of stops	Scheduled dwell times	Sectional Running Time
120%	Train type 31306: 3 vehicles	original	original	original

Level of Heterogeneity	No. of services	Speed limits	Driving style	Minimum dwell times
original	increased	increased	Default	original

Table 7-9: Predicted optimal factor levels

According to Table 7-9, the optimal S/N ratio is predicted to be -0.3325 dB. This number is attained from a regression analysis to fit the S/N ratio linearly.

Running the experiment with selected control factors and noise factors, the actual S/N ratio is -1.9926 dB. Compared to all other test results listed in Table 7-5, this S/N ratio is still much bigger. The best S/N ratio in the original experiments is -2.3607 dB, from Experiment 22. From the confirmation test, there is an improvement of 0.3681 dB in the response S/N ratio. It can be confirmed that the predicted optimal factor levels would produce the optimal performance.

		LEVEL	S/N (dB)
Initial Factor Levels (best S/N ratio)		A3B2C1D2E1F3G2H2J1K3	-2.3607
Optimal Factor Levels	Prediction	A2B2C1D1E1F1G2H2J1K1	-0.3325
	Experiment	A2B2C1D1E1F1G2H2J1K1	-1.9926

Table 7-10: Confirmation test results: overall S/N ratio

Breaking down to individual KPIs (Table 7-11), Accommodation and Resource Usage are improved most. There are also improvements in Journey Time, Connectivity and Energy Usage. Passenger Comfort remains the same, while Punctuality and Resilience are slightly reduced.

	Initial (dB)	Optimal (dB)	Improvement (dB)
AC	-3.7813	-1.8071	1.9742
JT	0	-0.2424	-0.2424
CN	-1.9500	-2.2457	-0.2958
PT	-6.2485	-5.1314	1.1171
RS	-3.8178	-3.8412	-0.0233
PC	-1.0231	-1.0231	0
EN	-0.9361	-0.5816	0.3545
RU	-0.1537	-0.6320	-0.4783

Table 7-11: Confirmation test results: KPI S/N ratios

7.6 Summary

This study investigated the factor significance of the performance evaluation process with the Taguchi Design of Experiment method. Ten 3-level control factors and two 2-level noise factors were selected. Based on an orthogonal array, 27 experiments were performed, showing the efficiency of the method. The outputs were evaluated for their QoS and KPI values. Based on experiments, the results are summarised as follows:

1. Based on signal-to-noise ratios, the factors ranked by significance are: number of services, sectional running time, configuration, seating capacity, number of stops, driving style, minimum dwell times, level of heterogeneity, scheduled dwell times and speed limits.

2. From the ANOVA analysis, four of the factors have shown large percentage contributions (>10%) on QoS: sectional running times, number of services, configuration and seating capacity. Level of heterogeneity, scheduled dwell times, and speed limits show little influence on the overall performance, which means modifying these factors is not effective or economical.

3. The significance of factors varies a lot for each KPI. Some factors contribute a lot to several KPIs, while on other KPIs the influence is minimal.

4. The confirmation test yields a better performance than the original 27 experiments, with the predicted optimal factor combination.

It is to be noted that the Taguchi method is only capable of searching for the optimal solution within the predefined factor levels.

The performance has been improved with a relatively small number of test runs, suggesting that the Taguchi method is both effective and efficient in identifying the importance of influencing factors to support performance improvement.

Chapter 8: Conclusions and discussions

8.1 Conclusions

In the railway industry, performance evaluation has long been a difficult subject. This is because of the complexity of railway systems, and the growing gap between traffic demand and the network capability. The passengers, government, infrastructure manager and train operating companies all require that the current performance be improved. Measuring and improving the performance of railway networks has thus become a main theme in railway planning and management.

In this thesis, the "Quality of Service" concept is adopted to develop a process for the performance evaluation and improvement based on a novel framework. Quality of Service is a measure of how well the service level delivered matches customer expectations. An analysis has been performed to assess the impact of the influencing factors on the KPIs and the overall QoS performance.

From the QoS framework, the key performance indicators for performance evaluation are defined as:

- Accommodation;
- Journey Time;
- Connectivity;
- Punctuality;
- Resilience;

- Passenger Comfort;
- Energy;
- Resource Usage.

The influencing factors which may contribute to performance changes are summarised in two categories:

- Strategic factors:
    - Rolling Stock;
    - Infrastructure;
    - Timetable;
    - Operational Rules;
- Tactical factors:
    - Traffic Management;
    - Operational Management;
    - Human Factors;
    - System Maintenance;
    - Environmental Factors.

The whole process of the QoS performance improvement approach can be summarised as below:

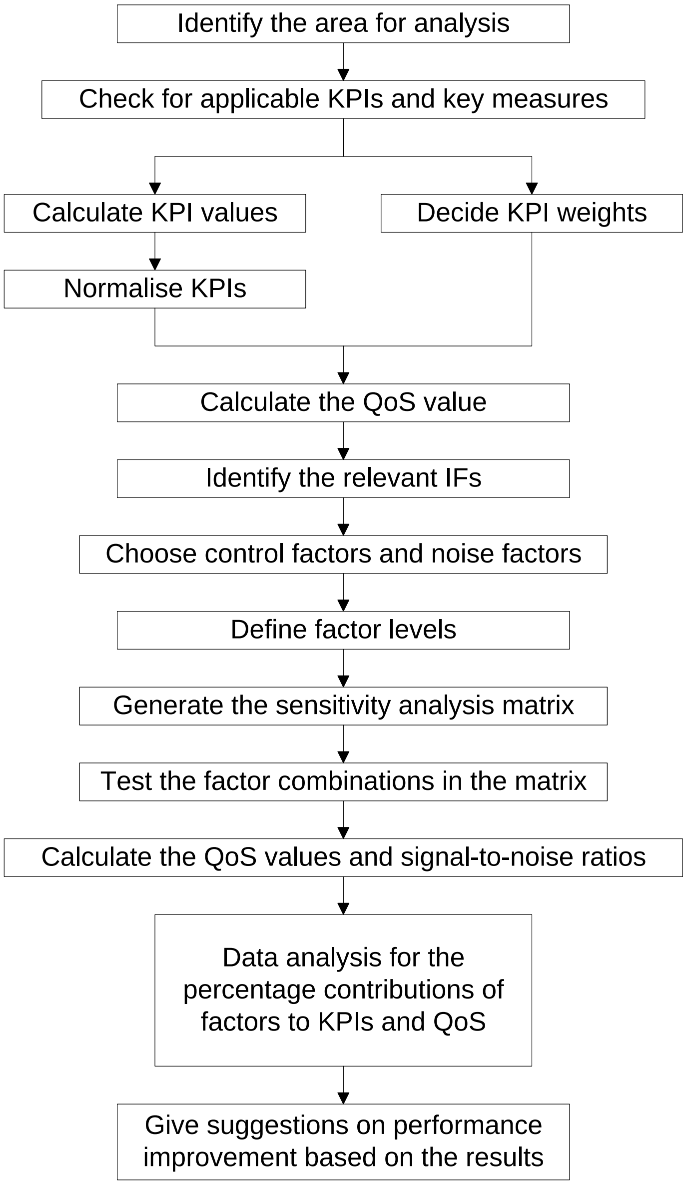


Figure 8-1: The QoS approach for performance improvement

With different characteristics of each network, it is necessary to identify the entities to be used in the QoS framework, specific to the scenario. In practice, the KPI weights vary a lot from scenario to scenario, and sometimes a number of KPIs may even be neglected in the evaluation. The choice of KPIs and their weights is relevant to the purpose of the project and the stakeholders involved. For example, if the infrastructure manager/timetable planner wants to investigate the influence of a timetable change, they might not be interested in passenger comfort or energy usage. On the other hand, if the project is aimed at attracting more passengers to use the network, passenger comfort would be valued a lot, while energy and resource usage might be given very low priority. The key measures used to quantify each KPI may also be modified to suit the scenario.

The QoS value is generated from the KPI weights and KPI values. These results can be applied for further analysis to provide decision support for performance improvement.

To establish the relationship between influencing factors and KPIs, the railway network can be treated as a black box, and by feeding enough data in, it is possible to get sufficient outputs that can be used to develop a mathematical expression of the relationship. This relationship can be expressed as a function of the outputs from the inputs. The input data would be combinations of influencing factors at different levels, and the outputs are the corresponding KPI and QoS values.

In theory, the more inputs applied and tested, the more accurate the function would be. To get the most accurate result, the influencing factors are changed one at a time. This would result in a large number of tests. However, in practice, only a limited number of inputs can be tested. This is due to the high cost of making modifications in the simulation. Some IFs are difficult to change in simulation software, such as infrastructure factors. Changes in IFs such as timetable patterns and conflict resolution may become time consuming when they are applied to a large portion of the simulation area. Besides, if there is a lot of influencing factors in the system, the number of testing scenarios would go up exponentially. It is often not feasible to put all these scenarios into simulation in a limited time span. It is necessary to apply Design of Experiment techniques to control the total number of tests. In this research the Taguchi method is used.

With the Taguchi analysis, it is possible to set a best level of QoS, against which performance could be assessed. Different levels of the influencing factors are tested through simulation, and the significance of each IF is produced in this process. The levels are selected with consideration of their feasibility in the studied scenario. The result from the analysis is the IFs are ranked by their percentage contribution to the KPIs and to the overall QoS performance. From these values we can get a best combination of influencing factor values that would produce the optimal performance. For a proposed performance improvement plan, it is also possible to make a prediction of the KPI and QoS results.

8.2 Discussions

When applying the quality of service approach in practice, a lot of pilot work is needed. Before quantification of performance, the KPIs would have to be chosen. In this thesis, eight KPIs are assessed. However, in a real life application, some other KPIs may also be appropriate to the scenario, including the ones that have been excluded in this study, such as safety. The decision maker has to take extra care in the selection and exclusion of KPIs.

According to different performance evaluation and improvement scenarios, the stakeholder panel would need to be widened. Apart from infrastructure managers, timetable planners, government, passengers and operators, other stakeholders may be relevant as well, such as the supply chain, funding agencies, etc. For the decision of KPI weights in focus group studies, it is necessary to include a full and representative range of stakeholder types.

During each focus group study, the terms used to describe KPIs need to be revised so that they can be understood by ordinary rail users.

Based on the findings of the study, it is reasonable to accept the research hypothesis that railway performance can be evaluated with a number of technical and operational parameters. The Quality of Service approach can be adopted for the rail industry to evaluate short, medium and long term performance improvement strategies. The process of short to medium term performance improvement has been discussed in the case study in this thesis. For medium to long term improvement plans in the system, such as infrastructure upgrades, it is to be noted that a number of other influencing factors would also be changed accordingly. The correlation of IFs need to be taken into consideration, which means a lot more test runs are required to reveal the relationship between IFs and performance outputs.

The results of the QoS analysis can be validated by testing different combinations of IFs in the simulation environment, and checking the KPI outputs. These outputs can be forwarded to the stakeholder panel to be evaluated. The feedback from the stakeholder panel would help to improve existing system upgrade plans.

8.3 Future work

8.3.1 Application of the QoS framework in the ON-TIME project

The research of quality of service has been influenced and developed by a group of infrastructure managers, operators, suppliers and academics working within the EC FP7 ON-TIME project. The QoS framework has been successfully applied in the FP7 ON-TIME (Optimal Networks for Train Integration Management across Europe) project for defining work and benchmarking outputs. This project aims at improving railway customer satisfaction through increased capacity and decreased delays. A case study has been carried out based on a Dutch passenger network to validate the processes developed in the working packages to reduce delays and improve performance (Nicholson and Roberts, 2014). The KPIs used for benchmarking purposes are Accommodation, Journey Time, Connectivity, Punctuality, Resilience, and Resource Usage. The outputs from the project show the possibility of applying the Quality of Service framework to better understand, manage and optimise railway operations to achieve better performance, and facilitate railway planning and real-time traffic management.

8.3.2 Evaluation of freight traffic

In European railways, the majority of the traffic comes from passenger services. Most railways carry passenger traffic, or mixed passenger/ freight traffic, while some lines run freight services only. In this thesis the focus has been put on passenger traffic, and if freight traffic is to be discussed, the KPI weights and key measures also need to be redefined. The reason for these changes is that the freight traffic is planned and managed differently from passenger traffic, with loose timetables and lower priorities compared to passenger services. In future research, it would be worthwhile investigating the interaction between passenger and freight traffic, and to develop unique KPIs and IFs for freight services.

8.3.3 Study about traffic demand

The evaluation and forecasting of traffic demand is an important theme in railway strategic planning. Apart from economic and political concerns, the current performance is also a main consideration. By checking how well the existing demand is accommodated with current performance, the forecast of future demand can be adjusted. On the other hand, demand evaluation and forecasting also helps in adapting performance improvement plans to suit future traffic. By studying the interaction between performance and traffic demand, more efficient and accurate strategic planning can be expected.

References

ABBINK, E., VAN DEN BERG, B., KROON, L. & SALOMON, M. 2004. Allocation of railway rolling stock for passenger trains. *Transportation Science,* 38**,** 33-41.

ABRIL, M., BARBER, F., INGOLOTTI, L., SALIDO, M. A., TORMOS, P. & LOVA, A. 2008. An assessment of railway capacity. *Transportation Research Part E: Logistics and Transportation Review,* 44**,** 774-806.

AEA TECHNOLOGY 2005. Capacity utilisation indices methodology.

AEATECHNOLOGY 2005. Capacity Utilisation Indices Methodology.

ALBRECHT, I. T., BRUNGER, O., DAHLHAUS, E. & GOVERDE, R. 2008. *Railway timetable & traffic: analysis, modelling, simulation*, Eurailpress.

AN, M., CHEN, Y. & BAKER, C. J. 2011. A fuzzy reasoning and fuzzy-analytical hierarchy process based approach to the process of railway risk information: A railway risk management system. *Information Sciences,* 181**,** 3946-3966.

ANANDA, J. & HERATH, G. 2005. Evaluating public risk preferences in forest land-use choices using multi-attribute utility theory. *Ecological Economics,* 55**,** 408-419.

ANTHONY, G., STEVEN, W. & MICHAEL, E. 2003. Overview of modern design of experiments methods for computational simulations *41st Aerospace Sciences Meeting and Exhibit.* American Institute of Aeronautics and Astronautics.

ANTONY, J. 2014. *Design of experiments for engineers and scientists*, Elsevier.

ARMSTRONG, J. & PRESTON, J. M. 2009. Automating the production of train graphs and the calculation of CUI values. *3rd International Seminar on Railway Operations Modelling and Analysis, Engineering and Optimisation Approaches.*

ATOC. 2012. *Capacity - ATOC's view* [Online]. ATOC. Available: http://www.atoc.org/clientfiles/File/capacity2012.pdf [Accessed 17/06 2012].

ATOC 2013. National rail network map.

ATOC. 2016. *Passenger Demand Forecasting Council* [Online]. Available: http://www.atoc.org/about-atoc/commercial-activities/passenger-demand-forecasting-council/ [Accessed 4th March 2016].

AZADEH, A., GHADERI, S. F. & IZADBAKHSH, H. 2008. Integration of DEA and AHP with computer simulation for railway system improvement and optimization. *Applied Mathematics and Computation,* 195**,** 775-785.

BARBER, F., ABRIL, M., SALIDO, M.A., INGOLOTTI, L., TORMOS, P., LOVA, A. 2007. Survey of automated Systems for Railway Management. Technical University of Valencia.

BARRON, F. H. & BARRETT, B. E. 1996. Decision Quality Using Ranked Attribute Weights. *Management Science,* 42**,** 1515-1523.

BEHN, R. D. 2003. Why measure performance? Different purposes require different measures. *Public administration review,* 63**,** 586-606.

BELL, M. L., HOBBS, B. F., ELLIOTT, E. M., ELLIS, H. & ROBINSON, Z. 2001. An evaluation of multi-criteria methods in integrated assessment of climate policy. *Journal of Multi-Criteria Decision Analysis,* 10**,** 229-256.

BURDETT, R. L. & KOZAN, E. 2006. Techniques for absolute capacity determination in railways. *Transportation Research Part B-Methodological,* 40**,** 616-632.

BURDETT, R. L. & KOZAN, E. 2009. Techniques for inserting additional trains into existing timetables. *Transportation Research Part B: Methodological,* 43**,** 821-836.

BURDETT, R. L. & KOZAN, E. 2010. A disjunctive graph model and framework for constructing new train schedules. *European Journal of Operational Research,* 200**,** 85-98.

BUREIKA, G. 2011. Multicriteria evaluation of operational effectiveness of freight diesel locomotives on Lithuanian railways. *Transport,* 26**,** 61-68.

CAVANA, R. Y., CORBETT, L. M. & LO, Y. 2007. Developing zones of tolerance for managing passenger rail service quality. *International Journal of Quality & Reliability Management,* 24**,** 7-31.
CHAPIN, A. & SCHMIDT, S. 1999. Do mergers improve efficiency? Evidence from deregulated rail freight. *Journal of Transport Economics and Policy***,** 147-162.
CHEN, L., ZHANG, Z., GONG, W. & LIANG, Z. 2015. Quantifying the effects of fuel compositions on GDI-derived particle emissions using the optimal mixture design of experiments. *Fuel,* 154**,** 252-260.
CHIOU, Y.-C., LAN, L. W. & YEN, B. T. H. 2010. A joint measurement of efficiency and effectiveness for non-storable commodities: Integrated data envelopment analysis approaches. *European Journal of Operational Research,* 201**,** 477-489.
CHOO, E. U., SCHONER, B. & WEDLEY, W. C. 1999. Interpretation of criteria weights in multicriteria decision making. *Computers & Industrial Engineering,* 37**,** 527-541.
CIRILLO, C., EBOLI, L. & MAZZULLA, G. 2011. On the asymmetric user perception of transit service quality. *International Journal of Sustainable Transportation,* 5**,** 216-232.
CORONADO, M., SEGAD ES, A. M. & ANDR S, A. 2015. Using mixture design of experiments to assess the environmental impact of clay-based structural ceramics containing foundry wastes. *Journal of Hazardous Materials,* 299**,** 529-539.
COSTA, J. P. & CL MACO, J. C. 1999. Relating reference points and weights in MOLP. *Journal of Multi-Criteria Decision Analysis,* 8**,** 281-290.
D'ARIANO, A., PRANZO, M. & HANSEN, I. A. 2007. Conflict resolution and train speed coordination for solving real-time timetable perturbations. *Intelligent Transportation Systems, IEEE Transactions on,* 8**,** 208-222.
D'AMBROSIO, S. & FERRARI, A. 2015. Potential of multiple injection strategies implementing the after shot and optimized with the design of experiments procedure to improve diesel engine emissions and performance. *Applied Energy,* 155**,** 933-946.
DE ONA, J., DE O A, R., EBOLI, L. & MAZZULLA, G. 2015. Heterogeneity in perceptions of service quality among groups of railway passengers. *International Journal of Sustainable Transportation,* 9**,** 612-626.
DEJAEGHER, B. & VANDER HEYDEN, Y. 2011. Experimental designs and their recent advances in set-up, data interpretation, and analytical applications. *Journal of Pharmaceutical and Biomedical Analysis,* 56**,** 141-158.
DELL'OLIO, L., IBEAS, A. & CECIN, P. 2011. The quality of service desired by public transport users. *Transport Policy,* 18**,** 217-227.
DEPARTMENT FOR TRANSPORT 2011. Transport Statistics Great Britain: 2011. 15 December 2011 ed.: Department for Transport.
DEPARTMENT FOR TRANSPORT. 2012. *Department for Transport input and impact indicators* [Online]. Department for Transport. [Accessed 10th Feburary 2016].
DEPARTMENT FOR TRANSPORT. 2016a. *Aims of rail franchising and what we expect from our delivery partners*

[Online]. Available: https://www.gov.uk/government/publications/rail-franchising-aims-and-expectations/aims-of-rail-franchising-and-what-we-expect-from-our-delivery-partners [Accessed 02/03 2017].
DEPARTMENT FOR TRANSPORT 2016b. Transport Statistics Great Britain 2015.
DONG, Y., XU, Y., LI, H. & DAI, M. 2008. A comparative study of the numerical scales and the prioritization methods in AHP. *European Journal of Operational Research,* 186**,** 229-242.
EDWARDS, B. 2013. *The modern station: new approaches to railway architecture*, Taylor & Francis.
ELLIOTT, M. A. 2010. Selecting numerical scales for pairwise comparisons. *Reliability Engineering & System Safety,* 95**,** 750-763.
EQUIP 2000. Extending the quality of public transport - Final Report. *4th RTD Framework Program.* Brussels.

EUROPEAN COMMITTEE FOR STANDARDIZATION 2002. Transportation – logistics and services – public passenger transport – service quality definition, targeting and measurement.
FIELDING, G. J., BABITSKY, T. T. & BRENNER, M. E. 1985. Performance evaluation for bus transit. *Transportation Research Part A: General,* 19**,** 73-82.
FIGUEIRA, J., GRECO, S. & EHRGOTT, M. 2005. *Multiple Criteria Decision Analysis: State of the Art Surveys,* Boston, USA, Springer Science + Business Media, Inc.
FINAN, J. S. & HURLEY, W. J. 1999. Transitive calibration of the AHP verbal scale. *European Journal of Operational Research,* 112**,** 367-372.
FIRST CAPITAL CONNECT 11 December 2011 to 8 December 2012. Train tmes.
FRANCO-SANTOS, M. & BOURNE, M. 2005. An examination of the literature relating to issues affecting how companies manage through measures. *Production Planning & Control,* 16**,** 114-124.
FRELING, R., LENTINK, R. M. & WAGELMANS, A. P. M. 2004. A decision support system for crew planning in passenger transportation using a flexible branch-and-price algorithm. *Annals of Operations Research,* 127**,** 203-222.
GIBSON, S., COOPER, G. & BALL, B. 2002. Developments in transport policy - the evolution of capacity charges on the UK rail network. *Journal of Transport Economics and Policy,* 36**,** 341-354.
GOLOB, T. F. & HENSHER, D. A. 2007. The trip chaining activity of Sydney residents: A cross-section assessment by age group with a focus on seniors. *Journal of Transport Geography,* 15**,** 298-312.
GONZ LEZ-GIL, A., PALACIN, R., BATTY, P. & POWELL, J. P. 2014. A systems approach to reduce urban rail energy consumption. *Energy Conversion and Management,* 80**,** 509-524.
GOVERDE, R. M. P. 2005. *Punctuality of railway operations and timetable stability analysis.*
GOVERDE, R. M. P. 2007. Railway timetable stability analysis using max-plus system theory. *Transportation Research Part B: Methodological,* 41**,** 179-201.
GRAFFICA. 2013. *Rail simulation* [Online]. Available: http://www.graffica.co.uk/rail-traffic-management/simulation/ [Accessed 11/05 2013].
GREEN, P. & SRINIVASAN, V. 1990. Conjoint analysis in marketing: New developments with implications for research and practice. *Journal of Marketing,* 54**,** 3-19.
GUINTO, D. R. 2008. *Multicriteria decision-making and the Comprehensive Everglades Restoration Plan.* Ph.D, The Johns Hopkins University.
GUNSELMANN, W. Technologies for increased energy efficiency in railway systems. Power Electronics and Applications, 2005 European Conference on, 2005. IEEE, 10 pp.-P. 10.
HANAFI, R. & KOZAN, E. 2014. A hybrid constructive heuristic and simulated annealing for railway crew scheduling. *Computers & Industrial Engineering,* 70**,** 11-19.
HANSEN, I. A., WIGGENRAAD, P. B. L. & WOLFF, J. W. Benchmark analysis of railway networks and undertakings. RailCopenhagen2013: 5th International Conference on Railway Operations Modelling and Analysis, 13-15 May 2013 2013 Copenhagen, Denmark. International Association of Railway Operations Research (IAROR).
HERRERA, F., HERRERA-VIEDMA, E. & VERDEGAY, J. L. 1996. A model of consensus in group decision making under linguistic assessments. *Fuzzy Sets and Systems,* 78**,** 73-87.
HO, T. K., TSANG, C. W., IP, K. H. & KWAN, K. S. 2012. Train service timetabling in railway open markets by particle swarm optimisation. *Expert Systems with Applications,* 39**,** 861-868.
HOBBS, B. F. 1980. A Comparison of Weighting Methods In Power Plant Siting. *Decision Sciences,* 11**,** 725-737.
HUDOVORNIK, G. & VREČER, F. 2015. Impact of the curing parameters on drug release from Eudragit RS and RL 30D coated pellets: Design of experiments. *Journal of Drug Delivery Science and Technology,* 30, Part A**,** 146-153.
HUISMAN, D., KROON, L. G., LENTINK, R. M. & VROMANS, M. J. C. M. 2005. Operations research in passenger railway transportation. *Statistica Neerlandica,* 59**,** 467-497.

HYD N, L. C. & B LOW, P. H. 2003. Who's talking: drawing conclusions from focus groups—some methodological considerations. *International Journal of Social Research Methodology,* 6**,** 305-321.

IKI, S., HORIGUCHI, I., SHIGEMATSU, M., SATA, T. & SUGIYAMA, K. 2012. Qualitative analysis of the perception and acceptability of pathogen transport among housewives using Focus Group Interviews. *Japanese Journal of Infectious Diseases,* 65**,** 403-409.

ISON, S., FROST, M. & WATSON, R. 2012. UK rail transport: a review of demand and supply. *Proceedings of the Institution of Civil Engineers-Transport,* 165**,** 225-234.

JAMES, T. E. 1991. Performance criteria for EM rail launchers with solid or transition armatures and laminated rails. *Ieee Transactions on Magnetics,* 27**,** 482-487.

JI, P. & JIANG, R. 2003. Scale transitivity in the AHP. *The Journal of the Operational Research Society,* 54**,** 896-905.

JIA, J., FISCHER, G. W. & DYER, J. S. 1998. Attribute weighting methods and decision quality in the presence of response error: a simulation study. *Journal of Behavioral Decision Making,* 11**,** 85-105.

KARLAFTIS, M. G. 2004. A DEA approach for evaluating the efficiency and effectiveness of urban transit systems. *European Journal of Operational Research,* 152**,** 354-364.

KEENEY, R. L. & RAIFFA, H. 1994. *Decisions with multiple objectives–preferences and value tradeoffs*, John Wiley & Sons, Ltd.

KERSTENS, K. 1996. Technical efficiency measurement and explanation of French urban transit companies. *Transportation Research Part a-Policy and Practice,* 30**,** 431-452.

KIRKWOOD, D. 2014. BRAVE user guide. University of Birmingham.

KLIEWER, N. & SUHL, L. 2011. A note on the online nature of the railway delay management problem. *Networks,* 57**,** 28-37.

KONTAXI, E. & RICCIA, S. 2012. Railway Capacity Handbook: A Systematic Approach to Methodologies. *Procedia - Social and Behavioral Sciences,* 48**,** 2689-2696.

KRUEGER, H. Parametric modeling in rail capacity planning. Simulation Conference Proceedings, 1999 Winter, 1999 1999. 1194-1200 vol.2.

KUHN, M. 1998. *Automatic route setting integrated in a modern traffic management system,* London.

KWAN, R. S. 2011. Case studies of successful train crew scheduling optimisation. *Journal of Scheduling,* 14**,** 423-434.

LAN, L. W. & LIN, E. T. J. 2005. Measuring railway performance with adjustment of environmental effects, data noise and slacks. *Transportmetrica,* 1**,** 161-189.

LAN, L. W. & LIN, E. T. J. 2006. Performance measurement for railway transport: Stochastic distance functions with inefficiency and ineffectiveness effects. *Journal of Transport Economics and Policy,* 40**,** 383-408.

LANDEX, A., KAAS, AH, SCHITTENHELM, B & SCHNEIDER-TILLI, J 2006. Evaluation of railway capacity. *Proceedings of Trafficdays*.

LEMONAKIS, N., SKALTSOUNIS, A.-L., TSARBOPOULOS, A. & GIKAS, E. 2016. Optimization of parameters affecting signal intensity in an LTQ-orbitrap in negative ion mode: A design of experiments approach. *Talanta,* 147**,** 402-409.

LEWIS, R. C. & BOOMS, B. The marketing aspects of service quality. AMA Proceeding, 1983 American Marketing Association Chicago. 99-104.

LONDON MIDLAND. 2016. *Performance* [Online]. Available: http://www.londonmidland.com/about-us/company/performance/ [Accessed 12/02 2016].

LU, M., NICHOLSON, G., SCHMID, F., DAI, L., CHEN, L. & ROBERTS, C. 2013. A framework for the evaluation of the performance of railway networks. *International Journal of Railway Technology* 2**,** 79-96.

LUCCHINI, L., CURCHOD, A. & RIVIER, R. 2001. Transalpine rail network: a capacity assessment model (CAPRES). *First Swiss Transport Research Conference.*

LUNDBERG, K., BALFORS, B. & FOLKESON, L. 2009. Framework for environmental performance measurement in a Swedish public sector organization. *Journal of Cleaner Production,* 17**,** 1017-1024.

MAHESHWARI, V., RANGAIAH, G. P., LAU, T. & SAMAVEDHAM, L. 2015. Application of design of experiments in hemodialysis: Optimal sampling protocol for β2-microglobulin kinetic model. *Chemical Engineering Science,* 131**,** 84-90.
MARWICK, P. 1977. *Impact of Freight/passenger Interference on Capacity And/or Cost,* Montreal.
MASKELIUNAITE, L. & SIVILEVICIUS, H. 2012. Expert evaluation of criteria describing the quality of travelling by international passenger train: technological, economic and safety perspectives. *Technological and Economic Development of Economy,* 18**,** 544-566.
MASSEY, O. T. 2011. A proposed model for the analysis and interpretation of focus groups in evaluation research. *Evaluation and Program Planning,* 34**,** 21-28.
MATEUS, R., FERREIRA, J. A. & CARREIRA, J. 2008. Multicriteria decision analysis (MCDA): Central Porto high-speed railway station. *European Journal of Operational Research,* 187**,** 1-18.
MINGOZZI, A., BOSCHETTI, M. A., RICCIARDELLI, S. & BIANCO, L. 1999. A set partitioning approach to the crew scheduling problem. *Operations Research,* 47**,** 873-888.
MIYATAKE, M. & KO, H. 2010. Optimization of train speed profile for minimum energy consumption. *IEEJ Transactions on Electrical and Electronic Engineering,* 5**,** 263-269.
MOHAJERI, N. & AMIN, G. R. 2010. Railway station site selection using analytical hierarchy process and data envelopment analysis. *Computers & Industrial Engineering,* 59**,** 107-114.
MOREIRA, N., GARCIA, L. & CATARRINHO, P. 2004. Network capacity. *Computers in Railway Six.*
NASH, A. & HUERLIMANN, D. 2004. Railroad simulation using OpenTrack. *In:* ALLAN, J., BREBBIA, C. A., HILL, R. J., SCIUTTO, G. & SONE, S. (eds.) *Computers in Railway Six.* Southampton: Wit Press.
NEELY, A. 2005. The evolution of performance measurement research: Developments in the last decade and a research agenda for the next. *International Journal of Operations & Production Management,* 25**,** 1264-1277.
NETWORK RAIL 2008a. East Coast Main Line route utilisation strategy.
NETWORK RAIL 2008b. Meeting the capacity challenge: the case for new lines. Network Rail.
NETWORK RAIL 2009. Route Plans 2009: Route 8 - East Coast Main Line. Network Rail.
NETWORK RAIL 2010. East Coast Main Line 2016 capacity review.
NETWORK RAIL 2012a. Our railway's future.
NETWORK RAIL. 2012b. *Planning and operating the railway* [Online]. Network Rail,. Available: http://www.networkrail.co.uk/structure-and-governance/planning-and-operating-the-railway/ [Accessed 3rd March 2016].
NETWORK RAIL 2014a. Anglia: route study draft for consultation. *Long Term Planning Process.*
NETWORK RAIL 2014b. How we performed. *Anual Reports and Accounts 2014.*
NETWORK RAIL 2015a. Calendar of milestone dates December 2016 and May 2017 timetables (production schedule). *In:* ALLEN, M. (ed.) *Capacity Planning.* Milton Keynes: Network Rail Operational Planning.
NETWORK RAIL 2015b. The Network Code.
NETWORK RAIL 2015c. The what, why and how of the GSM-R system.
NETWORK RAIL. 2016. *Working timetables* [Online]. Available: https://www.networkrail.co.uk/browseDirectory.aspx?dir=\Timetables\Working%20Timetable%20(WTT)&root=\Timetables\Working%20Timetable%20(WTT) [Accessed 26th February 2016].
NETWORK RAIL. 2017. *Our Vision* [Online]. Available: http://archive.nr.co.uk/aspx/122.aspx [Accessed 02/03 2017].
NETWORK RAIL, ATOC LTD & RAILWAY FREIGHT OPERATORS' ASSOCIATION 2014. Planning ahead Control Period 5 and beyond.
NICHOLSON, G. & ROBERTS, C. 2014. Analysis of the benchmarking. *ON-TIME Deliverable 5.3 Report.* Birmingham: University of Birmingham.
NIJKAMP, P. & DELFT, A. V. 1977. *Multi-Criteria Analysis and Regional Decision-Making*, Springer US.

NUZZOLO, A., RUSSO, F. & CRISALLI, U. 2001. A Doubly Dynamic Schedule-based Assignment Model for Transit Networks. *Transportation Science,* 35**,** 268-285.
NYSTROM, B. & SODERHOLM, P. 2010. Selection of maintenance actions using the analytic hierarchy process (AHP): decision-making in railway infrastructure. *Structure and Infrastructure Engineering,* 6**,** 467-479.
OFFICE FOR OFFICIAL PUBLICATIONS OF THE EUROPEAN COMMUNITIES 2008. Modern rail modern Europe.
OFFICE OF RAIL AND AND ROAD. 2017. *Train Protection* [Online]. Available: http://orr.gov.uk/what-and-how-we-regulate/health-and-safety/guidance-and-research/infrastructure-safety/train-protection [Accessed 13/02 2017].
OFFICE OF RAIL AND ROAD 2015. Network Rail Monitor Quarters 3-4 of Year 1 of CP5 12 October 2014 to 31 March 2015.
ON-TIME 2012a. Deliverable 1.2: A framework for developing an objective function for evaluating work package solutions (Cost function).
ON-TIME 2012b. Principles, definitions and requirements. University of Birmingham.
ON-TIME CONSORTIUM. 2012. *ONTIME description of work* [Online]. Available: http://www.ontime-project.eu/home.aspx [Accessed].
OUM, T. H. & YU, C. Y. 1994. Economic-efficiency of railways and implications for public-policy - a comparative-study of the OECD countries railways. *Journal of Transport Economics and Policy,* 28**,** 121-138.
PACHL, J. 2002. *Railway operation and control*, VTD Rail Pub.
PACHL, J., HANSEN, I. A. & ALBRECHT, T. 2008. *Railway timetable and traffic: analysis, modelling, simulation*, Eurailpress.
PARKINSON, T. & FISHER, I. 1996. *Rail Transit Capacity*, Transportation Research Board, National Research Council.
PARMENTER, D. 2015. *Key performance indicators: developing, implementing, and using winning KPIs*, John Wiley & Sons.
PAULLEY, N., BALCOMBE, R., MACKETT, R., TITHERIDGE, H., PRESTON, J., WARDMAN, M., SHIRES, J. & WHITE, P. 2006. The demand for public transport: The effects of fares, quality of service, income and car ownership. *Transport Policy,* 13**,** 295-306.
PEARSON, J. T., GOODALL, R. M. & PRATT, I. 1998. Control system studies of an active anti-roll bar tilt system for railway vehicles. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit,* 212**,** 43-60.
PEETERS, M. & KROON, L. 2008. Circulation of railway rolling stock: a branch-and-price approach. *Computers & Operations Research,* 35**,** 538-556.
POISTER, T. H. 2008. *Measuring performance in public and nonprofit organizations*, John Wiley & Sons.
ROBERTS, C., CHEN, L., BOUCH, C., LU, M., DAI, L., NICHOLSON, G., SCHMID, F., PURCELL, M., ACHILLE, V. & WAHLBORG, M. 2012a. Agreed Principles, Definitions and Requirements. *ON-TIME Deliverable 1.1 Report.* Birmingham: University of Birmingham.
ROBERTS, C., CHEN, L., LU, M., DAI, L., BOUCH, C., NICHOLSON, G. & SCHMID, F. 2012b. A framework for developing an objective function for evaluating work package solutions. *ON-TIME Deliverable 1.2 Report.* Birmingham: University of Birmingham.
ROBERTS, C., SCHMID, F., CONNOR, P., RAMDAS, V. & SHARPE, J. 2010. A new railway system capacity model.
ROBERTS, C., SCHMID, F., CONNOR, P., RAMDAS, V. & SHARPE, J. 2011. A new railway system capacity model.
ROM N, C., MART N, J. C. & ESPINO, R. 2014. Using stated preferences to analyze the service quality of public transport. *International Journal of Sustainable Transportation,* 8**,** 28-46.
ROSANTY, E. S., DAHLAN, H. M. & HUSSIN, A. R. C. Multi-Criteria Decision Making for Group Decision Support System. Information Retrieval & Knowledge Management (CAMP), 2012 International Conference on, 13-15 March 2012 2012. 105-109.

ROSS, P. J. 1996. *Taguchi techniques for quality engineering: loss function, orthogonal experiments, parameter and tolerance design*, McGraw-Hill.
ROY, B. 2005. Paradigms and Challenges. *In:* FIGUEIRA, J., GRECO, S. & EHRGOTT, M. (eds.) *Multiple Criteria Decision Analysis: State of the Art Surveys.* Boston, USA: Springer Science + Business Media, Inc.
SAATY, T. 2008. Relative measurement and its generalization in decision making why pairwise comparisons are central in mathematics for the measurement of intangible factors the analytic hierarchy/network process. *RACSAM - Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas,* 102**,** 251-318.
SAATY, T. L. 1980. *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation,* New York, McGraw Hill.
SAATY, T. L. 1990. How to make a decision: The analytic hierarchy process. *European Journal of Operational Research,* 48**,** 9-26.
SAATY, T. L. & BENNETT, J. P. 1977. Theory of analytical hierarchies applied to political candidacy. *Behavioral Science,* 22**,** 237-245.
SAYERS, T. M., JESSOP, A. T. & HILLS, P. J. 2003. Multi-criteria evaluation of transport options—flexible, transparent and user-friendly? *Transport Policy,* 10**,** 95-105.
SCHITTENHELM, B. & LANDEX, A. Development and application of Danish Key Performance Indicators for Railway Timetables. RailCopenhagen2013: 5th International Conference on Railway Operations Modelling and Analysis, 13-15 May 2013 2013 Copenhagen, Denmark. International Association of Railway Operations Research (IAROR).
SCHONER, B. & WEDLEY, W. 1989. Alternative scales in AHP. *In:* LOCKETT, A. G. & ISLEI, G. (eds.) *Improving Decision Making in Organisations.* Springer Berlin Heidelberg.
SIMONS, D., CLARYS, P., DE BOURDEAUDHUIJ, I., DE GEUS, B., VANDELANOTTE, C. & DEFORCHE, B. 2014. Why do young adults choose different transport modes? A focus group study. *Transport Policy,* 36**,** 151-159.
SIVILEVICIUS, H. & MASKELIUNAITE, L. 2010. The criteria for identifying the quality of passengers' transportation by railway and their ranking using AHP method. *Transport,* 25**,** 368-381.
SLANZI, D., DE LUCREZIA, D. & POLI, I. 2015. Querying Bayesian networks to design experiments with application to 1AGY serine esterase protein engineering. *Chemometrics and Intelligent Laboratory Systems,* 149, Part A**,** 28-38.
STATHOPOULOS, A. & MARCUCCI, E. 2014. De gustibus eisputandum est: Non-linearity in public transportation service quality evaluation. *International Journal of Sustainable Transportation,* 8**,** 47-68.
STEER DAVIES GLEAVE 2015. Estimates of station usage 2014-15 data.
STEWART, T. J. 1992. A critical survey on the status of multiple criteria decision making theory and practice. *Omega,* 20**,** 569-586.
STRATEGIC RAIL AUTHORITY 2003. Capacity utilisation policy: network utilisation strategy.
T RNQUIST, J. 2007. Railway traffic disturbance management—An experimental analysis of disturbance complexity, management objectives and limitations in planning horizon. *Transportation Research Part A: Policy and Practice,* 41**,** 249-266.
TAGUCHI, G. 2005. Taguchi's quality engineering handbook *In:* CHOWDHURY, S., TAGUCHI, S., WU, Y. & YANO, H. (eds.) *Quality engineering handbook.* Hoboken, N.J. : Livonia, Mich.: John Wiley & Sons

ASI Consulting Group.
TAKEUCHI, H., GOODMAN, C. & SONE, S. Moving block signalling dynamics: performance measures and re-starting queued electric trains. Electric Power Applications, IEE Proceedings-, 2003. IET, 483-492.
THORNES, J. E. & DAVIS, B. W. Mitigating the impact of weather and climate on railway operations in the UK. Railroad Conference, 2002 ASME/IEEE Joint, 2002. IEEE, 29-38.

TORRANCE, G. W., BOYLE, M. H. & HORWOOD, S. P. 1982. Application of Multi-Attribute Utility Theory to measure social preferences for health states. *Operations Research,* 30**,** 1043-1069.

TRANSPORT FOCUS 2016. Rail passenger satisfaction at a glance: Great Britain - Autumn 2015.

TRANSPORT FOCUS 2017. National rail passenger survey Autumn 2016 main report.

TRANSPORTATION RESEARCH BOARD 2013. Transit capacity and quality of service manual. 3 ed.

TRIANTAPHYLLOU, E. 2000. Multi-Criteria Decision Making Methods. *Multi-criteria Decision Making Methods: A Comparative Study.* Springer US.

TRIANTAPHYLLOU, E. 2013. *Multi-criteria decision making methods: a comparative study*, Springer Science & Business Media.

TRIANTAPHYLLOU, E., LOOTSMA, F., PARDALOS, P. & MANN, S. 1994. On the evaluation and application of different scales for quantifying pairwise comparisons in fuzzy sets. *Journal of Multi-Criteria Decision Analysis,* 3**,** 133-155.

TSIPORKOVA, E. & BOEVA, V. 2006. Multi-step ranking of alternatives in a multi-criteria and multi-expert decision making environment. *Information Sciences,* 176**,** 2673-2697.

TYE, H. 2004. Application of statistical 'design of experiments' methods in drug discovery. *Drug Discovery Today,* 9**,** 485-491.

UIC 1983. UIC leaflet 405-1: method to be used for the determination of the capacity of lines. International Union of Railways.

UIC 1996. UIC leaflet 405 OR: links between railway infrastructure capacity and the quality of operations. International Union of Railways.

UIC 2004. UIC leaflet 406: capacity. International Union of Railways.

UIC 2008. ETCS implementation handbook. Paris.

UK POWER NETWORKS SERVICES. 2015. *Infrastructure Services* [Online]. Available: http://www.ukpowernetworks.co.uk/internet/en/infrastructure-services/rail/ [Accessed 02/01 2016].

VIRGIN TRAINS EAST COAST. 2016. *The official view on our train services* [Online]. Available: https://www.virgintrainseastcoast.com/about-us/train-services-and-fleet-performance/ [Accessed 12/02 2016].

VOROPAI, N. I. & IVANOVA, E. Y. 2002. Multi-criteria decision analysis techniques in electric power system expansion planning. *International Journal of Electrical Power & Energy Systems,* 24**,** 71-78.

WAHLBORG, M. 2004. Banverket experience of capacity calculations according to the UIC capacity leaflet. *COMPUTERS IN RAILWAYS IX.* Dresden, Germany: WIT Press.

WALLENIUS, J., DYER, J. S., FISHBURN, P. C., STEUER, R. E., ZIONTS, S. & DEB, K. 2008. Multiple criteria decision making, multiattribute utility theory: recent accomplishments and what lies ahead. *Management science,* 54**,** 1336-1349.

WARDMAN, M. 2006. Demand for rail travel and the effects of external factors. *Transportation Research Part E: Logistics and Transportation Review,* 42**,** 129-148.

WATSON, R., GILLINGWATER, D. & BOODOO, A. 2003. Using network capacity effectively where train operators compete for paths - the UK experience. *World Congress on Railway Research.* Edinburgh, UK.

WINTERFELDT, D. V. & EDWARDS, W. 1986. Decision analysis and behavioral research. *Cambridge, England: Cambridge UniversityPress*.

WOODLAND, D. 2004. *Optimisation of automatic train protection system.* PhD, The University of Sheffield.

WU, Y., KLIMCHIK, A., CARO, S., FURET, B. & PASHKEVICH, A. 2015. Geometric calibration of industrial robots using enhanced partial pose measurements and design of experiments. *Robotics and Computer-Integrated Manufacturing,* 35**,** 151-168.

YU, M.-M. 2008. Assessing the technical efficiency, service effectiveness, and technical effectiveness of the world's railways through NDEA analysis. *Transportation Research Part A: Policy and Practice,* 42**,** 1283-1294.

YU, M.-M. & LIN, E. T. J. 2008. Efficiency and effectiveness in railway performance using a multi-activity network DEA model. *Omega,* 36**,** 1005-1017.

Appendix A: Calculation of criteria weights from the Focus Group study

1. Pairwise comparison forms: Which characteristic do you think is more important and to what extent?

1.1. Initial pairwise comparison forms

1.1.1. IM

A	B	more important? A or B	Scale (1-9)
Transport Volume	Journey Time	A	7
	Connectivity	A	8
	Punctuality	B	7
	Resilience	A	1
	Passenger Comfort	A	6
	Energy	A	4
	Resouce Usage	B	5
Journey Time	Connectivity	A	1
	Punctuality	B	5
	Resilience	B	7
	Passenger Comfort	A	4
	Energy	A	1
	Resouce Usage	B	6
Connectivity	Punctuality	B	6
	Resilience	B	5
	Passenger Comfort	A	4
	Energy	B	4
	Resouce Usage	B	7
Punctuality	Resilience	A	6
	Passenger Comfort	A	8
	Energy	A	5
	Resouce Usage	A	6
Resilience	Passenger Comfort	A	1
	Energy	A	5
	Resouce Usage	A	3
Passenger Comfort	Energy	B	3
	Resouce Usage	B	6
Energy	Resouce Usage	B	2

1.1.2. TOC

A	B	more important? A or B	Scale (1-9)
Transport Volume	Journey Time	A	6
	Connectivity	A	6
	Punctuality	A	5
	Resilience	A	5
	Passenger Comfort	A	7
	Energy	A	4
	Resouce Usage	A	1
Journey Time	Connectivity	A	4
	Punctuality	A	3
	Resilience	A	3
	Passenger Comfort	A	5
	Energy	B	3
	Resouce Usage	B	5
Connectivity	Punctuality	B	4
	Resilience	B	4
	Passenger Comfort	A	2
	Energy	B	5
	Resouce Usage	B	5
Punctuality	Resilience	A	3
	Passenger Comfort	A	3
	Energy	A	2
	Resouce Usage	B	3
Resilience	Passenger Comfort	A	4
	Energy	B	3
	Resouce Usage	B	3
Passenger Comfort	Energy	B	4
	Resouce Usage	B	4
Energy	Resouce Usage	B	2

1.1.3. Government

A	B	more important? A or B	Scale (1-9)
Transport Volume	Journey Time	A	7
	Connectivity	A	6
	Punctuality	A	1
	Resilience	B	4
	Passenger Comfort	A	8
	Energy	A	6
	Resouce Usage	B	3
Journey Time	Connectivity	B	3
	Punctuality	B	4
	Resilience	B	8
	Passenger Comfort	B	3
	Energy	B	7
	Resouce Usage	B	6
Connectivity	Punctuality	B	3
	Resilience	B	5
	Passenger Comfort	A	5
	Energy	A	3
	Resouce Usage	A	1
Punctuality	Resilience	B	3
	Passenger Comfort	A	5
	Energy	A	5
	Resouce Usage	B	5
Resilience	Passenger Comfort	A	7
	Energy	A	8
	Resouce Usage	A	3
Passenger Comfort	Energy	A	3
	Resouce Usage	B	6
Energy	Resouce Usage	B	4

1.1.4. Passenger

A	B	more important? A or B	Scale (1-9)
Transport Volume	Journey Time	A	5
	Connectivity	A	7
	Punctuality	B	3
	Resilience	A	2
	Passenger Comfort	A	3
	Energy	A	9
	Resouce Usage	A	9
Journey Time	Connectivity	A	2
	Punctuality	B	3
	Resilience	A	3
	Passenger Comfort	B	2
	Energy	A	9
	Resouce Usage	A	9
Connectivity	Punctuality	B	6
	Resilience	A	9
	Passenger Comfort	B	4
	Energy	A	9
	Resouce Usage	A	9
Punctuality	Resilience	A	6
	Passenger Comfort	A	2
	Energy	A	9
	Resouce Usage	A	9
Resilience	Passenger Comfort	B	5
	Energy	A	9
	Resouce Usage	A	9
Passenger Comfort	Energy	A	9
	Resouce Usage	A	9
Energy	Resouce Usage	A	1

1.2. Revised pairwise comparison forms

1.2.1. IM

A	B	more important? A or B	Scale (1-9)
Transport Volume	Journey Time	A	7
	Connectivity	A	8
	Punctuality	B	7
	Resilience	A	1
	Passenger Comfort	A	6
	Energy	A	4
	Resource Usage	B	5
Journey Time	Connectivity	A	1
	Punctuality	B	5
	Resilience	B	7
	Passenger Comfort	A	4
	Energy	A	1
	Resource Usage	B	6
Connectivity	Punctuality	B	6
	Resilience	B	5
	Passenger Comfort	A	4
	Energy	B	4
	Resource Usage	B	7
Punctuality	Resilience	A	6
	Passenger Comfort	A	8
	Energy	A	5
	Resource Usage	A	1
Resilience	Passenger Comfort	A	1
	Energy	A	5
	Resource Usage	A	3
Passenger Comfort	Energy	B	3
	Resource Usage	B	6
Energy	Resource Usage	B	2

1.2.2. Government

A	B	more important? A or B	Scale (1-9)
Transport Volume	Journey Time	A	7
	Connectivity	A	6
	Punctuality	A	1
	Resilience	B	4
	Passenger Comfort	A	8
	Energy	A	6
	Resource Usage	B	3
Journey Time	Connectivity	A	2
	Punctuality	B	4
	Resilience	B	8
	Passenger Comfort	B	3
	Energy	B	7
	Resource Usage	B	6
Connectivity	Punctuality	B	5
	Resilience	B	5
	Passenger Comfort	A	5
	Energy	B	2
	Resource Usage	A	1
Punctuality	Resilience	B	3
	Passenger Comfort	A	5
	Energy	A	5
	Resource Usage	B	5
Resilience	Passenger Comfort	A	7
	Energy	A	8
	Resource Usage	A	3
Passenger Comfort	Energy	B	3
	Resource Usage	B	6
Energy	Resource Usage	B	4

1.3. Final pairwise comparison form

A	B	more important? A or B	Scale (1-9)
Transport Volume	Journey Time	A	6
	Connectivity	A	7
	Punctuality	B	2
	Resilience	A	1
	Passenger Comfort	A	4
	Energy	A	5
	Resource Usage	B	2
Journey Time	Connectivity	A	2
	Punctuality	B	2
	Resilience	B	2
	Passenger Comfort	A	2
	Energy	B	3
	Resource Usage	B	6
Connectivity	Punctuality	B	5
	Resilience	B	2
	Passenger Comfort	A	2
	Energy	B	3
	Resource Usage	B	4
Punctuality	Resilience	A	2
	Passenger Comfort	A	5
	Energy	A	4
	Resource Usage	B	3
Resilience	Passenger Comfort	A	2
	Energy	A	3
	Resource Usage	A	2
Passenger Comfort	Energy	B	3
	Resource Usage	B	5
Energy	Resource Usage	B	3

2. Pairwise comparison matrices

2.1. The calculation of consistency ratios

The criteria weights are calculated as the eigenvector of the pairwise comparison matrix. Consistency ratio (CR) is calculated by comparing Consistency Index (CI) and Random Consistency Index (RI).

$$CI = \frac{\lambda_{max} - n}{n - 1}$$

$$CR = \frac{CI}{RI}$$

For matrix A, λ_{max} is the principal eigenvalue of A.
In this study, n = 8, and the corresponding RI = 1.41.
When the CR value is about 10% or less, the estimation of weights are considered accurate.

2.2. Initial matrices

2.2.1. IM

	TV	JT	CN	PT	RS	PC	EN	RU
TV	1.00	5.20	6.84	0.19	1.00	3.95	2.28	0.33
JT	0.19	1.00	1.00	0.33	0.19	2.28	1.00	0.25
CN	0.15	1.00	1.00	0.25	0.33	2.28	0.44	0.19
PT	5.20	3.00	3.95	1.00	3.95	6.84	3.00	3.95
RS	1.00	5.20	3.00	0.25	1.00	1.00	3.00	1.73
PC	0.25	0.44	0.44	0.15	1.00	1.00	0.58	0.25
EN	0.44	1.00	2.28	0.33	0.33	1.73	1.00	0.76
RU	3.00	3.95	5.20	0.25	0.58	3.95	1.32	1.00

CR = 19.39%

2.2.2. TOC

	TV	JT	CN	PT	RS	PC	EN	RU
TV	1.00	3.95	3.95	3.00	3.00	5.20	2.28	1.00
JT	0.25	1.00	2.28	1.73	1.73	3.00	0.58	0.33
CN	0.25	0.44	1.00	0.44	0.44	1.32	0.33	0.33
PT	0.33	0.58	2.28	1.00	1.73	1.73	1.32	0.58
RS	0.33	0.58	2.28	0.58	1.00	2.28	0.58	0.58
PC	0.19	0.33	0.76	0.58	0.44	1.00	0.44	0.44
EN	0.44	1.73	3.00	0.76	1.73	2.28	1.00	0.76
RU	1.00	3.00	3.00	1.73	1.73	2.28	1.32	1.00

CR = 4.88%

2.2.3. Government

	TV	JT	CN	PT	RS	PC	EN	RU
TV	1.00	5.20	3.95	1.00	0.44	6.84	3.95	0.58
JT	0.19	1.00	0.58	0.44	0.15	0.58	0.19	0.25
CN	0.25	1.73	1.00	0.58	0.33	3.00	1.73	1.00
PT	1.00	2.28	1.73	1.00	0.58	3.00	3.00	0.33
RS	2.28	6.84	3.00	1.73	1.00	5.20	6.84	1.73
PC	0.15	1.73	0.33	0.33	0.19	1.00	1.73	0.25
EN	0.25	5.20	0.58	0.33	0.15	0.58	1.00	0.44
RU	1.73	3.95	1.00	3.00	0.58	3.95	2.28	1.00

CR = 9.31%

2.2.4. Passenger

	TV	JT	CN	PT	RS	PC	EN	RU
TV	1.00	3.00	5.20	0.58	1.32	1.73	9.00	9.00
JT	0.33	1.00	1.32	0.58	1.73	0.76	9.00	9.00
CN	0.19	0.76	1.00	0.25	9.00	0.44	9.00	9.00
PT	1.73	1.73	3.95	1.00	3.95	1.32	9.00	9.00
RS	0.76	0.58	0.11	0.25	1.00	0.33	9.00	9.00
PC	0.58	1.32	2.28	0.76	3.00	1.00	9.00	9.00
EN	0.11	0.11	0.11	0.11	0.11	0.11	1.00	1.00
RU	0.11	0.11	0.11	0.11	0.11	0.11	1.00	1.00

CR = 30.46%

2.3. Revised matrices

2.3.1. IM

	TV	JT	CN	PT	RS	PC	EN	RU
TV	1.00	5.20	6.84	0.19	1.00	3.95	2.28	0.33
JT	0.19	1.00	1.00	0.33	0.19	2.28	1.00	0.25
CN	0.15	1.00	1.00	0.25	0.33	2.28	0.44	0.19
PT	5.20	3.00	3.95	1.00	3.95	6.84	3.00	1.00
RS	1.00	5.20	3.00	0.25	1.00	1.00	3.00	1.73
PC	0.25	0.44	0.44	0.15	1.00	1.00	0.58	0.25
EN	0.44	1.00	2.28	0.33	0.33	1.73	1.00	0.76
RU	3.00	3.95	5.20	1.00	0.58	3.95	1.32	1.00

CR = 18.80%

2.3.2. Government

	TV	JT	CN	PT	RS	PC	EN	RU
TV	1.00	5.20	3.95	1.00	0.44	6.84	3.95	0.58
JT	0.19	1.00	1.32	0.44	0.15	0.58	0.19	0.25
CN	0.25	0.76	1.00	0.33	0.33	3.00	0.76	1.00
PT	1.00	2.28	3.00	1.00	0.58	3.00	3.00	0.33
RS	2.28	6.84	3.00	1.73	1.00	5.20	6.84	1.73
PC	0.15	1.73	0.33	0.33	0.19	1.00	0.58	0.25
EN	0.25	5.20	1.32	0.33	0.15	1.73	1.00	0.44
RU	1.73	3.95	1.00	3.00	0.58	3.95	2.28	1.00

CR = 10.64%

2.4. Final matrix

	TV	JT	CN	PT	RS	PC	EN	RU
TV	1.00	3.95	5.20	0.76	1.00	2.28	3.00	0.76
JT	0.25	1.00	1.32	0.76	0.76	1.32	0.58	0.25
CN	0.19	0.76	1.00	0.33	0.76	1.32	0.58	0.44
PT	1.32	1.32	3.00	1.00	1.32	3.00	2.28	0.58
RS	1.00	1.32	1.32	0.76	1.00	1.32	1.73	1.32
PC	0.44	0.76	0.76	0.33	0.76	1.00	0.58	0.33
EN	0.33	1.73	1.73	0.44	0.58	1.73	1.00	0.58
RU	1.32	3.95	2.28	1.73	0.76	3.00	1.73	1.00

CR = 5.49%

Appendix B: Case Study scenario data

1. Stations in simulation

Number	Station ID	Full name
1	ALEXNDP	Alexandra Palace
2	BAYFORD	Bayford
3	BOWESPK	Bowes Park
4	BRKMNPK	Brookmans Park
5	CRHL	Crews Hill
6	CUFFLEY	Cuffley
7	DRYP	Drayton Park
8	ENFC	Enfield Chase
9	ESSEXRD	Essex Road
10	FNPK	Finsbury Park
11	GORDONH	Gordon Hill
12	GRPK	Grange Park
13	HADLYWD	Hadley Wood
14	HATFILD	Harringay
15	HFDN	Hatfield
16	HIGHBY	Highbury & Islington
17	HRGY	Hertford North
18	HRNSY	Hornsey
19	KNGX	Kings Cross
20	MRGT	Moorgate
21	NBARNET	New Barnet
22	NEWSGAT	New Southgate
23	OKLGHPK	Oakleigh Park
24	OLDST	Old Street
25	PALMRSG	Palmers Green
26	POTRSBR	Potters Bar
27	WELHAMG	Welham Green
28	WLWYNGC	Welwyn Garden City
29	WNMHILL	Winchmore Hill

2. All origin-destination pairs

Origin	Destination
KNGX	WLWYNGC
WLWYNGC	KNGX
KNGX	FNPK
HRGY	KNGX
MRGT	GORDONH
MRGT	HFDN
HFDN	MRGT
GORDONH	MRGT
WLWYNGC	MRGT
MRGT	WLWYNGC
MRGT	HRGY

3. Rolling stock types

Class	No. of trains	Description	Brake (m/s^2)	Max speed (km/h)	Max acceleration (m/s^2)	No. of carriages per unit	No. of Units	Total No. of seats
10009	2	HST 9 Car (Class 1)	0.67	201	1	9	1	624
10410	19	Intercity 225	0.67	201	1	10	1	537
18001	1	Class 180 10-car Class 1	0.78	201	1	5	2	620
22151	3	222 5 Car (Class 1)	0.78	201	0.588	5	1	242
31301	1	313 3 Car (Class 5)	0.78	121	1	3	1	196
31303	2	313 3 Car (Class 2)	0.78	121	1	3	1	231
31306	28	313 6 Car (Class 2)	0.78	121	1	3	2	462
31316	2	313 6 Car (Class 5)	0.78	121	1	3	2	392
31704	4	317 4 Car (Class 2)	0.78	161	1	4	1	292
31706	2	317 8 Car (Class 2)	0.78	161	1	4	2	584
32142	1	321 4Car (Class 2)	0.78	161	1	4	1	299
36504	8	365 4 Car (Class 1)	0.78	161	1	4	1	263
36508	12	365 8 Car (Class 1)	0.78	161	1	4	2	526

4. Experiment parameters

Parameter	Level 1	Level 2	Level 3
Seating capacity	All trains: 100% seats	All trains: 120% seats	All trains: 80% seats
Configuration	Change number of units for train type 31306: 2 units	Change number of units for train type 31306: 3 units	Change number of units for train type 31306: 1 unit
No. of stops	Original number of stops	Add stops at PORTSBR and OKLGHPK, for fast services with few stops	Reduce stops at GORDNH and OLDST, for slow services with a lot of stops
Scheduled dwell times	Original scheduled dwell times	Add 10% to the original scheduled dwell times for all trains	Reduce 10% from the original scheduled dwell times for all trains
Sectional running times	Original running times	Add 10% to the original sectional running times for all trains	Reduce 10% from the original sectional running times for all trains
Level of Heterogeneity	Original timetable heterogeneity	Take out stops and reduce running times, to make train graphs more uniform	Add stops and increase running times to add variability to train runs
No. of services	Original number of services (75 trains in total)	Add services in the “spare space” of the timetable (85 trains in total)	Take out services at busy times (63 trains in total)
Speed limits	Original speed limits	Increase speed limits by 10% across the whole simulation area	Reduce speed limits by 10% across the whole simulation area
Driving style	Default driver, running to the line speed	Start coasting if the train runs 60s early compared to the timetable	Runs to 80% of line speed, accelerate to line speed if 60s early, and brake when 30s late
Minimum dwell times	Original minimum dwell times	Add 10% to the original minimum dwell times for all trains	Reduce 10% from the original minimum dwell times for all trains
Conflict resolution	Reroute trains if they are stuck	Cancellation trains that get stuck	
Random dispatcher delays	No random dispatcher delays	Add random delays of 0-15 seconds at each dispatching station	

Appendix C: Publications during PhD research

[1] LU, M., CHEN, L., NICHOLSON, G., SCHMID, F. & ROBERTS, C. 2013. Defining Railway Capacity through Quality of Service. In: HANSEN, I. A. (ed.) 5th International Conference on Railway Operations Modelling and Analysis RailCopenhagen2013. Copenhagen, Denmark.

[2] LU, M., NICHOLSON, G., SCHMID, F., DAI, L., CHEN, L. & ROBERTS, C. 2013. A framework for the evaluation of the performance of railway networks. International Journal of Railway Technology 2, 79-96.

Defining Railway Capacity through Quality of Service

Menglei Lu, Lei Chen, Gemma Nicholson, Felix Schmid and Clive Roberts
Birmingham Centre for Railway Research and Education, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom,
E-mail: railoperations@contacts.bham.ac.uk

Abstract
The capacity problem is among one of the most significant concerns in the rail industry. Even though capacity is viewed as a very important problem by railways worldwide, different stakeholders view capacity in different ways. The result is there being no existing capacity definition to successfully reflect the complex relationship between market demand and railway management. There is therefore still doubt about how to identify remaining capacity, making it very hard to identify which approaches to use when attempting to make capacity improvements on railway lines. Furthermore, different stakeholders have varying perceptions of which are the most important factors affecting capacity. In this paper, the author analysed the existing capacity definitions and measures, and compared conventional capacity improvement measures. The Quality of Service Model is proposed as a new generic definition of capacity, and the static and dynamic solutions for improving railway capacity are discussed.

Keywords
Railway Capacity, Quality of Service

1 Introduction

In recent years, railways worldwide are experiencing unprecedented growth of both market demand and capacity supply. In the last decade, the British railway has been one of the busiest and fastest growing railways in the world. In the year 2010, more than 1.35 billion passenger journeys were made, with over 21,500 passenger services operated each weekday (ATOC, 2012). To meet the fast growing demand, the government has invested £7,606 million into rail industry in 2010, taking 33% of the total public spending on transport (Department for Transport, 2011).

While the rail industry is making every effort to meet the market demand, there is still a significant capacity gap (Network Rail, 2008b). The capacity problem is becoming one of the most significant concerns of many railways worldwide (Roberts et al., 2011).

Even though capacity is viewed as a very significant problem by the railway industry, different stakeholders view capacity in different ways. This results in there being no single agreed definition of the nature of railway capacity and its management. There is therefore still doubt about how to identify remaining capacity, making it very hard to identify which approaches to use when attempting to make capacity improvements on railway lines. Furthermore, different stakeholders have varying perceptions of which are the most important factors affecting capacity.

In this paper, the author analysed the existing capacity definitions and measures, and compared conventional capacity improvement measures. A new generic definition of capacity is proposed, and the static and dynamic solutions for improving railway capacity are discussed.

2 Definition of railway capacity

Railway capacity is a diverse concept. The capacity of a railway network is determined by a wide range of factors from civil, electrical, mechanical and environmental engineering (Roberts et al., 2010). All these factors require sustainable planning, operation and management. However, although the term capacity is frequently quoted, there is not a standard definition or agreed measurement. The concept of railway capacity is not clear.

Due to differences in requirements, capacity is viewed differently from different stakeholders (Table 1).

Market (customer needs)	**Infrastructure planning**	**Timetable planning**	**Operations**
Expected number of train paths (peak) Expected mix of traffic and speed (peak) Infrastructure quality need Journey times as short as possible Translation of all short- and long-term market-induced demands to reach optimised load	Expected number of train paths (average) Expected mix of traffic and speed (average) Expected conditions of infrastructure Time supplements for expected disruptions Maintenance strategies	Requested number of train paths Requested mix of traffic and speed Existing conditions of infrastructure Time supplements for expected disruptions Time supplements for maintenance Connecting services in stations Requests out of regular interval timetables (system times, train stops, etc.)	Actual number of trains Actual mix of traffic speed Actual conditions of infrastructure Delays caused by operational disruptions Delays caused by track works Delays caused by missed connections Additional capacity by time supplements not needed

Table 1: Different views of capacity (UIC, 2004)

This results in a wide variety of capacity definitions. Popular definitions range from volume of traffic to infrastructure occupation (Marwick, 1977) (Krueger, 1999) (Albrecht et al., 2008) (UIC, 2004) (Moreira et al., 2004).

To assess the capacity of a specific railway system, Krueger (1999), Woodland (2004) and Parkinson and Fisher (1996) have given more detailed definitions. They classified railway capacity into Theoretical Capacity, Practical Capacity, Train Capacity, etc.

However, with all these views, there is no single agreed definition or measure of railway capacity. It is agreed that railway capacity depends on the way it is utilised (UIC, 2004). There is still no clear explanation of the term *Capacity* in the rail world.

3 Introduction of existing capacity measures

Due to the differences in capacity definitions, there are two main types of capacity measures. The most common definition of railway capacity reveals the highest volume capability of a network. It is addressed by the UIC 405 method, which has already been superseded. The other type of definition concentrates on capacity consumption or utilisation. UIC 406 (the succeeding assessment to UIC 405) and many other methods all support this definition.

3.1 UIC 405 Method

UIC 405 was created by the International Union of Railways (UIC, 1983) to determine the capacity of lines. Although it is not in use today, it is still of great importance. UIC 405 provided a uniform method for comparison, as it directly expresses the capacity of a line in terms of trains per hour.

The UIC 405 basic formula is:

$$\mathrm{L} = \frac{\mathrm{T}}{\mathrm{t_{fm}} + \mathrm{t_r} + \mathrm{t_{zu}}} \qquad (4)$$

where L is the capacity of a line section in number of trains in period T (the reference period in minutes), t_{fm} is the average duration of minimum train headway time (minutes), t_r is the extra time margin (minutes) and t_{zu} is an additional time (minutes).

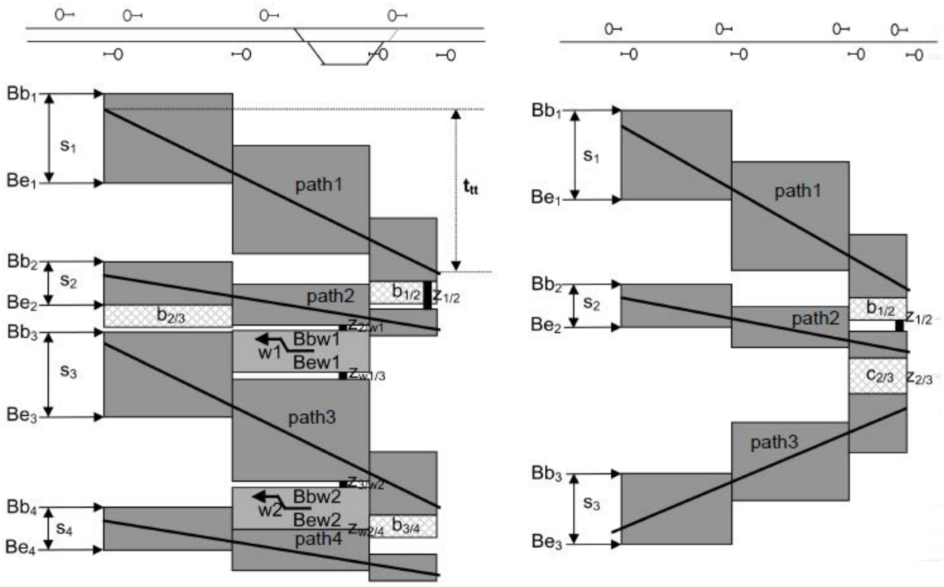
UIC 405 gives an explicit explanation of the relevant parameters based on a timetable. It is widely used in supporting railway network capacity evaluation.

The document has been applied by Swiss Federal Railways in developing CAPACITY, a computing tool used for quickly analysing different long-term scenarios and to determine bottlenecks for the whole of the Swiss railway network. UIC 405 was also used to construct CAP1 (for one direction flow) and CAP2 (for bi-directional flow) capacity models. Railway planning software such as VIRIATO also includes modules based on UIC 405 (Moreira et al., 2004).

3.2 UIC 406 Method

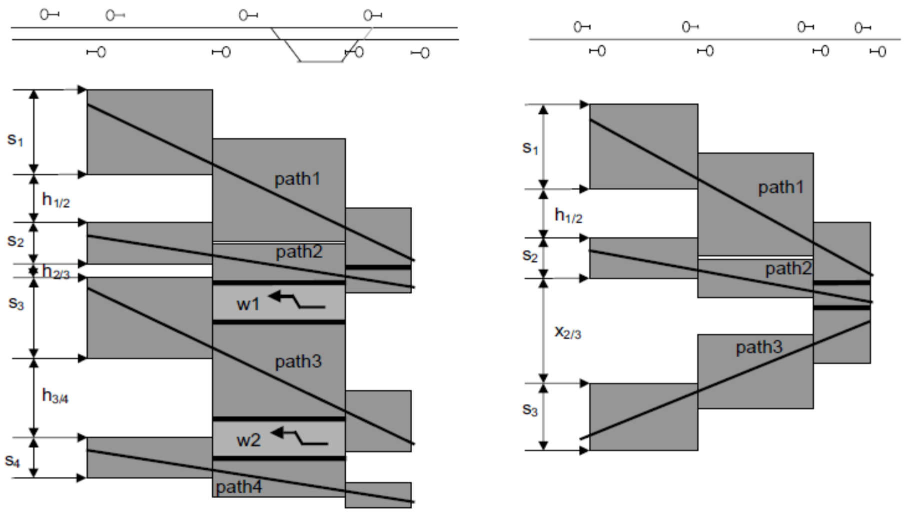
With the rising volume of border-crossing traffic in Europe and increasing demands for quality and quantity, the UIC 406 (UIC, 2004) method was developed to evaluate railway infrastructure capacity. Railway capacity is assessed though the capacity consumption.

The calculation of UIC 406 can be done using a timetable compression method, which is to push single train paths together up to the minimum theoretical headway, without any buffer time (Figure 1).



w_1, w_2 are white blocks or indirect infrastructure occupation times

Timetable shares within a timetable



$h_{1/2}$, $h_{2/3}$, $h_{3/4}$ are heterogeneity times

Timetable shares after compression

Figure 1: UIC 406 timetable compression (UIC, 2004)

The timetable compression method allows calculation by graphical analysis, using suitable tools, or by analytical calculation. It also makes the UIC 406 method easily transferrable to railway simulation software, such as RAILSYS.

The UIC 406 method has been used widely in many European countries (Wahlborg, 2004, Landex, 2006, UIC, 2004) to assess railway capacity. By far, it has been applied on about 3,000 km of lines on several European networks (UIC, 2004).

3.3 British Method: Capacity Utilisation Index

The Capacity Utilisation Index (CUI) was established in the late 1990s and has been used in Britain for assessing the capacity of lines, not junctions/locations. It is based on planning headways.

The idea of the CUI calculation is to push the trains on the train graph as close as the planning headway allows.

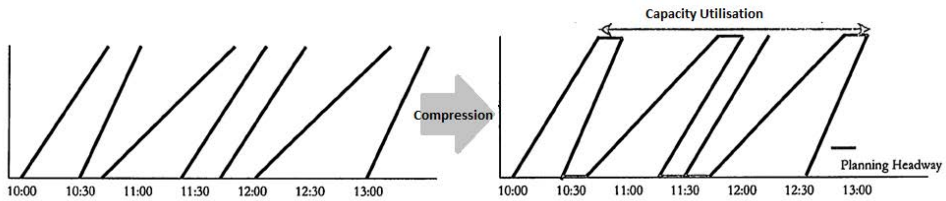


Figure 2: Timetable compression for CUI calculation (AEATechnology, 2005)

The mathematical expression is:

$$\text{Capacity Index} = \sum (\max(\text{Journey Time Differential}_{i,i-1}, 0) + \text{Headway}) / \text{Timeband Length} \qquad (5)$$

$Journey\ Time\ Differential_{i,i-1}$ is the difference in journey times between each pair of scheduled trains (i and i-1).

As suggested by SRA's Capacity Utilisation Policy (Strategic Rail Authority, 2003), generally CUI should not exceed 75%.

3.4 Analysis of Existing Capacity Measures

UIC 405, UIC 406 and the CUI are all widely used capacity measures. They are all timetable-related, yet there are many differences between them.

UIC 405 is measured in tph (trains per hour), while UIC 406 and the CUI are expressed as percentages. This is mainly because of the different aims of each method. UIC 405 measures the number of trains that can be run on a network whereas UIC 406 and the CUI are both for the determination of the extent of 'spare' capacity on a route.

UIC 406 and the CUI both use a timetable compression method. The UIC 406 analysis is based on the occupation times of individual signal blocks, whereas the CUI approach is based on the planning headways specified in Network Rail's 'Rules of the Plan' for entire route sections (2009). Thus the CUI calculation is simpler and more general than UIC 406.

Another difference is that UIC 406 can be applied to links. The CUI is not for junctions or stations, only for plain lines. UIC 406 and the CUI are general methods which give the possibility to study the relation between the number of train paths, infrastructure and capacity utilisation (a percentage value).

It can be concluded from that all measures are influenced by the timetable. UIC 405 and UIC 406 are extremely timetable-dependent, while the CUI is only relevant to the planning headway time. As planning headway is generally changed only when there is a modification of the infrastructure, CUI is much harder to change than UIC 406. This makes the CUI quite limited.

UIC 405, UIC 406 and CUI have all been successfully applied for capacity evaluation throughout many existing railway networks. However, none of the measures have considered the capacity of trains. The number of passenger seats or the freight tonnes the network can transfer is not reflected by any of the measures.

Passenger demands such as punctuality, connectivity, comfort and information needs are not taken into consideration. The Railway Undertakers' concerns for longer trains and trains with more seats are not taken into consideration either. For train length there are threshold values, for example, when a single track technical station is shorter than the freight train or when a platform is shorter than a passenger train.

The dynamic influence factors are not considered either in these measurements. A modified version of capacity definition and measurement is required to meet both system requirements and passenger satisfaction. In Section 0, a new generic definition of capacity and the static and dynamic solutions for improving railway capacity are discussed.

4 Capacity improvement solutions

Railway capacity is highly dependent on the way it is used. It varies with changes in traffic characteristics, infrastructure and operations. Railway resources are complex, difficult to change, and technically constrained, while market demands are dynamic, heterogeneous, and ambiguous. To find a balance between capacity demand and supply, a number of solutions to improve capacity have been raised.

On existing networks, efforts for capacity improvement are often focused on several areas that the railway industry believes to be most significant. In the UK, to accommodate demand for capacity whilst ensuring performance, the SRA (Strategic Rail Authority, 2003) provided several options:

- Increase load factors (where crowding is not an issue);
- Lengthen trains;
- Improve train path take up arrangements;
- Change pattern and mix of train services (timetables focussed on achieving higher throughput rather than highly diverse services);
- Reduce timetable 'fragility' (e.g. more robust plans for crew and stock movements);
- Better train regulation (revisit prioritisation rules, class regulation practices and use of passing facilities by passenger services).

These measures have been proven to improve capacity utilisation in the UK rail network (Strategic Rail Authority, 2003).

Studies by Pachl (2002) and Abril, Barber et al. (2008) also suggested several solutions for improving the capacity of a railway line or a part of a network. These solutions include changing parameters regarding the infrastructure, rolling stock, timetable, and operations.

Capacity improvement solutions can be classified into three categories: short term, medium term and long term solutions.

In the short term, by taking measures such as tariff change and better staff management, market demand can be controlled, and the staff system can have better performance. These measures can lead to a temporary improvement on capacity, and is great for peaks caused by certain events, such as the Olympics. However, they may cause dissatisfaction from market, staff and trade unions. They are not compatible with the ever-increasing traffic volume, either.

Medium term solutions involve adapting the following parameters:

- Rolling Stock parameters, such as number of seats, reliability and availability;
- Traffic parameters, such as adding new lines, changing priority, changing train mix, etc.;
- Operating parameters, including modifying timetable, better delay management, etc.

These solutions are the most popular for existing railway networks. They can maximize the amount and utilisation of the capacity of the current network. They are generally quick, and not very expensive, yet they require much more thinking.

In the long term, by modifying infrastructure parameters (e.g. the signalling system, track layout and junction structure), the capacity of the rail system can be improved significantly. These solutions, however, take the highest amount of time and investment thus needs the most planning time and effort.

Additionally, many other subsystems within the railway system affect railway capacity and its usability, such as the power supply, train door arrangements, junction characteristics and passenger management capabilities (Roberts et al., 2010). The impact of these subsystems is often not considered in conventional measures as they are hard to quantify, although in certain cases they can have a significant impact.

5 Defining 'Quality of Service'

In order to address the key requirements of all stakeholders, a general high-level measure of 'Quality of Service' is developed and defined as a new model of railway capacity. This model tries to provide solutions for the market, infrastructure planning, timetabling and operations (as emphasised in UIC 406). This research has been influenced and developed together with a group of infrastructure managers, operators, suppliers and academics working within the EC FP7 ON-TIME project.

Quality of Service is a combination of the capability and dependability of the railway system. The aim is to develop a comprehensive concept that covers the main attributes of the railway system, considering parameters as diverse as traffic volume, journey time, connectivity, punctuality, resilience, passenger comfort, energy, and resource usage. Mathematical expressions are developed for the various parameters. A high level view of the Quality of

Service measure can be found in Figure 3.

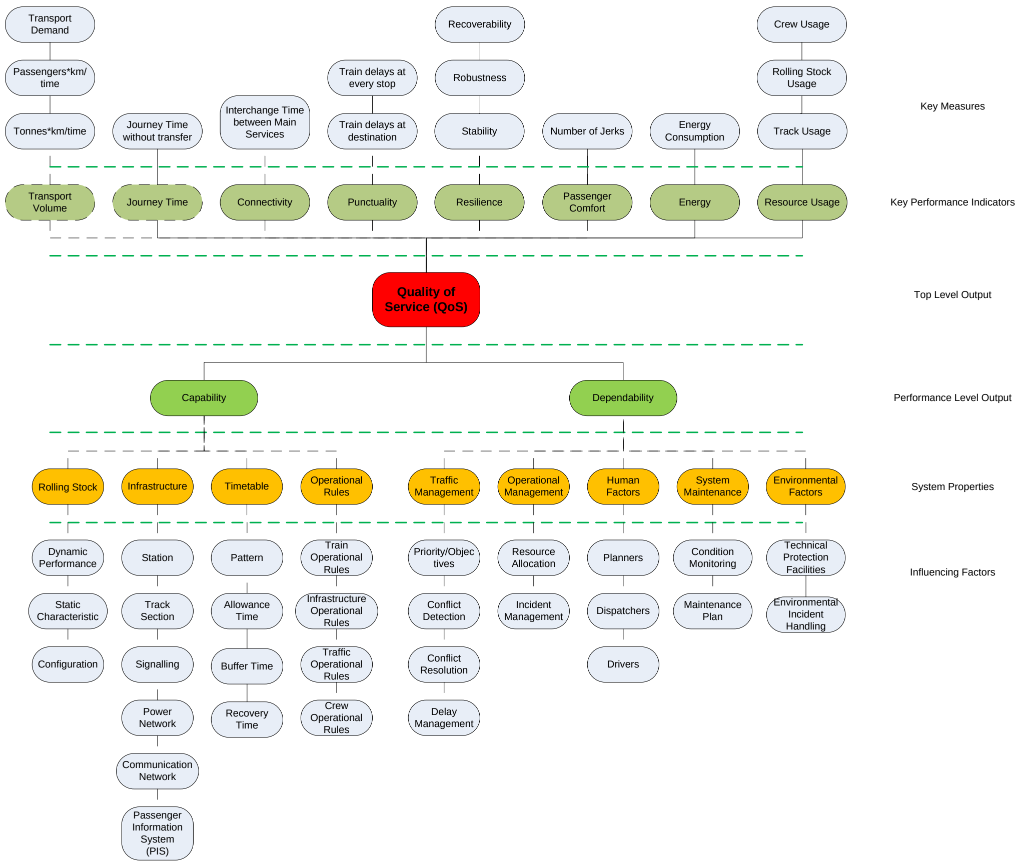


Figure 3: Factors considered in the Quality of Service measure

From Figure 3, it can be seen that from the point of view of the railway system, the factors affecting Quality of Service can be broken down into capability and dependability. Capability covers all the “static” components that are relatively hard to change, which define the underlying ability of the railway system to perform its function, such as rolling stock type, infrastructure systems, timetables and operational rules. Dependability includes all the “dynamic” components of the system, which determine the performance of the system given the static components, such as, traffic management strategies, operational management, human factors, maintenance strategies and environmental factors. Dynamic components can be modified over a relatively short term with moderately low cost.

An illustration of the Key Performance Indicators and System properties is given in this section. By modifying the Influencing Factors, the Quality of Service of the system can be improved, which is shown by the Key Performance Indicators values optimised.

5.1 Key Performance Indicators

Transport Volume

Transport Volume is defined as the volume of products (passengers or cargo) that can actually be delivered by a transport system over specific infrastructures.

Journey Time

In general, journey time is considered as the total practical consumed time for a passenger to complete his trip, without transfer. Both sectional and total journey times need to be taken into account.

Connectivity

At a given interchange, the passenger or cargo interchange time between any two services is the timetabled time in seconds between a passenger/cargo arriving on the first service and departing on the second service.

Punctuality

Punctuality is the characteristic of being able to complete a required task or fulfil an obligation before or at a previously designated time; it is also an important measure of the performance of train operations.

Resilience

Resilience is defined as the ability of a system to withstand stresses, pressures, perturbations, unpredictable changes or variations in its operating environment without loss of functionality. Three levels of resilience are defined:

- Stability: the ability to recover without active train rescheduling.
- Robustness: the ability to recover with active train rescheduling/ordering.
- Recoverability: the ability to recover with operational management measures such as train cancellation, rolling stock re-allocation etc.

Passenger Comfort

In railway transport, to guarantee a good level of comfort for passengers on their journeys, many factors need to be considered. Among these, the smoothness of the train driving performance has a significant impact. To improve passenger comfort, the number of jerks needs to be reduced.

Energy

Generally, energy consumption in the railway system includes energy consumed both by running rolling stock and infrastructure such as stations, signalling systems, etc. The most important part of energy used is that needed to move the trains.

Resource Usage

The resources used in the railway system include three main aspects: track usage, rolling stock usage and crew usage.

5.2 System Properties

Capability: Rolling stock

Rolling stock comprises all the vehicles running on the railway network. The main features of rolling stocks that can affect Quality of Service are as follows:

- Dynamic Performance: braking, acceleration, resistance, traction force, etc.

- Static Performance: length, mass, adhesion, maximum speed, etc.
- Configuration: the way rolling stocks are formed is also very important. This may affect the length, mass, traction, etc. and thus affect the running dynamics.

Capability: Infrastructure

Infrastructure is a vital component of the railway system. It has great influence on the train run, energy utilisation and potentially passenger comfort. Infrastructure is mainly made up of station, track section, signalling, power network, communication network, and the Passenger Information System (PIS).

Capability: Timetable

A railway timetable is a detailed plan of trains departing and arriving at stations.

- Pattern: the time slots arrangement pattern for all the trains in the nominal timetable, e.g. the mixture patterns for fast trains and slow trains. This would affect the resilience of the nominal timetables.
- Allowance time, Buffer time, Recovery time.

Capability: Operational rules

Operational Rules are the short-term requirements in practice. The main aspects are Train operational rules, infrastructure operational rules, traffic operational rules and crew operational rules.

Dependability: Traffic management

Traffic Management controls the movement of rolling stocks in short term practice. the content of traffic management includes priority, conflict detection, conflict resolution and delay management.

Dependability: Operational management

- Resource allocation: Rolling stock, crew and other resource planning.
- Incident management: This deals with faults and break-downs in the operation.

Dependability: Human factors

There are a number of human factors in the system, such as planners, dispatchers and drivers. They introduce uncertainties into the practice.

Dependability: System maintenance

The equipment and facilities are essential foundations of the railway system. The condition of them should be observed and maintained regularly.

Dependability: Environmental factors

Environmental factors such as wind, rain, snow and lightning are the source of many railway accidents. Technical Protection Facilities and emergency plans are applied to protect against these environmental factors.

5.3 Application of the Quality of Service Model

The Quality of Service model covers all aspects of the railway system that should be considered in the capacity planning process. It is an indication of the comprehensive performance of the railway system. It covers Transport volume, Journey time, Connectivity, Punctuality, Resilience, Passenger comfort, Energy and Resource usage. The railway systems are expected to be optimal in terms of all the indicators, however, trade-offs need to be made in practice due to the various constraints in real life railway operations.

The Key Performance Indicators are an intuitive reflection of the Quality of Service of a railway system. They provide an overall assessment of the Capability and Dependability of the system. By giving weightings to the Key Performance Indicators, the Quality of Service model can be developed into a comprehensive capacity measure.

The Quality of Service measure is a function of all the KPIs listed in

Figure . The formula for the calculation is:

$$Q = f(TV, JT, CN, PT, RS, PC, EG, RU) \qquad (6)$$

For each KPI, an objective function composed of certain key terms and weighting functions, subject to the necessary constraints, must be developed. Each objective function is loosely defined as a constrained optimisation problem, but could be replaced by a series of unconstrained optimisation problems with appropriate penalty functions.

The determination of weighting functions needs to consider the main objective trade-off in practice. The final weightings will be tested with a simulation based case study in RAILSYS.

6 Conclusions

Railway capacity is a diverse concept. With different requirements in practice, it is viewed differently from each stakeholder. However, the existing definitions and measures of capacity are often focused only on traffic volume or infrastructure occupation, and they are not enough for the improvement of the overall service level. Therefore the author introduced Quality of Service as the new standard to evaluate the performance of a railway network.

The Quality of Service model is an indication of the comprehensive performance of the railway system. It is the result of the interaction between the static and dynamic elements of the railway system. By changing the influencing factors of the railway system, Quality of Service can be improved.

The Quality of Service of a railway system can be assessed using an objective function of the Key Performance Indicators. Individual weightings are given to each KPI to form a single formula to determine the Quality of Service value. Trade-offs can be made for specific scenarios. The development of the Quality of Service measure will be assisted and testified by both analytical and simulation based studies.

References

[1] ATOC. (2012, 17/06). Capacity - ATOC's view. Available: http://www.atoc.org/clientfiles/File/capacity2012.pdf

[2] DfT. Transport Statistics Great Britain: 2011 [Online].

[3] NR, "Meeting the capacity challenge: The case for new lines," Network Rail2008.

[4] C. Roberts, et al., "A new railway system capacity model," 2011.

[5] C. Roberts, et al., "A new railway system capacity model," 2010.

[6] UIC, "UIC Leaflet 406: Capacity," ed: International Union of Railways, 2004.

[7] P. Marwick, Impact of Freight/passenger Interference on Capacity And/or Cost. Montreal, 1977.

[8] H. Krueger, "Parametric modeling in rail capacity planning," in Simulation Conference Proceedings, 1999 Winter, 1999, pp. 1194-1200 vol.2.

[9] I. T. Albrecht, et al., Railway Timetable & Traffic: Analysis, Modelling, Simulation: Eurailpress, 2008.

[10] N. Moreira, et al., "Network capacity," in Computers in Railway Six. vol. 15, ed, 2004, pp. 35-43.

[11] D. Woodland, "Optimisation of Automatic Train Protection System," PhD, Department of Mechanical Engineering, The University of Sheffield, Sheffield, 2004.

[12] T. Parkinson and I. Fisher, Rail Transit Capacity: Transportation Research Board, National Research Council, 1996.

[13] UIC, "UIC Leaflet 405-1: Method to be Used for the Determination of the Capacity of Lines," ed: International Union of Railways, 1983.

[14] M. Wahlborg, "Banverket Experience of Capacity Calculations According to the UIC Capacity Leaflet," presented at the COMPUTERS IN RAILWAYS IX Dresden, Germany, 2004.

[15] A. Landex, Kaas, AH, Schittenhelm, B & Schneider-Tilli, J, "Evaluation of railway capacity," Proceedings of Trafficdays, 2006.

[16] AEATechnology, "Capacity Utilisation Indices Methodology," 2005.

[17] SRA, "Capacity Utilisation Policy: Network Utilisation Strategy," Strategic Rail Authority2003.

[18] J. Armstrong and J. M. Preston, "Automating the production of train graphs and the calculation of CUI Values," presented at the 3rd International Seminar on Railway Operations Modelling and Analysis, Engineering and Optimisation Approaches, 2009.

[19] J. Pachl, Railway operation and control: VTD Rail Pub., 2002.

[20] M. Abril, et al., "An assessment of railway capacity," Transportation Research Part E: Logistics and Transportation Review, vol. 44, pp. 774-806, 2008.

A Framework for the Evaluation of the Performance of Railway Networks

M. Lu, G. L. Nicholson, F. Schmid, L. Dai, L. Chen and C. Roberts

Birmingham Centre for Railway Research and Education, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom

Abstract

The infrastructure of many European railways is becoming increasingly saturated, while the operators face strong demands to increase services and carry more passengers and freight. The European FP7 project 'ON-TIME' is working to achieve a step change improvement in railway performance across Europe through improvements to timetabling, real time traffic management, operational management of large scale disruptions and driver advisory systems. In this paper, the authors compare and analyse the existing capacity definitions and measures and propose a new quality of service based concept as a generic framework of railway network performance. The key factors affecting quality of service are identified and analysed. Two examples show the influence of the system properties on a performance indicator of resilience to delays.

Keywords: railway capacity; quality of service; performance evaluation.

1 Introduction

Railways are playing an ever more important role in stimulating economic growth, reducing road congestion and limiting the climate change impact of transport networks (Ison et al., 2012). Due to the rapidly growing demand for passenger and freight transportation, railways in Europe are experiencing increasingly intensive use of their infrastructures and train services. However, many main line railways are already highly saturated, leading to poor performance that manifests itself in extensive delays and disturbances (Watson et al., 2003, Network Rail, 2008b, Goverde, 2007). Thus, improving railway network performance and traffic capacity are becoming highly significant concerns in the railway industry. In the present paper, railway network performance is a term that combines the quantity and quality of the operational behaviour that result from the characteristics of the railway infrastructure and the nature of its utilisation.

Even though the terms capacity and performance are widely used, different stakeholders have diverse requirements that lead to different criteria for the evaluation of the capability of a network. For example, the passenger and freight transport markets expect journey times to be as short as possible, the infrastructure planners require track access for maintenance, timetable planners request connections between services in stations, while operators aim to minimise delays since late arrivals result in penalties on many networks (UIC, 2004).

Railway network performance is determined by a complex relationship between infrastructure, rolling stock, line structure, timetable structural data and requirements for service quality (Lucchini et al., 2001). Therefore, railway businesses use a wide variety of railway capacity definitions, ranging from a measure of the volume of traffic to the utilisation, that is, the proportion of available infrastructure time slots that are occupied by trains (Albrecht et al., 2008, UIC, 2004, Burdett and Kozan, 2006, Moreira et al., 2004). More detailed definitions have also been created to assess the capacity of specific railway systems (Woodland, 2004, Kontaxi and Riccia, 2012, Krueger, 1999). However, these existing definitions and measures of capacity are often focused on the particular needs of the infrastructure owner, timetable planner or the train service provider, mentioned above.

In order to address the key requirements of all parties, a novel framework, conceptually based on ‘quality of service’, is proposed in this paper for the evaluation of railway networks.

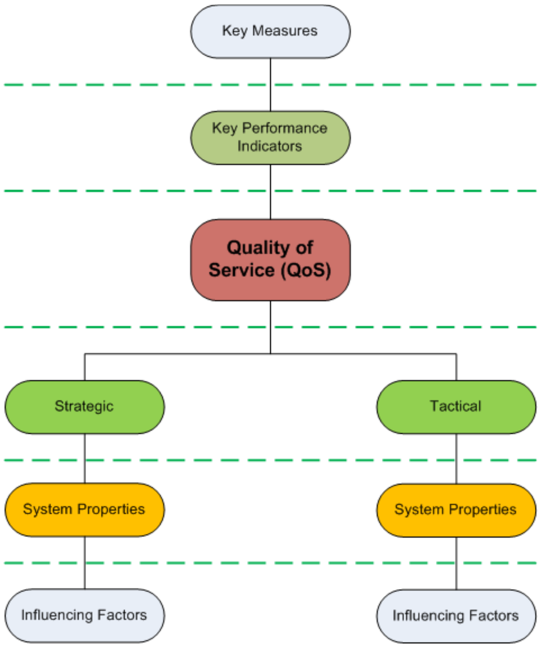


Figure 1. Outline of the quality of service framework.

In this framework the quality of service concept is used to represent the expectations of the performance of railway systems, covering the factors affecting the network performance from the perspectives of all stakeholders, such as operators, timetable planners, infrastructure managers, passengers and the government. The factors are divided into strategic and tactical categories. The former are factors that are expensive and time consuming to change while the latter are relatively easily changed in shorter timescales. By modifying these factors, the values of the key performance indicators (KPIs) of the framework will change, thus indicating that the performance of the system has changed. The outline of the framework is given in Figure 1. The quality of service of a system is evaluated with the KPIs, each assessed with several key measures. The bottom boxes show the influencing factors within the system, categorised as strategic factors and tactical ones, as described above. The full framework with all the KPIs, key measures and influencing factors listed will be given in Figure 2 in Section 4.

Changes in performance can be measured using an evaluation function for the measurement of the quality of service of the railway system. The precise format of the quality of service framework will be described and a discussion of the performance evaluation function will be given in the fourth section. The framework can be applied to compare and evaluate the impact of changes to subsystem properties of a railway network, such as its rolling stock, infrastructure, timetable or operational and traffic management methods. In the fifth section of the paper two examples are provided that demonstrate the quantification of the key measures of system resilience, one of the key performance indicators of the quality of service framework.

This research has been influenced and developed by a group of infrastructure managers, operators, suppliers and academics working within the EC FP7 ON-TIME project (ON-TIME Consortium, 2012).

2 Background information

The existing network capacity definitions and measures fall into two main categories, namely, the traffic volume capability of a network and the level of utilisation thereof.

The traffic volume capability of a network is addressed by the UIC 405 (UIC, 1983) method, which measures the number of trains that can be run per hour on a railway line, with given timetable data. Capacity consumption or utilisation is defined by the Capacity Utilisation Index (CUI) (AEA Technology, 2005) and UIC 406 (the notional successor standard to UIC 405) (UIC, 2004). The two measures both use a timetable compression method. The UIC 406 analysis is based on the occupation times of individual signal blocks, whereas the CUI approach is based on the planning headways specified in Network Rail's 'Rules of the Plan' for entire route sections (2009).

UIC 405, UIC 406 and the CUI provide widely used capacity measures. All have been applied for capacity evaluation on existing railway networks (Moreira et al., 2004, Wahlborg, 2004, Landex, 2006, Strategic Rail Authority, 2003, Gibson et al., 2002), mainly in Europe, yet they reflect the point of view of either the infrastructure planner or the train service provider.

None of the measures (UIC 405, UIC 406, and CUI) consider passenger requirements such as punctuality, reliability, connectivity, comfort and information provision, nor do they reflect operators' needs for longer trains and trains with more accommodation for passenger and goods. The tactical influencing factors in operational management (e.g., delays, imposition of temporary speed limits and inadequate traffic management strategies) are not considered either in these measurements.

To meet not only infrastructure and timetable planning requirements, but also demands from the market and operations, modified definitions of railway capacity and measurements of network performance are needed. Therefore, the strategic and tactical solutions for improving railway capacity and a new method of measuring railway network performance are discussed in this paper.

3 Enhancing railway network performance

The performance of a railway network, in terms of the quantity and quality that it offers to its users, is affected by many factors, such as traffic characteristics, infrastructure and operational methods. Therefore, there is a wide variety of potential solutions for improving the performance and capacity of a railway line or part of a network (UIC, 1996, Roberts et al., 2010, Strategic Rail Authority, 2003), among which improved approaches towards scheduling is an important concern (Burdett and Kozan, 2010, Burdett and Kozan, 2009). Depending on how easily such solutions can be implemented, they can be classified as 'strategic' or 'tactical'.

The 'strategic' solutions involve modifying the more fixed parameters of the network or system (see Figure 1) by improving components and features that require a long time or significant investment, if a change is to be achieved. Several aspects are involved, listed here in ascending order of difficulty of intervention:

1. Timetable parameters: traffic pattern, buffer times, recovery times, etc;
2. Rolling stock parameters: quantity of passenger and freight accommodation[2], train performance in terms of acceleration and braking rates;
3. Infrastructure parameters: upgrades to the signalling system, track alignment, junction layouts, station tracks and new lines, etc.

Solutions that address these aspects of network performance may lead to a significant increase in the quantity and quality of network capability.

The 'tactical' solutions are those that are relatively easy to implement in terms of time and monetary investment. The factors changed are included in the category of tactical factors in Figure 1. These are the short to medium term solutions.

In the short term, measures can be taken to improve the quantity and quality of existing traffic capacity and to influence the market demand. These measures include tariff changes, minimal timetable modifications and better staff management. The short term solutions can lead to a temporary improvement in network performance, particularly during peaks caused by certain events, such as sport games and exhibitions. However, dissatisfaction may be caused amongst passengers, staff and trade unions. Generally, the measures are also not compatible with the need to accommodate an ever-increasing traffic volume.

In the medium term, solutions can involve, again in increasing order of difficulty:

1. Adjustment of allowances and minor timetable changes such as small train path amendments;
2. Changing operating parameters: delay management, etc;
3. Changing traffic parameters: priorities, train mix, etc.

The medium term solutions are often considered the best approaches for network performance improvement, especially with a limited budget. With little change to the characteristics of the current railway networks, these solutions can maximise the quantity and quality of the existing network. Yet, on already overloaded networks, the effects may be minor.

Within the ON-TIME project, emphasis has been placed on "investigating new ways of managing existing capacity that will allow more services to operate more reliably than is currently the case" (ON-TIME, 2012b). The research is focused on four main themes; all related to the tactical components of network performance enhancement.

- Theme 1: Development of timetables that accommodate minor disruptions and improve traffic flow in bottleneck areas;
- Theme 2: Implementation of real time traffic management that allows traffic to recover rapidly from perturbations and therefore improves service dependability;

2 This is the rolling stock's maximum capacity to carry passengers or freight, in terms of available passenger seats and standing room, and available freight container volume subject to a permitted maximum tonnage.

- Theme 3: Research into operational management methods to recover from large scale disruptions;
- Theme 4: Development of advanced driver advisory systems that provide real time information and advice to drivers and that enhance energy efficiency.

The first three themes are based on the notion of resilience to describe the ability of a railway system to withstand perturbations and unexpected variations within its operating environment. Within the ON-TIME project, service resilience in response to a delaying incident is divided into three levels: (i) a set of interacting train services and the associated timetable are defined as stable when the system recovers to an on time state without active train rescheduling. (ii) If active train rescheduling or reordering measures have to be taken to return a system to normal operation after a delaying incident, this set of interacting services is defined as robust. (iii) The requirement to implement major operational management measures, including train cancellation and rolling stock reallocation, to return a system to an on time state describes a set of interacting train services that are recoverable in response to a major delaying incident.

Research conducted within the four themes is resulting in railway network quality improvements that allow the use of a greater quantity of the potentially available capacity.

A quality of service based framework is proposed as a new generic approach to the measurement and enhancement of railway network performance. An introduction to this framework and examples of the quantification of the resilience KPI can be found in the following sections.

4 Quality of service framework

4.1 Framework description

Railway network performance in terms of quality of service is achieved by a combination of strategic factors and tactical factors that influence the way in which a railway system or network supports commercial objectives. The focus of this paper is to provide a comprehensive framework that covers the main attributes of the railway system, considering parameters as diverse as accommodation, journey time, connectivity, punctuality, resilience, passenger comfort, energy, and resource usage. A detailed overview of the quality of service framework is shown in Figure 2.

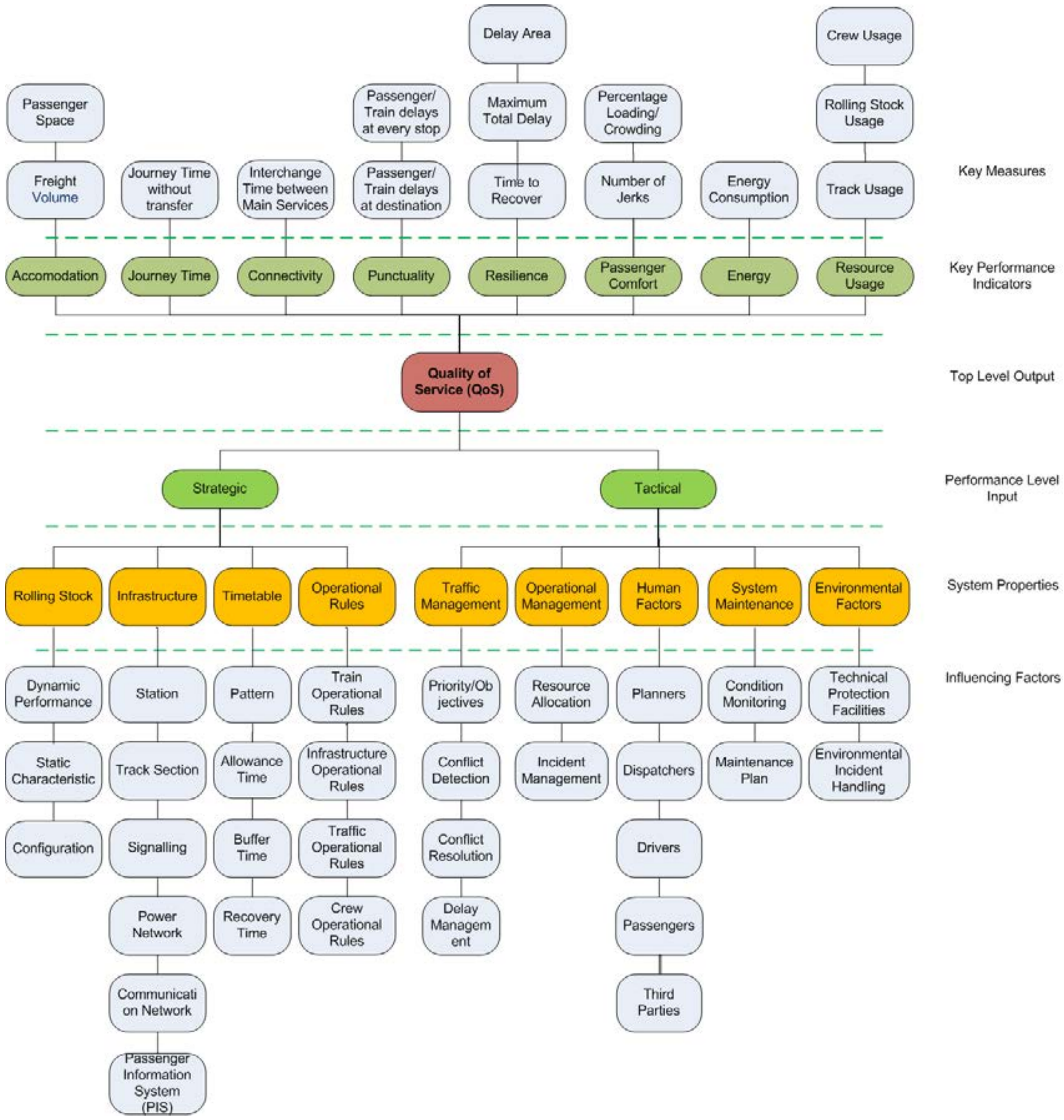


Figure 2. The quality of service framework.

The inputs of this framework are shown in the bottom boxes. At the performance level, these factors are categorised as 'strategic' factors on the left and 'tactical' factors on the right. Each category contains several system properties, which can be further broken down into the influencing factors with the system.

The direct outputs of the framework are managed through the key performance indicators (shown in the top boxes) that are monitored by means of the measures derived from a railway's operations. One of the KPIs, namely resilience, is selected for detailed discussion in this paper. This KPI is proposed as a new way to measure the recovery of a railway system from delays.

4.2 Resilience KPI and its graphical representation

The resilience KPI is evaluated by considering the delays to all trains that travel within a given area, during a given time period. The delay of each train is calculated with reference to scheduled timings at stations (both for arrivals and departures) and at selected signals or similar timing points. These delays to each train are then plotted against time. Additionally, the system delay is calculated as the sum of the delay values of each train at its most recent timing point (see Figure 3). This is a continuous, event-driven function. The key measures of resilience are based on this system delay curve; they are: i) its maximum value, (the maximum total delay [seconds] during the time period under consideration); ii) the time to recover [seconds], the time taken for the system delay curve to return below a threshold value after a delay was introduced and iii) the delay area [seconds^2] calculated as the area under the system delay versus time curve.

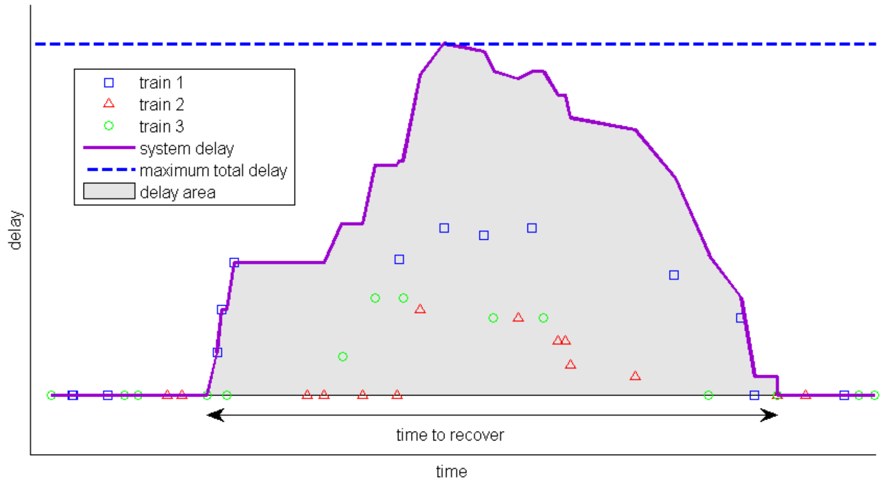


Figure 3. Schematic of the key measures used to evaluate the resilience KPI using an example with three delayed trains.

4.3 Performance evaluation function for the quality of service framework

The KPIs are independent measures of the railway system performance, yet they are related to each other. A change in the influencing factors may lead to changes in more than one KPI in the system, and the changes are not always in the same direction. For example, an increase in the allowance times in the timetable may lead to improved punctuality, resilience and potentially greater passenger comfort, yet the journey time, energy consumption and resource usage may be increased, meaning these three KPI values become less favourable. In the ideal situation, all the KPIs would be optimal, but this can never be fulfilled in practice, as trade-offs need to be made, due to the complexity and constraints of practical railway operations.

For each KPI, an objective function composed of its key measures and appropriate weighting functions, subject to the necessary constraints, must be developed. The formulation of a combination of the normalised KPI objective functions in, for example, a weighted linear combination, will allow this framework to be used to rank or rate alternatives (e.g. to evaluate a set of alternative timetables produced in Theme 1 of On Time). The knowledge and judgement of a group of expert stakeholders will be used to aid the transformation of the framework into a format that can be used for the evaluation of railway systems. They must choose the weights by expressing a priori their preferences and opinions of the relative importance of the KPIs and come to a consensus. A systematic weight selection method such as the analytic hierarchy process (Saaty, 1980), a pairwise comparison method, can be used to facilitate this. The essential features of the chosen weight selection method are transparency and traceability, and additionally to ensure that the selected weights accurately preserve the qualitative preferences of the stakeholders.

For specific scenarios, the KPIs considered may vary. Different weightings may also be given to each KPI depending on its importance in the scenario under consideration. For the ON-TIME project, the format of the performance evaluation function and its associated weightings will vary depending on the emphasis of each of the themes:

- For Theme 1 the KPIs that need to be considered are passenger and goods accommodation, journey time, connectivity, resilience, energy, and resource usage;
- In Theme 2 the KPIs covered are journey time, connectivity, punctuality, resilience, energy, and resource usage;
- Theme 3 considers passenger and goods accommodation, journey time, connectivity, punctuality, resilience, resource usage;
- Finally, the performance evaluation function for Theme 4 will address journey time, connectivity, punctuality, resilience, passenger comfort, energy, and resource usage.

5 Evaluation of delay scenario example key measures

In this section, two examples are presented that demonstrate the quantification of the key measures for the resilience KPI. This demonstrates the manner in which the relevant key measures can be used to compare different operational management solutions in response to a given delaying incident.

The examples are based on the simulated running of trains over a section of the East Coast Main Line (ECML) in the United Kingdom (see Figure 4) during the morning peak period. The area considered is chosen to be representative of a section of the network under control of a single traffic control centre. The simulated rolling stock characteristics and types closely match those running on the physical network. The timetable used is derived from the public published timetable (First Capital Connect, 11 December 2011 to 8 December 2012) and is implemented by means of Graffica's Hermes simulator (Graffica, 2013).

Figure 4. The section of the East Coast main line considered in examples 1 and 2.

In example 1, the delaying incident considered is the failure of a signal just before Welwyn Garden City station (approximately 32 km north of London) for 10 minutes between 07:24 and 07:34 on a weekday. The second example considers a partial loss of power to a train travelling along the East Coast Main Line from the north towards London. This train's power is permanently reduced, with the incident beginning at 07:24, as it arrives at the signal before Welwyn Garden City station.

In each example, three responses to the delaying incident are presented; first, no operational measures are taken and the trains pass junctions in the order in which they arrive (first come first served), while, in the second and third cases, two different operational interventions are made. These are summarised in Table 1.

(1)		**Scenario 1**	**Scenario 2**
Scenario Description	**Time**	Between 07:24 and 07:34 on a weekday	Beginning at 07:24 on a weekday
	Place	Signal just before Welwyn	Just before Welwyn Garden City

		Garden City station	station
	Event	Signal failure for 10 minutes	Power loss to service TS158
Operational Response	**A**	**OR1a**: No operational intervention; first come, first served.	**OR2a**: No operational intervention; first come, first served.
	B	**OR1b**: The first affected train is rerouted past a different platform through Welwyn Garden City station (it was not due to make a stop there).	**OR2b**: Train is cancelled at next station once it is at least 3 minutes late.
	C	**OR1c**: The first two affected trains are rerouted past a different platform through Welwyn Garden City station (neither was due to make a stop there).	**OR2c**: Train is cancelled at next station once it is at least 5 minutes late.

Table 1. Description of delay scenario examples and operational responses tested.

In each of Figures 5 to 10, trains that are delayed as a consequence of the initial delaying incident are plotted. The delays are calculated at each timing point (station arrival, departure and designated signals) that the trains pass. Additionally, the delays of all trains in the simulation area are summed to give the overall system delay against time.

		Operational Response		
KPI	**Key Measure**	**OR1a**	**OR1b**	**OR1c**
Resilience	Maximum total delay (s)	1877	1136	762
	Time to recover (mm:ss)	51:44	33:00	42:46
	Delay area (× $10^6 s^2$)	2.571	1.200	0.874

Table 2. The key measures for example 1; refer to Figures 5 to 7.

5.1 Analysis for Example 1

As mentioned before, in this scenario a signal on the approach to Welwyn Garden City station from the north fails for ten minutes beginning at 07:24. This situation first affects train TS158 travelling from the north towards London, as it is the first service to arrive at that signal. In Figure 5, showing the development of delays without operational interventions (OR1a), it can be seen that train TS158 develops a delay of 600 seconds as a direct consequence of the signal failure. Through the course of its journey, the delay to TS158 increases slightly decreasing by a small amount.

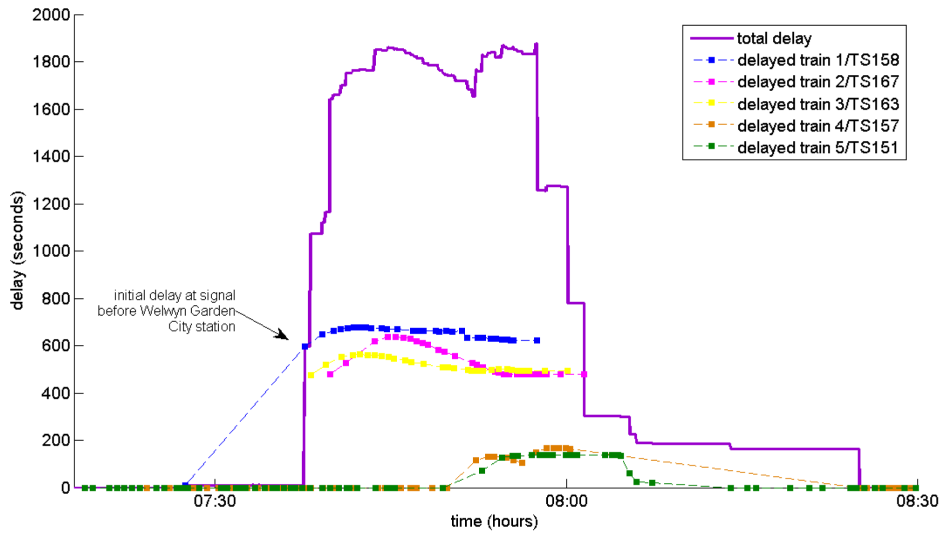


Figure 5. Example 1: signal failure before Welwyn Garden City station. No operational measures are taken (OR1a).

The operational measures taken in this example are first the diversion of TS158 along an alternative track and platform at Welwyn Garden City station (OR1b) and second the diversion of both TS58 and TS163 through Welwyn Garden City station (OR1c).

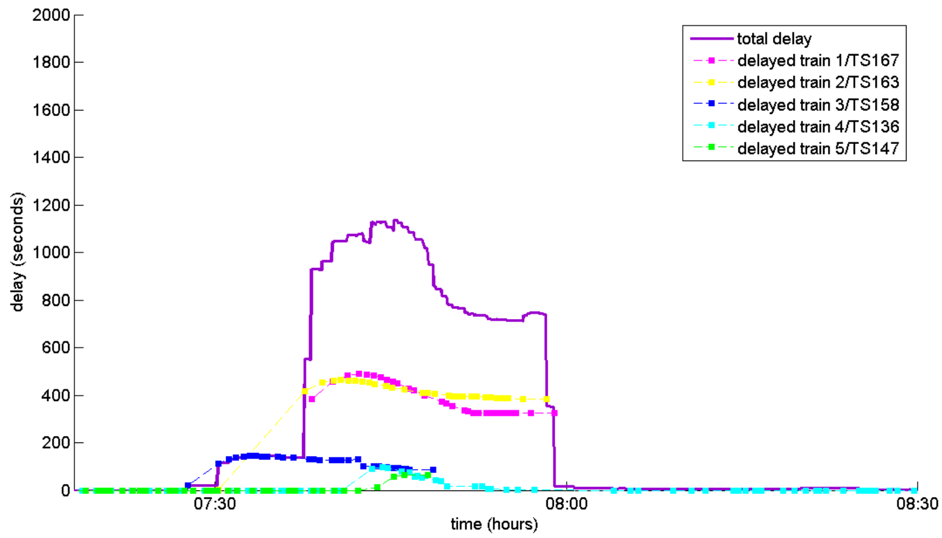


Figure 6. Example 1: signal failure before Welwyn Garden City station. Train TS158 rerouted through Welwyn Garden City station (OR1b).

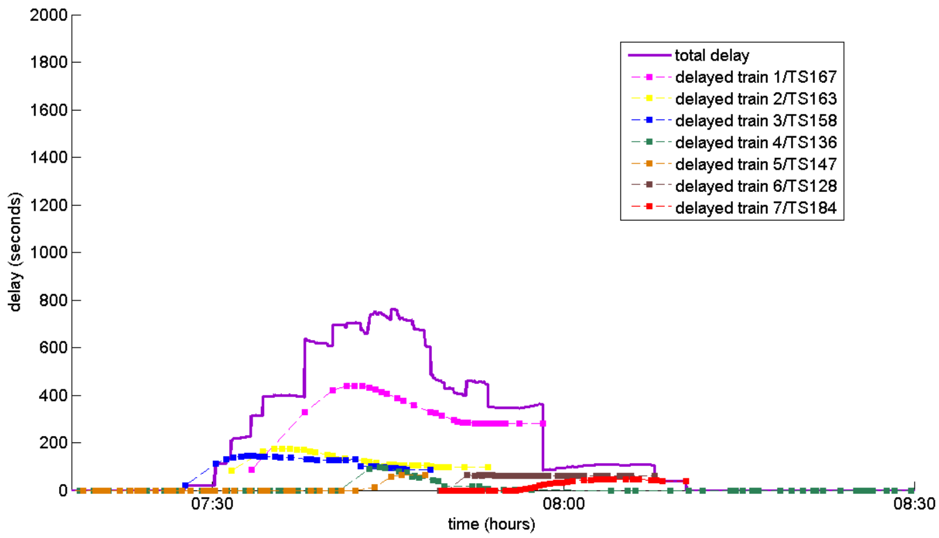


Figure 7. Example 1: signal failure before Welwyn Garden City station. Trains TS158 and TS 163 rerouted through Welwyn Garden City station (OR1c).

It can be seen from Figures 5 to 7 that, as a result of the initial delaying incident and subsequent interventions, 5, 5 and 7 trains were delayed during their journeys in the absence of interventions (OR1a) and with interventions (OR1b and OR1c), respectively. Without operational interventions (OR1a), three trains (TS158, TS167, TS163) had a delay of at least 300 seconds for the majority of their journeys within the control area; with interventions OR1b and OR1c this was reduced to 2 (TS167, TS163) and 1 (TS167) trains, respectively, those that were not diverted through Welwyn Garden City station. The overall severity of the delay situation within the control area is encompassed by the key measures of the resilience KPI: maximum total delay, time to recover and delay area. Without intervention, it is clear that the overall delay situation was the most severe, with the highest maximum total delay of 1877 seconds, longest time to recover (0:51:44) and largest delay area, 2.571×10^6 s^2 of the three operational responses tested. The resilience key measure values clearly reflect that OR1a is the worst performing of the three. Taking operational responses OR1b and OR1c reduce the effect of the delaying incident compared to no operational response. With intervention OR1b, the time to recover was shorter than with OR1c (0:33:00 compared to 0:42:46), but the maximum total delay (1136 seconds versus 762 seconds) and delay area (1.200×10^6 s^2 versus 0.874×106 s^2) were significantly greater. It is therefore less clear which of OR1b and OR1c was the better response to the delaying incident

If the preference within the system under consideration is for the system to return to an on time state as quickly as possible, it would be considered that OR1b performed better than OR1c. However, OR1c introduces less severe delays to other trains and less delay in total, which could be considered to better meet the needs of a system.

The resilience objective function will be formed of a weighted sum of the resilience key measures. An assessment based on the knowledge and opinions of a range of expert representatives of relevant stakeholders will be carried out to determine the weighting values. This format will allow the resilience KPI to be used to fully quantify and compare different operational responses to situations in terms of resilience.

		Operational Response		
KPI	**Key measure**	**OR2a**	**OR2b**	**OR2c**
Resilience	Maximum total delay (s)	2417	593	1165
	Time to recover (mm:ss)	n/a	27:08	30:11
	Delay area ($\times 10^6 s^2$)	3.099	0.408	0.859

Table 3. The key measures for example 2; refer to Figures 8 to 10.

5.2 Analysis for Example 2

In this example the initial delaying incident is the partial loss of power to train TS158. It can be seen in Figure 8 that this causes the particular train to become more and more delayed throughout its journey and to cause delays to 6 other trains. The maximum total delay is 2417 seconds, while the delay situation has not recovered by 08:30. Therefore values for time to recover and delay area cannot be calculated in the same way as described in Section 5 and in the previous example. However, time to recover and delay area key measures must still be included in the performance evaluation function.

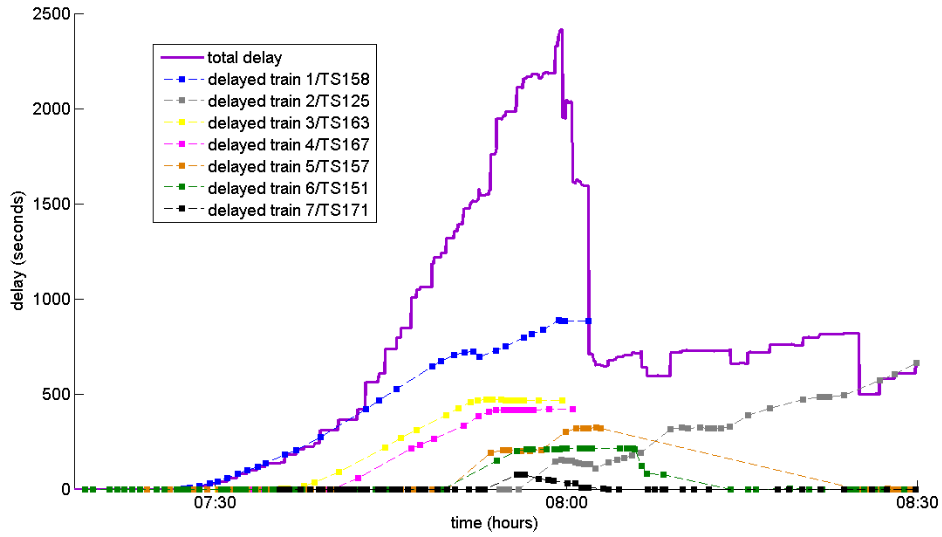


Figure 8. Example 2: Loss of power to TS158 beginning at 07:24. No operational measures are taken (OR2a).

The delay area may be calculated as the area under the system delay curve between the start of the delay at 07:24 until the end of the time period under consideration, in this case 08:30, when the delay area is 3.099×106 s^2. It is suggested that the time to recover may be recorded as the time from the start of the incident until the end of the time period under consideration (here, 08:30) with a penalty applied for non recovery.

In the final performance evaluation function, a significant penalty must be included for the cancellation of a train service in order to represent the significance and disruption that this action causes, due to resources appearing in the wrong place and the inevitable disruption to customers. This will be incorporated into the accommodation and resource usage KPIs.

When operational measures are taken, namely the cancellation of TS158 at the nearest station when it is at least 3 minutes late and 5 minutes late (for OR2b and OR2c, respectively), it can be seen that the overall delay situation considerably improves (see Figures 9 and 10). The maximum total delay is 593 seconds, the time to recover 0:27:08 and the delay area 0.408×106 s^2 when the train is cancelled at New Barnet station (OR2b), while the key measure values are 1165s, 0:30:11 and 0.859×106 s^2 when the train is cancelled 2 stations later on its journey, at New Southgate station.

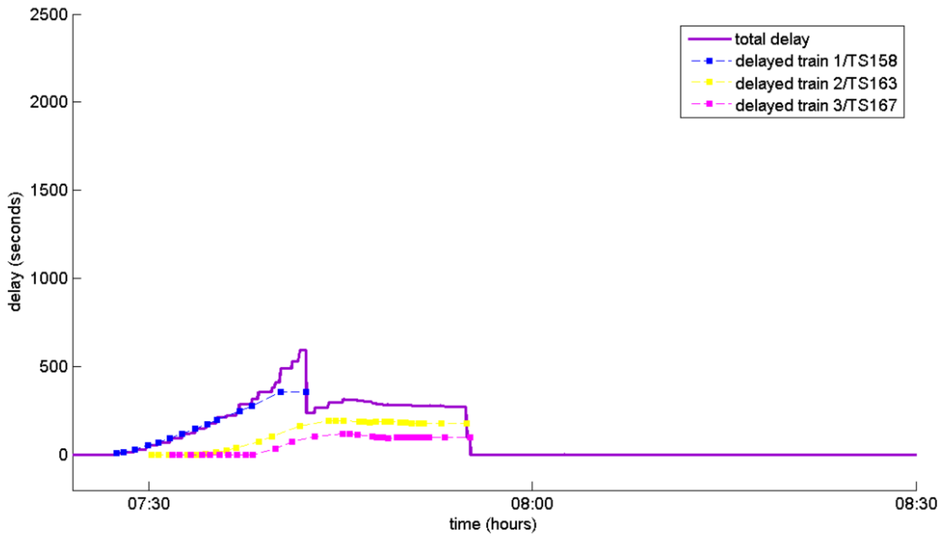


Figure 9. Example 2: Loss of power to TS158 beginning at 07:24. TS158 cancelled at New Barnet station (OR2b).

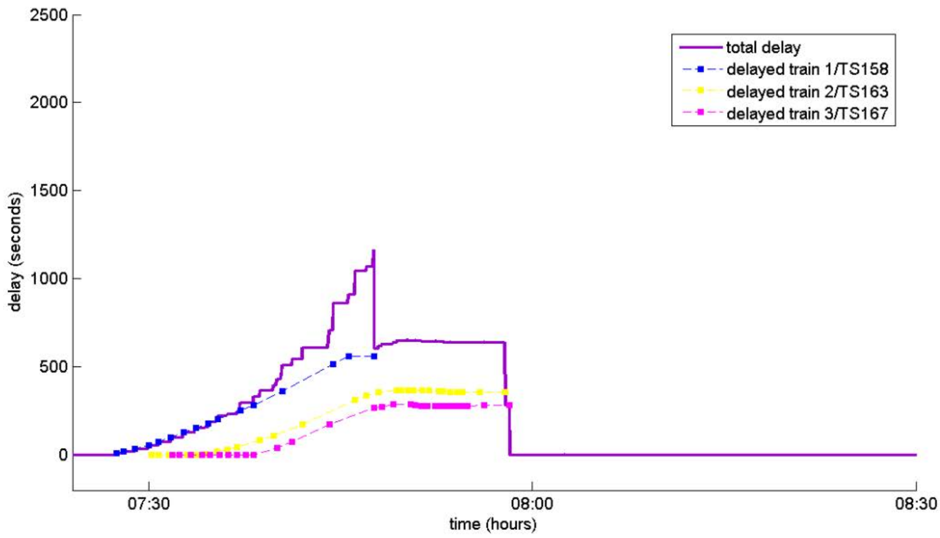


Figure 10. Example 2: Loss of power to TS158 beginning at 07:24. TS158 cancelled at New Barnet station (OR2c).

6 Conclusion

A quality of service framework is proposed as a framework for determining the performance of a railway system from the perspective of all relevant parties. The framework provides a comprehensive coverage of all factors influencing railway network performance and, thus, capacity. Conventional capacity measures (UIC 405, UIC 406, CUI) have a narrower view based respectively on the needs of infrastructure and timetable planning. Changes made to a railway network's properties in an effort to increase its performance may be measured using the eight key performance indicators of the quality of service framework. These KPIs will form the basis of a performance evaluation function that will be used to quantify and compare the interventions proposed to respond to a delaying incident in order to deduce which approach has the most favourable effect on network performance. The function will be used within the European FP7 ON-TIME project which is working to find new methods to maximise capacity, make better use of existing capability and manage delays through improved timetabling, real time traffic management for small delays, operational interventions to large delays and driver advisory systems.

Two examples demonstrating the quantification of the key measures of resilience have been presented in the paper. Different operational responses to two delaying incidents were modelled using railway simulation software. In each scenario, different levels of resilience are achieved with the different operational responses taken.

The first example, the failure of a signal for ten minutes, would potentially be dealt with using real-time operational interventions (Theme 2 of ON-TIME). Therefore, within the ON-TIME project the journey time, punctuality, connectivity, energy and resource usage key measures would have to be considered in addition to the resilience KPI.

In the second example, a train lost power and was cancelled at different points during its journey. The necessity to cancel the train places this situation within Theme 3 of the ON-TIME project and, therefore, the final evaluation function would not consider the energy usage as in the previous example, but would include the evaluation of the number of seats available to passengers, the accommodation.

The quantification of the remaining key performance indicators (see Figure 2) and the weightings of each of the key measures with respect to one another must be established in order to fully evaluate the network performance using the quality of service framework presented here.

Future work will focus on the application of the framework for the evaluation of the performance of current systems as well as the comparison and evaluation of system improvement approaches. This research provides a possible framework for optimising railway system upgrading plans. This framework will be applied in the benchmarking of timetables, traffic management and operational management strategies within the ON-TIME project.

Acknowledgements

The authors would like to thank the ON-TIME partners for their support in the research that underpins the present paper.

References

[1] S. Ison, M. Frost, and R. Watson, "UK rail transport: a review of demand and supply," Proceedings of the Institution of Civil Engineers-Transport, vol. 165, pp. 225-234, Aug 2012.
[2] R. Watson, D. Gillingwater, and A. Boodoo, "Using network capacity effectively where train operators compete for paths - the UK experience," presented at the World Congress on Railway Research, Edinburgh, UK, 2003.
[3] Network Rail, "Meeting the capacity challenge: the case for new lines," Network Rail2008.
[4] R. M. P. Goverde, "Railway timetable stability analysis using max-plus system theory," Transportation Research Part B: Methodological, vol. 41, pp. 179-201, 2007.
[5] UIC, "UIC leaflet 406: capacity," ed: International Union of Railways, 2004.
[6] L. Lucchini, A. Curchod, and R. Rivier, "Transalpine rail network: a capacity assessment model (CAPRES)," presented at the First Swiss Transport Research Conference, 2001.
[7] I. T. Albrecht, O. Brunger, E. Dahlhaus, and R. Goverde, Railway timetable & traffic: analysis, modelling, simulation: Eurailpress, 2008.
[8] R. L. Burdett and E. Kozan, "Techniques for absolute capacity determination in railways," Transportation Research Part B-Methodological, vol. 40, pp. 616-632, Sep 2006.
[9] N. Moreira, L. Garcia, and P. Catarrinho, "Network capacity," in Computers in Railway Six. vol. 15, ed, 2004, pp. 35-43.
[10] D. Woodland, "Optimisation of automatic train protection system," PhD, Department of Mechanical Engineering, The University of Sheffield, Sheffield, 2004.
[11] E. Kontaxi and S. Riccia, "Railway Capacity Handbook: A Systematic Approach to Methodologies," Procedia - Social and Behavioral Sciences, vol. 48, pp. 2689-2696, 2012.
[12] H. Krueger, "Parametric modeling in rail capacity planning," in Simulation Conference Proceedings, 1999 Winter, 1999, pp. 1194-1200 vol.2.
[13] ON-TIME Consortium. (2012). ONTIME. Available: http://www.ontime-project.eu/home.aspx
[14] UIC, "UIC leaflet 405-1: method to be used for the determination of the capacity of lines," ed: International Union of Railways, 1983.
[15] AEA Technology, "Capacity utilisation indices methodology," 2005.
[16] J. Armstrong and J. M. Preston, "Automating the production of train graphs and the calculation of CUI values," presented at the 3rd International Seminar on Railway Operations Modelling and Analysis, Engineering and Optimisation Approaches, 2009.
[17] M. Wahlborg, "Banverket experience of capacity calculations according to the UIC capacity leaflet," presented at the COMPUTERS IN RAILWAYS IX Dresden, Germany, 2004.
[18] A. Landex, Kaas, AH, Schittenhelm, B & Schneider-Tilli, J, "Evaluation of railway capacity," Proceedings of Trafficdays, 2006.
[19] Strategic Rail Authority, "Capacity utilisation policy: network utilisation strategy," 2003.

[20] S. Gibson, G. Cooper, and B. Ball, "Developments in transport policy - the evolution of capacity charges on the UK rail network," Journal of Transport Economics and Policy, vol. 36, pp. 341-354, May 2002.
[21] UIC, "UIC leaflet 405 OR: links between railway infrastructure capacity and the quality of operations," ed: International Union of Railways, 1996.
[22] C. Roberts, F. Schmid, P. Connor, V. Ramdas, and J. Sharpe, "A new railway system capacity model," 2010.
[23] R. L. Burdett and E. Kozan, "A disjunctive graph model and framework for constructing new train schedules," European Journal of Operational Research, vol. 200, pp. 85-98, 2010.
[24] R. L. Burdett and E. Kozan, "Techniques for inserting additional trains into existing timetables," Transportation Research Part B: Methodological, vol. 43, pp. 821-836, 2009.
[25] ON-TIME, "Principles, definitions and requirements," University of Birmingham2012.
[26] T. L. Saaty, The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation. New York: McGraw Hill, 1980.
[27] First Capital Connect, "Train tmes," 11 December 2011 to 8 December 2012.
[28] Graffica. (2013, 11/05). Rail simulation. Available: http://www.graffica.co.uk/rail-traffic-management/simulation/