ISBN 978-82-326-3742-3 (printed ver.)
ISBN 978-82-326-3743-0 (electronic ver.)
ISSN 1503-8181

Norwegian University of Science and Technology

Doctoral thesis

Anette Østbø Sørensen

Doctoral theses at NTNU, 2019:69

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Engineering
Department of Mechanical and Industrial
Engineering

Doctoral theses at NTNU, 2019:69

Anette Østbø Sørensen

# Exploiting available data sources for ex-post evaluation of railway projects

## Case illustrations with traffic and mobile phone data

Norwegian University of Science and Technology

Anette Østbø Sørensen

# Exploiting available data sources for ex-post evaluation of railway projects

## Case illustrations with traffic and mobile phone data

Thesis for the Degree of Philosophiae Doctor

Trondheim, March 2019

Norwegian University of Science and Technology
Faculty of Engineering
Department of Mechanical and Industrial Engineering

NTNU
Norwegian University of Science and Technology

**NTNU**
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Engineering
Department of Mechanical and Industrial Engineering



ISBN 978-82-326-3742-3 (printed ver.)
ISBN 978-82-326-3743-0 (electronic ver.)
ISSN 1503-8181

Doctoral theses at NTNU, 2019:69

Printed by NTNU Grafisk senter

# Preface

This thesis is prepared in partial fulfilment of the requirements for the degree of Philosophiae doctor (PhD) at the Department of Mechanical and Industrial Engineering under the Faculty of Engineering Science and Technology at the Norwegian University of Science and Technology (NTNU). The work was conducted from September 2014 until November 2018.

The PhD research has been carried out in close collaboration with my main supervisor, Professor Nils O. E. Olsson at the Department of Mechanical and Industrial Engineering. My co-supervisors are research scientist Andreas Dypvik Landmark at SINTEF Digital and senior research scientist Agnar Joahnsen at SINTEF Building and Infrastructure.

My interest and curiosity in utilising available data sources to support processes in companies started growing during the final year of my master's degree studies and my work with the Department of Technology Management at SINTEF.

While working on this PhD, I have experienced both successful and challenging periods. One of the more demanding challenges I found throughout these studies was moving between disciplines. With my background in industrial mathematics, I am not unfamiliar with numbers, but I was less familiar with the project management discipline. The interdisciplinarity made the PhD work both challenging and rewarding in the struggle to recognise each disciplines' integrity - in the way they think and the way they write.

# Acknowledgements

This thesis would not have become a reality without those who have contributed to the research and supported me throughout the process.

First and foremost, I would like to thank my supervisor, Nils Olsson, for invaluable guidance and support during the project period, for his positive engagement, for connecting me to people in the industry and research field, and for invaluable discussions and guidance through the final stages of writing the dissertation.

I would also like to thank my co-supervisors, Andreas Landmark, for constructive discussions on opportunities and methods, and Agnar Johansen, for informative talks and for his interest in my work. And both for the helpful discussions and feedback about writing the dissertation.

In addition, I wish to thank several co-authors who have played an important role in this research. Thanks to Johannes Bjelland at Telenor, Siva Ekambaram, Heidi Bull-Berg and Andreas Seim at SINTEF, and Mohsin Akhtar for their cooperation. I have learned a lot through working with them.

I would like to thank Telenor, BaneNor and other railway operators for granting me access to data.

Thanks are also due to Johan Nyström and Ida Kristoffersson at the Swedish National Road and Transport Research Institute (VTI) Stockholm, Johan for organising the VTI-seminar and Ida for playing the role of opponent and providing valuable feedback on the study I presented.

I am very grateful for the opportunity I had to stay at the Construction Industry Institute (CII) in Austin, Texas, in the United States, and for the opportunity to attend the CII Seminar while I was in Austin. I would therefore like to thank Stephen Mulva, director of CII; Daniel Oliveira, research associate at CII; and Terri Buvia, Cynthia King, Hyeon-Yong Park and the other co-workers and students at the CII office.

I want to thank SINTEF Technology and Society for providing me with office space. Thanks also go to Tor Nicolaisen at BaneNor for providing me a desk at the offices at Marienborg in the early phase of the PhD work.

I also want to thank my fellow PhD researchers at NTNU and colleagues at SINTEF Digital, Department of Technology Management, who have made the time at the office more enjoyable.

I am also thankful to all my friends in Trondheim with whom I have spent time outside of the office, especially Karen Synnøve for our coffee breaks during the first years, and Sandra, Marcin and Tomas.

I am very grateful to my family, my parents Elin and Arne, and my sister

Benedikte for their unwavering support and encouragement and for always showing interest in my work and giving me energy.

Finally, I want to thank my boyfriend Are Bertheussen for his support and encouragement, for providing me with good distractions and for helping me to worry less.

Trondheim, November 2018
*Anette Østbø Sørensen*

# Summary

Technological developments increase the amount of available data sources in all industries. This thesis focuses on how we can exploit these data sources in ex-post project evaluation to support evaluators with new insight. The focus has been on railways and railway projects.

A project should be evaluated to establish to what degree the intended effects on the users and the society are achieved. The ex-post evaluation should be carried out a few years after major infrastructure projects are completed. One matter that should be evaluated is tactical success, which is measured in terms of effectiveness. Effectiveness measures if the project achieved its goals and is typically assessed based on the change of state before and after. However, evaluators experience problems in getting hold of essential data on the pre- and post-situation. For the evaluators, data collected from the railway operations are mainly available as aggregated numbers from performance and evaluation reports.

Typical goals of railway projects are to reduce travel time and increase capacity and demand. Two measures of particular interest are therefore punctuality and traffic volume. Punctuality is whether the traffic runs according to the timetable, and statistical numbers on punctuality are well-developed measures. However, a deeper evaluation of punctuality can be provided by analysing delay propagation, and that has been more difficult to obtain a good measure on. The number of travellers is an important performance measure that has been difficult for evaluators to obtain good data on, often because train operators consider such data confidential business information. In addition, there is a challenge regarding varying quality and coverage of the available data on the number of travellers.

Data of relevance to ex-post project evaluation are generated and collected from both the construction phase and the operation phase. The data collected from railway operations are useful when evaluating the tactical and strategic success of a project. The conclusions drawn from this evaluation can provide useful insight and learning to the strategic planning and concept development of future projects. The author has provided two practical examples of how data generated and collected during railway operations can be exploited to obtain relevant information for the evaluation. This was done through thorough investigation into two measures of particular interest for ex-post evaluation of railway projects, i.e. delay propagation and number of travellers.

Delay propagation was analysed based on traffic data, and a method was developed to find cases of knock-on delay on single tracks. The tool allows an attempt at indicating the direction of knock-on effects. In addition, the method traces the propagation of delay from one train to the next to find the networks of dynamic delay propagations. Mobile phone data were investigated as an alternative source of the number of travellers. The possibility of using mobile phone data is interesting because it is independent of the railway operators. The study showed that it is possible to combine mobile phone data with railway infrastructure and train traffic data. The findings show a potential for utilising mobile phone data to collect the number of travellers on the railway.

The data collected from technologies used during the construction project are useful when evaluating the project execution. This includes the experience report, which is the internal form of knowledge sharing, and the evaluation of efficiency, which examines time, cost and quality. These evaluations are important for learning from similar projects during the construction phase and can contribute to faster decisions.

Ex-post project evaluation has traditionally been qualitative, but because of technological developments, more data from the railway operations are available. It is still a problem to get hold of data, especially when it is about private information, while other data sources have become more available. The two empirical studies are good examples of how such data can provide useful insight into the ex-post evaluation of railway projects.

# List of Publications

This dissertation consists of the following publications, which are referred to in the text by their corresponding Roman numerals.

I. Sørensen, A. Ø., Landmark, A. D., Olsson, N. O. E. & Seim, A. A. (2017). Method of analysis for delay propagation in a single-track network. *Journal of Rail Transport Planning & Management*, 7(1), 77-97. doi:10.1016/j.jrtpm.2017.04.001.

II. Sørensen, A. Ø., Bjelland, J., Bull-Berg, H., Landmark, A. D., Akhtar, M. M. & Olsson, N. O. E. (2018). Use of mobile phone data for analysis of number of train travellers. *Journal of Rail Transport Planning & Management*, 8(2), 123-144. doi:10.1016/j.jrtpm.2018.06.002.

III. Sørensen, A. Ø., Olsson, N. O. E., Akhtar, M. M. & Bull-Berg, H. (2019). Approaches, technologies and importance of analysis of number of travellers. *Urban, Planning and Transport Research*, 7(1), 1-18. doi:10.1080/21650020.2019.1566022.

IV. Sørensen, A. Ø., Olsson, N. O. E., & Ekambaram, A. (2015). Evaluation and learning—Experiences from a construction project in Norway. *Procedia Economics and Finance*, 21, 510-517. doi:10.1016/S2212-5671(15)00206-3.

V. Sørensen, A. Ø., Olsson, N. O. E. & Landmark, A. D. (2016). Big Data in Construction Management Research. I: *Proceedings of the CIB World Building Congress 2016, Volume III. Building up business operations and their logic. Shaping materials and technologies*. Tampere University of Technology, ISBN 978-952-15-3743-1, 405-416.

## Related publications not included in the thesis

- Ekambaram, A., Sørensen, A. Ø., Bull-Berg, H. & Olsson, N. O. E. (2018). The role of big data and knowledge management in improving projects and project-based organizations. *Procedia Computer Science*, 138, 851-858. doi:10.1016/j.procs.2018.10.111.

- Olsson, N. O. E., Sørensen, A. Ø. & Leikvam, G. (2015). On the need for iterative real estate project models. *Procedia Economics and Finance*, 21, 524-531. doi:10.1016/S2212-5671(15)00208-7.

# Contents

# List of Figures

# Chapter 1

# Introduction

This chapter presents the background and motivation for the research, the aim and research questions, research scope, process and publications, and the structure of the thesis.

## 1.1 Background

Access to good relevant data is a common challenge in project evaluations (Samset, 2003; Frumin, 2010), and evaluators experience problems in getting hold of essential data on the pre- and post-situation (Volden & Samset, 2013). The situation is similar in Norway and Sweden, where the train operators consider some relevant data confidential business information, which may seem like a general problem for railways. In ex-post evaluation, the effectiveness of railway projects is traditionally evaluated by the four measures of travel time, frequency, punctuality and volume of traffic. Travel time and frequency are easy to measure based on the timetable (Volden & Samset, 2013), and punctuality is in itself well developed (e.g. Olsson & Haugland, 2004; Goverde, 2005; Palmqvist et al., 2017). However, a deeper evaluation of punctuality by analysing delay propagation, which also is a measure of robustness, has been more difficult to obtain, as has volume of traffic, also referred to as ridership. The focus of this dissertation is therefore on utilising available data sources to support ex-post evaluation of railway projects, which creates new opportunities in measuring delay propagation and ridership.

Ex-post evaluation should be carried out a few years after major transportation infrastructure projects are completed. These evaluations are typically done based on the situation at one point in time. The ex-post evaluation is central to establishing to what degree the intended effects on the users and society are achieved (Volden & Samset, 2013). It is also useful because it provides information that can be utilised in upcoming projects. The success of a project is evaluated by determining the project's efficiency, effectiveness, impact, relevance and sustainability (OECD, 2000; Samset, 2003). Tactical success is measured by effectiveness, which is the extent to which the goal, also described as the tactical objective, has been achieved, unlike impact, which is all other positive and negative changes and effects caused by the project. Key issues in evaluations are data quality and availability (Samset, 2003; Parthasarathi & Levinson, 2010; Olsson et al., 2015b). New types of data can add both precision and new perspectives to evaluations (Tanaka, 2015) and have potential to complement traditional qualitative evaluation methods with a quantitative approach (Olsson et al., 2015b).

People want a robust and reliable railway with trains that arrive on time. Railway projects are therefore often carried out to improve the performance of the railway service. Evaluation of effectiveness is typically based on the change of state before and after the project is carried out (Volden & Samset, 2013). Thus, evaluating a railway project's effectiveness usually involves evaluating the performance of railway service after the project compared to how the railway performed before the project.

Typical goals of railway projects are to reduce travel time and increase capacity and demand (Volden & Samset, 2013; Harris et al., 2016). This is why the four measures traditionally used to evaluate the effectiveness of a railway project are travel time, frequency, punctuality and volume of traffic (Volden & Samset, 2013). Infrastructure projects have in addition goals related to safety, environment and economic development (Olsson et al., 2015b). Controlling whether the project achieved reduced travel time is easily done by studying the timetable. Improved capacity will be evident through increased frequency and improved punctuality (Volden & Samset, 2013). Punctuality is a necessary prerequisite for increasing capacity utilisation (Goverde, 2005). Frequency, measured as the number of trains using the tracks, is easy to check by studying the timetable. Measures of punctuality and ridership are more difficult to obtain good data on and are usually only available to the evaluators as aggregated numbers through the performance and evaluation reports, as illustrated in Figure 1.1.

In general, the performance of a railway first of all involves measuring the efficiency of the railway service. General measures of efficiency are related to time and volume (Olsson et al., 2015b). Time is typically about travel time reduction, while volume includes measures of the number of passengers and volume of freight. The number of travellers is a measure of demand for transportation services, which is important information for planning and evaluations. Railway operations are often evaluated by key performance indicators based on ridership, such as the number of passenger journeys made, passenger kilometres, average passenger trip length and average passenger volume (Vuchic, 2005). Achieving an efficient railway system is essential for quality of operation. Most performance evaluation methods applied in railways are therefore directly or indirectly based on quality of operation (Hansen & Pachl, 2014). Quality of transport service and passenger satisfaction are defined by factors such as punctuality, regularity, reliability, comfort, predictability, safety and service (Goverde, 2005; Hansen & Pachl, 2014; Harris et al., 2016). The criteria of punctuality and regularity are comparative measures between planned and actual train services. They examine if the customers got what they paid for at the agreed-upon time (Harris et al., 2016). Punctuality is one of the most important quality factors in railway operations (Seco & Gonçalves, 2007; Harris et al., 2016).

All railway companies undertake checks, reporting and analysis to evaluate the operation and see if the railway service is being delivered according to the timetable (Harris et al., 2016). Data are generated and collected from the railway operations,

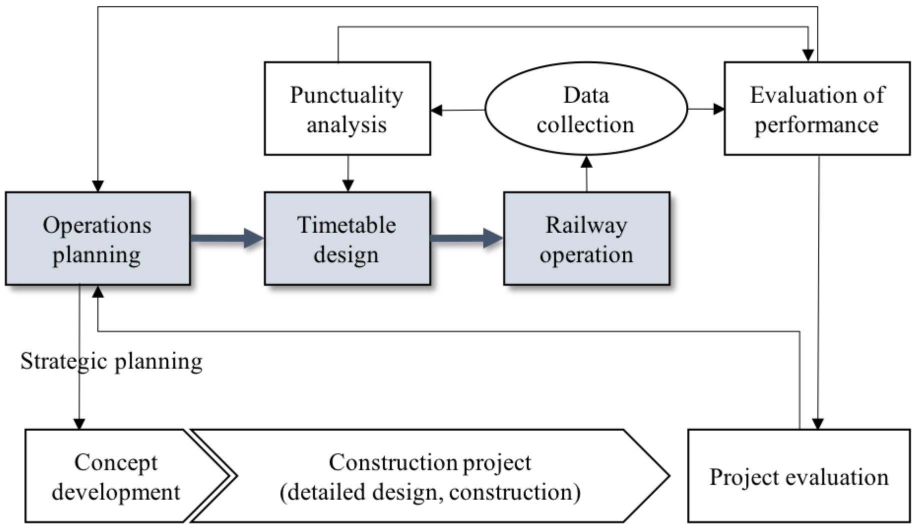


**Figure 1.1: Illustration of the evaluation and feedback loop in railway operations.**

including, but not limited to, punctuality data, as illustrated in Figure 1.1. Figure 1.1 is inspired by Goverde's figure on feedback loops in robust timetable design (Goverde, 2005, Figure 1.3), but it is modified to include the evaluation processes and projects. The information acquired through this evaluation of performance is, among other things, fed back to timetable design and operations planning, as Figure 1.1 shows. The Norwegian National Rail Administration (former Jernbaneverket, now Bane NOR) publishes annual official railway statistics that include punctuality numbers and aggregated data on number of travellers, passenger kilometres, and number of sold single tickets and monthly tickets.

As argued, punctuality and ridership are important measures when evaluating railway projects. Punctuality in railways means that train traffic runs according to the timetable (Harris et al., 2016). The definition is usually expressed as the percentage of trains with a delay of less than a certain time in minutes and often on arrival at the final destination. It is in itself a well-developed measure (Hansen, 2001; Goverde, 2005). The cause of low punctuality performance is, among other things, an insufficient degree of robustness of timetable design (Hansen & Pachl, 2014). Furthermore, delays on the railway section that has been improved through a railway project may have been caused by delays on other parts of the railway (Volden & Samset, 2013), which is a result of delay propagation. Delay propagation reflects the degree of robustness of timetable design and the stability of train operation (Yuan & Hansen, 2007). Statistical reports do not provide measures of delay propagation. In addition, a more robust railway with fewer deviations is a typical objective of railway projects. Approaches to studying delay propagation include analytical models, statistical models and simulations. Simulations have been used extensively and are commonly applied to, for instance, studying the impact on random primary delays (Radtke & Bendfeldt, 2001; Goverde, 2010; Warg, 2013; Lindfeldt, 2015). The Norwegian railway network primarily consists of a star-shaped single-track network, but few models for calculating secondary delays on single-track lines have been published (Handstanger, 2009).

The benefit of reduced travel time and increased capacity is dependent on the volume of traffic (Volden & Samset, 2013). A characteristic feature of railway transport, with both passengers and freight, is the advantage of traffic density. The costs per passenger or per tonne of freight are reduced with increased numbers of passengers or volumes of goods (Harris et al., 2016). Thus, the evaluation should answer the question of whether the railway is used, and more important, whether the railway project resulted in an increase in traffic volume. However, it is difficult for the evaluators to get good data on ridership. The number of travellers is traditionally measured by manual counting at chosen stations on each railway line, and the practicalities for storing and summarising these data are well established (e.g. Vuchic, 2005). In addition, different travel behaviour surveys have been carried out. Using surveys and manual counts provides transport organisations with a reasonable snapshot of existing demand on their transport system. Several automatic passenger counting systems have been developed to provide increased resolution.

When evaluating the effectiveness of railway projects, it is important to have good measures of ridership and a more thorough study of punctuality. The fact that access to good relevant data can be a challenge when evaluating railway investments may seem like a paradox, when the volume of data and different sensor technologies are maturing (Frumin, 2010; Olsson et al., 2015b). Because evaluators usually only have access to aggregated data on punctuality and ridership from performance and evaluation reports, this dissertation uses available datasets to analyse these two parameters. This provides the evaluators with data directly from the data collection. New technologies generate and store large amounts of data about the use, status or performance of equipment. For this reason, there has been a huge interest in and discussion about the consequences of digitalization in all industries and research fields.

This dissertation focuses on data that are available and can be obtained from existing sources. Relevant data that are generated from existing data-capture infrastructures can be categorised based on how they are collected or generated, as suggested by Olsson et al. (2015b). The four categories are internet activity, movements, physical environment and commercial activity. Specifically related to evaluation of railway projects, two types of data are interesting for this study. First is data generated from movements, which include data from signal systems and GPS. Such data can provide indicators of punctuality and actual travel time and have good availability and low cost. Data generated from movements also include data from systems for counting the number of travellers, such as Automatic Passenger Count systems. Train movements in Norway are registered in a database either automatically through the signalling system or manually by the train dispatchers at the respective stations. In recent years, more and more automatic registration systems have been developed (Veiseth, 2009). The practices in other countries may vary. For instance, according to Palmqvist et al. (2017), in Sweden, train movements are derived from the track blocking and signalling systems, while in the Netherlands, they are based on the train describer system described by Kecman & Goverde (2013). As a consequence, the format and structure of the collected data may differ between countries, but it should be possible to restructure them.

The second type of data interesting for this study is generated from the physical environment, including mobile phone data from base stations, data from weather, temperature or snowfall, and data on physical infrastructure from sensors in the infrastructure. Mobile phone data from base stations are expected to be relatively similarly structured and formatted between countries and companies. Data generated through new technologies can be useful to evaluate the effect of the project. For instance, in urban planning, mobile phone data can provide a detailed picture of the mobility patterns of travellers and thus of how the project impacted these patterns in a positive or negative way.

Access to new types of data may have a wide range of application within evaluation. Data can be used to support triangulation and quality assurance; complement and enhance existing evaluation parameters; provide new evaluation parameters; provide quantitative data on the conditions previously based on qualitative assessments; and illustrate effects that have not been possible to visualise previously (Olsson et al., 2015b). This dissertation focuses on utilising the available underexploited data sources that gives opportunities for new methods of analysis to complement and enhance existing evaluation parameters. The research is comprised of studies on how accessible data can be useful in evaluating the effectiveness of railway projects to support the measures of delay propagation and ridership.

## 1.2 Aim, objectives and research questions

Through this thesis, the author wishes to obtain new insights to support evaluators in ex-post evaluations of railway projects with available data sources. This is the aim of the research. The research objectives describe what the research will do to achieve this aim. The key objectives of this dissertation are to

- analyse suitable datasets from existing data-capture-infrastructures,
- identify and test methods to utilise these datasets, and
- discuss the relevance to project evaluation.

The research questions describe what factors and relationships will be investigated and what we need to find out. The research questions of this dissertation are the following.

RQ1 How can one exploit available data sources in railway to obtain information relevant for project evaluation,

a) specifically to evaluate delay propagation with traffic data?

b) specifically with mobile phone data to evaluate ridership?

RQ2 How are such empirical analyses useful for learning from project evaluation in a life cycle perspective?

## 1.3 Research scope

In a project evaluation, the evaluators want a broad collection of data that forms part of the assessment of the socio-economic benefit, including wider economic benefits. A wide range of data collection methods is needed to ensure the quality of the evaluation (Samset, 2003). The question of socio-economic profitability is considered the final test of whether the investment was successful (Volden & Samset, 2013). As mentioned in Section 1.1, in the work of evaluating railway projects, there are four areas to assess—travel time, frequency, punctuality and traffic volume. These measures of effectiveness are closely related to socio-economic benefit. The author has addressed two areas of great interest, delay propagation and number of travellers. Digitalization provides a better basis for better analysis. There is a lot of data one could have collected. Regarding data on the number of travellers, other alternatives are elaborated in Publication III. Regarding a deeper analysis of punctuality, railway planners and infrastructure managers are interested in understanding and quantifying the influences of delay causes, not just statistics. As the availability of train traffic data increases, there are large opportunities for analytical use of these data to gain an improved understanding of delays and their causes. As already mentioned, the author has done a thorough investigation into the two areas of delay propagation and number of travellers in Publication I and Publication II (see Figure 1.3).

The scope of the research is described through the specific data used for the research, limitations of the research, and delimitations of the research. The empirical data used in this dissertation are presented in Figure 1.2 with an overview of which publications the data are used in. The white boxes represent sources that are relatively static, while the grey databases are sources with continuous filling. For Publication IV, data on punctuality and ridership were obtained as aggregated data from performance and evaluation reports. In addition, timetables were used, and data were gathered from interviews. The work of this thesis seeks to improve the evaluation measures by going directly to the high resolution data, as opposed to using aggregated data from performance and evaluation reports. To analyse the punctuality measure in Publication I, traffic data from railway operations (TIOS) and delay cause registrations were used. To study ridership in Publication II, mobile phone data were obtained, including handset counts on mobile phone base stations, location data for base stations and automatic passenger counts on trains. In addition, location data for railway tracks and train stations were used for Publication II.

The limitations in the data analysis are elaborated in Publications I and II. Another limitation of this thesis is access to relevant data. Specifically, the process to gain permission for using the mobile phone data required some time and thus indirectly formed this part of the research.

The delimitations of this thesis include, first of all, that the studies in this thesis

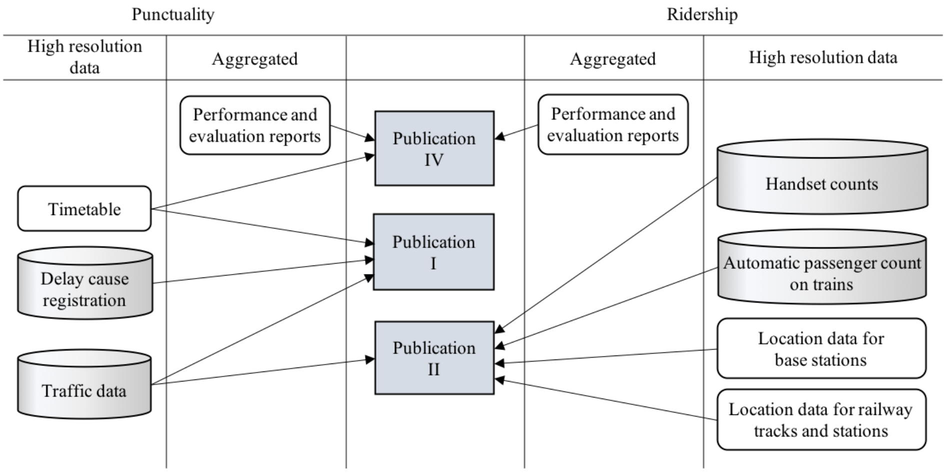


**Figure 1.2: Illustration of the empirical data used in this dissertation.**

are based on data from Norwegian railway operations and Norwegian companies. However, the data, structure and the methods developed are generalizable to other countries. Second, this thesis does not focus on choosing the right project concept, but rather evaluating whether the project achieved the project goals. However, such information can be relevant when choosing or evaluating project concepts. Finally, traffic data and mobile phone data used in this thesis are available as a consequence of technological developments and digital transformations in railways. However, other areas of digitalization in the railway are not considered.

## 1.4 Research process and publications

This dissertation is built on five peer-reviewed publications in scientific journals and conference proceedings. The procedure of this PhD work consists of development of the project plan, complete courses, development of each of the publications and development of the dissertation. The project plan has been adapted during the PhD period, which is necessary to respond to both challenges and opportunities (Croom, 2010). The research questions have been modified since the author started on the PhD as a natural consequence of continual literature review (Croom, 2010). The aim and objectives have been more or less the same since the project description was written. The dissertation connects the five publications and argues that they together answer the research questions and the aim. Figure 1.3 illustrates how the publications answer the research questions. Publication I answers research question RQ1a, and Publications II and III answer research question RQ1b. Research question RQ2 is answered by Publications IV and V, and partially by Publications I, II and III. The three journal papers have been subject to extensive peer review and have been revised based on the reviewers' comments. The two conference papers were also subject to peer review.

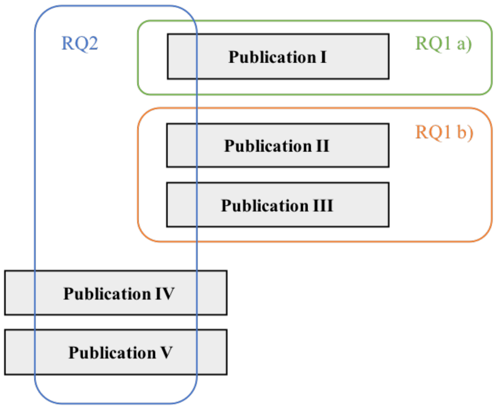


**Figure 1.3: Illustration of how the publications answer the research questions.**

## 1.5 Structure of the dissertation

Chapter 2 of the dissertation presents a literature review of relevant research areas. Chapter 3 describes the research methods and design. In Chapter 4, the research findings of each of the publications are presented. Chapter 5 offers a conclusion and suggests further work. The conference papers and journal articles are attached in the appendix.

# Chapter 2

# Literature Review

This chapter presents a literature review of the main topics that are relevant in this dissertation. The aim of this dissertation is to obtain new insights to support evaluators in ex-post evaluations of railway projects with available data sources.

The main aspects of project evaluation are introduced in Section 2.1. Evaluating a railway project's effectiveness examines whether the project achieved the goals, such as reduced travel time, increased capacity and increased demand, which are measures of railway performance. Section 2.2 therefore presents characteristics of railway operation. It introduces the operations planning, including the importance of a stable and robust timetable design to limit the impacts of small disturbances and achieve a reliable train operation, and factors that affect capacity utilisation. Furthermore, it investigates how larger disturbances in the railway operations are handled with dispatching. Finally, the focus shifts to evaluation of the performance of railway operations and different measures used to assess the quality of operation. These include punctuality, which is one of the most important quality factors in railway operations; delay propagation and published methods on how to analyse delay propagation, which reflects the degree of robustness of timetable design and the stability of train operation; and finally, the importance of ridership in key performance indicators.

Section 2.3 presents the increase in data from digitalization and new technologies. First, recent developments are introduced, including the importance to businesses, challenges and privacy issues; developments and data collection in railways; how railway companies currently focus on digital transformation; methods and technologies currently used to collect data on train movement and ridership; and publications on mobile phone data in mobility research.

Section 2.4 gives a concise summary of the literature review and elaborate upon the research gaps addressed in this dissertation.

## 2.1 Project evaluation

An important part of finalizing a project is evaluation and transfer of experience (Rolstadås et al., 2014), illustrated in Figure 1.1 on page 3. An internal evaluation should therefore be carried out to identify lessons from the project. External evaluation is done to see how the project actually performed. Scriven (1991) defines evaluation as the process of determining the merit, worth or value of something. OECD (2000) defined evaluation as a systematic and objective assessment of an ongoing or completed project, program or policy, its design, implementation and results. Ex-post evaluation is evaluation of an intervention after it has been completed, as opposed to an ex-ante evaluation that is performed before implementation (OECD, 2000). An ex-ante evaluation aims to clarify the prerequisites of the project, define its goals and expected effects and contribute to planning (Andersen et al., 2007). In contrast, the intention of an ex-post evaluation is to identify the factors of success or failure, assess the sustainability of results and impacts and draw conclusions that may inform other projects (OECD, 2000).

According to Cracknell (1989), the logical framework was developed in the United States during the 1960s, and was adopted by several foreign aid agencies. It has later been adopted for use in project management in general, and has proven particularly useful for analysing public investments. The framework is used to assess the hierarchy of objectives during project preparation, but it has also a role in ex-post evaluation. Objective-oriented evaluation will be based on assessment of the logic of the objectives (Volden & Samset, 2013). The evaluator should assess whether there is a logical connection and a causal relationship between purpose (strategic objective), goal (tactical objective), output (operational objective) and input of the project.

An evaluation should determine the project's efficiency, effectiveness, impact, relevance and sustainability (OECD, 2000; Samset, 2003). Each of these evaluation criteria evaluates the success of a project on different levels. Efficiency is a measure of how the project performed, meaning whether the outputs are delivered effectively, i.e. on time, on budget and with agreed-upon quality. Efficiency measures operational success, and it is this area that encompasses the most easily accessible information (Volden & Samset, 2013). However, the most important aspects to evaluate are the effects on tactical and strategic levels, which, in contrast, are more demanding to evaluate. In addition, important user groups as well as the financing party and society as a whole are interested in achieving the objectives on tactical and strategic levels (Volden & Samset, 2013).

Success in the tactical perspective is measured in terms of effectiveness. Eikeland (2001) relates effectiveness to how the results of a project contribute to added value for owners and users. In OECD terms, effectiveness measures the realisation of the project's objectives (OECD, 2000). Whether the project achieved the goal means to which extent the anticipated positive consequences of the project have been achieved (Samset, 2003). This is the perspective of the target group, as well as the project owner or financing party that in many types of projects might have a perspective similar to that of the users. Ideally, the assessment of effectiveness should be compared to the hypothetical predicted state if the project had not been carried out. However, typically the change of state before and after is used as a basis to assess the effectiveness (Volden & Samset, 2013). Typical goals of railway projects are to reduce travel time, increase capacity on the network, increase the demand and hence revenue, and/or to reduce operational costs (Volden & Samset, 2013; Harris et al., 2016). Infrastructure projects have in addition goals related to safety and the environment (Olsson et al., 2015b).

Success in the strategic perspective is measured in terms of the project's impact, relevance and sustainability. The concept of impact includes both positive and negative consequences, whether foreseen or not (Samset, 2003). A broad assessment of impact is essential in a comprehensive evaluation. Typical measures used to evaluate the impact of infrastructure investments, such as new double tracks, railway tunnels, new timetables or fare changes, are based on ridership (Vuchic & Newell, 1968). Relevance is about whether the objectives still correspond to valid priorities in society and the users' needs. Sustainability concerns whether the positive effects of the project are likely to continue after the project has been concluded. Sustainability in this model is in large part about the life cycle of projects, which consists of several phases, from definition to completion. In UK, the definite model for the building design and construction process is the Royal Institute of British Architects (RIBA) Plan of Work (RIBA, 2013). The RIBA model has inspired a Norwegian phase model for construction, described in Knotten et al. (2016). The Norwegian model uses a time frame similar to the RIBA model, but highlights the perspectives of owner, user, supplier and society. The Norwegian model has the following phases: strategic definition; concept development; concept design; detailed design; construction; handover; use and facility management; and completion. The effects of a project may be positive for some groups and negative for others, which is why it is important to identify the stakeholders (Harris et al., 2016). Stakeholders of railway projects include the financing party, train operator, passengers taking the train, people living close to the tracks and the general public (Samset, 2003; Hansen & Pachl, 2014). When evaluating sustainability, the evaluator should consider factors such as environmental impact, economic and financial aspects, socio-economic aspects and choice of technology (Samset, 2003). For instance, public traffic such as railway can be argued to be environmental friendly, both in terms of the climate and for the local community (Volden & Samset, 2013). Furthermore, when evaluating economic and financial aspects of railway projects, the evaluator should consider the costs of railway operations and maintenance.

Other types of evaluation methods, apart from the objective-oriented one, are also used to evaluate projects, such as socio-economic analysis. A socio-economic analysis focuses on measuring the effects and impacts generated from a project in monetary value (Volden & Samset, 2013). Many perceive questions about socio-economic profitability the final test of whether a project investment is successful. The socio-economic effects are often calculated based on the change of state before and after.

Project evaluation results in two types of learning. Learning that can be extracted from evaluating the operational objectives is called singe-loop learning. The term was introduced in the 1970s by psychologist Chris Argyris and philosopher Donald Schøn (Volden & Samset, 2013). Single-loop learning provides knowledge about how to improve project execution. In investment projects, this type of learning can contribute to better time and cost control, risk management and contract strategy. The second type of learning is double-loop learning, which extracts experiences from the results of the project and provides knowledge about whether the project was useful to society. The effects and benefits of a project are often not realised for some time after its completion. Evaluations should therefore be conducted a few years after a project is terminated, typically after three to five years (Andersen et al., 2007; Volden & Samset, 2013). It will then be possible to discern a lot about the strategic success of the project. Double-loop learning concerns the project concept and whether it provided the benefits to users and society as assumed. Both single- and double-loop learning are part of the life cycle of investment projects. A study on sharing lessons learned in construction companies in UK carried out by Carrillo et al. (2013) suggests that sharing lessons learned leads to learning for similar projects in the future, avoiding making mistakes and repeating success, providing a competitive edge over other companies and learning lessons for consecutive stages of ongoing projects.

Furthermore, the actual evaluation approach includes a learning aspect. Evaluations should build on acquired knowledge and experiences from previous evaluations so that these approaches are adapted to the specific case (Andersen et al., 2007).

## 2.2 Railway operations

This section identifies characteristics of railway operations, including operations planning, and key concepts for achieving reliable train operation, traffic control in railway operations and performance measures to evaluate the railway operations.

### 2.2.1 Operations planning

Because of the high dependencies between the actors in a railway system, train movements must be planned in great detail. Furthermore, traffic needs to be supervised and coordinated continuously. Harris et al. (2016) provide a thorough introduction to railway operations planning. Operations planning usually consists of three levels that are organised by decreasing length of planning horizon, and mostly with an increasing level of detail. These are strategic (long-term), tactical (medium-term), and operational (short-term) planning. Moreover, the operational planning process is comprised of key elements ranging from infrastructure (e.g. track development and capacity to store rolling stock), timetable as illustrated in Figure 1.1 on page 3 (including turnaround times), train service (e.g. capacity on different routes and extra peak trains) and personnel planning (e.g. personnel resources, rolling stock and maintenance planning) (Harris et al., 2016).

Several aspects of railway operations make it essential to take a long-term perspective. First, timetables, plans for rolling stock operation and personnel plans may change twice each year. In contrast, trains and infrastructure have a considerably longer lifespan than that of any plan for using them (Harris et al., 2016). Flexibility and adaptability are therefore important to consider when infrastructure and trains are acquired. Second, long-term planning normally takes place about every fifth year and is about bigger changes to timetables that are often connected to rolling stock acquisition or larger changes to infrastructure (Harris et al., 2016), as illustrated in

Figure 1.1. Hence, the train planning task needs to consider the land-use planning of the areas around the railway several years in advance. Other factors that need to be taken into account include providing sufficient transport capacity at appropriate speeds, marketability of the service and the laws of physics and those related to health and safety of the workforce. Inevitably, trade-offs will be required between several variables, such as operating cost, seating provision, reliability, service frequency and train length (Harris et al., 2016).

Optimal use of available resources in railway operations "means a balance between an efficient resource usage and appropriately-dimensioned reserves and slack in plans designed to ensure recovery capability and not to make them too vulnerable to disturbances" (Harris et al., 2016). Plans that are excessively tight may end up costing more for disturbance handling, overtime and sick leave, while resource utilisation that is too low means increased costs and poor utilisation of public and company resources (Harris et al., 2016).

**Timetable design — stability and robustness**

The function of scheduling is mainly to coordinate the train paths in the planning process for optimum use of the infrastructure (Hansen & Pachl, 2014). Scheduling ensures the predictability of train traffic and produces timetable data for passenger information. Furthermore, it is essential for traffic control, locomotive and rolling stock usage and crew scheduling.

The railway traffic can never follow the original schedule deterministically because it depends on many internal and external constraints and forces. The design and construction of the timetables aim to develop a consistent, stable and robust model of train operation that support the railway personnel and the technical equipment efficiently (Hansen & Pachl, 2014). Furthermore, market demands must be reflected in the timetables, and the timetables must generate economic and social benefits.

To achieve reliable train operation, timetables must be robust and stable, with built-in slack to avoid and reduce delays and delay propagation (Goverde, 2005; Hansen & Pachl, 2014). Train delays propagate in densely frequented station areas and networks and can only fade out in time if sufficient buffer times exist in the timetable (Hansen, 2001). The robustness of the timetable determines the effectiveness of schedule adherence after disruptions (Goverde, 2005). In the planning process, it is therefore important to construct the timetables in such a way that small disturbances in real-time operation have only a limited impact on punctuality. Mathematical optimisation models form the basis for the advanced timetable tools that are currently used in the railway industry (Hansen & Pachl, 2014). Mathematical optimisation models and techniques can generate solutions themselves, guided by appropriately set parameters that define the objective function(s).

Timetable stability is a key performance issue associated with the ability the railway system has to recover from delays (Hansen & Pachl, 2014). Stability refers to the self-regulating behaviour of the timetable after disruption, that is, any delay on the network can be settled after some time without adjusting the timetable. Hansen & Pachl (2014) distinguish between two types of stability: local stability of an open system and global stability of a closed system. The issue of timetable stability is rapidly gaining attention as the European railway infrastructure has become increasingly saturated where a train that is slightly delayed may cause a domino effect of consecutive delays over the entire network (Hansen & Pachl, 2014; Andersson et al., 2014). As a consequence, analytical methods and simulation models have been developed to support the planning process. Timetable stability can be analysed using network-wide simulation (Middelkoop & Bouwman, 2002) or analytically using max-plus algebra (Goverde, 2005, 2010). The max-plus approach determines whether the timetable is stable and also identifies the critical processes in the railway traffic network (Hansen & Pachl, 2014). These models can evaluate solutions generated by the mathematical optimisation models.

**Capacity utilisation**

The minimum time interval necessary between trains is called the timetable headway, and it determines the maximum number of trains per hour, i.e. the theoretical capacity (Goverde, 2005; Harris et al., 2016). The timetable headway is influenced by the signalled headway derived from the braking distances of the critical train (that require the longest distance to brake); the type and spacing of signals; variations in run times between trains; station dwell times; time buffer between trains to avoid knock-on delay effects; and meeting traffic on single-track lines (Harris et al., 2016, p. 95).

The theoretical capacity of the railway system is determined by characteristics of the infrastructure (e.g. railway layout, track speeding limits and signalling system) and the rolling stock (e.g. braking and acceleration capacity, maximum speed and train composition). The capacity of the elements of the railway has an impact on the capacity of the railway system, such as junctions, passing loops, termini and variations in train performance (Harris et al., 2016). Capacity also depends on how the traffic is organised in, for instance, the timetable. The effective capacity of railway infrastructure is therefore defined as the maximum number of trains per unit of time that can be operated given the traffic pattern, operational characteristics or timetable structure (Goverde, 2005). The percentage of (effective) capacity per time unit is called capacity utilisation (Goverde, 2005).

Achieving higher line utilisation (hence economies of density) is a key target of the railway planner and a fundamental underlying economic principle (Harris et al., 2016). Increasing line utilisation, i.e. operating more trains in a given infrastructure network, reduces the amount of slack in the timetable, which means less ability to recover from service irregularities and smaller time windows for engineering work and infrastructure inspection between trains (Goverde, 2005). Furthermore, persistent poor performance reduces effective capacity and thus prevents increasing the quality of service or performance engineering work within tightly designed slots. Punctuality and reliability are therefore necessary prerequisites for increasing capacity utilisation (Goverde, 2005). A way of improving train service is thus to focus on better punctuality, which will give more capacity because fewer free slots are needed in the timetable to handle delays. According to Harris et al. (2016), one way of improving punctuality is to improve the reliability of existing infrastructure and rolling stock, which has previously been underestimated. Efforts to improve this reliability involve providing more resources for operation and maintenance by undertaking small projects (such as moving signals slightly or installing stop boards and requiring train drivers to use them) and by improving competence.

**Simulation**

Simulations are used to evaluate the timetable and analyse timetable stability. There is a substantial number of software tools. The significant research that led to these developments and the refinement of the algorithms was done in the late 1990s and in the first decade of the 21st century (Hansen & Pachl, 2014). These tools appear to be very complete and reliable. Consequently, in the last few years, research works have focused on assessing the quality of outputs and calibrating the stochastic inputs (e.g. Bešinović et al., 2013). Harris et al. (2014) did a "forensic" work to understand the reasons for and durations of delays at stations in the Oslo area, with the intention that the data can be used as input into operational simulations.

Simulation is typically undertaken to test and analyse possible scenarios. It is common to apply simulation when studying the relationships between key factors, such as train homogeneity or capacity utilisation and the impact on random primary delays (e.g. Lindfeldt, 2015). In the strategic long-term planning process, simulations are used to make the best possible decisions regarding the next major timetable change or investment in infrastructure or rolling stock. In the operational planning process, simulation is used to (Hansen & Pachl, 2014) undertake analysis of buffer times and allowances as an input to timetable construction; test draft timetables for conflicts between trains; modify timetables to accommodate engineering work; model train movements within maintenance depots; and predict conflicts in real time to support on the day decisions to amend the train service in reaction to delays. Furthermore, other research has used simulation to analyse train station passenger flow (for improvement, planning and design) (Li, 2000) and to evaluate the impact of different dispatching choices (Pellegrini et al., 2016).

### 2.2.2 Railway operations and dispatching

Robust timetables should be able to deal with small disruptions. However, unforeseen events or large and distributed delays may cause timetable perturbation. In busy and heavily utilised railway networks, deviations from the planned path of a single train can easily propagate as a secondary delay to other trains with a logistical connection or that run over the same infrastructure. Such perturbations and delay propagations can be prevented or reduced by suitable dispatching actions (D'Ariano et al., 2007; Hansen & Pachl, 2014). To resolve conflicts, the dispatcher can take several actions to reduce delays, such as changing train orders at a given junction or station, modifying routes or dwell times at scheduled stops, or even cancelling a train path (D'Ariano et al., 2007).

Dispatchers are predominantly guided by past experiences (D'Ariano et al., 2007), but the information made available to signallers, dispatchers and network supervisors in traffic control centres has improved significantly (Hansen & Pachl,

2014) due to the introduction of electronic interlocking, associated with computer screen monitoring of the actual set-up of routes, track occupancy and sectional clearances for each train. Train describer records keep track of train positions based on information from the signalling and interlocking systems and have been used in several studies (see, for instance, Goverde & Meng, 2012). The Norwegian National Rail Administration records traffic data that describe train movements through the network, with arrival and departure times based on when trains pass home and departure signals. These are some of the data sources from which data are collected during railway operations, as illustrated in Figure 1.1, page 3.

Traffic monitoring and train delay prediction are approaches to support traffic controllers by monitoring train positions, predicting train running times and detecting train path conflicts well in advance (Hansen & Pachl, 2014). Historic train describer data can be used for building and calibrating data-driven railway traffic models to learn about running and dwell times in various traffic conditions (Hansen & Pachl, 2014). The results of such analyses can then be used to predict the future train trajectories on the network. A predictive traffic model estimates the future traffic state and propagation of delays, i.e. forecasts the progress of trains and possible conflicting train paths and their consequences (Hansen & Pachl, 2014).

Hansen suggested in 2001 that a decision support system for automatic piloting would improve railway punctuality and enhance the performance of railway operations (Hansen, 2001). Later research has shown that such decision support systems for dispatchers can predict potential conflicting routes and resolve them in real time (D'Ariano et al., 2007). Estimates could determine the appropriate speed, acceleration or braking to assure seamless crossing. The model by D'Ariano & Pranzo (2009), which predicts future evolution of railway traffic, is a real-time optimisation module that can detect and solve train conflicts while minimizing the propagation of delays. Jánošíková et al. (2014) propose a model to find the optimal platform assignment when incoming trains are delayed that can be used to support the dispatching control of real-time traffic.

### 2.2.3 Evaluation of performance

To run a competitive railway, it is essential to deliver reliable, punctual and fast transport of passengers and goods at minimal cost (Hansen & Pachl, 2014). Railway companies therefore strive to improve their networks, timetables and quality of operation. Thus, it is important to have performance evaluation of train traffic to identify problem areas and improvement activities (Harris et al., 2016). Performance evaluations are also important to show the effect of public investment in railways. Thus, all railway companies undertake checks, reporting and analysis to see if their train service is being delivered according to the timetable (Harris et al., 2016). The information acquired through this process is, among other things, fed back into production plans (see Figure 1.1 on page 3) (Harris et al., 2016). This information is input when plans are revised or made and is therefore an interface to the planning process.

In performance evaluation, it is customary to look at effectiveness and efficiency.

Effectiveness is the degree to which objectives are achieved and the extent to which targeted problems are solved (Zidane & Olsson, 2017), i.e. are we doing the right thing? The effectiveness of the railway is expressed by the service rendered, such as the number of passengers, passenger kilometres, tonnes lifted and tonne kilometres and the associated revenues (Hansen & Pachl, 2014), without reference to costs. Efficiency signifies a level of performance that describes a process that uses the lowest amount of input to create the greatest amount of output (Zidane & Olsson, 2017), i.e. are we doing things right? Efficiency is an important factor in determining productivity. Measures of the efficiency of railway service include the number of passenger journeys made or passenger kilometres per employee and per train, respectively, including the average revenue per employee, per train and per train kilometres, respectively (Hansen & Pachl, 2014). There is an established tradition for economically based performance measurements in the railway industry, both at an aggregated and a more detailed level (Harris et al., 2016). Other common studies in performance evaluation and benchmarking examine punctuality, in which the aggregated results get particular attention, e.g. percentage of train on time. Furthermore, statistical figures for railway operations of various countries are generally used for comparisons, in which there is a tendency to also include more "soft" indicators, such as percentage of customers satisfied.

Quality of operation is an essential requirement for achieving an efficient railway system (Hansen & Pachl, 2014). Most performance evaluation methods applied to rail transport are therefore directly or indirectly based on quality of operation. The quality of train service can be examined by metrics of the quality required by the customer, the quality planned through the goals, the quality produced by train operating companies or the quality perceived by the customers (Harris et al., 2016). Several stakeholders, or target groups, have different basic aims in a system of performance evaluation (e.g. see Hansen & Pachl, 2014, p.275, Figure 14.1). Stakeholders include the passenger, the operator, the infrastructure manager, the public authority and the general public, and the aims of the different stakeholders are all interdependent. For the passenger, who is the customer of the train operator, the aim is fast and cheap transport on schedule (high transportation service level). For the train operator, who is a customer of the infrastructure manager, who uses public funds for urban and regional public transportation services and offers railway transportation services, the aim is high quality transportation performance (to sell as many passenger kilometres as possible). For the infrastructure manager, who uses public funds for infrastructure investments and offers the use of infrastructure by trains, the aim is high quality operation performance (to sell as many train kilometres as possible). For the public authority, which is responsible for infrastructure development and investments and enables urban and regional railway services, the aim is the best possible usage of the provided resources. For the general public, which is affected by construction and operation of the railway system, the aim is avoiding negative environmental effects and thrifty use of public fiscal resources.

Quality of transport service and passenger satisfaction is defined by factors such as punctuality, reliability, comfort, predictability, safety and service (Goverde, 2005; Hansen & Pachl, 2014; Harris et al., 2016). The performance indicators considered most important for passengers are reliability and punctuality (Seco & Gonçalves, 2007). Reliability is important in the railway for short-term profit, but more importantly for long-term market development. To build a customer base over time, good punctuality is important so that customers see the train as a reliable transport mode (Harris et al., 2016). This is equally important for freight because customers need to be able to rely on the goods arriving at the destination in accordance with the agreement. Good railway performance is therefore often measured by the two key criteria of punctuality and regularity. These are comparative measures between planned and actual train services that measure if the customer got what he or she paid for and got the transport at the agreed-upon time (Harris et al., 2016).

### Punctuality analysis

Punctuality is one of the most important quality factors in railway operations, probably the most important after safety (Seco & Gonçalves, 2007; Harris et al., 2016). Punctuality in railway operations means that train traffic runs according to the timetable (Harris et al., 2016). The definition is usually expressed as the percentage of the trains (that arrive at/depart from/pass a location) with a delay less than a certain time in minutes (Hansen, 2001). There is no standard definition for the threshold value, but most European railway companies use five minutes (Hansen & Pachl, 2014). In Norway, local trains are considered punctual if they arrive at the destination less than four minutes late, and less than six minutes for freight and long-distance trains (BaneNOR, 2015). Train delays at stations refer to both arrival and departure events. Since a certain amount of slack time is generally incorporated in the scheduled running and dwell times, the difference between a train's actual arrival time and the schedule can be negative. A comprehensive punctuality analysis should include the level of train punctuality and the mean and standard deviation of the delays (Hansen & Pachl, 2014). Detailed delay analysis enables identification of the influencing factors, which helps operators improve the level of train punctuality. Furthermore, suitable strategies to reduce delays, improve punctuality and save energy can be applied based on statistical analysis of train delays. Statistical analysis of train delays and movements can also enable a deep insight into their variability. In addition, the reliability of operations can be measured as a percentage of scheduled trips completed.

Punctuality analysis presented in performance and evaluation reports published by train operating companies (see Figure 1.1, page 3) often operate with aggregated measures of punctuality (see Figure 1.2, page 7), in which train delays below the threshold value are not considered delays (Hansen, 2001; Goverde, 2005; Yuan, 2006). This can be unfortunate because the punctuality measure only gives limited information about the extent of train delays. The size of the delay does not necessarily reflect the impact on the quality of train operations, and small delays can be considered as bad as large delays. Furthermore, the reports often only consider delays on arrival at the final destination. However, punctuality at the final destination may not reflect the punctuality of the whole trip. Passengers may experience considerable delays at earlier stations, while the high level of punctuality may be the result of sufficient recovery time on the previous route section. Additional reporting points connected to major stations should therefore be included.

Historically, data on train dwell times and arrival and departure times at stations were found in manual records. In modern rolling stock, accurate and detailed data of train movements are available through GPS and train event recorder data, including precise arrival and departure times at stations (Hansen & Pachl, 2014). Furthermore, automatic track occupancy and release records from existing signalling and safety systems offer a way to collect data for analysing track blocking times and real capacity utilisation (e.g. D'Ariano & Pranzo, 2009; Goverde & Meng, 2012). The assessment by Harris et al. (2016) of developments in punctuality work in Norway over the last ten years reveals great improvements in data availability and skills to analyse punctuality information. These improvements include data on time and causes, operational appropriate analysis and strategic analysis. The causes of delays of more than four minutes and additional delay to already delayed trains are registered by the train dispatcher. There are also some improvements in communicating with others, while the largest potential lies in incorporating punctuality considerations in strategic work and directing resources and means towards activities that give the best punctuality for money spent.

Passengers can be delayed even though the train is punctual within the threshold, for instance, if the passenger misses a connection. To evaluate how passengers are impacted by train delays during their travel, traffic data can be combined with the passenger number. For instance, if passenger counting devices are available or a seat reservation is obligatory, actual train occupation rates can be used for estimating the time loss passengers experience (Hansen & Pachl, 2014).

Higher quality of timetabling and more reliable operation has been enabled by implementation of advanced tools for timetabling, simulation and construction during the past decades (Hansen & Pachl, 2014). On the other hand, the achieved levels of reliability and train punctuality do not indicate higher performance. According to Hansen & Pachl (2014), the disappointing train punctuality performance of many networks is hypothesised to only partly be due to high traffic volumes and capacity usage. The main cause is an insufficient degree of optimisation, accuracy and robustness of the current practice of timetabling, which overlooks the feedback of detailed operational performance data (see Figure 1.1). Propagation of delays reflects the degree of robustness of timetable design and the stability of train operations (Yuan & Hansen, 2007).

### Delay propagation

Delays can be categorised as primary or secondary based on the cause. Primary delays are caused by specific incidents like infrastructure failures, rolling stock failures, weather or the absence of personnel (Harris et al., 2016). Secondary delays are knock-on delays caused by other delayed trains through interdependencies in the timetable and infrastructure (Conte, 2007). A train suffering a knock-on delay may cause further knock-on delays on other trains, known as dynamic delay propagation (Yuan & Hansen, 2007). Flier et al. (2009) point to the need to develop algorithms to describe and quantify dynamic delay propagation based on real-time data.

The occurrence of primary delays is directly observable and relatively well understood (Goverde & Hansen, 2001; Olsson & Haugland, 2004), although sometimes challenging to quantify accurately. Propagation of delays from these primary delays reflects the degree of robustness of timetable design and the stability of train operations (Yuan & Hansen, 2007). Delays caused by delay propagation, i.e. secondary or knock-on delays, were shown by Yuan & Hansen (2007) to increase exponentially with capacity utilisation.

Infrastructure managers are working to understand and quantify the influence of individual infrastructure component failures and other delay causes on train traffic in near real-time. This capability will enable smarter and more cost-effective infrastructure maintenance. Successful development of the capability requires a detailed understanding of the occurrence of delays and their propagation through the system. Such insight may also support other efforts to improve punctuality. As the availability of train traffic data increases, there are large opportunities to use them analytically to gain an improved understanding of delays and their causes. It can be important to identify exact allocation of cause when it affects monetary compensation, although many delays may be multi-causal (Harris et al., 2016).

Methods used to study knock-on delays and dynamic delay propagation have been developed within analytical models, statistical methods and simulations. In 1970, Potthoff provided equations to describe and analyse delay propagation (Potthoff, 1970). The propagation factor on single tracks is a function of the initial delay, headway time, buffer time between trains and the number of sections on the single track. The equation indicates that secondary delays increase with increasing numbers of sections on the track. Schwanhäußer (1974) proposed an analytical model to calculate the secondary delay on a railway line. The overall secondary delay is calculated by the summation of the secondary delays multiplied by the frequency of primary delay.

Simulations have been used extensively to study delay propagation (Radtke & Bendfeldt, 2001; Goverde, 2010; Warg, 2013), and it is common to apply simulation techniques when studying the relationship between key factors such as train homogeneity or capacity utilisation and the impact on random primary delays (e.g. Lindfeldt, 2015). Simulation is an adequate method to reveal delay data for future timetabling scenarios, as shown by Warg (2013). For instance, Radtke & Bendfeldt (2001) simulate traffic with introduced primary delays to compare different scenarios with alternative timetables and infrastructures. The algorithm presented by Goverde (2010) models delay propagation in a large-scale railway network that includes single-track lines. The model detects the propagation of delays based on a given railway network, a periodic timetable and a set of initial delays. The results are visualised in the network and show the magnitude of the delays and how they propagate and die out over time.

Regression analysis has been used to identify interrelationships between train arrivals and departures, as proposed by Goverde & Hansen (2001). The input is real-time data from the Netherlands and time stamps for events connected to entering and leaving a station for individual train movements along the route. Others use analytical stochastic models, such as Yuan & Hansen (2007), to estimate the propagation of train delays at platform tracks and junctions. Their model also adopts recursive substitutions to estimate dynamic delay propagation. Cule et al. (2011) use real-world data to search for patterns of delayed trains in the Belgian railway network and find all trains that are frequently delayed for a given time period.

Flier et al. (2009) uses real-world data to detect two important types of delay dependencies on large-scale railways: knock-on delays caused by sharing of the same infrastructure and delays caused by a late connection between two trains. The algorithmic methods solve maximum optimisation problems and find correlations and patterns in systematic dependencies. Such findings can be statistically analysed, for instance, to assess the significance of the dependencies and for timetable planners to detect often-occurring secondary delays and adjust the timetable to reduce the dependencies. Similarly, information on route conflict based on train describer records, as suggested by Goverde & Meng (2012), can also be valuable input for adjusting timetables. They introduce a tool that determines chains of route conflicts, which are used to automatically identify and analyse structural and serious route conflicts due to timetable flaws or capacity bottlenecks.

Several of the aforementioned papers use real-time data, also called train detection data. For instance, Yuan & Hansen (2007) validate their model by means of train detection data. D'Ariano & Pranzo (2009) use it for designing a traffic management system, while Schwanhäußer (1974), Goverde & Hansen (2001), Flier et al. (2009) and Cule et al. (2011) use it in post-evaluation.

While models exist for calculating waiting time and secondary delays on double-track lines, few models for single-track lines have been published (Handstanger, 2009). In Norway, single-track methods are of particular interest, given that the Norwegian railway primarily consists of a star-shaped single-tracked network (approximately 7 % double/multi-track). Of the literature presented here, only the method by Goverde (2010) is used on networks incorporating single tracks. Studies of the Swedish railway are relevant because it also consists of a significant proportion of single-track lines. Lindfeldt (2007) evaluates the single-track railway with a focus on crossing times as the time loss that occurs in crossing situations on single-track lines, compared to double-track lines where no such time loss occurs. An analytical model calculates the crossing times based on infrastructure, train properties, delays of two crossing trains and the timetable. Lindfeldt (2010) presents a program that calculates descriptive parameters that can be used to identify weaknesses in the network.

### Ridership

The number of travellers, also referred to as ridership, is a measure of demand for transportation service, which is important information for planning and evaluation. With updated ridership information, planners should be able to get a detailed, continuous and accurate vision of the travel behaviour of their customers. When evaluating railway operations, as well as the effectiveness of railway projects, the railway company wants to know whether the railway is used, particularly since the benefit of reduced travel time and increased capacity is dependent on the volume of traffic (Volden & Samset, 2013). Boyle (1998) identifies four main reasons why ridership data are collected. First, ridership is reported to external funding and oversight agencies. Second, it monitors trends over time. Third, ridership is a key performance indicator at various levels of the transportation system. Finally, ridership data help identify locations with the greatest boarding and alighting activity, which is important not only for it own purpose, but also because the safe management of the railway depends upon it. The purpose on monitoring trends and travel behaviour over time is also highlighted by Vuchic (2005). Such data show passenger volumes on different sections of a line, the maximum number of passengers on different lines and when the maximum occurs, along with information on variations in passenger volume. Other issues that call for data on ridership on rains include fare equipment location optimisation, fare policy change and train schedules (Li, 2000).

Train ridership is influenced by several factors, including fares, transit time, transit comfort characteristics and feeder accessibility of transit, price and service characteristics of the competing modes, seasonal variations and monthly working day variations and the socio-economic conditions of the service areas in the medium or long term (Doi & Allen, 1986). Planners need to understand the drivers of demand, including those that cannot be controlled (Harris et al., 2016).

Passenger counting is the key measuring parameter associated with ridership. Different measurement types and ridership estimation techniques are applied for different network levels. The selection of the appropriate network level is dependent on the particular use and issue being addressed (Boyle, 1998; Gordillo, 2006). Boyle (1998) and Gordillo (2006) identify uses of passenger counting and ridership calculation based on the way data are measured. The uses are different for each measurement type, and the operator normally uses multiple types to fulfil different purposes. The different levels are system level, route level, trip level, station level and origin-destination level.

Demand for railway travel is typically expressed in number of travellers. Vuchic (2005) lists a set of relevant key performance indicators related to ridership:

- average passenger trip length as total passenger kilometres divided by number of passengers,
- average passenger volume measured as total passenger kilometres divided by line length,
- coefficient for flow variations to indicate the degree to which passenger volume peaks along a line,
- coefficient of passenger exchange to describe what proportion of passengers is exchanged along a line,
- riding habit to measure how much of a population in an area utilises the transport in question (such as commuter railway), and
- market share to describe use of a particular type of transportation in relationship to total travel volume in the same market.

Traditionally, multiple methods are used to calculate the demand between an origin and destination point. The most common is the O-D matrix, representing the origin (O) and destination (D) of a route (Frias-Martinez et al., 2012), that characterizes the transitions of a population between different geographical regions. The most commonly used method for populating these matrices is user surveys. Strengths of traditional surveys are that they include important information about the respondent such as age and gender and also information about the purpose of the trip (Alexander et al., 2015). A major problem with user surveys is declining response rates (Schoeni et al., 2013), which may introduce bias into the samples. Such surveys are typically not done with a higher frequency than yearly, but may also be conducted less frequently and not necessarily on a regular basis. Consequently, this method may possess low frequency, high cost, varying data quality, low precision and susceptibility to errors. Alternatives to traditional transport surveys include an origin-only automatic fare collection system, as proposed by Zhao et al. (2007), and mobile phone data (e.g. Jiang et al., 2013). For instance, Calabrese et al. (2013) showed that it is possible to use mobile phone data to describe people's movement patterns when they analysed the mobile phone records of a million users in Boston to describe transportation needs.

The Norwegian National Rail Administration annually publishes reports on official railway statistics in Norway (e.g. BaneNOR, 2015). The railway statistics include aggregated data on number of travellers (illustrated in Figure 1.2 on page 7), passenger kilometres and number of sold single tickets and monthly tickets. The practice for measuring ridership is manual counting at chosen stations on each railway line.

There are also other challenges concerning obtaining data on ridership. First, train operators consider such data confidential business information, especially in high resolutions. Second, the data that actually are available are of varying quality and coverage. Several studies highlight the unreliability of ridership data (including Fowkes et al., 1985).

## 2.3 Digital transformation and available data

This section presents recent developments of technologies and the discussions and research regarding the increasing amount of available data sources. Furthermore, the section presents developments and data collection within the railway.

### 2.3.1 Recent developments

New technologies generate and store large amounts of data about the use, status or performance of the equipment. Digital data are now in every sector, every economy, every organisation and every user of digital technology (Manyika et al., 2011). The reasons for the large increase in digital data are many: more technology with sensors, more people interacting with information and sharing information (The Economist, 2010). Since the capabilities of digital devices are rising and prices are dropping, sensors and gadgets are digitising lots of information that was previously unavailable.

A large amount of digital data is digital "exhaust data", i.e. data that are created as by-products of other activities as companies and organisations go about their business and interact with individuals (Manyika et al., 2011).

The rapid increase of personal devices is a prominent reason for the growing availability of information (De Mauro et al., 2016). The digital sensors that these devices are equipped with, such as cameras, audio recorders and GPS locators, are what makes digitisation possible. Network connection lets data be collected, transformed and organised as information (De Mauro et al., 2016). The Internet of Things (IoT) refers to sensors and actuators embedded in physical objects and connected by networks to computers (Manyika et al., 2011). They have also been described as artificial objects, equipped with unique identifiers, that interact with each other to achieve common goals, without human interaction (De Mauro et al., 2016). The Internet of Things also represents a promising source of information (De Mauro et al., 2016).

Wu et al. (2014) describes the sources as autonomous, with distributed and decentralised control, resulting in data sources being able to generate and collect information without involving or relying on any centralised control. Consequently, data are generated with a higher frequency than being recorded, updated or measured monthly and weekly to more frequent updates such as daily, hourly or continuously (Courtney, 2012). Furthermore, the generation of data happens with a high velocity, e.g. from the sensor that generates the data to the data that are available for analysis. Access to real-time or almost real-time information makes it possible for a company to be much more agile than its competitors (McAfee & Brynjolfsson, 2012).

Digital data are generated from a variety of sources, including retail transactions, security cameras, internally registered data in the organisation, time stamps, GPS tracking, sensor data from instrumented machinery and metavalues of documents. Consequently, this generates a high variety of data such as texts, pictures, videos and GPS location, and the data are measured and captured in more detail, such as location, time and metadata (Russom, 2011; Waller & Fawcett, 2013). The expanding variety of forms is an important feature of the data that get produced and utilised nowadays (De Mauro et al., 2016). Traditional alphanumerical tables are being overtaken by the growing availability of less structured data sources such as video, pictures and text produced by humans (Russom, 2011). The data are represented by heterogeneous and diverse dimensionalities, which, according to Wu et al. (2014), is because different information collectors prefer their own schemata or protocols for data recording, and the nature of different applications also results in diverse data representation. This type of data therefore yields both structured and unstructured datasets (Russom, 2011; Waller & Fawcett, 2013). The data available to companies are often unstructured (Davenport et al., 2012).

Because the data are generated with such high variety and velocity, it is likely that the datasets will be large, containing a few terabytes to many petabytes. On a worldwide basis, the total amount of digital data created and replicated each year is expected to increase exponentially up to 2020 (Tien, 2013).

### Big Data

Volume, velocity and variety (the three Vs) are the characteristics typically used to describe "Big Data", which has recently become a commonly used business term (Russom, 2011; Courtney, 2012). The characteristics of the three Vs were first described by Laney (2001), but he did not mention Big Data explicitly. There have been several definitions of Big Data. Manyika et al. (2011) interprets Big Data as "datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyse" within a tolerable elapsed time (Tien, 2013; Waller & Fawcett, 2013). Other definitions will probably be needed because Big Data is becoming part of commonly used software tools. However, Manyika et al. (2011) underscores that this definition is purposely subjective and incorporates a moving definition of how big a dataset must be to be considered Big Data, assuming that as technology advances over time, the size that qualifies as Big Data will also increase. Furthermore, they note that the definition can vary by sector, depending on what kinds of software tools are commonly available and what sizes of datasets are common in a particular industry.

Other characteristics suggested to describe Big Data include "veracity", "value" and "smart". "Veracity" refers to a quality of the datasets, while "value" refers to the goal of using the datasets (Courtney, 2012). George et al. (2014) points out discussions among practitioners that "big" is no longer the defining parameter, but rather "smart", including the fine-grained nature of the data.

De Mauro et al. (2016) discuss how the focus of the definitions varies, from the attributes of the data to technology needs to overcome the thresholds to social impact. They define the term thus: "Big Data is the Information asset characterized by such a High Volume, Velocity and Variety to require specific Technology and Analytical Methods for its transformation into Value ".

### Information and knowledge from data

According to the data-information-knowledge-wisdom hierarchy, "information appears as data that are structured in a way to be useful and relevant to a specific purpose" (Rowley, 2007). Knowledge is information combined with understanding, experience and values and is typically divided into explicit and tacit knowledge. Tacit knowledge is embedded in the individual, while explicit knowledge resides in documents, databases and other recorded formats. Information becomes a knowledge asset that can create value for firms (De Mauro et al., 2016).

The main reason to carry out data analysis is to derive information from data, knowledge from information, and wisdom from knowledge (Tien, 2013). The purpose of data analysis is to provide new information and knowledge for decision-making. It can be used to make more precise predictions, and it follows that better predictions yield better decisions (Jagadish, 2015). McAfee & Brynjolfsson (2012) found that the more companies characterized themselves as data-driven, the better they performed on objective measures of financial and operational results. More and more business activity is digitised, and new sources of information are available (McAfee & Brynjolfsson, 2012). Big Data is now relevant for leaders across every sector, and its application will benefit customers of products and services (Manyika et al., 2011). From a macro-perspective, Big Data-informed decision-making is expected to have a similar positive effect on efficiency and productivity as information and communications technology (ICT) have had, according to Hilbert (2013). However, it is expected to add to the existing effects of digitisation (Hilbert, 2013).

As a consequence of the vast amount of digital data, some say that “data” and “information” are increasingly difficult to tell apart, and they use these terms interchangeably (The Economist, 2010). Indeed, given enough raw data, algorithms and powerful computers can now reveal new insights that would previously have remained hidden.

According to Manyika et al. (2011), data can create significant value for the world economy, enhancing the productivity and competitiveness of companies and the public sector and creating substantial economic surplus for consumers.

#### Challenges

With the advent of large-scale data collections or warehouses, the problem of data rich, information poor (DRIP) has become pervasive (Tien, 2003). As Manyika et al. (2011) points out, the topic of information overload has been widely studied by academics ranging from neuroscientists to economists. In contrast, Big Data somewhat mitigates the DRIP problem with the Big Data approach, which has unleashed information in a manner that can support informed, although not necessarily defensible or valid, decisions or choices (Tien, 2013). A consequence of the increase in volume is an increase in complexity and the relationships within the data (Wu et al., 2014). The key is to take the complex (nonlinear, many-to-many) data relationships, along with the evolving changes, into consideration to discover useful patterns. Access to large chunks of information presents challenges of estimation, integration and validation (Toole et al., 2015). On the one hand, it is attractive because it can be extracted instantaneously at a low cost, and the available samples are long running and cover multiple aspects simultaneously. On the other hand, the sheer size of the data means that they lack the contextual demographic information pertaining to privacy, resolution of data and inherent noise (Toole et al., 2015). Another challenge is about bias from incomplete observations. In the age of Big Data, data collection is not restricted to samples, but basically to collecting as much potentially related data as possible. Thus, questions about the representability of the data collected are raised. In practically all cases that rely on external data, the analysers have to assume incomplete observations and thus a bias (Lugmayr et al., 2016). In interpreting the results of data analysis, the challenge is therefore to take into account this bias and, if possible, assess its impact on the results (Lugmayr et al., 2016).

Large amounts of data also cause technological challenges for companies, including the critical issue of storing extensive amounts of data and issues linked to computational requirements that an average IT system might not be able to grant (De Mauro et al., 2016), a need for more complex methods than the usual statistical procedures, availability of competence (Manyika et al., 2011; Olsson et al., 2015b) and managerial talent. In a project management context, the needed skills include identification of relevant data sources, data collection, analysis and interpretation

(Olsson et al., 2015b). Other issues include the need to ensure the right infrastructures, that the incentives and competition are in place to encourage continued innovation, that the economic benefits to users, organisations, and the economy are properly understood, and that safeguards are in place to address public concerns about Big Data Manyika et al. (2011).

#### Privacy issues

Both from an ethical and legal point of view, it is important to protect personal information and respect people's privacy. Data that do not include personal information are basically unproblematic, both as individual data sources and the combination of several sources. Combinations of different data sources in which persons are the link between the various data are more problematic. Data from different sources that include personal data can be combined without revealing personal information, but this can be challenging. Anonymity in datasets is typically achieved by aggregation, in which each group includes so many persons that individuals cannot be identified.

Ensuring data security and protecting privacy are becoming harder as the information multiplies and is shared ever more widely around the world (The Economist, 2010). Studies that, for instance, utilise call detail records (CDR) data, which are not anonymous, are required to protect privacy through measures such as anonymizing the data (i.e. removing personal identification), only using the minimum of information needed for the studies, only presenting aggregated results and not focusing the analysis on individual phones (Becker et al., 2013). With pseudo-anonymized data (i.e. the ID is replaced with a code), the record must be pre-processed to reduce the probability of re-identification. A common procedure is to decrease time resolution or increase space granularity (Bianchi et al., 2016).

Norwegian law states that collected personal information should only be used for the specific purpose for which it was originally collected (Drageide, 2009). As a consequence, any use of CDR data that goes beyond billing requires active consent from the subscriber. The privacy principles in Norway, which are supervised by The Norwegian Data Protection Authority (Datatilsynet), follow the European Union's (EU) General Data Protection Regulation (GDPR), which was enforced in May 2018. GDPR is a regulation by the European Parliament, the Council of the European Union and the European Commission. The intention is to strengthen and unify data protection for all individuals within the European Union (EU). Core guiding principles behind GDPR's design include consent and tranpsarency about the use of personal data.

### 2.3.2 Data collection and developments in railways

There is a need for and currently a focus on digital transformation in railways (Pieriegud, 2018), for instance, within safety (Parkinson & Bamford, 2016), freight transport (Green, 2017) and maintenance (Tute, 2018). Several undertakings and programs focus on this transformation. For instance, the Shift2Rail Joint Undertaking is a public-private partnership in the rail sector, established under Horizon 2020, with the rationale of pooling and coordinating research and innovation efforts at the

EU level. Another example is Britain's Digital Railway program, which seeks to harness technological advances to make capacity enhancement a more cost-efficient process and to optimise operations by integrating key systems (Templeton, 2017). Furthermore, the Network Rail's Offering Rail Better Information Services (ORBIS) program, launched in 2012, aims to digitise maintenance of the UK's railway infrastructure (The Institution of Engineering and Technology, 2012). Data were often siloed in disparate systems located in different parts of the business. The program is designed to collect, join and exploit accurate asset data on the rail infrastructure to enable decision support tools for managing the infrastructure assets. The Geo-RINM viewer is an example of an achievement of the ORBIS program. The viewer is used in preparation and maintenance work at Network Rail in the UK (Tute, 2018). The viewer produces high-resolution images that give employees better visibility of tracks, level crossings or overhead lines. This saves time on track because the teams can plan and familiarise themselves with a site before leaving the office.

Within risk assessment and safety, the research by Parkinson & Bamford (2016) showed that Big Data could potentially help mitigate accidents when the causes are systematic and complex. They classify data related to safety incidents. For instance, information from condition-based monitoring from sensors that would provide digital information includes vibration, machine vision, heat, displacement, strain, humidity and particle ingress.

### Data collection on train movement

Historically, data on train dwell times and arrival and departure times at stations were recorded manually. Introduction of electronic interlocking, associated with computer screen monitoring of the actual set-up of routes, track occupancy and sectional clearances for each train, has improved information available to signallers, dispatchers and network supervisors (Hansen & Pachl, 2014). In modern rolling stock, accurate and detailed data on train movements are available through GPS and train event recorder data, including precise arrival and departure times at stations (Hansen & Pachl, 2014). Automatic track occupancy and release records from existing signalling and safety systems offer a way to collect data for analysing track blocking times and real capacity utilisation (e.g. D'Ariano & Pranzo, 2009; Goverde & Meng, 2012). Train describer records track train positions based on information from the signalling and interlocking systems (e.g. Kecman & Goverde, 2013).

The Norwegian National Rail Administration records traffic data that describe the train movements through the network, in which the arrival and departure times are based on when trains pass home and departure signals (Olsson et al., 2015a). The data include train information, stations and codes explaining causes for any delays in excess of the defined margins for the various classes of traffic. The data are recorded in a national database called TIOS ("Trafikkinformasjon og oppfølgingssystem", translated as "Traffic Information and Follow-up System"). The time registrations are generally based on automatic registrations and are considered accurate measurements of when trains enter and exit different sections of the network, with a high resolution (measured and recorded to the second).

### Data collection on ridership

Surveys and manual passenger counts are well established in railway transit services to obtain information on passenger volume and load counts (Vuchic, 2005). Boyle (1998) found that other applied methods among US transport agencies were electronic registering fareboxes (EFR), on-board surveys, vehicle operator trip cards, estimates from passenger revenue, checkers and hand-held units, automatic passenger counting (APC) and smart cards. Manual observations have been the main data source for the Norwegian railways. In addition, different travel behaviour surveys have been conducted. Using surveys and manual counts provides transport organisations with a reasonable snapshot of existing demand on their transport system. Increasing the resolution of manually obtained information is typically costly and requires personnel resources. Accuracy has been a major concern in any data collection efforts (Boyle, 1998). The use of manual techniques can result in errors both in the collection and registration phases. Such errors tend to be random in nature.

The most complete passenger counts can be obtained from fully controlled stations, according to Vuchic (2005). Fare collection systems are platforms for collecting passenger fares and controlling access to the transportation service. There is a trend towards automating the fare collection process through automated fare collection (AFC) systems. In addition to collecting fares, such systems can track not only the number of passengers, but also the entry and exit points for travels. Frumin (2010) use such entry and exit data and develop a methodology for building an unbiased estimate of existing travel patterns on the London Underground.

Automatic passenger counting (APC) is gaining popularity. An APC is an electronic device that accurately records boarding and alighting data on transit vehicles such as trains. Sensors are located in the doorways to a vehicle. When a person passes, the sensors count movements and determine if the person is entering or exiting the vehicle. APC systems are installed on a few vehicles operating in the Norwegian railway system. The data are used to calculate the total number of passengers on each train at different points in time. Barabino et al. (2014) addresses the challenges of using APC to measure ridership on busses, which include matching data to the bus stops, tackling anomalies and building intelligible performance reports.

APC units are not needed on every vehicle in the fleet. Boyle (1998) found that the eight agencies that made regular use of APCs equipped about 10 percent of their fleets with APC units and rotated these units throughout the routes in the system. APCs can be based on different technologies, such as infrared lights above the doorways of a vehicle (Chu, 2010), treadle mats on the vehicle's steps (Boyle, 1998) and closed-circuit television (CCTV) and intelligent people counters to log numbers of travellers getting on and off a vehicle (Saponara et al., 2016). On-board CCTV is frequently installed on trains for surveillance and safety. The technology can also be applied for counting people. All these methods require physical equipment.

Passenger distribution in the urban Copenhagen rail network is tracked with a combination of electronic weighing equipment (EWE) and APC (Nielsen et al., 2014). Nielsen et al. (2014) show that EWE-based monitoring can provide estimates with higher accuracy than infrared sensor technology. EWE is installed in many modern trains because it supplies data for the braking system. This information can be used to estimate the number of passengers in the trains because the weight of a train is a function of the number of passengers in it at any particular time. Provided that the weighing equipment is installed on a train, this can be a cost-efficient way of estimating ridership. The weighing system also has potential to provide a complete sample of weight, and thus of ridership.

In 2012, Virginia Railway Express looked into the latest payment technologies to pilot at key stations, as well as technology to verify ticket purchases and use along with ticket history (Henry & Grant, 2012). The technology options included mobile ticketing, radio frequency tags and near field communication (NFC). The ticket database would be used for analytical purposes such as ridership, travel patterns, boarding at particular stops, client use of facilities by time of day and other information that enables the providers to better plan their services. A key component is the ability to support smartphone technology, including mobile devices utilising the prevalent versions of mobile operating systems.

Travel time and ridership can be detected using GPS traces. Higuchi et al. (2015) identifies a number of innovative use forms based on mobile devices, including several technologies that typically are found in smartphones such as GPS, Wi-Fi and Bluetooth. These approaches can capture in detail individual travel behaviour, but are limited by their sample sizes (e.g. number of volunteers) and scaling difficulties (Holleczek et al., 2014). Also, low penetration of smartphones on a global scale and limited access to GPS-related information from telecom operators because of user privacy policies hinder this from being an effective mode for calculating travel times going forward. Moreover, it is possible to detect transportation mode based on the GPS sensor on mobile devices and knowledge of the underlying transportation network (Stenneth et al., 2011). Proximity sensing can be based on logging of Bluetooth units. This is applied by the Norwegian Road Authority (Olsson et al., 2015b).

### Mobile phone data in mobility research

Much research has focused on developing methods to extract meaningful information about human mobility from mobile phone traces and on understanding its limitations (Alexander et al., 2015). Several studies have shown that mobile phone data can be used to describe people's movement patterns (Calabrese et al., 2011; Järv et al., 2012; Holleczek et al., 2014; Xu et al., 2016; Kujala et al., 2016). However, most studies have not focused on rail in particular, but typically address travel in general. Mobile phone datasets allow deriving a statistical analysis of human activities at a fine level of detail (Leo et al., 2016). Furthermore, mobile phone data can be used to derive good estimates of dynamic quantities such as travel times, train occupancy levels and origin-destination flows for transportation studies (Aguiléra et al., 2014). For this very reason, mobile phone data can be utilised in estimating the commuting patterns and travel times for individuals.

Three main types of mobile phone data are collected using passive collection: call detail records (CDR) data, probes data and Wi-Fi data (Larijani et al., 2015). CDRs contain anonymized traces of a user at approximate locations when the phone communicates with a cell phone tower. Studies have shown that CDR data can be used to study the habits and mobility patterns of mobile users (Bianchi et al., 2016; Zhao et al., 2016), to study users' movements (Leo et al., 2016) and to calculate commuting matrices with a very high level of accuracy (Frias-Martinez et al., 2012). Studies have also looked at utilising mobile data to estimate intra-city travel time (Kujala et al., 2016) and found that mobile data could be employed as a real-time traffic monitoring tool in the near future (Järv et al., 2012).

The approach typically consists of performing some kind of trip extraction to extract the movements relevant for traffic analysis from the raw cellular network data (e.g. Doyle et al., 2011; Calabrese et al., 2011; Iqbal et al., 2014; Alexander et al., 2015). There is no obvious definition of what a movement/trip is, so trip extraction algorithms vary a lot between authors (Gundlegård et al., 2016).

An origin-destination matrix can be computed based on the extracted trips (e.g. Calabrese et al., 2011; Larijani et al., 2015). For instance, the method presented by (Alexander et al., 2015) estimates average daily origin-destination trips from triangulated mobile phone records of millions of anonymized users. The CDR records are converted into cluster locations and are inferred to be home, work or other depending on observation frequency, day of the week and time of day. The aggregation of OD flows gives an estimation of the number of cell phone users who are travelling, but only those of the operator that provided the data (Gundlegård et al., 2016). This can only yield information about how the travel demand distributes relatively between different OD pairs. To estimate the total travel demand in terms of the number of people travelling, authors use different scaling factors (e.g. Calabrese et al., 2011; Iqbal et al., 2014; Alexander et al., 2015; Toole et al., 2015).

Several authors have also tried to reconstruct the specific travel mode and route that a user took for a trip, which is challenging. However, as Larijani et al. (2015) showed, detection of the trip segments in which people take the metro is promising because underground tunnels are serving by dedicated base stations (Gundlegård et al., 2016). Xu et al. (2016) used a large-scale mobile phone dataset to estimate demand for bicycle trips in a city.

Calabrese et al. (2013) discussed three challenges in using mobile data. First, demographic information about individuals was not available due to privacy concerns. Second, mobile users were not necessarily representative of the whole population. Third, the data were not formatted for this type of analysis. To address the first challenge, they used aggregated data in which users were collected into groups corresponding to the most detailed level of economic and demographic data available. The second issue introduces sample bias among the population. To validate representativeness, they calibrated the data based on information from security inspections of the vehicles, which included mileage condition, to determine if the estimated mileages seemed realistic.

There are both advantages and disadvantages of using mobile phone data. For instance, CDR data contain the approximate locations when the phone communicates with a cell phone tower, hence providing an inexact and incomplete picture of daily trips. Furthermore, mobile phone data cannot provide information about the traveller, like age, income or purpose of trip, as a survey would (Alexander et al., 2015). On the other hand, mobile phone data are automatically collected, which make them more frequent and economical than, for instance, a survey. In addition, because mobile phone data can be gathered over a longer time period, they can capture information such as variations in travellers' daily travel behaviour (Alexander et al., 2015).

## 2.4 Summary and research gaps

Ex-post project evaluation should assess the project's operational, tactical and strategic success. Tactical success is measured in terms of effectiveness and gauges to what extent the goals were achieved. Evaluation of effectiveness contributes to double-loop learning, which considers whether the project provided the benefits to users and society as assumed. Larger changes to the railway infrastructure are planned for in strategic long-term planning. Typical goals of railway projects are to reduce travel time, increase capacity and increase demand. Because the effects and benefits of a project are not realised until after a few years, an ex-post project evaluation should be executed three to five years after the project is completed.

To achieve a reliable train operation, timetables must be robust and stable, with built-in slack to avoid and reduce delays and delay propagation. Achieving higher line utilisation is a key target of the railway planner, but operating more trains in a given infrastructure network reduces the amount of slack in the timetable. In a highly saturated infrastructure, a train that is slightly delayed may cause a domino effect of consecutive delays over the entire network. Thus, poor performance reduces effective capacity. Consequently, punctuality and reliability are necessary prerequisites for increasing capacity utilisation.

The railway companies strive to improve their networks, timetables and quality of operation to run competitive railways. Common performance indicators used to evaluate railway operations are metrics based on the number of passengers and volume of freight and on punctuality and reliability. These are also typical measures to evaluate the effectiveness of railway projects. Punctuality is one of the most important quality factors in railway operations. Detailed delay analysis enables identification of the influencing factors, which helps operators improve the level of train punctuality. Disappointing train punctuality is partially due to high traffic volumes and capacity usage, but the main cause is an insufficient degree of optimisation, accuracy and robustness of timetable design. The current practice of timetabling overlooks the feedback of detailed operational performance data. Propagation of delays reflects the degree of robustness of timetable design and the stability of train operation. Analysis of delay propagation provides a deeper evaluation of punctuality and helps infrastructure managers understand what causes delays in train traffic. The number of travellers provides information on demand and travel behaviour, which are important for both planning and evaluation. The most commonly used method for collecting numbers on ridership is surveys, but a major problem with user surveys is declining response rates, which may introduce bias into the samples.

More and more business activity is digitised, and new sources of information are available. The data are measured and captured in more detail, such as location, time and metadata. The purpose of data analysis is to provide new information and knowledge for decision-making. Within railway research, there is currently a focus on digital transformation. New technological developments increase the amount of available data sources within railway operations. Traffic data, which describe trains' movements through the network, are generally based on automatic registrations of high resolution. Several publications have explored different technologies for collecting data on ridership. Furthermore, much research focused on using mobile phone data to extract origin-destination trips and travel mode. However, most of these studies address travel in general and do not focus on rail in particular.

# Chapter 3

# Research design

This chapter presents the research design and methodology of this dissertation and the research methods applied to answer the research questions.

Research design is a plan for how an investigation will be executed, from the first vague thoughts to the finished product. The research design has to be a result of the research questions (Grenness, 1997). The overall aim of this dissertation is to obtain new insights to support evaluation of railway projects with new data, as discussed in Section 1.2. The first research question (RQ1) is how one can exploit digitalization and new data in railway to obtain information relevant for project evaluation, specifically with traffic data and mobile phone data. The second research question (RQ2) is how such empirical analyses are useful for project evaluation and learning in a life cycle perspective.

A common challenge in project evaluations is access to good relevant data. Four measures are traditionally used in evaluating the effectiveness of railway projects: travel time, frequency, punctuality and volume of traffic. Travel time and frequency are easy to find by studying the timetable. Punctuality in itself is well developed, but a deeper analysis of punctuality through delay propagation and measures of volume of traffic is more difficult to obtain good data on. The focus of this dissertation is therefore on utilising available data sources, which offers new opportunities for measuring delay propagation and ridership.

A fundamental principle of all good research is that we must be able to claim that there is a coherence between theory, problem and research approach (Grenness, 1997). There are several types of research (Grenness, 1997; Fellows & Liu, 2003; Neville, 2007), but the three main types are exploratory, descriptive and explanatory. Exploratory design is generally recommended when the problem description is unclear, previous knowledge is limited, and we are not able to make a clear hypothesis. Descriptive design is first and foremost used when the task is to describe variables and relationships between them. It is often based on a relatively clear hypothesis about what such relationships look like. Explanatory design is used when we wish to measure the effects of various stimuli, that is, cause and effect relationships.

This research is exploratory. Exploratory research is concerned with a new problem or topic about which little is known (Phillips & Pugh, 2010). The research aims to push the frontiers of knowledge by examining what theories and concepts are appropriate, developing new ones if necessary, and determining whether existing methodologies can be used. This thesis aims to obtain new insights to support railway project evaluation. The research questions presented in Section 1.2 are formulated to exploit digitalization and new data in railways to obtain information relevant for project evaluation. Specifically, this thesis focuses on two evaluation measures, ridership and a deeper analysis on punctuality through delay propagation. Each measure entails exploratory research in the form of a new way of measuring something. Exploratory research is concerned with the development of theory from data in a process of continuous discovery (Davies, 2011).

The data used to analyse delay propagation were traffic data. Traffic data have been available for some time, and recently, there has been a development towards more and more automatic registrations (Veiseth, 2009). The measure of delay propagation, however, has been more difficult to obtain. The measure of ridership has been available from different types of approaches and sources with different levels of quality, but the approach of using mobile phone handset counts is new.

The research of this thesis is neither a pure basic nor a pure applied research. Basic research aims to expand general knowledge of process in which the results are universal principles. On the other hand, the intention of applied research is to improve the understanding of a particular problem. It is designed from the start to apply its findings to a particular situation or problem (Neville, 2007). The research of this thesis is applied because it utilises specific data sources from railway operations to find ways to apply them in evaluation of railway projects. The particular problem is about railways in Norway. However, the research is basic in that it aims to be generalizable to other countries.

The next sections present the research methodology, the research methods applied to answer the research questions and a discussion of reliability, validity and generalization.

## 3.1 Research methodology

The research methodology encompasses three dimensions (Creswell, 2009; Croom, 2010): 1) the research approach, either inductive or deductive; 2) the philosophical position; and 3) the research strategy, i.e. qualitative or quantitative design. There is often a strong relationship between the research topic, the philosophical position employed and the methodology used in the publications (Pannirselvam et al., 1999; Croom, 2010).

### 3.1.1 Research philosophy

The different philosophical approaches are delineated by several core assumptions concerning ontology (reality), epistemology (knowledge), human nature and methodology (Holden & Lynch, 2004). These assumptions are consequential to each other, meaning the researchers' views on ontology affect their epistemological belief, which consequently affects the choice of methodology.

At the core of the study of knowledge and truth (i.e. epistemology and philosophy) is the dilemma that truth may be conceived as either objective or subjective (Croom, 2010). The issue of whether truth and fact are objective or subjective is at the heart of the philosophical debate and therefore central to methodological considerations (Habermans, 1972, cited by Croom, 2010). In the objectivist approach, the philosophy of epistemology or knowledge is positivism (Holden & Lynch, 2004). The positivist stance is about empirical validation, i.e. a belief in objective reality. The emphasis is on observable facts, derived from valid reliable measurement, to provide a rational explanation for a phenomena (Neville, 2007). The approach seeks to provide results and conclusions that are replicable (verifiable) and generalizable (Bryman, 1988:40-41, cited in Croom, 2010). Constructivism takes an opposite stance in which the researcher considers all observations and analysis to be dependent upon the researcher as a participant (Croom, 2010).

The use of mixed methodologies has gained popularity among researchers, especially in the social sciences (Creswell, 2009). However, as Croom (2010) states, many writers take the debate further and advocate the use of several methodologies from different paradigms (i.a. Hassard, 1988; Mingers, 2001). The argument is that such an approach is desirable to "make the most effective contribution in dealing with the real worlds" (Mingers & Brocklesby, 1997).

The aim of this dissertation was to obtain new insights to support evaluators in ex-post evaluations of railway projects with available data sources. The purpose was therefore to utilise data that were automatically generated from movements and infrastructures. Hence, the research answering research questions RQ1a and RQ1b, with, respectively, traffic data and mobile phone data, is positioned in positivism because the data are objective and observable facts. On the other hand, in the field of project management and project evaluation, the philosophical assumption has typically been constructivism. The author recognises that to research within this field, it is necessary to gather assessments from people as well. The case study of Publication IV to answer research question RQ2 is positioned in constructivism. The author therefore has one foot in each philosophy. Furthermore, the interpretation of RQ1a and RQ1b and the process of placing them in a larger context slide from positivist thinking into constructivist thinking.

### 3.1.2 Inductive vs deductive

The research approach applied in this dissertation is both deductive and inductive. Induction builds on empirical data and is described as research that is exploratory (Tjora, 2013). This dissertation consists of two major empirical studies that resulted in Publications I and II (see Figure 3.1). These two studies on empirical data, i.e. on punctuality data and mobile phone data, are inductive. The case study of Publication IV is also inductive. Induction extracts universal, general conclusions based on empirical facts through use of reasoning (Grenness, 1997).

However, gathering and arranging data are impossible without a theory that steers the investigation (Grenness, 1997). Deduction builds on logic and is used in research that builds on theory (Tjora, 2013). Theory is used to provide a basis for the research, i.e. for the empirical analyses. The literature studies conducted are therefore part of a deductive approach to provide a basis for the empirical analysis.

### 3.1.3 Quantitative and qualitative research

Research can be quantitative or qualitative. The field of railway operations traditionally applies quantitative methods, while within project evaluation, the traditional research design is qualitative. This dissertation attempts to build a bridge between these and show quantitative methods that are applicable in project evaluation.

Good social science should employ those methods that best help answer the research questions, and often a combination of qualitative and quantitative methods will be the most appropriate (Flyvbjerg, 2006; Creswell, 2009). To answer research question RQ1 on how to exploit available data sources to obtain information relevant for project evaluation, the main research design was quantitative with empirical analyses. In quantitative research, objective theories and concepts are examined and tested through the clear delineation of variables (Creswell, 2009; Croom, 2010). These variables, which are typically measured on instruments, are observable, tangible and clearly defined.

In qualitative research, the data are typically collected in the participant's setting (Creswell, 2009). Central to qualitative methods is the influence of the researcher's interpretation, subjective perception and interaction, and the methods attempt to account for the significance of this influence (Croom, 2010). This is at various degrees present in the process of defining, collecting and analysing evidence. To answer research question RQ2 on how such empirical analyses are useful for learning from project evaluation from a life cycle perspective, the research design was qualitative, with a case study, interviews and document analyses. Research question RQ2 is also answered by the results of research question RQ1.

## 3.2 Research approach and research methods

The research of this dissertation is positioned between positivism and constructivism and applies both qualitative and quantitative methods. As Croom (2010) argue, there is no clear link between the epistemology and the choice of methods in social science research studies, and from a technical perspective, there is no clear dichotomy between qualitative and quantitative methods. The research should rather be problem-driven and not methodology-driven (i.a. Flyvbjerg, 2006). An important issue is that the researcher chooses the most appropriate methods for the investigation of the research questions (Croom, 2010).

The term "research method" refers to the technique of data collection and analysis rather than the interpretation of empirical findings (Croom, 2010). The overall research approach is concerned with the main issues of why one collected certain data; what data one collected; where one collected it; how one collected it; and how one analysed it (Neville, 2007). The research design consists of both strategic considerations and tactical decisions (Grenness, 1997). The tactical decisions are answered after the research design has been determined. The approach of this

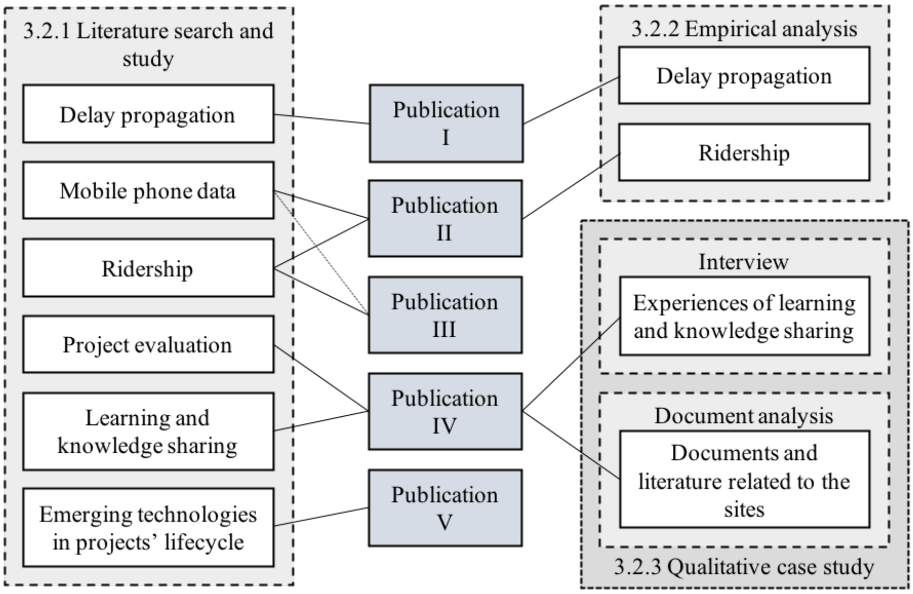


**Figure 3.1: Research design and methods in the thesis.**

dissertation was to attain relevant data for chosen cases. The tactical decisions were about deciding what these relevant data would be and determining what time periods, railway sections and railway projects to analyse.

The research methods and tactical decisions applied to answer each research question are described in the next sections. An overview of the research methods used in each of the publications of this dissertation is illustrated in Figure 3.1. The empirical data attained for each of the publications are presented in Figure 1.2 on page 7.

### 3.2.1 Literature search and study

Investigating previous studies on the research topics is necessary and a natural part of all research. Reviewing the literature can bring to light the limitations in current knowledge, focus the research questions and narrow the scope of the research (Croom, 2010). Furthermore, examining existing publications helps establish the authority and legitimacy of the research, ensure the researchability of the topic and clarify the possible contributions of the research. Thus, literature studies were conducted prior to and during the research.

The research questions demanded literature studies on several research topics, as illustrated in Figure 3.1. Two of the literature studies resulted in separate publications, i.e. Publication III and Publication V. The focus of the literature search has been on academic/scientific journals, books and conference proceedings, databases including the Norwegian library database (Bibsys), Scopus (which is the largest abstract and citation database of peer-reviewed literature) and Google Scholar.

Publication I is about delay propagation on single-track networks using traffic data. Thus, the search keywords included “delay propagation”, “knock-on delay”, “secondary delay”, “single-track”, “double-track” and “multi-track”, “real-time data” and “real-world data”.

Publication II concerns mobile phone data for analysis of the number of train travellers. The search keywords therefore included “mobile phone data”, “CDR”, “GPS”, “Wi-Fi” and “Bluetooth”, “travel mode”, “mobility”, “movement” and “origin-destination”. The literature search also involved the importance of ridership analysis, which is also a part of Publication III, as illustrated in Figure 3.1. Publication III was a literature review of common and new approaches and technologies on ridership and the importance of analysis of the number of travellers. Hence, the search keywords included “ridership”, “number of travellers and passengers”, “smart card”, “app”, “mobile phone” and “counting systems” in combination with transport and railway.

Publication IV concerns evaluation and learning experiences. The search keywords included “project evaluation”, “learning” and “knowledge sharing”.

Publication V was a literature review on Big Data in a life cycle perspective of construction projects within construction management research. The search keywords are described in the publication, but the main ones include “Big Data”, “construction project and industry”, and “management”. The previous research was reviewed and structured into the time perspective of a typical construction project based on a Norwegian model inspired by the RIBA Plan of Work model (RIBA, 2013; Knotten et al., 2016). The phases are strategic definition; concept development; concept design; detailed design; construction; handover; use and facility management; and disposal. The selected literature is categorised based on this life cycle perspective. In addition to the time perspective, we found a need for adding a second dimension and discuss a number of alternatives.

### 3.2.2 Empirical analysis

This section presents the research methods of the empirical studies.

#### Analyse delay propagation with traffic data

The research carried out to answer research question RQ1a resulted in Publication I, as illustrated in Figure 1.3 on page 8. Research question RQ1a asks how we can exploit traffic data to obtain new information on delay propagation that is relevant for project evaluation.

Performance measurement of railway traffic in several countries, including Norway, is based on two types of punctuality data. The first is traffic data that show at what time a train arrived or departed a defined point in the railway system. These registrations are commonly based on data from the signalling system. These data are typically of relative high quality, especially if the registrations include seconds and not just minutes. The second dataset is related to delay causes. Train dispatchers, or similar personnel, typically register these data manually. These data include a higher degree of subjectivity. Because the coding frequently includes a tagging of responsibility for a delay, it can cause heated discussions about root cause and responsibilities for delays. Reviews of the registrations have shown that some of the coding can be questioned. However, the performance measurement that is based on the delay coding typically receives more attention than the performance measurement based on pure time registrations. It is therefore a paradox that of the two available datasets, the best data are the least utilised, and the data of lower quality get much more attention. Traffic data is therefore interesting as a data source, either to support, or possibly even replace, manual registrations.

The approach of the study was to develop a method based on traffic data to analyse delay propagation. The data used in the study presented in Publication I were collected from three different sources, illustrated in Figure 1.2 on page 7. The main data source was of course traffic data, which are daily records containing time stamps for when each train passes home and departure signals (recorded to the second), along with train information and stations. The data are recorded by the Norwegian National Rail Administration (Jernbaneverket at the time of the study, now Bane NOR) in a national database called TIOS ("Trafikkinformasjon og oppfølgingssystem", translated as "Traffic Information and Follow-up System"). The data describe the movements of the trains through the network. The time registrations are generally based on automatic registrations and are considered accurate measurements of when trains enter and exit different sections of the network, with a high resolution (measured and recorded to the second).

The second source for data for this study was delay causes. The Norwegian railway authorities register delay causes with codes explaining causes for any delays in excess of the defined margins for the various classes of traffic. The causes of delays of more than four minutes and additional delay to already delayed trains are registered by the train dispatcher. The registrations are not automatic and thus are subject to the personal judgement of the train dispatcher. These registrations are considered less accurate than the pure time registrations.

The third source was the timetables for the case line. Both traffic data and timetables were available as graphic timetables in addition to tables. Graphic timetables illustrate the traffic data on a single-track railway line in a time-distance graph as illustrated in Figure 3.2, where the vertical axis refers to the stations along a railway line and the horizontal axis refers to time. The lines represent the traffic data, where the black lines describe the scheduled times and the red lines are actual times.

Norway's rail system is comprised primarily of single tracks in a star-shaped network. A lot of research has been conducted on double tracks, but less on single tracks. It is therefore of interest to study a single-track line. The railway line chosen was a single-track railway line that functions as the main connection between eastern Norway and Trøndelag and further north for passenger and freight traffic. The line is the Dovre Line (Dovrebanen) running from Eidsvoll station to Trondheim station. The line has a mixture of regional and long-distance passenger traffic as well as freight traffic, all on the same infrastructure. The railway line has 28 stations and stops for passenger traffic. The time period of the study and the data collection is

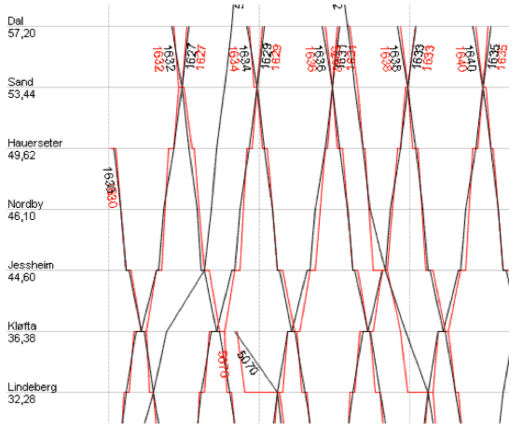


**Figure 3.2: Example of a graphic timetable on a single-track railway line.**

five months, from January to May 2015. This time period was selected because it is after schedule changes in December 2014 and before changes in the schedule in June 2015.

The research was a study of dynamic delay propagation using real-time train traffic data. The study included three parts. First, the research sought to develop an algorithm that uses traffic data to detect cases of knock-on delays, revealing dependencies between two trains on a single-track railway line. The set of conditions in the algorithm is thoroughly described in Publication I. Second, the research sought to automatically trace the propagation of knock-on delays from one case to the next, describing and quantifying the dynamic delay propagations. The algorithm that was developed was implemented recursively in an online analysis tool based on PostgreSQL 9.2 and the statistical software R 3.1. The implementation includes a combination of the algorithm, an extraction and aggregation framework for accessing the track-circuit information and some additional code for visualising the results in graphic timetable illustrations. The first prototype of the algorithm was implemented in Java 1.6.0.

The third part of the research focused on propagation factors to compare the results to manual delay cause registrations and to theoretical calculations of expected propagation factors. A propagation factor is the sum of delays in a dynamic delay propagation divided by the initial delay. The study also investigated the behaviour of the propagation factors found with the proposed algorithm. The method is described in detail in Publication I.

When applying the algorithm on the case sample, a delay limit of 239 seconds was chosen, coinciding with the official threshold for delay (< 4 minutes). This makes it possible to compare the results with the delay cause registration. The details of the data analyses are described in the publication.

### Study ridership with mobile phone data

The research carried out to answer research question RQ1b resulted in Publications II and III, as illustrated in Figure 1.3 on page 8. The research consisted of a literature review and an empirical study. Research question RQ1b asks how we can exploit mobile phone data to obtain new information on ridership that are relevant for project evaluation.

The process for evaluators to obtain data on ridership poses at least two major challenges. First, train operators consider such data confidential business information, especially in high resolution (Vigren, 2017). Second, the data that actually are available vary in quality and coverage. Data from on-board sensors and ticketing systems are typically managed by the transportation providers. On the other hand, surveys, payment statistics and mobile phone data may be available to external evaluators. Surveys have the advantage that they include important information about the respondent and the purpose of the trip. However, declining response rates (Schoeni et al., 2013) may introduce bias into the samples, in addition to the low frequency and high cost. Mobile phone data appear to be an interesting option because of their high frequency and detail and the potential to track complete journeys.

The procedure of the empirical study (Publication II) was to test approaches of how to apply handset counts from base stations to obtain information on ridership. Data used in the empirical study were collected from several sources, as illustrated in Figure 1.2 on page 7. The main data source for this study was mobile phone data. The mobile phone data are handset counts, which are the total number of handsets connected to a base station. Handset counts for a certain base station serve as an indicator of the number of people using the mobile network in the coverage area of the base station at the exact time the data are collected. The data were obtained from selected base stations. The datasets are described in detail in Publication II. For the purpose of this study, a script was made that extracts these data with certain intervals. When a phone turns up on another base station, it will no longer be counted on the previous one.

The mobile phone data used in this study are pure counts by cell by time unit, which are anonymous data that are not covered by privacy legislation and cannot be used to identify a person. Both from an ethical and legal point of view, it is important to protect personal information and respect people's privacy. To make sure that the study did not violate private privacy, the data used were approved by privacy representatives.

A second source of data used in this study was traffic data. The source is the TIOS database in which the time registrations generally are based on automatic registrations, as described above. The traffic data were used to connect trains running on the tracks with the time stamps of the handset counts. Location data for the base stations and for railway tracks and train stations were used to connect the mobile phone data with the train traffic data.

The last source of data used in the empirical study was passenger counts from trains based on automatic passenger counting (APC). The APC system is installed on a sample of the vehicles on the railway line that was studied. The APC data that were made available were collected from the trains on two of the railway lines that pass one of the studied train stations. The APC system registers the number of people who board and alight through each train door on every station by means of sensors in the doorways. The dataset for this study is the calculated total number of passengers on each train when the train is leaving the station.

The railway line chosen as a case in the empirical study was a railway line going into a city in Norway, where people commute daily to work from towns on the outskirts of the city. The analyses look at five base stations located near the railway tracks and between train stations on the selected railway line. The selected base stations are located so that it is likely that a large share of the mobile telephone traffic is generated by train passengers.

Data were collected in three time periods: (i) eight consecutive days in the spring of 2016, (ii) three consecutive days in the fall of 2016 and (iii) nine consecutive days in the spring of 2017. Traffic data were made available for all three time periods. The mobile phone data in time period i were collected with five-minute time intervals between each collection time. In time period ii, the mobile data were collected with one-minute time intervals between each collection time. This dataset is less complete than the dataset with five-minute collection intervals. In time period iii, the mobile data were collected with one-minute time intervals between each collection time. This dataset is complete. APC data were made available for time periods i and iii.

The research was a study of the potential for using handset counts from selected base stations to find numbers on ridership. The initial step was to compare the times of collection of handset counts with when trains were passing the base station. The method is described in detail in Publication II. With this method as a basis, the analysis included five steps, all of which are thoroughly described in Publication II. The first step was graphic inspection of peaks in handset counts in relationship to trains passing the base station. The trends at different times of the day were studied separately. Moreover, the train direction was added, and analyses for trains travelling in different directions were considered. The second step was analysis of data resolution by comparing datasets of five- and one-minute collection time intervals. In the third step, an algorithm was developed for a uniform evaluation of the different trends, and the data were categorised into multiple categories based on the time and direction information. Statistical analysis of the output values of the proposed algorithm was conducted. Furthermore, the study investigated distributional characteristics of the data to study the differences between direction (trains travelling towards or away from the city) and time of day (morning and afternoon rush hours). The fourth step was extracting the peaks in handset counts and analysing the peaks using the proposed algorithm, correlated with trains passing the base station. The fifth step was comparison and validation with APC data. The availability of the actual passenger count from the train operator served as a mechanism for validation of the ratio between the number of passengers and number of events on the adjacent cell sites.

### 3.2.3 Qualitative case study

Research question RQ2 asks how such empirical analyses are useful for project evaluation and learning in a life cycle perspective. A qualitative case study was carried out as a step to answering RQ2, which resulted in Pubication IV (see Figure 1.3 on page 8).

In the study, a qualitative case study research approach was used, as described by Yin (2009). Case study is a research method used for collecting data. It is useful in getting a comprehensive contextual view of a particular phenomenon. Information relating to the site was obtained from three main sources: literature related to the sites and other relevant documents, interviews and on-site inspection.

Case study data were collected in a case-study protocol that includes collected documentation, transcribed notes from interviews and codification of results to fit the applied evaluation framework. As Bowen (2009) points out, documents have immense value in case study research. In particular, information was obtained from literature related to the sites and other relevant documents to evaluate the case project. For the interviews, an interview guide from the SINTEF project SpeedUp was adapted. Three people were interviewed who had key roles in the case project.

The research mainly studied learning in one project, using multiple sources. However, the research also addressed how this project interacted with other projects. The railway project chosen as a case for this study was a major public investment executed by the Norwegian National Rail Administration (Jernbaneverket at the time of the study, now Bane NOR). The project was the construction of the railway tunnel through Gevingåsen and the connecting railway between Hommelvik and Hell on the Nordland line. The building started in the spring of 2009, and was finished in the fall of 2011.

## 3.3 Reliability, validity and generalization

There are two main quality criteria within research: the reliability of the results of the study/investigation and the validity of the results (Grenness, 1997).

Reliability is the extent to which a measurement procedure yields the same answer however and whenever it is carried out. It is about whether the measurements that lead to the final results have been executed with such a precision that we can trust them (Kirk & Miller, 1986; Grenness, 1997).

Validity is the extent to which the measurement procedure gives the correct answers, i.e. provides information that answers the research questions (Grenness, 1997; Tjora, 2013). A study can therefore never be valid without also being reliable. Hence, reliability is a condition for validity that is essential, but not sufficient.

### 3.3.1 Reliability

Reliability is the extent to which a measurement procedure would produce consistent findings if it was repeated on another occasion or was replicated by another researcher (Kirk & Miller, 1986).

Factors that could affect the reliability of a study include bias or errors regarding the analytical procedures. During the research, the author have made some choices that are accounted for in the publications. The author believes these will not affect the conclusion.

Other factors include bias or errors regarding data collection techniques and the measurement instruments. The data used in the empirical studies of Publications I and II are easily available, i.e. traffic data and handset counts (see Figure 1.2). In particular, traffic data are commonly used, for instance, in punctuality analysis. To support project evaluations with relevant and useful insight, it is relevant to check such data sources.

Furthermore, regarding data collection techniques and the measurement instruments, a possible direction of this dissertation could have been to not take these data sources for granted, but rather to do a deeper investigation into the reliability of the data sources and measurement instruments. The author is aware that the distance between the point of measure and the train station may vary and the railway company has therefore made corrections in the traffic data (Olsson et al., 2015a). However, this has not been the focus of this dissertation.

Within positivistic traditions, the ideal is neutral or objective observers (Tjora, 2013). In qualitative research, having neutral or objective observers is not as easy, and it is therefore important to give an account of how the researcher's position may affect the work. It is also important that the researcher give an account of what information is a result of the data generation and what is the researcher's own analyses (Seale, 1999). In qualitative research, there are two ways to improve reliability, according to Moisander & Valtonen (2006). The first is to make the research process transparent, and the second is to pay attention to theoretical transparency. Both have been applied in the qualitative case study, in which the authors described the theoretical basis for the proposed model and the research process. Furthermore, all three authors were involved in analysis of empirical data and results.

### 3.3.2 Validity

Validity is about whether the findings of the research can draw valid answers to the questions the study tries to answer (Grenness, 1997; Tjora, 2013). Do the findings provide information about and actually answer what was formulated in the aim and research questions? The methodology should therefore be chosen based on the research questions (Tjora, 2013).

First, assuring that the research is deeply rooted in other relevant research is the most important step to achieving a high degree of validity (Tjora, 2013). The research of this dissertation is built on several studies in research areas relevant to the topics researched, as illustrated in Figure 3.1.

The performance measures studied in this thesis do not provide a complete picture of the railway operations, but each is an important part of an evaluation. In this regard, the thesis has a poor validity. However, this was not the goal of the study, but rather to do a thorough investigation into two important parameters.

Validity is ensured in each of the studies, and quality assurance was done with other data sources.

The details of the data sources and the data analyses of this dissertation are thoroughly described in the publications. Being open about the practice of the research by giving an account of the choices made for methods on data generation and analysis is part of the effort to strengthen validity (Tjora, 2013).

The suitability of the data sources is also a part of validity. The author has done a thorough investigation into two parameters with empirical studies. The two main data sources of these studies (i.e. Publications I and II) are traffic data and handset counts. When evaluating railway operations, for their own purposes or to evaluate the effectiveness of railway projects, it is desirable to find out whether the railway is used, i.e. the number of travellers, and how punctuality, travel time and frequency change. Low punctuality may lead to delay propagations, which then may reduce punctuality further. Delay propagation is therefore a measure of the robustness of the timetable.

Traffic data describe the trains' movement through the network and are therefore a suitable source to analyse delay propagation of the actual railway operations. However, GPS data also describe the trains' movement and could therefore also have been used. As mentioned in the discussion of reliability, traffic data are more easily available and commonly used. The author has therefore used traffic data to investigate the appropriateness of this data source. The study of delay propagation (i.e. Publication I) includes a comparison to other data and therefore contributes to quality assurance.

Evaluators want to find out whether the railway is used, which is determined by measures of the number of travellers. This measure provides information on demand and is important for both project evaluation and evaluation of railway operations. Several different data sources and approaches were used to obtain numbers on traffic volume, as described in the literature review and Publication III. The author has used mobile phone data (i.e. handset counts) to investigate the appropriateness of this data source to obtain data on ridership. The benefit is that these data are already anonymized, compared to CDR data.

### 3.3.3 Generalization

Generalizability is a study's validity beyond the cases that were investigated (Tjora, 2013).

The study on delay propagation investigated a single-track railway line in Norway. Regarding generalization, the method can be applied to any single-track lines in Norway. The traffic data applied in the method to analyse delay propagation may differ from other countries, but the data are expected to be similar to the extent that they can be adapted to the method. The approach is therefore generalizable to other countries with single-track lines. The method cannot be easily adapted to double- or multi-track railways.

The study on mobile phone data to measure ridership investigated one railway line with data from one telecom operator in Norway. Regarding generalization, the mobile phone data used in the study are expected to be similar in other countries and companies, and the method should therefore be applicable in other countries and with data from other telecom operators. Other railway lines can also be analysed with this approach, provided that the base stations are selected based on the criteria discussed in Publication II.

The method and data source investigated on mobile phone data can be used on other transportation modes.

In this thesis, the studies are used to gain knowledge for ex-post evaluation of railway projects. These approaches of empirical analyses can also be applied prior to a project to find out what kind of actions are needed to improve railway operations, for instance, what section or junction on the line to focus on. The analyses can also be applied in performance evaluation of railway operations.

The analysis methods applied in the qualitative study are generalizable. Regarding the findings, one would have to run the analyses before one could say something about generalizability.

# Chapter 4

# Research findings

This chapter presents the key findings of each of the publications.

## 4.1 Publication I

The first publication is titled "Method of analysis for delay propagation in a single-track network". Delay propagation is a key factor in punctuality of rail services. Propagation of delays reflects the degree of robustness of timetable design and the stability of train operations. Traditional methods of descriptive statistics and tallying uncover little about the correlation between trains and presumed knock-on effects. The tool developed and presented in this paper is a first attempt at indicating the direction of knock-on effects in single-track traffic. The method opens up the possibilities of analysing larger time periods.

The algorithm developed in this study was applied to the traffic data from the Dovre Line and the selected five-month time period. The algorithm is thoroughly described in the paper. It is a recursive implementation of a set of conditions that (i) detects cases of knock-on delays, revealing dependencies between two trains on a single-track railway, and (ii) finds the networks of dynamic delay propagations by tracing the propagation of knock-on delays from one interaction between trains to the next. The program is not a simulation tool and cannot directly be used to predict outcomes of future changes to schedules. Instead, the program analyses real-world data gathered from actual traffic, finds delays and knock-on delays that occurred and identifies the networks of dynamic delay propagations. The application of the algorithm shows it successfully finds cases of knock-on delay and networks of dynamic delay propagation. From the analysis of a five-month period, the results showed that the proposed method makes it possible to find which train station and what time of day are most vulnerable for delay propagation. This is illustrated by Figure 4.1. The method can also identify the trains that are most often the ones that cause the first knock-on delay in a dynamic delay propagation, as well as the trains that most often are affected by dynamic delay propagation. The results of the five-month period also showed that more than 70 % of the dynamic delay propagations consist of two trains, meaning the train that was initially delayed plus one train that experienced a knock-on delay.

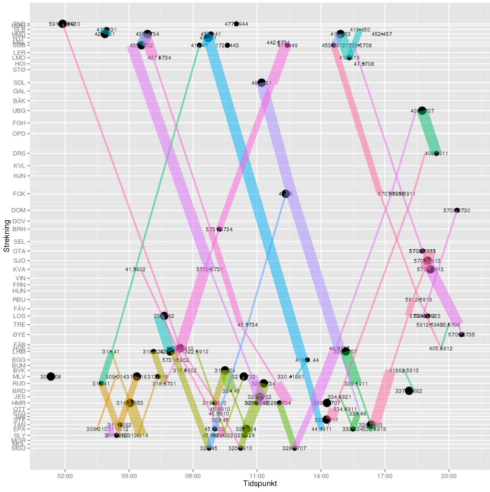


**Figure 4.1: Delay propagations on the Dovre Line in the time period January 1 to May 31, 2015, with the condition that the delay propagations occurred a minimum of seven times. The size of the dots and the thickness of the lines indicate relatively how many times the incidence occurred.**

The propagation factor was compared to both delay cause registrations and to a theoretical calculation of the expected propagation factors, which is the ratio between total delays of all the trains in a dynamic delay propagation and the initial delay of the delayer train. The comparison between propagation factors based on the proposed method and on delay cause registrations is provided in Figure 4.2. The observed propagation factors based on the proposed method are represented by the blue bars, and propagation factors based in delay cause registrations from the same five-month period are represented by the green bars. The mean propagation factor found with the method is about 2, which means that an initial primary delay is expected to generate secondary delays of the same magnitude as the initial delay. This corresponds with the finding that over 70 % of dynamic delay propagations consist of two trains. Figure 4.2 shows that the propagation factors based on delay cause registration were lower than with the proposed method, with a weight below 2. The study also found that the relationship between the initial delay and the propagation factor is in general $1 : n$, where $n$ is the number of trains in the dynamic delay propagation, including the delayer train. This result is specifically for dynamic delay propagations with less than seven trains and initial delays of around four to eight minutes.

The method therefore provides a picture of how the knock-on delay and dynamic delay propagations actually behave in the rail network. It can therefore be applied to analyse and compare the pre- and post-situation regarding delay propagation for a project evaluation. The method can support identification of which parts of a line, and at what time of day, delay propagation is most severe. Therefore, regarding project evaluation, perhaps it can reveal if the delay propagations have decreased

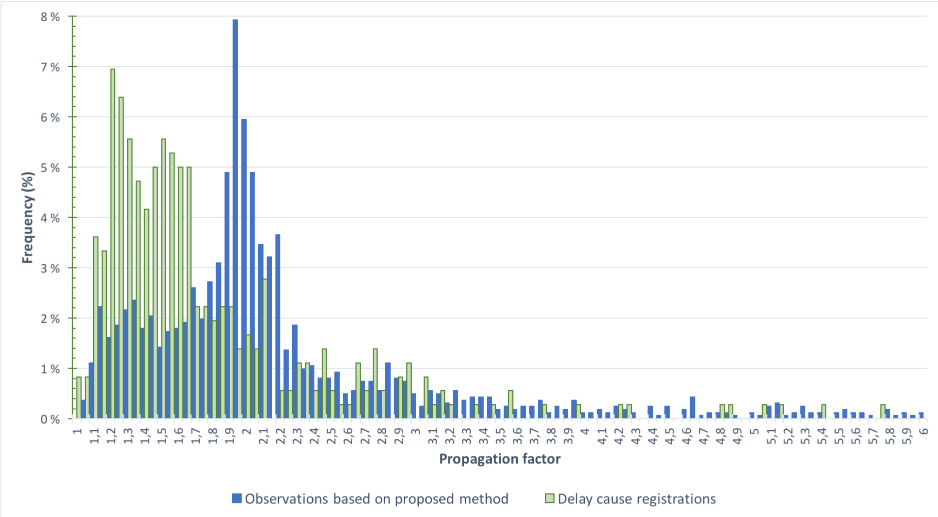


**Figure 4.2: Comparison of the distribution of propagation factors from the five-month period based on the proposed method and delay cause registrations.**

or shifted as a result of the project.

## 4.2 Publication II

The second publication is titled "Use of mobile phone data for analysis of number of train travellers". Studies on mobile phone data are suggested and tested for mobility analysis in the literature. The research in Publication II suggests handset counts from base stations as a data source to obtain numbers on ridership for evaluators. The study shows a correlation between handset counts and trains passing mobile base stations and that this source should be further investigated. The study shows that it is possible to combine mobile phone data with railway infrastructure and train traffic data. It also discusses preliminary results and methodological and technical challenges.

Handset counts were obtained from one telecom operator for selected mobile phone base stations and compared with timetable data and APC. The selected base stations are located so that it is likely that a large share of the mobile phone traffic is generated by train passengers. The study found a connection between the train passing and changes in the handset counts, and the number of units connected to a base station was found to correspond relatively well with the trains that pass close to the base stations.

The findings show that one-minute collection time intervals are more favourable than five-minute intervals. As illustrated by Figure 4.3, the frequency in peaks is better captured with the one-minute collection time intervals, as are peaks that are connectable to trains passing the base station.

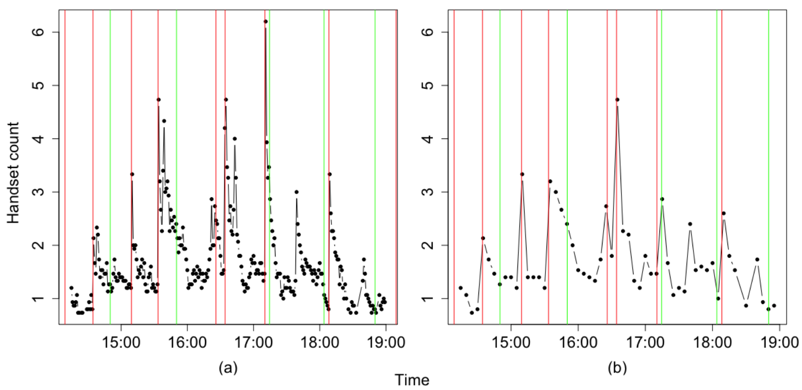


**Figure 4.3: Handset count with one-minute collection time intervals in (a) and the one-minute interval handset count aggregated to five-minute intervals in (b). Times when trains are passing are represented by the vertical green lines (towards the city) and red lines (away from the city).**

These preliminary results show a connection between the train passing and changes in the handset counts. However, it is also evident that a lot of the trains passing the base station are not detectable in the handset counts. Furthermore, some peaks seem to occur even though no trains were passing the base station.

The strengths and weaknesses of formats for presenting and analysing the results were discussed. The findings showed that it is better to separate the data between which directions the trains are travelling in when analysing them, because more people are travelling towards the city in the morning and away from the city in the evening than during the rest of the day.

The handset counts were compared to the number of travellers as measured by on-board APC equipment. The ratio of handsets over passengers varies, as shown in Figure 4.4, indicating that a large-scale calibration is needed, using more data than we had available, to increase the accuracy of handset counts as indicators of the number of travellers. A ratio between the handset count and APC data appears promising in utilising handset count to calculate train ridership, with ratios around one in the rush hours.

## 4.3 Publication III

The third publication is titled "Approaches, technologies and the importance of analysis of number of travellers". The number of travellers is an important performance measure in railways and thus for project evaluation. The benefit of, for instance, reduced travel time is dependent on traffic volume because more traffic volume decreases the cost per passenger or per tonne of freight. This is a review of how the number of travellers on trains is measured, including technologies and practices for measuring actual ridership. The studied technologies and approaches for ridership

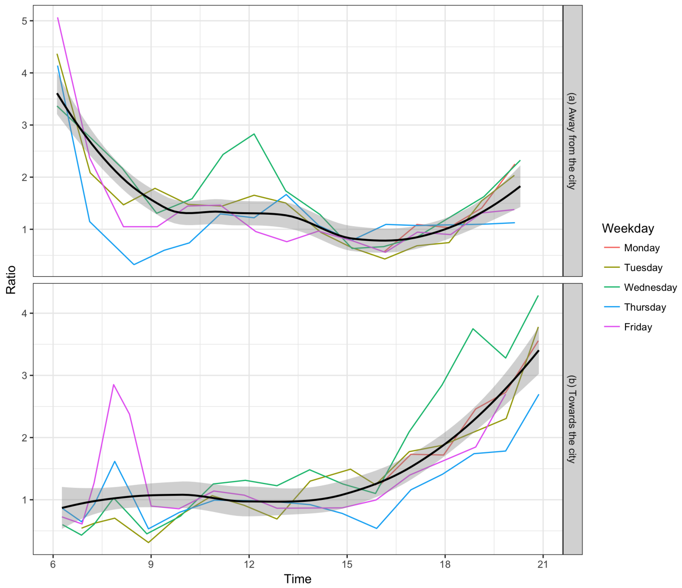


**Figure 4.4: Ratio of handset counts to the number of passengers (APC) for trains going towards the city (b) and away from the city (a). The black lines represent a smooth curve fitted by loess regression with confidence bands.**

measurement are summarised in a table that provides an overview of each indicator on ridership, its source, part of trip, data owner, strength and weakness. Because the traditional methods have been described as sporadic and inefficient because of small sample sizes in relationship to the target and a poor level of accuracy, new technologies are of special interest, especially mobile phone data.

Several publications and practical work have been done on estimating ridership. The review identifies several technologies that can be applied for measuring ridership on trains. The technologies and approaches include (1) manual counts and surveys, (2) on-board sensors such as door passing, weight, CCTV and Wi-Fi-use, (3) ticketing systems, ticket sales or ticket validation and (4) tracking of travellers for larger parts of the journey, for instance, by mobile phones and payments.

Regarding the part of the trip, it was found that most technologies measure either the number of travellers continuously or at entry and exit points at train stations. Use of on-board sensors is established. However, it mainly measures ridership on individual rolling stock units. It is less suitable for measuring multi-mode journeys or even transfers between lines in a train system. Ticketing systems can measure journeys within the ticketing system, which can cover both transfers between lines and between travel modes within the same public transport system. However, it requires ticket validation at entry and preferably also at exit. Tracking of travellers for larger parts of journeys can be done by traditional surveys. They are costly and dependent on response rates. They also measure people's stated preferences and impressions of their travel time. Payment tracking requires access to payment statistics. It only records electronic payments and only records visits to commercial areas. Finally, mobile phone data appear to be an interesting option. They can track complete journeys, with accuracy on the level of coverage of base stations or even more accuracy if apps allowing for GPS tracking are used.

Evaluators experience difficulty in obtaining good data on the number of travellers, which several studies have pointed out. This is often because train operators consider such data confidential business information. Furthermore, the data on ridership that are available vary in quality and coverage.

Data from on-board sensors and ticketing systems are both managed by the public transportation providers and may not necessarily be available to others. In contrast, surveys, payment statistics and mobile phone data may be available to stakeholders outside the public transportation system such as evaluators, which can be an advantage. Furthermore, mobile phone data appear to be an interesting option because they can track complete journeys that include several modes of transportation. New technologies, especially mobile phone data, are therefore of special interest in future uses of ridership data for evaluations and quality assessments.

## 4.4 Publication IV

The fourth publication is titled "Evaluation and learning—Experiences from a construction project in Norway". The paper studied the relationship between evaluation and learning, considering the approaches to learning and knowledge sharing in a construction project in Norway, and discussed the role of evaluation in relationship to experience transfer. The qualitative case study approach illustrated how project evaluations typically employ aggregated data as a source to evaluate a project, as illustrated in Figure 1.2.

In this paper, the term "learning" was considered a process in which knowledge sharing plays a significant role. The term "knowledge sharing" included sharing of information, reflection and sense-making by the individuals and possible interactions between the individuals who were involved in the sharing process. This definition therefore incorporates three categories of knowledge. Knowledge-as-data points out the explicit and objective characteristics of knowledge. Knowledge-as-meaning deals with reflection and sense-making. Knowledge-as-practice incorporates tacit characteristics of knowledge.

The paper described different approaches to evaluation and learning and proposed the model illustrated in Figure 4.5. The model distinguishes between an internal and external perspective when it comes to evaluation of projects, and between a structured and an informal perspective when it comes to learning and knowledge sharing. The structured approach typically focuses on knowledge-as-data, such as searchable repositories. The informal approach focuses on human interaction and includes storytelling and ad-hoc experience transfer. With the focus on this model, the paper also describes enablers and barriers of learning and knowledge sharing. The model provides a structured illustration of the connection between project evaluation and learning. Thus, the model would be useful for identifying and applying learning mechanisms for both internal and external evaluation of projects.

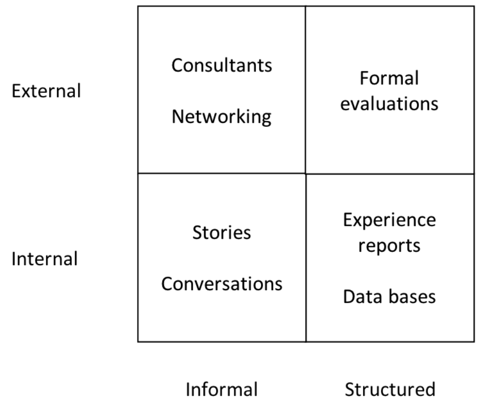


**Figure 4.5: Illustration of approaches to knowledge sharing in different perspectives.**

Based on this model, traditional project evaluations can be categorised as externally structured. Evaluation by an external evaluator has the benefit that the evaluators are outsiders. They look at the project from a wider perspective and a different value system. The external evaluator is more likely to ask questions about assumptions and interpretations that the project is based upon. Regarding learning, this type of evaluation appears to not necessarily be an important tool for learning because external-formal evaluations were not used in the studied case. Internally structured approaches are important, mainly because of a high demand for experience reports. The process of an internal evaluation incorporates reflection during and at the end of the project that will lead to learning. Furthermore, the study found that external informal learning was important. One example was when consultants shared experiences in their home organisation.

## 4.5 Publication V

The fifth publication is titled "Big Data in construction management research". "Big Data" has become a common term in most businesses. However, there is little published management scholarship that tackles the challenges of using Big Data, or even that explores the opportunities for new theories and practices that Big Data might bring about. There is a need for further discussion of the possible implications Big Data can have for construction management research. The reviewed literature shows practical examples of uses and potential uses of Big Data in construction management research. The literature review illustrates the increase in different sources of data that are becoming available as a consequence of emerging technologies and digitalization, within the life cycle of construction projects.

The findings show there is a lot of literature that can be used in projects. The identified studies describe Big Data applications and theory. Thematically, they fall into three broad categories: 1) new construction equipment that generate, share and store data about use; 2) data from internal IT systems such as planning, procurement and building information modelling (BIM) can be utilised; and 3) people generate an increasing flow of information that can be useful if handled with care. In combination, this addresses the life cycle from concept to decommissioning.

The construction process can be studied in a number of dimensions. Documented and potential use of Big Data related to construction projects reviewed in this study was structured in the time perspective of a typical construction project. The research found that data generated from new technologies can be useful for several phases in the life cycle of a construction project from the development of concept to the actual construction to the operation phase and finally completion. The review found that the construction phase appears to have received the most research attention. Furthermore, several studies are applicable to more than one phase of a construction project. The findings show a potential for increased use of Big Data methods and applications within construction. While some data and applications have been analysed in isolation previously, there is a potential to combine different types of data.

# Chapter 5

# Concluding discussion

Ex-post project evaluation should be carried out a few years after major infrastructure projects are completed to establish to what degree the intended effects on the users and the society have been achieved. An evaluation should evaluate tactical success in terms of effectiveness, which measures if the project achieved its goals. Evaluating effectiveness requires data from the railway operations, but evaluators experience problems in getting hold of essential data on the pre- and post-situation. Relevant data are mainly available as aggregated numbers from performance and evaluation reports. In contrast, technological developments increase the amount of available data sources. The aim of this thesis has been to obtain new insights to support ex-post evaluation of railway projects with available data sources.

This chapter presents a discussion of the key contributions of the dissertation and offers final conclusions and suggestions for further work.

## 5.1 Traffic data to analyse delay propagation (RQ1a)

*How can one exploit available data sources in railway to obtain information relevant for project evaluation, specifically to evaluate delay propagation with traffic data?*

Research question RQ1a is answered by Publication I. The research of Publication I shows a practical example of how traffic data (i.e. data describing train movements) can be analysed to describe delay propagation.

Because of continual developments in technologies and the adoptions of these in the railway operations, data describing train movement have evolved from manual records to more accurate and detailed databases. There is a focus on digital transformation in the railway research. Data generated during railway operations can be used to evaluate the effectiveness of railway projects in ex-post evaluations. Traffic data, which are a result of the digitalization in railways, were applied in this dissertation to analyse delay propagation.

Publication I presented an algorithm for quantifying delay propagation. The algorithm was applied to railway traffic data, showing that traffic data can be successfully used to analyse delay propagation in a single-track railway network. The algorithm and software developed in this study indicate how delays and delay propagations behave on single-track railways.

Delay propagation reflects the degree of robustness and stability and hence provides a deeper analysis of punctuality. Punctuality is an important measure used in evaluation of railway projects, as well as the performance measure in railway operations. Delay propagation is a measure that is useful to more thoroughly describe the effects and impact of railway projects.

## 5.2 Mobile phone data to measure ridership (RQ1b)

*How can one exploit available data sources in railway to obtain information relevant for project evaluation, specifically with mobile phone data to evaluate ridership?*

Research question RQ1b is answered by Publications II and III. The research of Publication II shows that handset counts have the potential to provide data on ridership.

Data generated during the railway operations can be used to evaluate the effectiveness of railway projects in ex-post evaluations. The number of travellers is an important measure in project evaluation and in the performance of railway operations. The literature shows there are new developments in technologies and practices for collecting ridership, each of which has strengths and weaknesses. However, such data are not only directly generated from the railway operations, but can also be mobile phone data from mobile base stations. Base stations are existing data-capture infrastructure in close proximity to the railway tracks. The literature shows use of mobile phone data in different mobility research, including examples on travellers with different mode.

The research of Publication II investigated the potential for using handset counts to measure train ridership. It showed that it is possible to combine mobile data with railway infrastructure and train traffic data. The study found a connection between the trains passing the base stations and changes in the handset counts.

The main implication of the findings is that mobile phone data can potentially be used for ridership analyses. The results show a potential for utilising mobile phone data to collect the number of travellers on the railway. Several technical and methodological challenges need to be addressed for such an approach to become suitable for use in project evaluation. However, the research is an initial step towards this and shows promising results. The research contributed to a preliminary approach that can inspire further research on the use of mobile phone handset counts as a source to analyse ridership.

## 5.3 Empirical analyses for learning and evaluation in a life cycle perspective (RQ2)

*How are such empirical analyses useful for learning from project evaluation in a life cycle perspective?*

The findings show that technological developments increase the amount of available data sources within railway operations and construction projects. These data sources are interesting to consider for project evaluation because they might contribute to structured approaches of knowledge sharing. Traditionally, ex-post project evaluation has been mostly qualitative, but the author has shown how available data sources can be processed and analysed.

The findings show that it is possible to follow up on railway projects in a life cycle perspective. The studies of this thesis can be seen in two dimensions, qualitative and quantitative, and in the life cycle perspective. The studies involve two of the project's phases, i.e. construction and operation. The construction phase is covered in Publication IV with a qualitative approach (i.e. the experience reports). The operation phase is investigated in Publications I and II with quantitative approaches. The measures that were thoroughly investigated, i.e. delay propagation and ridership, are measures that are relevant in projects' "use" phase.

The data collected from the railway operations are useful when evaluating the tactical and strategic success of a project in a formal project evaluation (see Figure 5.1). Such evaluation is an external and structured approach of knowledge sharing, which includes evaluation of effectiveness, impact, relevance and sustainability and is important as a quality assurance. Learning and knowledge sharing are important parts of the final phase of railway projects. Assessment of the sustainability of results and impact and the conclusions drawn from the project can provide useful insight into the strategic planning and concept development of future projects (see Figure 5.1).

The technologies utilised in construction projects generate data that are useful for evaluating the project's execution (see Figure 5.1). Such evaluations include, first, experience reports, which are an internally structured approach for knowledge sharing and have been in high demand, and second, the evaluation of efficiency (i.e. cost, time and quality) in a formal project evaluation (external and structured). In the early phase of a project, experience reports together with the evaluation of efficiency are important to learn from similar projects for the project's execution, as found in Publication IV. This is part of the single-loop learning and can contribute to making faster decisions.

The author has focused on the upper part of Figure 5.1 and provided two practical examples of how data generated and collected during the railway operations can be exploited to obtain relevant information for project evaluation. The studies on train traffic data and mobile phone data are good examples of how data can provide new and useful insights to the evaluation of railway projects.

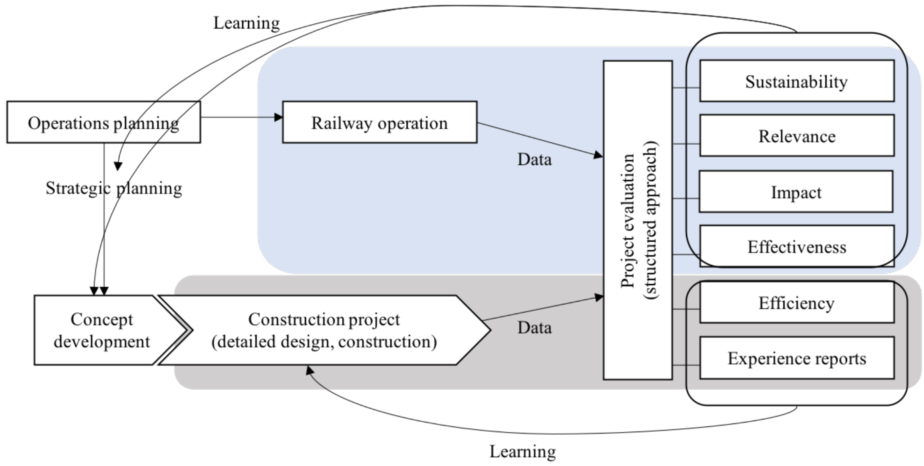


**Figure 5.1: Where the data source originates from, which part of the evaluation they are used for, and the learning loop.**

## 5.4 Final remarks

The aim of this thesis was to obtain new insights to support ex-post evaluation of railway projects. The author obtained data from existing data-capture infrastructures and showed that it is possible to combine several data sources.

The author gained some experiences from executing the studies. For instance, the mobile phone data contained some irregularities and noise that caused unclear patterns where peaks not always corresponded with the trains. With access to train traffic data, the author showed that it is possible to do something fun with it rather than just calculating statistics on punctuality. The author learned that traffic data have become more easily available. Getting hold of data has been a problem, and is still partially, especially when it is about private information. In contrast, other sources of data has started to become more available. Nevertheless, the challenge is to exploit these data sources.

## 5.5 Further work

This chapter presents suggestions for further research within this research topic.

Regarding the life cycle of projects, this dissertation has studied two of the phases in a project, i.e. construction phase and use phase. One could also investigate the possibilities of applying quantitative methods in the other phases of projects.

This dissertation investigated a method to use traffic data to analyse delay propagation on a single-track railway network. Because the railway network in Norway consists of both single and double tracks connected in a star-shaped network around the capital, a similar method for double tracks should be studied. This would provide a more complete evaluation of delay propagation on the network. Furthermore, future iterations would also have to better account for and handle delays in the junctions between single- and multi-track infrastructure, as well as better handling delays induced in first-/final-stop (where head codes change). In addition, further development of the algorithm should include validation with competing methods of single-track delay propagation, including theoretical models for further calibration and validation of the conditions considered knock-on delays.

Regarding using mobile phone data to obtain information on number of travellers, the study is at an early stage. The method is not complete and ready to be used, but rather an early investigation into the possibilities and the suitability of the data source. The format provides a good starting point for looking into the concept of utilising mobile phone data in ridership evaluation. Further work should therefore involve a further development of the method, including calibrating over a longer time period and over more trains. The study can also be expanded to explore base stations located in tunnels. Railway tunnels should be a good indicator that the connected handset belongs to someone travelling on the railway. Furthermore, an investigation into travels from an origin to a destination could be interesting.

Regarding the data sources, further research could include a deeper investigation into the reliability of the data sources and measurement instruments.

Regarding further work on evaluation, one could interview more people, both those who worked on the project execution, as was done in this thesis, and people who worked on earlier phases of the project, such as strategic planning and concept development. In addition, it would be interesting to use this type of quantitative analysis on more project evaluation and over longer time periods.

In construction projects, there are constantly new data from new technologies. In a further study, one could use such new data in construction projects and connect them to planning and control. A further work could draw inspiration from the type of analyses the author and others have done on transportation, for instance, in a study of logistics.

Regarding new data that are generated, it will in future be pertinent to pay attention to GDPR, particularly on how the balance will develop. It may require more confidentiality, i.e. more pressure on personal privacy. Some types of data might have improved availability, while the protection will increase on other types of data.

# Bibliography


Aguiléra, V., Allio, S., Benezech, V., Combes, F. & Milion, C. (2014). Using cell phone data to measure quality of service and passenger flows of Paris transit system. *Transportation Research Part C: Emerging Technologies*, 43, 198–211. ISSN 0968-090X.

Alexander, L., Jiang, S., Murga, M. & González, M. C. (2015). Origin–destination trips by purpose and time of day inferred from mobile phone data. *Transportation Research Part C: Emerging Technologies*, 58, Part B, 240–250. ISSN 0968-090X. doi:10.1016/j.trc.2015.02.018.

Andersen, B., Bråthen, S., Fagerhaug, T., Nafstad, O., Næss, P. & Olsson, N. (2007). *Effektvurdering av store statlige investeringsprosjekter (Impact Assessment of Major Public Investment Projects).* Report, NTNU.

Andersson, E. V., Peterson, A. & Törnquist Krasemann, J. (2014). Quantifying railway timetable robustness in critical points. *Journal of Rail Transport Planning & Management*, 3(3), 95–110. ISSN 2210-9706. doi:10.1016/j.jrtpm.2013.12.002.

BaneNOR (2015). *Railway Statistics 2015.* Report, The Norwegian National Rail Administration. URL `www.banenor.no`.

Barabino, B., Di Francesco, M. & Mozzoni, S. (2014). An offline framework for handling automatic passenger counting raw data. *Intelligent Transportation Systems, IEEE Transactions on*, 15(6), 2443–2456. ISSN 1524-9050. doi:10.1109/TITS.2014.2315573.

Becker, R., Cáceres, R., Hanson, K., Isaacman, S., Loh, J. M., Martonosi, M., Rowland, J., Urbanek, S., Varshavsky, A. & Volinsky, C. (2013). Human mobility characterization from cellular network data. *Communications of the ACM*, 56(1), 74–82. ISSN 0001-0782.

Bešinović, N., Quaglietta, E. & Goverde, R. M. P. (2013). A simulation-based optimization approach for the calibration of dynamic train speed profiles. *Journal of Rail Transport Planning & Management*, 3(4), 126–136. ISSN 2210-9706. doi:10.1016/j.jrtpm.2013.10.008.

Bianchi, F. M., Rizzi, A., Sadeghian, A. & Moiso, C. (2016). Identifying user habits through data mining on call data records. *Engineering Applications of Artificial Intelligence*, 54, 49–61. ISSN 0952-1976. doi:10.1016/j.engappai.2016.05.007.

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative research journal*, 9(2), 27–40. ISSN 1443-9883. doi:10.3316/QRJ0902027.

Boyle, D. K. (1998). *Passenger Counting Technologies and Procedures*. Report, Transportation Research Board.

Calabrese, F., Di Lorenzo, G., Liu, L. & Ratti, C. (2011). Estimating origin-destination flows using mobile phone location data. *IEEE Pervasive Computing*, 10(4), 36–44. doi:10.1109/mprv.2011.41.

Calabrese, F., Diao, M., Di Lorenzo, G., Ferreira, J. & Ratti, C. (2013). Understanding individual mobility patterns from urban sensing data: A mobile phone trace example. *Transportation Research Part C: Emerging Technologies*, 26, 301–313. ISSN 0968-090X. doi:10.1016/j.trc.2012.09.009.

Carrillo, P., Ruikar, K. & Fuller, P. (2013). When will we learn? Improving lessons learned practice in construction. *International Journal of Project Management*, 31(4), 567–578. ISSN 0263-7863. doi:10.1016/j.ijproman.2012.10.005.

Chu, X. (2010). *A guidebook for using automatic passenger counter data for National Transit Database (NTD) reporting*. Report, National Transit Resource Center.

Conte, C. (2007). Identifying dependencies among delays. PhD thesis, Georg-August-Universität zu Göttingen.

Courtney, M. (2012). Puzzling out big data. *Engineering and Technology*, 7(12), 56–60. doi:10.1049/et.2012.1215.

Cracknell, B. (1989). Evaluating the effectiveness of the logical framework system in practice. *Project appraisal*, 4(3), 163–167. ISSN 0268-8867.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*. SAGE Publications, Incorporated, 3rd edition. ISBN 141296556X.

Croom, S. (2010). Introduction to research methodology in operations management. In C. Karlsson (ed.), *Researching operations management*, pp. 42–83. Routledge. ISBN 9781135854027. doi:10.4324/9780203886816.

Cule, B., Goethals, B., Tassenoy, S. & Verboven, S. (2011). Mining train delays. In *International Symposium on Intelligent Data Analysis*, pp. 113–124. Springer.

Davenport, T. H., Barth, P. & Bean, R. (2012). How 'big data' is different. *MIT Sloan Management Review*, 54(1). ISSN 15329194.

Davies, P. (2011). Exploratory research. In *The SAGE Dictionary of Social Research Methods*. SAGE Publications, Ltd. doi:10.4135/9780857020116.

De Mauro, A., Greco, M. & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. *Library Review*, 65(3), 122–135. ISSN 0024-2535. doi:10.1108/LR-06-2015-0061.

Doi, M. & Allen, W. B. (1986). A time series analysis of monthly ridership for an urban rail rapid transit line. *Transportation*, 13(3), 257–269. ISSN 1572-9435. doi:10.1007/BF00148619.

Doyle, J., Hung, P., Kelly, D., McLoone, S. & Farrell, R. (2011). Utilising mobile phone billing records for travel mode discovery. URL `http://eprints.maynoothuniversity.ie/3649/`.

Drageide, V. (2009). Towards privacy management of information systems. Master thesis, University of Bergen.

D'Ariano, A. & Pranzo, M. (2009). An advanced real-time train dispatching system for minimizing the propagation of delays in a dispatching area under severe disturbances. *Networks and Spatial Economics*, 9(1), 63–84. ISSN 1566-113X. doi:10.1007/s11067-008-9088-1.

D'Ariano, A., Pranzo, M. & Hansen, I. A. (2007). Conflict resolution and train speed coordination for solving real-time timetable perturbations. *IEEE Transactions on Intelligent Transportation Systems*, 8(2), 208–222. ISSN 1524-9050. doi:10.1109/TITS.2006.888605.

Eikeland, P. T. (2001). *Teoretisk analyse av byggeprosesser (Theoretical analysis of construction processes),*. Report Project number P10602.

Fellows, R. & Liu, A. (2003). *Research Methods for Construction*. Blackwell Science, 2nd edition. ISBN 0-632-06435-8.

Flier, H., Gelashvili, R., Graffagnino, T. & Nunkesser, M. (2009). Mining railway delay dependencies in large-scale real-world delay data. In *Robust and online large-scale optimization*, pp. 354–368. Springer. doi:10.1007/978-3-642-05465-5_15.

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry*, 12(2), 219–245. doi:10.1177/1077800405284363.

Fowkes, A. S., Nash, C. A. & Whiteing, A. E. (1985). Understanding trends in inter-city rail traffic in great britain. *Transportation Planning and Technology*, 10(1), 65–80. ISSN 0308-1060.

Frias-Martinez, V., Soguero, C. & Frias-Martinez, E. (2012). Estimation of urban commuting patterns using cellphone network data. In *Proceedings of the ACM SIGKDD International Workshop on Urban Computing*, pp. 9–16. ACM. ISBN 1450315429.

Frumin, M. S. (2010). Automatic data for applied railway management: passenger demand, service quality measurement, and tactical planning on the London Overground Network. Master thesis, Massachusetts Institute of Technology.

George, G., Haas, M. R. & Pentland, A. (2014). Big data and management. *Academy of Management Journal*, 57(2), 321–326. doi:10.5465/amj.2014.4002.

Gordillo, F. (2006). The value of automated fare collection data for transit planning: an example of rail transit od matrix estimation. Master thesis, Massachusetts Institute of Technology. URL `http://hdl.handle.net/1721.1/38570`.

Goverde, R. M. (2005). Punctuality of railway operations and timetable stability analysis. PhD thesis, TU Delft, Delft University of Technology.

Goverde, R. M. (2010). A delay propagation algorithm for large-scale railway traffic networks. *Transportation Research Part C: Emerging Technologies*, 18(3), 269–287. ISSN 0968-090X. doi:10.1016/j.trc.2010.01.002.

Goverde, R. M. P. & Hansen, I. A. (2001). Delay propagation and process management at railway stations. In *5th World Conference on Railway Research (WCRR 2001), Köln, November 25-29*. WCRR.

Goverde, R. M. P. & Meng, L. (2012). Advanced monitoring and management information of railway operations. *Journal of Rail Transport Planning & Management*, 1(2), 69–79. ISSN 2210-9706. doi:10.1016/j.jrtpm.2012.05.001.

Green, A. (2017). Technology and digitalisation at forefront of freight development. *International Railway Journal*, 57(7), 50. URL `https://search.proquest.com/docview/1926952207?accountid=12870`.

Grenness, T. (1997). *Innføring i vitenskapsteori og metode*. Tano Aschehoug, Oslo. ISBN 8251836190. URL `https://urn.nb.no/URN:NBN:no-nb_digibok_2008071400110`.

Gundlegård, D., Rydergren, C., Breyer, N. & Rajna, B. (2016). Travel demand estimation and network assignment based on cellular network data. *Computer Communications*, 95, 29–42. ISSN 0140-3664. doi:10.1016/j.comcom.2016.04.015.

Handstanger, A. C. T. (2009). Scheduled waiting time from crossing on single track railway lines. PhD thesis, Norwegian University of Science and Technology. URL `http://hdl.handle.net/11250/231388`.

Hansen, I. A. (2001). Improving railway punctuality by automatic piloting. In *Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE*, pp. 792–797. IEEE. ISBN 0780371941.

Hansen, I. A. & Pachl, J. (2014). *Railway Timetabling & Operations: Analysis, Modelling, Optimisation, Simulation, Performance Evaluation*. Eurailpress/DVV Media Group, Hamburg, 2nd edition. ISBN 3777104621.

Harris, N. G., Haugland, H., Olsson, N. & Veiseth, M. (2016). *Introduction to Railway Operations Planning*. A & N Harris, London. ISBN 0952999765.

Harris, N. G., Mjøsund, C. S. & Haugland, H. (2014). Improving railway performance in Norway. *Journal of Rail Transport Planning & Management*, 3(4), 172–180. ISSN 2210-9706. doi:10.1016/j.jrtpm.2014.02.002.

Hassard, J. (1988). Overcoming hermeticism in organization theory: An alternative to paradigm incommensurability. *Human Relations*, 41(3), 247–259. ISSN 0018-7267. doi:10.1177/001872678804100304.

Henry, C. & Grant, S. (2012). Implementing new automated ticketing technology at Virginia Railway Express. In *2012 Joint Rail Conference*, pp. 449–454. American Society of Mechanical Engineers.

Higuchi, T., Yamaguchi, H. & Higashino, T. (2015). Mobile devices as an infrastructure: A survey of opportunistic sensing technology. *Journal of information processing*, 23(2), 94–104. ISSN 1882-6652.

Hilbert, M. (2013). Big Data for development: From information-to knowledge societies. doi:10.2139/ssrn.2205145.

Holden, M. T. & Lynch, P. (2004). Choosing the appropriate methodology: Understanding research philosophy. *The marketing review*, 4(4), 397–409. ISSN 1469-347X. doi:10.1362/1469347042772428.

Holleczek, T., Yu, L., Lee, J. K., Senn, O., Ratti, C. & Jaillet, P. (2014). Detecting weak public transport connections from cellphone and public transport data. In *Proceedings of the 2014 International Conference on Big Data Science and Computing*, p. 9. ACM. ISBN 1450328911.

Iqbal, M. S., Choudhury, C. F., Wang, P. & González, M. C. (2014). Development of origin–destination matrices using mobile phone call data. *Transportation Research Part C: Emerging Technologies*, 40, 63–74. ISSN 0968-090X.

Jagadish, H. V. (2015). Big Data and Science: Myths and Reality. *Big Data Research*, 2(2), 49–52. ISSN 2214-5796. doi:10.1016/j.bdr.2015.01.005.

Jiang, S., Fiore, G. A., Yang, Y., Ferreira Jr, J., Frazzoli, E. & González, M. C. (2013). A review of urban computing for mobile phone traces: current methods, challenges and opportunities. In *Proceedings of the 2nd ACM SIGKDD international workshop on Urban Computing*, p. 2. ACM. ISBN 1450323316.

Jánošíková, L., Kavička, A. & Bažant, M. (2014). Optimal operation scheduling and platform track assignment in a passenger railway station. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 228(3), 271–284. ISSN 0954-4097. doi:10.1177/0954409712472275.

Järv, O., Ahas, R., Saluveer, E., Derudder, B. & Witlox, F. (2012). Mobile phones in a traffic flow: a geographical perspective to evening rush hour traffic analysis using call detail records. *PloS one*, 7(11), e49171. ISSN 1932-6203. doi:10.1371/journal.pone.0049171.

Kecman, P. & Goverde, R. M. (2013). Process mining of train describer event data and automatic conflict identification. *Computers in Railways XIII: Computer System Design and Operation in the Railway and Other Transit Systems*, 127, 227. ISSN 1845646169. doi:10.2495/CR120201.

Kirk, J. & Miller, M. L. (1986). *Reliability and Validity in Qualitative Research*. Qualitative Research Methods. SAGE Publications. ISBN 9781412985659. doi: 10.4135/9781412985659.

Knotten, V., Hosseini, A. & Klakegg, O. J. (2016). "Next Step": A New Systematic Approach to Plan and Execute AEC Projects. In *Proceedings of the CIB World Building Congress 2016, Volume III. Building up busbiess operations and then logic. Shaping materials and technologies*, pp. 484–495.

Kujala, R., Aledavood, T. & Saramäki, J. (2016). Estimation and monitoring of city-to-city travel times using call detail records. *EPJ Data Science*, 5(1), 6. ISSN 2193-1127.

Laney, D. (2001). 3-D data management: Controlling data volume, velocity and variety. URL `http://goo.gl/wH3qG`.

Larijani, A. N., Olteanu-Raimond, A.-M., Perret, J., Brédif, M. & Ziemlicki, C. (2015). Investigating the mobile phone data to estimate the origin destination flow and analysis; case study: Paris region. *Transportation Research Procedia*, 6, 64–78. ISSN 2352-1465. doi:10.1016/j.trpro.2015.03.006.

Leo, Y., Busson, A., Sarraute, C. & Fleury, E. (2016). Call detail records to characterize usages and mobility events of phone users. *Computer Communications*, 95, 43–53. ISSN 0140-3664.

Li, J. P. (2000). Train station passenger flow study. In *Proceedings of the 32nd conference on Winter simulation*, pp. 1173–1173. Society for Computer Simulation International. ISBN 0780365828.

Lindfeldt, A. (2010). A study of the performance and utilization of the swedish railway network. In *CETRA 2010*. ISBN 978-953-6272-38-9.

Lindfeldt, A. (2015). Railway capacity analysis — Methods of simulation and evaluation of timetables, delay and infrastructure. PhD thesis, KTH Royal Institute of Technology.

Lindfeldt, O. (2007). Quality on single-track railway lines with passenger traffic: Anlytical model for evaluation of crossing stations and partial double-tracks. Licentiate thesis, KTH Royal Institute of Technology.

Lugmayr, A., Stockleben, B. & Scheib, C. (2016). A comprehensive survey on Big-Data research and its implications—what is really 'new' in Big Data?-it's cognitive Big Data! In *PACIS*, p. 248. AIS Electronic Library (AISeL). URL `http://aisel.aisnet.org/pacis2016/248`.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. & Byers, A. H. (2011). *Big data: The next frontier for innovation, competition, and productivity*. Report, McKinsey Global Institute.

McAfee, A. & Brynjolfsson, E. (2012). Big data: the management revolution. *Harvard business review*, (90), 60–6, 68, 128. ISSN 0017-8012.

Middelkoop, D. & Bouwman, M. (2002). Testing the stability of the rail network. *WIT Transactions on The Built Environment*, 61. ISSN 1853129135.

Mingers, J. (2001). Combining is research methods: towards a pluralist methodology. *Information systems research*, 12(3), 240–259. ISSN 1047-7047.

Mingers, J. & Brocklesby, J. (1997). Multimethodology: Towards a framework for mixing methodologies. *Omega*, 25(5), 489–509. ISSN 0305-0483. doi:10.1016/S0305-0483(97)00018-2.

Moisander, J. & Valtonen, A. (2006). *Qualitative marketing research: A cultural approach*. Sage. ISBN 1847878040.

Neville, C. (2007). *Introduction to research and research methods*. Bradford: Effective Learning Service. Bradford University School of Management.

Nielsen, B. F., Frølich, L., Nielsen, O. A. & Filges, D. (2014). Estimating passenger numbers in trains using existing weighing capabilities. *Transportmetrica A: Transport Science*, 10(6), 502–517. ISSN 2324-9935. doi:10.1080/23249935.2013.795199.

OECD (2000). *OECD glossary of evaluation and results based management terms*. Report, Organization for Economic Co-operation and Development. URL `http://www.oecd.org/dac/2754804.pdf`.

Olsson, N., Halse, A. H., Hegglund, P. M., Killi, M., Landmark, A. D., Seim, A. R., Sørensen, A. Ø., Økland, A. & Østli, V. (2015a). *Punktlighet i jernbanen—hvert sekund teller*. SINTEF akademisk forlag. ISBN 8253614845.

Olsson, N. O. E., Bull-Berg, H. & Walker, D. (2015b). Use of big data in project evaluations. *International Journal of Managing Projects in Business*, 8(3). ISSN 1753-8378. doi:10.1108/IJMPB-09-2014-0063.

Olsson, N. O. E. & Haugland, H. (2004). Influencing factors on train punctuality––results from some Norwegian studies. *Transport Policy*, 11(4), 387–397. ISSN 0967-070X. doi:10.1016/j.tranpol.2004.07.001.

Palmqvist, C.-W., Olsson, N. & Hiselius, L. (2017). Some influencing factors for passenger train punctuality in Sweden. ISSN 2153-2648.

Pannirselvam, G. P., Ferguson, L. A., Ash, R. C. & Siferd, S. P. (1999). Operations management research: an update for the 1990s. *Journal of Operations Management*, 18(1), 95–112. ISSN 0272-6963. doi:10.1016/S0272-6963(99)00009-1.

Parkinson, H. J. & Bamford, G. (2016). The potential for using big data analytics to predict safety risks by analyzing rail accidents. In *3rd International Conference on Railway Technology: Research, Development and Maintenance, Cagliari, Sardinia, Italy*, pp. 5–8.

Parthasarathi, P. & Levinson, D. (2010). Post-construction evaluation of traffic forecast accuracy. *Transport Policy*, 17(6), 428–443. ISSN 0967-070X.

Pellegrini, P., Marlière, G. & Rodriguez, J. (2016). A detailed analysis of the actual impact of real-time railway traffic management optimization. *Journal of Rail Transport Planning & Management*, 6(1), 13–31. ISSN 2210-9706. doi:10.1016/j.jrtpm.2016.01.002.

Phillips, E. & Pugh, D. (2010). *How to get a PhD: A handbook for students and their supervisors*. McGraw-Hill Education (UK). ISBN 0335242022.

Pieriegud, J. (2018). Digital transformation of railways. URL `https://shift2rail.org/wp-content/uploads/2018/04/DIGITAL_TRANSFORMATION_RAILWAYS_2018_web.pdf`.

Potthoff, G. (1970). *Verkehrsströmungslehre. 1. Die Zugfolge auf Strecken und in Bahnhöfen*. Transpress, Verlag für Verkehrswesen, Berlin.

Radtke, A. & Bendfeldt, J.-P. (2001). Handling of railway operation problems with railsys. In *Proceedings of the 5th World Congress on Rail Research*.

RIBA (2013). RIBA Plan of Work 2013. URL `https://www.ribaplanofwork.com`.

Rolstadås, A., Olsson, N., Johansen, A. & Langlo, J. A. (2014). *Praktisk prosjektledelse. Fra idé til gevinst*. Fagbokforlaget, Bergen. ISBN 978-82-450-1690-1.

Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of information science*, 33(2), 163–180. ISSN 0165-5515.

Russom, P. (2011). *Big data analytics*. Report, TDWI Research.

Samset, K. (2003). *Project evaluation: Making investments succeed*. Akademika Pub. ISBN 8251918405.

Saponara, S., Pilato, L. & Fanucci, L. (2016). Exploiting CCTV camera system for advanced passenger services on-board trains. In *Smart Cities Conference (ISC2), 2016 IEEE International*, pp. 1–6. IEEE. ISBN 1509018468.

Schoeni, R. F., Stafford, F., McGonagle, K. A. & Andreski, P. (2013). Response rates in national panel surveys. *The Annals of the American Academy of Political and Social Science*, 645(1), 60–87. ISSN 0002-7162. doi:10.1177/0002716212456363.

Schwanhäußer, W. (1974). Die bemessung der pufferzeiten im fahrplangefüge der eisenbahn. Thesis, Verkehrswissenschaftliches Institut der Rheinisch-Westfälischen Technischen Hochschule.

Scriven, M. (1991). *Evaluation thesaurus*. Sage th Publications, CA, 4th edition. ISBN 978-0803943643.

Seale, C. (1999). *The Quality of Qualitative Research.* Introducing Qualitative Methods. England, London: SAGE Publications Ltd United Kingdom, London: SAGE Publications Ltd, London. ISBN 9780761955986. doi:10.4135/9780857020093.

Seco, A. J. M. & Gonçalves, J. H. G. (2007). The quality of public transport: relative importance of different performance indicators and their potential to explain modal choice. *Urban Transport XIII: Urban Transport and the Environment in the 21st Century*, 96, 313–325.

Stenneth, L., Wolfson, O., Yu, P. S. & Xu, B. (2011). Transportation mode detection using mobile phones and GIS information. In *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 54–63. ACM. ISBN 1450310311.

Tanaka, M. (2015). Prospective study on the potential of Big Data. *Quarterly Report of RTRI*, 56(1), 5–9. doi:10.2219/rtriqr.56.5.

Templeton, D. (2017). Britain's digital railway seeks to harness technical innovation. *International Railway Journal*, 57(6), 46. URL `https://search.proquest.com/docview/1917279431?accountid=12870`.

The Economist (2010). A special report on managing information: Data, data everywhere. *The Economist.* URL `http://www.economist.com/node/15557443#print`.

The Institution of Engineering and Technology (2012). ORBIS - Network Rail's offering rail better information services. URL `https://www.theiet.org/sectors/information-communications/topics/connected-data/files/case-study-1.cfm?type=pdf`.

Tien, J. M. (2003). Toward a decision informatics paradigm: a real-time, information-based approach to decision making. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 33(1), 102–113. ISSN 1094-6977. doi:10.1109/TSMCC.2003.809345.

Tien, J. M. (2013). Big Data: Unleashing information. *Journal of Systems Science and Systems Engineering*, 22(2), 127–151. doi:10.1007/s11518-013-5219-4.

Tjora, A. (2013). *Kvalitative forskningsmetoder i praksis.* Gyldendal, Oslo, 2nd edition.

Toole, J. L., Colak, S., Sturt, B., Alexander, L. P., Evsukoff, A. & González, M. C. (2015). The path most traveled: Travel demand estimation using big data resources. *Transportation Research Part C: Emerging Technologies*, 58, 162–177. ISSN 0968-090X.

Tute, R. (2018). How advancing technology is allowing network rail to view rail like never before. URL `http://www.infrastructure-intelligence.com/article/jan-2018/how-advancing-technology-allowing-network-rail-view-rail-never`.

Veiseth, M. (2009). Forbedring av punktlighet i jernbanedrift: Utvikling av punktlighetsprosessen. PhD thesis, Norges teknisk-naturvitenskapelige universitet.

Vigren, A. (2017). Competition in public transport : Essays on competitive tendering and open-access competition in Sweden. Doctoral thesis, comprehensive summary, KTH Royal Institute of Technology. URL `http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-217267`.

Volden, G. H. & Samset, K. F. (2013). *Etterevaluering av statlige investeringsprosjekter. Konklusjoner, erfaringer og råd basert på pilotevaluering av fire prosjekter.* Report 30. URL `http://hdl.handle.net/11250/2456760`.

Vuchic, V. R. (2005). *Urban transit: operations, planning, and economics*, volume 21. John Wiley & Sons, Portland, United States. ISBN 0471632651.

Vuchic, V. R. & Newell, G. F. (1968). Rapid transit interstation spacings for minimum travel time. *Transportation Science*, 2(4), 303–339. ISSN 0041-1655.

Waller, M. A. & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics*, 34(2), 77–84. doi:10.1111/jbl.12010.

Warg, J. (2013). Economic evaluation of capacity with simulations. In *Proceedings of the 10th World Congress of Railway Research.*

Wu, X., Zhu, X., Wu, G. Q. & Ding, W. (2014). Data mining with big data. *IEEE Transactions on Knowledge and Data Engineering*, 26(1), 97–107. ISSN 1041-4347. doi:10.1109/TKDE.2013.109.

Xu, Y., Shaw, S.-L., Fang, Z. & Yin, L. (2016). Estimating potential demand of bicycle trips from mobile phone data—an anchor-point based approach. *ISPRS International Journal of Geo-Information*, 5(8), 131. doi:10.3390/ijgi5080131.

Yin, R. K. (2009). *Case study research : design and methods*, volume vol. 5 of *Applied social research methods series.* Sage, Thousand Oaks, Calif, 4th edition. ISBN 9781412960991.

Yuan, J. (2006). Stochastic modelling of train delays and delay propagation in stations. PhD thesis, Delft University of Technology.

Yuan, J. & Hansen, I. A. (2007). Optimizing capacity utilization of stations by estimating knock-on train delays. *Transportation Research Part B: Methodological*, 41(2), 202–217. ISSN 0191-2615.

Zhao, J., Rahbee, A. & Wilson, N. H. M. (2007). Estimating a rail passenger trip origin-destination matrix using automatic data collection systems. *Computer-Aided Civil and Infrastructure Engineering*, 22(5), 376–387. ISSN 1467-8667. doi:10.1111/j.1467-8667.2007.00494.x.

Zhao, Z., Shaw, S.-L., Xu, Y., Lu, F., Chen, J. & Yin, L. (2016). Understanding the bias of call detail records in human mobility research. *International Journal of Geographical Information Science*, 30(9), 1738–1762. ISSN 1365-8816. doi: 10.1080/13658816.2015.1137298.

Zidane, Y. J.-T. & Olsson, N. O. (2017). Defining project efficiency, effectiveness and efficacy. *International Journal of Managing Projects in Business*, 10(3), 621–641. ISSN 1753-8378.

# Appendix: Publications

# Publication I

Sørensen, A. Ø., Landmark, A. D., Olsson, N. O. E. & Seim, A. A. (2017). Method of analysis for delay propagation in a single-track network. *Journal of Rail Transport Planning & Management*, 7, 77-97. doi:10.1016/j.jrtpm.2017.04.001.

Journal of Rail Transport Planning & Management 7 (2017) 77–97

Contents lists available at ScienceDirect

Journal of Rail Transport Planning & Management

journal homepage: www.elsevier.com/locate/jrtpm

# Method of analysis for delay propagation in a single-track network

Anette Østbø Sørensen [a,*], Andreas Dypvik Landmark [b], Nils O.E. Olsson [a], Andreas Amdahl Seim [b]

[a] *Norwegian University of Science and Technology, Dept. of Mechanical and Industrial Engineering, N-7491 Trondheim, Norway*
[b] *SINTEF Technology and Society, Postboks 4760 Sluppen, N-7465 Trondheim, Norway*

A R T I C L E I N F O

*Article history:*
Received 12 January 2017
Accepted 24 April 2017
Available online 15 May 2017

*Keywords:*
Knock-on delay
Dynamic delay propagation
Single-track railway lines
Real-time punctuality data
Propagation factor

A B S T R A C T

Delay propagation is a key factor in punctuality of rail services. Propagation of delays reflects the degree of robustness of timetable design and the stability of train operations. This paper studies delay propagation for trains on single-track railways as the Norwegian rail network comprises primarily single-track operations. The analysis is based on real-time punctuality data. We present an algorithm and results from application of the algorithm. The algorithm is a recursive implementation of a set of conditions that (i) detects cases of knock-on delays, revealing dependencies between two trains on single-track railway, and (ii) finds the networks of dynamic delay propagations by tracing the propagation of knock-on delays from one interaction between trains to the next. Finally, we compare our results to delay cause registrations and theoretical calculations of the expected propagation factors. The propagation factors based on our model for a five-month period are typically around two, and over 70% of dynamic delay propagations consist of two trains. This indicates that each delay minute generates an additional delay of 1 min for other trains, which is higher than the present delay registrations. The presented method and analysis can be applied in punctuality improvement work, including timetable analysis.



## 1. Introduction

Infrastructure managers are working to understand and quantify the influence of individual infrastructure component failures and other delay causes on train traffic in near real-time. This capability will enable smarter and more cost-effective infrastructure maintenance. Successful development of the capability requires a detailed understanding of the occurrence of delays and their propagation through the system. Such insight may also support other efforts to improve punctuality. As the availability of train traffic data increases, there are large opportunities for analytical use of this data to gain an improved understanding of delays and their causes.

The occurrence of primary delays is directly observable and relatively well understood (Olsson and Haugland, 2004; Goverde and Hansen, 2001), although sometimes challenging to quantify accurately. Propagation of delays, from these primary delays, reflects the degree of robustness of timetable design and the stability of train operations (Yuan and Hansen,

* Corresponding author.
*E-mail address:* anette.o.sorensen@ntnu.no (A.Ø. Sørensen).

http://dx.doi.org/10.1016/j.jrtpm.2017.04.001
2210-9706/

2007). Delays caused by delay propagation are also called secondary or knock-on delays. They have been shown to increase exponentially with capacity utilization (Yuan and Hansen, 2007).

While there are models for the calculation of waiting time and secondary delays on double track lines, few such models for single track lines have been published (Handstanger, 2008). A train suffering a knock-on delay may cause further knock-on delays on other trains, known as dynamic delay propagation (Yuan and Hansen, 2007). While Flier et al. (2009) point to the need to develop algorithms to describe and quantify dynamic delay propagation based on real-time data, we are not aware of previous research to develop such algorithms. Thus, we conducted a study of dynamic delay propagation using real-time train traffic data. Our purpose was three-fold. First, we sought to develop an algorithm that uses real-time data to detect cases of knock-on delays, revealing dependencies between two trains on a single-track railway line. Second, we sought to automatically trace the propagation of knock-on delays from one case to the next, describing and quantifying the dynamic delay propagations. Third, we compared our results to manual delay cause registrations, and to theoretically calculated propagation factors. The algorithm and method is used in a case study to show examples of dynamic delay propagations, and to compare the events of knock-on delays over a period of five months.

## 2. Literature review

Samuel (1961) highlighted early on the need for statistics in railway management. Several advances have been made since then. Studies of delays, and delay propagation can be based on simulations, empirical data, theoretical approaches, or a combination of these. In our study, we are interested in studies of delay propagation in general, and on single-track lines in particular.

Simulations have been used extensively to study delay propagation (Lindfeldt, 2015; Warg, 2013; Radtke and Bendfeldt, 2001; Goverde, 2010) and it is common to apply simulation techniques when studying the relationship between key factors, such as train homogeneity or capacity utilization and the impact on random primary delays. Lindfeldt's (2015) work is a recent example of such an approach. Warg (2013) has shown that simulation is an adequate method to reveal delay data for future timetabling scenarios. Radtke and Bendfeldt (2001) simulate traffic with introduced primary delays, to be able to compare different scenarios with alternative timetables and infrastructures. Goverde (2010) presents an algorithm that models delay propagation in a large-scale railway network that includes single-track. The model detects the propagation of delays, based on a given railway network, a periodic timetable and a set of initial delays. The results are visualized in the network, and show the magnitude of the delays and how they propagate and die out over time.

Several of the aforementioned papers, use real-time data in their studies. Yuan and Hansen (2007) validate their model by means of train detection data. D'Ariano and Pranzo (2009) use it for the design of a traffic management system. Schwanhausser (1974), Cule et al. (2011), Goverde and Hansen (2001) and Flier et al. (2009) use it in post-evaluation.

Yuan and Hansen (2007) propose an analytical stochastic model which estimates the propagation of train delays at platform tracks and junctions. Their model also adopts recursive substitutions to estimate the dynamic delay propagation. Cule et al. (2011) use real-world data in their search for patterns of delayed trains in the Belgian railway network. The method finds all trains that are frequently delayed for a given time period. Goverde and Hansen (2001) use regression analysis to identify interrelationships between train arrivals and departures. They use real-time data from the Netherlands and time-stamps for events connected to entering and leaving a station. The events correspond to individual train movements along the route.

Flier et al. (2009) uses real-world data to detect two important types of delay dependencies on large-scale railways. These dependencies are knock-on delays caused by sharing of the same infrastructure, and delay caused by late connection between two trains. The algorithmic methods solve maximum optimization problems, and find correlations and patterns in systematic dependencies. Systematic dependencies refer to events (of delay dependencies) that occur on a regular basis. It was suggested that the results could then be statistically analysed, e.g., to assess the significance of the dependencies, and used for timetable planners to detect often-occurring secondary delays and adjust the timetable to reduce the dependencies. As a third step in an evaluation process, Flier et al. (2009) suggested extending their approach to global dependencies, i.e. to trace back the propagation of delays along the route of trains, possibly yielding a network of delay propagations. There are also other examples of articles presenting research on delay propagation on railway tracks (e.g. Conte, 2007; Hansen, 2001; D'Ariano et al., 2007).

Delay propagation can also be reduced through analysis of the operating railway, for instance, for decision support for traffic controllers, as proposed by D'Ariano and Pranzo (2009). The model predicts future evolution of railway traffic based on actual track occupation, characteristics of the signalling system, and train characteristics. The decision kernel of the system is a real-time optimization module, responsible for detecting and solving train conflicts while minimizing the propagation of delays. Others, such as Harris et al. (2013) have also demonstrated the use of structured methods to evaluate and improve performance in dispatching and on station punctuality. While primarily related to studies on multi-tracked infrastructure, Harris et al. (2013) describes a method and explore a case of minor delays incurred on station under the context of major timetable changes.

Schwanhausser (1974) proposed an analytical model to calculate the secondary delay on a railway line. The overall secondary delay is calculated by the summation of the secondary delays multiplied by the frequency of primary delay. Potthoff (1970) provides equations for delay propagation. The propagation factor on single tracks is a function of the initial delay, headway time, buffer time between trains and the number of sections on the single track. The equation is valid for primary delays that are longer than the buffer time on the line. The equation indicates that secondary delays increase with increasing numbers of sections on the track.

For our study, single-track methods are of particular interest, given that the Norwegian railway primarily consists of a star-shaped single-tracked network (approximately 7% double/multi-track). Of the literature presented here, only the method by Goverde (2010) is used on networks containing single-tracks. Studies of the Swedish railway are relevant as it also consists of a significant proportion of single-track lines. Lindfeldt (2010) presents a program that analyses the performance and utilization on the entire Swedish railway network, including single- and double-tracks with passenger and freight trains. The program calculates several descriptive parameters based on data related to infrastructure, train properties such as length and weight, timetables and delays. The parameters can be used to identify weaknesses in the network, and the paper investigates the correlation between the calculated parameters and the delay situation. Lindfeldt (2007) evaluates the single-track railway, with a focus on crossing times, where crossing time means the time loss that occurs in crossing situations on single-track lines, compared to double-track lines where no such time loss occurs. Lindfeldt (2007) presents an analytical model to calculate the crossing times, based on infrastructure, train properties, delays of two crossing trains and the timetable.

There are known issues with using train detection time-stamps as proxies for arrival and departure times. This has been discussed by Goverde (2005) and others. The core of this discussion is outside the scope of this paper, but we have used the departure time-stamps as-is without applying any deterministic correction terms or other adjustments. As such, we consider the measurements to be comparisons under equal terms when comparing trains passing the same departure signal. This will introduce a bias in comparison to scheduled times, i.e. for the departure time from the platform of a passenger service, which will consistently be some time prior to passing the departure signal. The registration system adjusts the times by applying an offset time at stations where the distances between the signals and the platform stops are particularly long. A train detected at the departure signal at zero seconds compared to the schedule would have had to close doors and depart the platform before the scheduled time.

## 3. Material and methods

### 3.1. Norwegian rail network and punctuality data

The Norwegian rail network is primarily single-track operations with crossings at stations and certain critical junctions. Norway has 3857 km of railway in regular traffic (64% electrified, by length) (Jernbaneverket, 2015), but only 256 km (7%, by length) of double/multi-tracked lines. NSB, the largest train operator transported 73 million passengers in 2015. The network is star-shaped around the capital, and the multi-tracked lines traverse the capital east-west/north-south. Long distance, regional and freight trains run mainly on single tracked lines. Many local commuter trains in Oslo run on both single and double tracks.

The majority of rail traffic (in number of departures and passenger volume) runs through Oslo Central Station, and capacity utilization generally decreases with distance to the central hub. However, to ensure smooth operation in the busiest section, punctual operation and strict management of delay in the lines feeding into the hub is paramount to the overall stability of the system. Hence the need for the current study on the propagation of delays between crossing traffic in single-track operations.

The Norwegian National Rail Administration (Jernbaneverket, now Bane NOR) records punctuality data (passing of home and departure signals, recorded to the second) daily, for each train, along with train information, stations and codes explaining causes for any delays in excess of the defined margins for the various classes of traffic. The data are recorded in a national database called TIOS ("Trafikkinformasjon og oppfølgingssystem", translated as "Traffic information and follow-up system"). These records include scheduled and actual arrival and departure times for each train at every station, train number and operating company, and class information (e.g., freight, running empty, or passenger train). The data describe the movements of the trains through the network, and are the basis of the method developed in this paper. The time registrations are generally based on automatic registrations and are considered as accurate measurements of when trains enter and exit different sections of the network, with a high resolution (measured and recorded to the second). There are some deviations in the measurement, as the timetable refers to the station platforms, while measurements are made at the home and departure signals. However, these deviations apply to all trains in the studied time period.

In addition to pure time data, Norwegian railway authorities register delay causes. When a train is more than 4 min delayed or additionally delayed, train dispatchers are required to register a delay cause. The registrations are not automatic, and thus subject to the personal judgement of the train dispatcher. These registrations are considered as less accurate than the pure time registrations. However, one objective of the work presented in this paper is to test approaches to improve the delay cause information, utilising the relatively high-quality time data that are available.

### 3.2. Algorithm for detecting delay propagation

Based on this data source we developed and implemented an algorithm for quantifying suspected delay propagation in the aforementioned traffic data in an online analysis tool, based on PostgreSQL 9.2 and the statistical software R 3.1. The implementation includes a combination of the algorithm developed, and an extraction and aggregation framework for accessing the track-circuit information, and some additional code for visualizing the results in Graphical timetable illustrations. The first prototype of the algorithm was implemented in Java 1.6.0.

The graphical timetable in Fig. 1 illustrates the punctuality data on a single-track railway line, where the vertical axis refers to the stations along a railway line and the horizontal axis refers to time. Black lines describe the scheduled times, while red lines are actual times. The crossing of two black or two red lines at a station illustrate crossings. When the red line lies to the left of the black line corresponding to the same train, it means the train is earlier than scheduled. Likewise, when the red line lies to the right of the black line the train is delayed, and this illustrates a primary delay when it first occurs and a secondary delay if the delay is caused by a crossing with a delayed train.

Section 3.2.1 presents the conditions for a crossing to be identified as a knock-on delay. The algorithm is a recursive implementation of these conditions, which makes it possible to identify dynamic delay propagations. When train *A* delays train *B* in accordance with the conditions presented in Section 3.2.1, the program makes a similar search on the second train to find whether train *B* delays other trains later on its route. If the program finds that train *B* delays a third train *C*, the search continues on train *C*, etc. The search on each train continues until the train in question is less than a limit of *y* seconds delayed. As soon as a train is delayed by less than *y* seconds, we conclude that the propagation of this delay ends. After the train has caught up with its delay, a potential new delay later on the route will be explained by a different source. The limit *y* should not be confused with the limit *x* that will be introduced in the next section.

### 3.2.1. Conditions

By studying time-distance graphs, relations and patterns may be detected between trains at crossing situations. Observations of when one train delays another train, gives the basis to some conditions which have to be satisfied for knock-on delays to occur. The following conditions describe the situation when a train *A* causes a knock-on delay on a second train *B* at a random station *j* on a single-track railway line.

1. The first condition to be satisfied is that train *A* is arriving at station *j* with a delay of more than *x* seconds, where *x* is a predefined limit chosen by the analyser. It is natural that *x* is larger than zero, indicating a delay, even though an early train may also cause a delay on other trains e.g., if the crossing is moved to another station.
2. The second condition is that train *A* is crossing train *B* at station *j*, meaning that train *A* and *B* have overlapping station times. Since train *A* is the delayer, it arrives between the actual arrival time and the actual departure time of train *B*, ensuring that train *B* is on the station at the time of the crossing and thus has to wait for train *A*. This condition is the foundation of the next two conditions, which describe how the crossing takes place.
   a. The delay-condition is that train *A* ***arrives*** in the time space between the scheduled and actual departure of train *B*, independent of the departure of train *A*.
   b. Or, that train *A* ***departs*** in the time space between scheduled and actual departure of train *B.*
3. The last condition to be satisfied is that train *B* is delayed with more than *x* seconds at departure from station *j.*

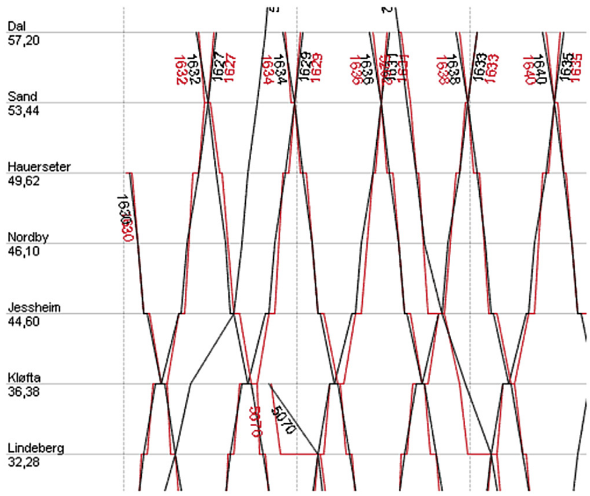


**Fig. 1.** Example of a graphical timetable on a single-track railway line. The axes refer to the stations along a railway line and time. The lines represent the punctuality data, where the black lines describe the scheduled times and the red lines are actual times. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

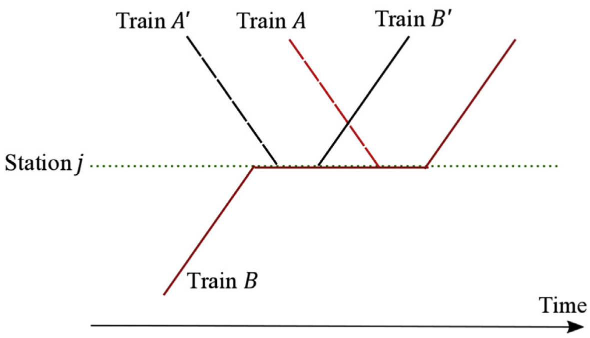


**Fig. 2.** Condition 2a.

Conditions 1–3 describe the requirements for a knock-on delay on a single-track railway line, and are depicted in Fig. 2 and Fig. 3. Given that $t$ is actual time and $s$ is scheduled time, and let $a$ denote arrival and $d$ denote departure, the conditions may then be summarized by the following inequalities:

1. $t_a(A) - s_a(A) > x$,
2. $t_a(B) < t_a(A) \& \, t_a(A) < t_d(B)$,
   a. $t_a(A) > s_d(B)$,
   b. $t_d(A) < t_d(B) \& \, t_d(A) > s_d(B)$,
3. $t_d(B) - s_d(B) > x$.

Condition 1 + 2a + 3 states that train $A$ is delaying train $B$ because of late arrival at the crossing station. If train $A$ arrives after the scheduled departure of train $B$, then train $B$ has to wait until train $A$ has arrived, but it is independent of the departure time of train $A$.

Condition 1 + 2b + 3 describes a case when train $A$ is arriving before train $B$ is scheduled to depart, but because of a late departure for train $A$, train $B$ is also delayed. As a result of train $A$ arriving before the scheduled departure of train $B$, this condition will generally only detect small knock-on delays. The reasons for such occurrences can be the times for interlocking, signalling and reaction, or if train $A$ is quite long (e.g. a freight train), and train $B$ cannot pass onto the single-track section vacated by train $A$ until all of train $A$ has cleared the overlap.

Fig. 4 illustrates a delay interaction on a single-track railway, beginning with train 5932, spreading to include train numbers 5719, 5718, 406, and 405.

### 3.3. Calculation of propagation factor

Potthoff (1970) provides equations to calculate propagation factors for dynamic delay propagations. The theory has been reproduced by Jernbaneverket (2015) in a Norwegian translation. The propagation factor on single tracks is a function of the initial delay, headway time, buffer time between trains and the number of sections on the single track. This section provides a brief description of the equation for single tracks.

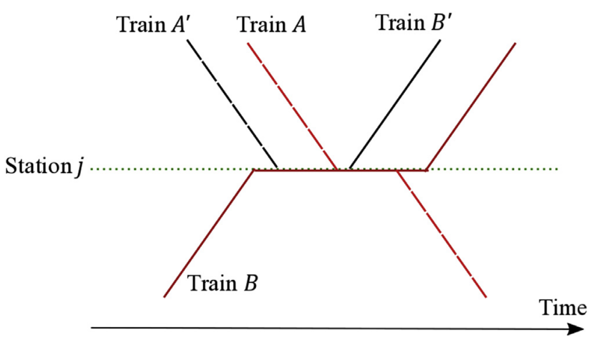


**Fig. 3.** Condition 2b.

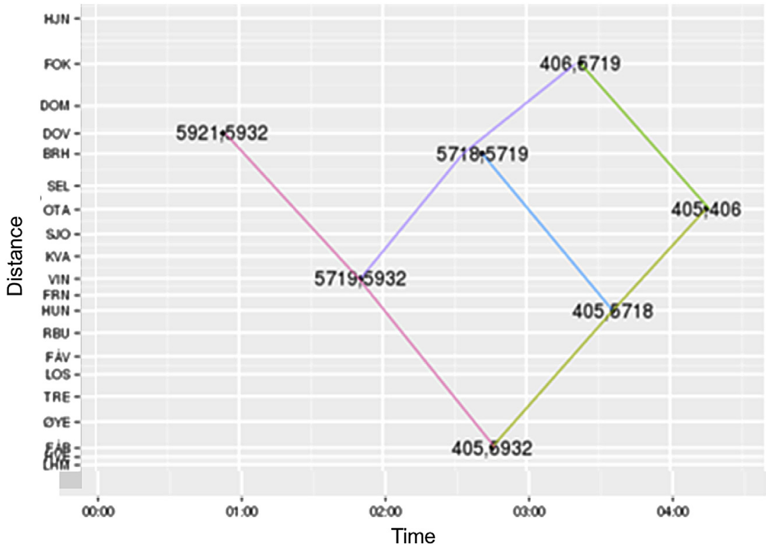


**Fig. 4.** Illustration of delay interaction in a graphical timetable. Numbers show train numbers that interact in a chain of delays. The x-axis shows time, the y-axis refers to stations along the line.

The line is assumed to consist of $a$ equal sections. Delays are denoted $p$ and indexed by the order number in the delay propagation, such that $p_1$ is the initial delay of the delayer train. The sum of delays in a dynamic delay propagation can approximately be expressed by:

$$\sum_{i=1}^{n} p_i = p_1 y(p_1), \tag{1}$$

with the propagation factor $y(p_1)$ given by

$$y(p_1) = \frac{\frac{p_1}{t_b} + a - \frac{(a-1)t_b}{p_1}}{2}, \tag{2}$$

where $t_b$ is the buffer time and $n$ is the number of trains in the dynamic delay propagation, including the delayer train.

The equations are derived with the assumption that the initial delay is larger than the buffer time, i.e. $p_1 > t_b$. We use Equation (2) to calculate expected values for propagation factors, for comparison with our empirical results. In addition, we calculate the propagation factor based on delay cause registrations. We can then present estimates of the propagation factor based on two datasets, as well as theoretical calculations.

### 3.4. Other notations – delay and additional delay

We can denote $T_i$ as the train number of the $i$th train in a delay propagation, such that $T_1$ corresponds to train $A$ and $T_2$ corresponds to train $B$, from the notation used in Section 3.2.1. The delay of train $T_i$ for $i \geq 2$ at the crossing station, as illustrated in Fig. 5, is given by $p_i = t_d(T_i) - s_d(T_i)$, (see notation in Section 3.2.1 and Section 3.3). The sum of delays of all affected trains in a dynamic delay propagation is then given by $\sum_{i=2}^{n} p_i$ from $i = 2$. Additional delay for a train, $T_i$ for $i \geq 2$, is given by $p_i^{add} = p_i^j - p_i^{j-1}$, where we let $j\,[= j(T_i)]$ be the crossing station for train $T_i$. The sum of additional delays is then given by $\sum_{i=2}^{n} p_i^{add}$.

## 4. Results

This section presents the results of a case study of a single-tracked railway line in Norway. The algorithm was applied on the punctuality data from the Dovre Line for a selected time period. The output of this study was further analysed, as well as compared to theory from the literature. The analyses and results consist of three parts. The first part presents the results of the

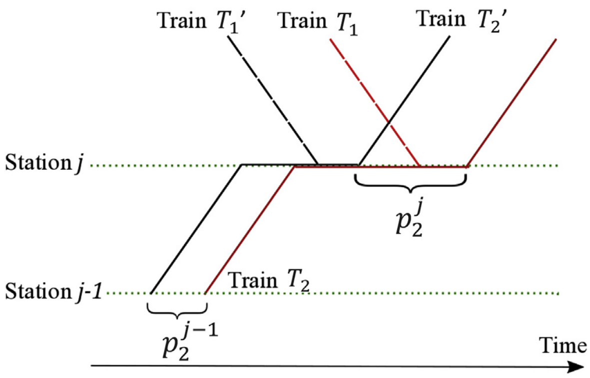


**Fig. 5.** Illustration of the delay $p_2^j$ and the additional delay $p_2^{add} = p_2^j - p_2^{j-1}$.


**Fig. 6.** The Dovre line highlighted with the main stations.

dynamic delay propagations that occurred on one given day. The second part presents the occurrences of dynamic delay propagations in a five-month period, as well as analyses of the results. The third part focuses on delay propagation factors, based on the algorithm, delay cause registrations and theoretical calculations. In analysis of both the one-day and the five-month case sample, a delay limit of 239 s was chosen for the variables $x$ and $y$, coinciding with the official threshold for delay (<4 min). This makes it possible to compare the results with the delay cause registration. However, note that the

**Table 1**
Stations on the Dovre line, and appurtenant station code.

| Station | Station code | Station | Station code |
|---|---|---|---|
| Eidsvoll | EVL | Kvam | KVA |
| Minnesund | MSU | Sjoa | SJO |
| Morskogen | MOR | Otta | OTA |
| Strandlykkja | SLY | Sel | SEL |
| Espa | EPA | Brennhaug | BRH |
| Tangen | TAN | Dovre | DOV |
| Steinsrud | STE | Dombås | DOM |
| Sørli | SRI | Fokstua | FOK |
| Stange | STG | Vålåsjø | VÅL |
| Ottestad | OTT | Hjerkinn | HJN |
| Hamar | HMR | Kongsvoll | KVL |
| Jessnes | JES | Drivstua | DRS |
| Brumunddal | BRD | Oppdal | OPD |
| Rudshøgda | RUD | Fagerhaug | FGH |
| Moelv | MLV | Ulsberg | UBG |
| Bergsvika | BVK | Berkåk | BÅK |
| Brøttum | BUM | Garli | GAL |
| Bergseng | BGG | Soknedal | SDL |
| Lillehammer | LHM | Støren | STØ |
| Hove | HVE | Hovin | HOI |
| Fåberg | FÅB | Lundamo | LMO |
| Øyer | ØYE | Ler | LER |
| Tretten | TRE | Søberg | SØB |
| Losna | LOS | Melhus | MEL |
| Fåvang | FÅV | Nypan | NYP |
| Ringebu | RBU | Heimdal | HMD |
| Hundorp | HUN | Selsbakk | SLB |
| Fron | FRN | Trondheim | TND |
| Vinstra | VIN | | |

threshold can be set at any value, and is not limited to the one we used here. The railway line used as a case study is introduced in the following section.

### 4.1. *Case description*

The Dovre Line (Dovrebanen) is a single tracked railway line running from Eidsvoll station to Trondheim station, as shown in Fig. 6. The line is the main connection between Eastern Norway and Trøndelag and further north for passenger and freight traffic. The line sees a mixture of regional and long-distance passenger traffic as well as freight traffic all combined on the same infrastructure. In 2014, the southern section of the Dovre Line, south of Lillehammer, had a capacity utilization of 86–100% during the peak-hour (Jernbaneverket, 2014). This is a high capacity utilization with little spare capacity available on the section (UIC , 2013). The same stretch had a total of 4809 h of delays registered in 2014, of which 3070 were assigned "primary" faults (infrastructure and rolling stock malfunction) and 1214 h were recorded as traffic related (not directly related to a technical incident).

The railway line has 28 stations and stops for passenger traffic. In the time frame of our case study, the length of the line was 485.6 km and had only 4 km of double tracks (Jernbaneverket, 2014). Key points on the route have centralized collection of statistics from track circuit devices (shown in Table 1).

In the fall of 2015, a 17-km new double tracked section was opened from Minnesund station to Kleverud station, which is a small part of the section between Eidsvoll and Tangen. We have therefore chosen the time frame from January to May 2015, after the schedule changes in December 2014 and before the new double tracks and the schedule changes in June 2015.

### 4.2. *Analysis of one day*

In this section, the algorithm is used to analyse the dynamic delay propagation that occurred on one given day. The chosen case is March 5th, 2015, which is an arbitrary and regular day in the sense of delay propagations (the punctuality on the Dovre Line in March 2015 was 92% (Jernbaneverket Trafikkdivisjonen, 2016)). The result is illustrated in two different ways; with a tree structure and by a graphical timetable illustration.

The tree structures in Fig. 7 show three examples of dynamic delay propagations that were identified on the Dovre Line on March 5th, 2015. The identified delayer, or initial train, is found in the root node. Each node contains the information of the train number, and the child nodes contain the station at which the crossing and the knock-on delay occurred. The colour indicates the direction of train travel, i.e. blue indicates that the train was travelling in the direction from Eidsvoll to Trondheim, while grey indicates the opposite direction, from Trondheim to Eidsvoll. The branches represent the crossings,

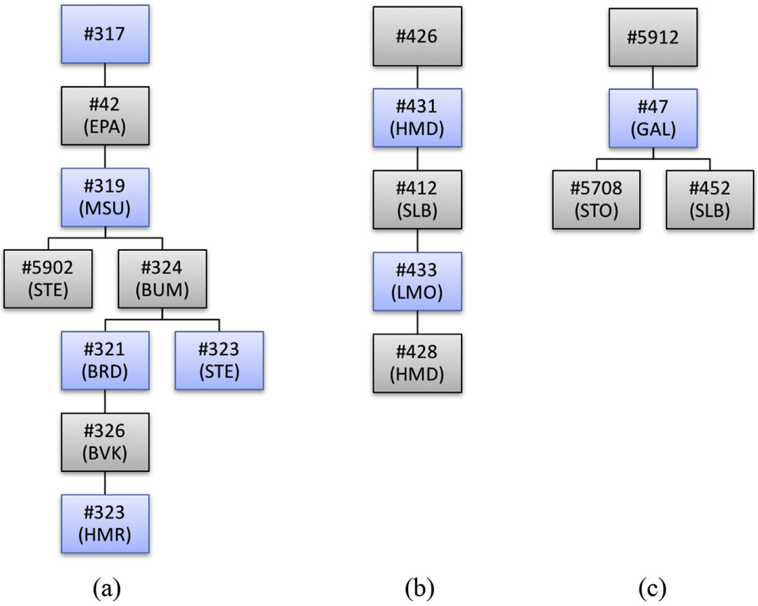


**Fig. 7.** Examples of dynamic delay propagations illustrated by tree structures, identified on the Dovre line on March 5th, 2015, showing both single- and multi-branched trees. Blue indicates train travel in the direction from Eidsvoll to Trondheim, and grey indicates the direction from Trondheim to Eidsvoll. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

where the parent node causes a knock-on delay on the child node, starting from the root node and spreading down through the tree structure.

The delay may propagate in several ways, as indicated by Fig. 7. The simplest structure is when the delayer train and all the parent nodes delay one train each, resulting in a single-branched tree, as in Fig. 7 (b). The more complex trees are multi-branched, where one or several of the parent nodes delay more than one train, as shown in Fig. 7 (a) and (c). The structure of the tree of a dynamic delay propagation will, among other factors, depend on the buffer times and density of the train traffic on the railway line.

The graphical timetable illustration in Fig. 8 shows all delay propagations that occurred on March 5th, 2015. In the graphical timetable illustration, the stations are placed geographically along the vertical axis and the time is represented on the horizontal axis. Each dot represents a crossing that resulted in a knock-on delay between the two train numbers connected to the dot. The lines follow trains that are involved in knock-on delays on more than one station. This type of presentation can be used to illustrate key stations and the time of day when delays propagate.

### 4.3. *Analysis of a five-month period*

In the previous section, we analysed the dynamic delay propagation that occurred on one given day. In this section, we analyse the occurrences of dynamic delay propagations in a five-month period, from January 1st until May 31st, 2015 (the punctuality on the Dove Line in this period was 90.5% (Jernbaneverket Trafikkdivisjonen, 2016)). The first step is to present how the dynamic delay propagation appears, and how it is distributed between the stations and train numbers. The propagation factor is then calculated and our findings are compared with the theory of Potthoff (1970).

#### 4.3.1. *Graphical timetable illustration*

The graphical timetable illustration in Fig. 9 shows dynamic delay propagations that occurred on the Dovre line in the time period January 1st until May 31st, 2015. The thickness of the lines and the size of the dots indicates relatively how many times an incidence occurred in the time period, where a thick line or a large dot means the delay propagation was identified on several days. Several of the delay propagations are so unique that they only happened once during the five-month period. But to draw attention to the fact that there is a pattern where the same delay propagations happen repeatedly, Fig. 9 was constructed with the condition that the delay propagations occurred a minimum of seven times in the time period. This type of presentation can be used to illustrate key trains, crossings and time of the day when delays propagate.

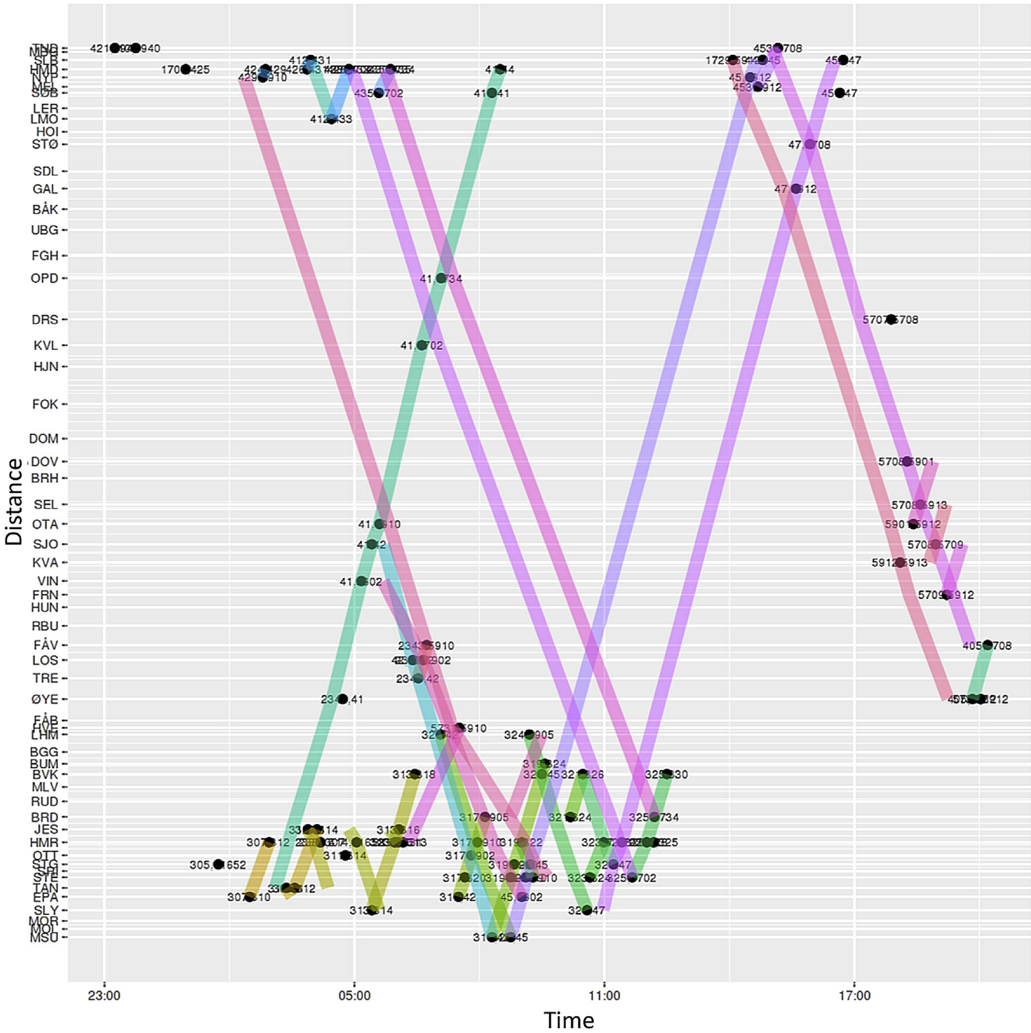

**Fig. 8.** Graphical timetable illustration of all delay propagations on the Dovre line on March 5th, 2015.

### 4.3.2. *Distribution of dynamic delay propagation for trains and stations*

In the examples presented in Fig. 7, there are four, five and nine trains involved in the dynamic delay propagation. It is informative to consider which type of delay is most common. In the five-month period there were 1616 incidences of dynamic delay propagation, including single knock-on delays. From Fig. 10, it is clear that most of the dynamic delay propagations consist of less than seven trains. The most commonly occurring events are single knock-on delays that only involve the delayer train and the one train delayed.

Fig. 11 illustrates how the knock-on delays are distributed between the train stations. The figure shows the number of times a crossing led to a knock-on delay. With reference to the tree structure in Fig. 7, the blue bars, which correspond to all knock-on delays, are the count of all the branches in all the dynamic delay propagations identified in our case sample. The red bars, which correspond to the initial knock-on delays, are the count of only the branches that lead from the root nodes.

The output showed that, in total, 131 unique train numbers had been involved in at least one knock-on delay during the case period. More than a quarter of them were rarely involved, that is, had been involved less than five times. However, one-fifth were involved in more than 60 knock-on delays, either being affected by a knock-on delay or as the delayer train. For each train we can look at the number of events where the train is the delayer train, and number of events where the train received a secondary delay from a knock-on delay. Sorting both of these numbers from maximum to minimum values, we examined the top 20 trains in each list (21 trains are in the list of affected trains because the train in ranks 20 and 21 had the same value). In Table 2 these trains are divided into their train category; local trains, intercity trains, long distance trains and freight trains.

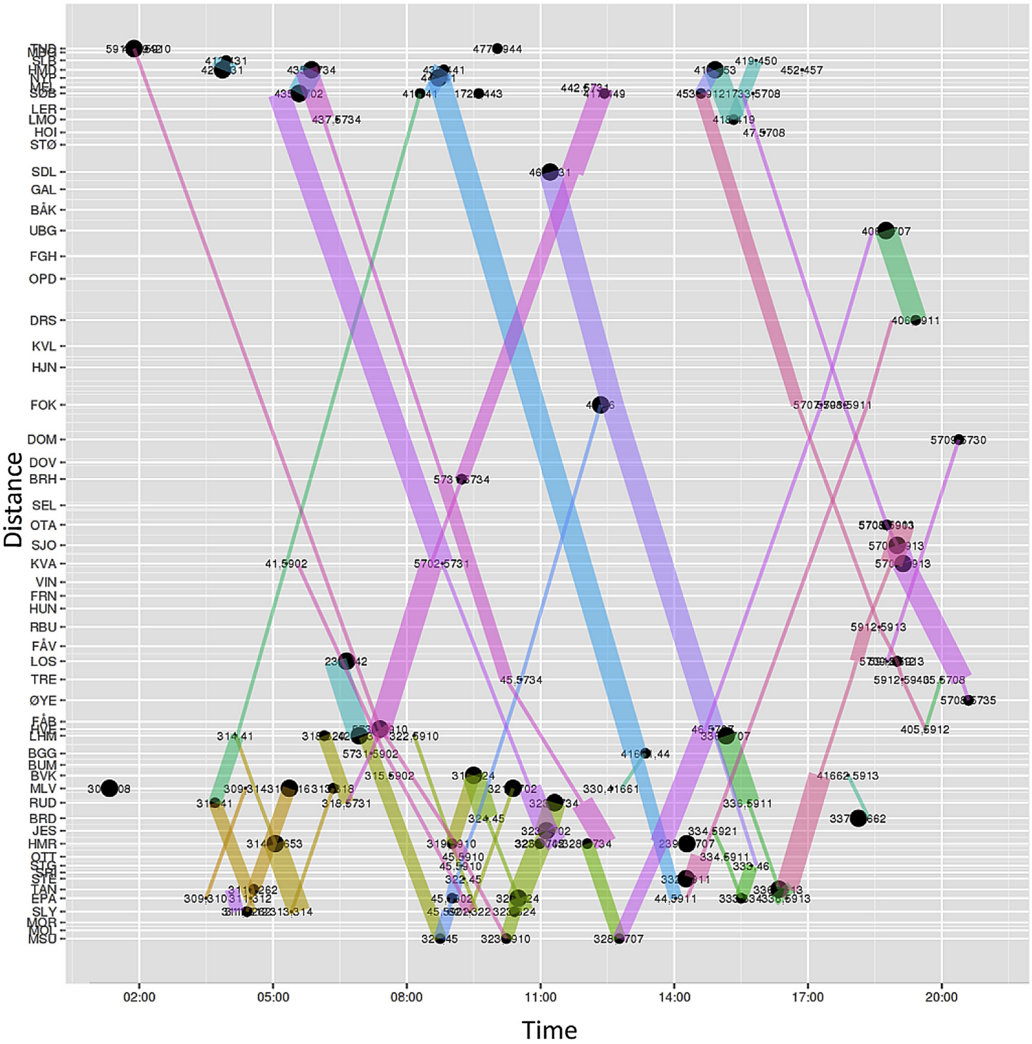


**Fig. 9.** Delay propagations on the Dovre Line in the time period January 1st to May 31st, 2015, with the condition that the delay propagations occurred a minimum of seven times. The size of the dots and the thickness of the lines indicate relatively how many times the incidence occurred.

We found that ten of the trains were represented among the top 20 in both lists, as seen in the bottom row in Table 2. It is clear that intercity trains and freight trains are well represented.

#### *4.3.3. Identifying the ratio between delayer train and affected train*

It is informative to investigate the relation between the number of events where a train is the initial train, and the number of events where a train received a secondary delay. For each train number, we calculated the ratio #*delayer train*/#*affected train*. Fig. 12 shows the twenty trains that had the largest ratio, greater than one. Trains with a ratio larger than one are initiating more delays than they are receiving, i.e. they are the delayer train more often than the affected train. In Fig. 12, this is shown by the red bar being higher than the blue bar. However, because small numbers also can give a large ratio, we excluded trains that had #*delayer train* less than six. Three of the trains in Fig. 12 are actually the top three trains that were most often the initial train in delay propagations (#45, #42 and #44).

Fig. 13 shows all trains that had a ratio less than 0.2, and equivalent to Fig. 12, excludes trains with #*affected train* less than six. These are trains that are the affected train more often than the delayer train. This ratio is a means to find trains that should be further investigated, for instance train 5730 in Fig. 13 appears to have been affected virtually every day.

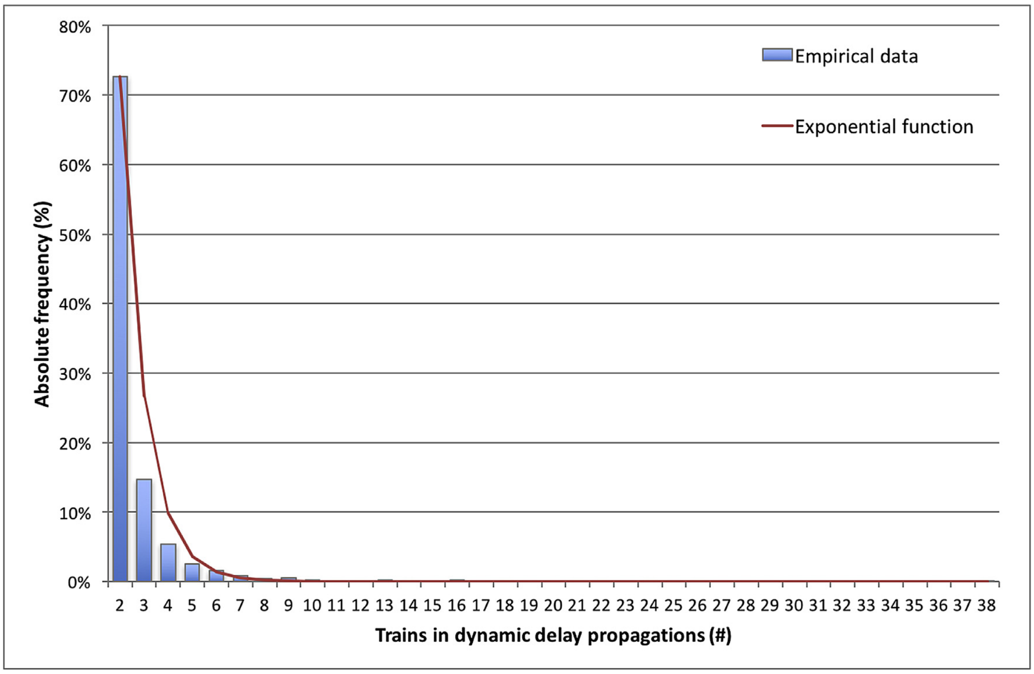


**Fig. 10.** Number of trains involved in dynamic delay propagations, identified on the Dovre Line between January 1st and May 31st, 2015. The x-axis represents the number of trains in each dynamic delay propagation, including the initial train. The y-axis represents the absolute frequency as the percent of the total number of events. The graph is compared to the exponential function $1174\exp(2-x)$ (converted to percent).

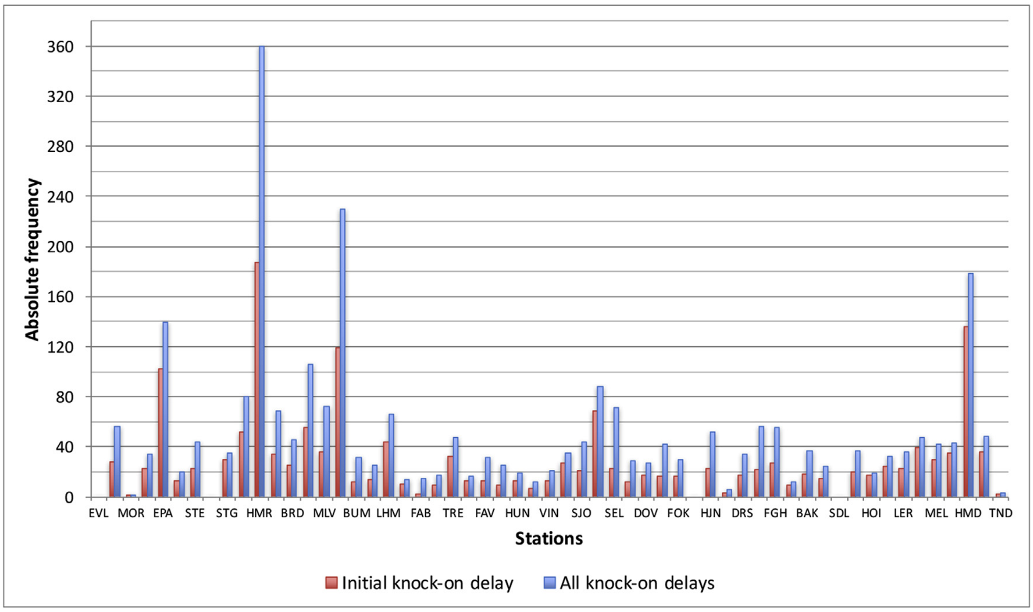


**Fig. 11.** Distribution of knock-on delays between the stations. Referring to the tree structure, the initial knock-on delays (red) are the count of branches only leading from the root node. All knock-on delays (blue) are the count of all branches. The x-axis shows stations along the line from south to north. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

**Table 2**
The trains with highest frequency as the delayer train and as the affected train, distributed over train categories.

| Train categories | Delayer train | Affected train | Delayer train and Affected train | Total |
|---|---|---|---|---|
| Local trains | 1 | 0 | 0 | 1 |
| Intercity trains | 4 | 5 | 3 | 12 |
| Long distance trains | 3 | 1 | 3 | 7 |
| Freight trains | 2 | 5 | 4 | 11 |
| Total | 10 | 11 | 10 | 31 |

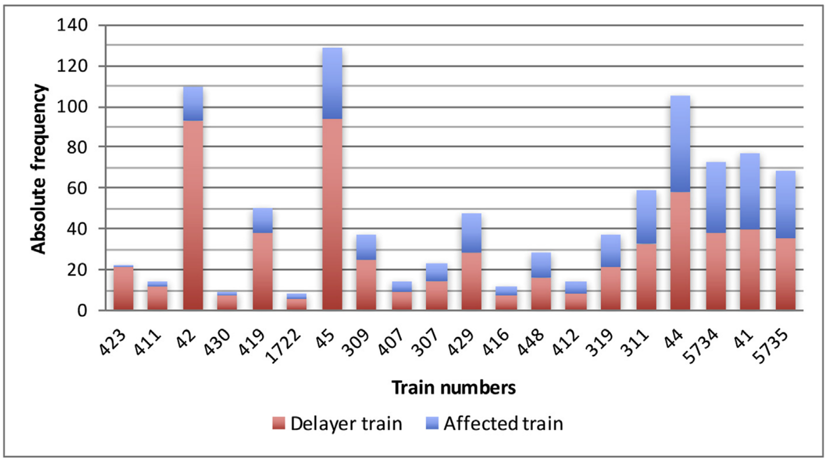


**Fig. 12.** Number of events where a train is the delayer train (red) in dynamic delay propagations, and number of events where a train received a secondary delay (blue) from a knock-on delay (meaning it was the affected train in dynamic delay propagations), where #delayer train/#affected train is larger than one, provided that #delayer train is larger than five. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

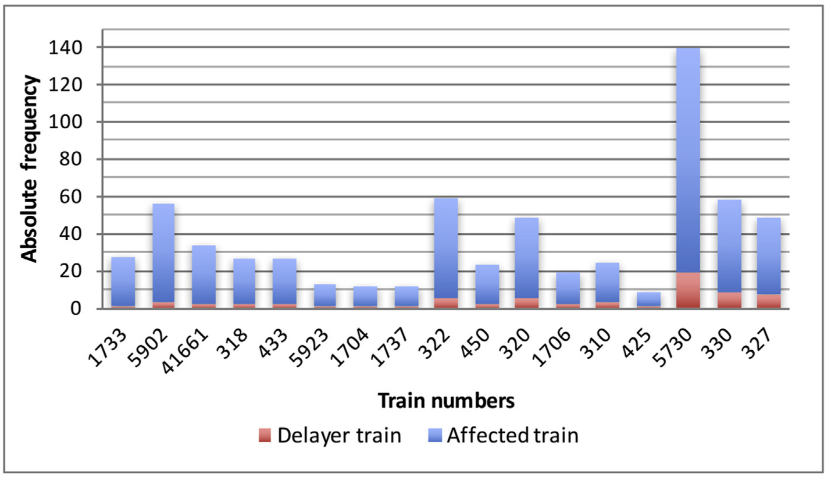


**Fig. 13.** Number of events where a train is the delayer train (red) in dynamic delay propagations, and number of events where a train received a secondary delay (blue) from a knock-on delay (meaning it was the affected train in dynamic delay propagations), where #delayer train/#affected train is less than 0.2, provided that #affected train is larger than five. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

### 4.4. Propagation factors

The propagation factor is defined as the sum of delays for all trains involved in a dynamic delay propagation divided by the initial delay, as used by Potthoff (1970) and presented in Section 3.3. Our studies are based on a relatively long (about 500 km),

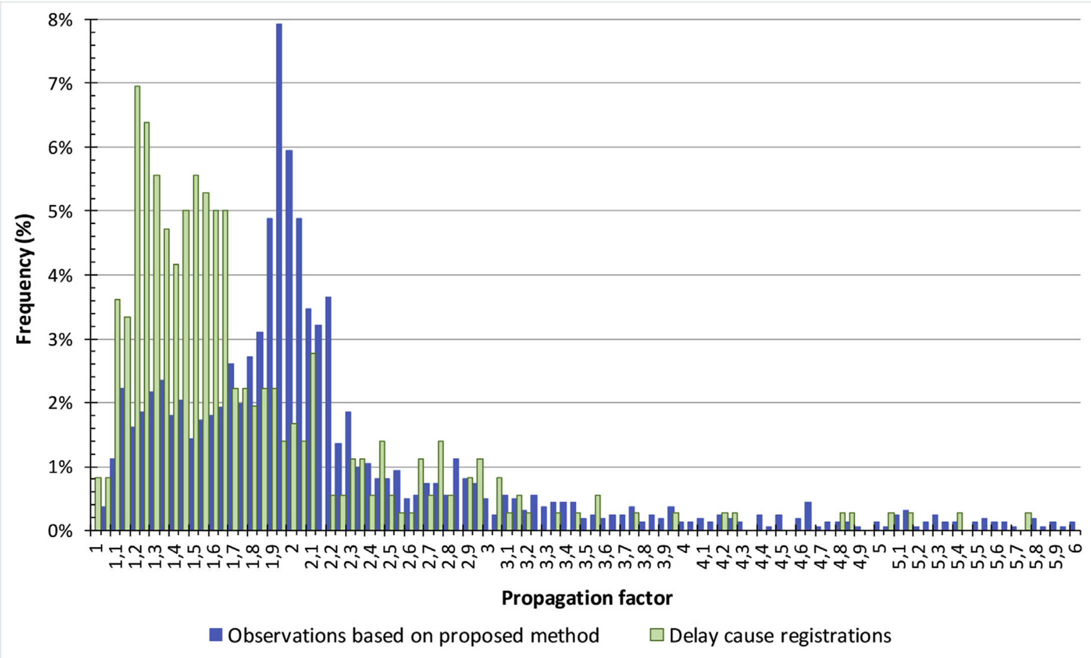


**Fig. 14.** Comparison of the distribution of propagation factors based on the five-month case and the delay cause registrations. The observed propagation factors based on the method proposed in this paper are represented by the blue bars, while the propagation factors based on the delay cause registrations are represented by the green bars. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and heterogeneous line. To replicate Potthoff's estimations would be highly dependent on whether we treat the line as one long line with many sections, or divide it into sub-lines with fewer sections. There are no stations along the line with extra long station stops.

#### *4.4.1. Distribution of the propagation factors based on the proposed algorithm*

We calculated the propagation factor $y(p_1)$ from Equation (1) for the five-month case sample. The mean propagation factor in the sample is 3.3, while the median is 2. The distribution of propagation factors are illustrated by the blue bars in Fig. 14, and are typically between 1.1 and 2.2 with a peak around 2. The highest propagation factor registered in the sample is 93.9. However, 92% of the propagation factors lie in the interval 1 to 6.

#### *4.4.2. Distribution of the propagation factors based on delay cause registrations*

The propagation factors presented in Fig. 14 by the green bars are based on delay cause registration for the same railway line. Delay registrations use a code system consisting of nine categories. Delays coded as dispatching and station stops are mainly considered as secondary delays, while delays related to infrastructure and rolling stock are considered as primary delays. However, categories can include both primary and secondary delays. For instance, a primary delay caused by more passengers than anticipated on a station would also be registered under the code station stops. Though this will affect the results to a small degree, we calculated the propagation factors based on the two mentioned codes of secondary delay. We find that the propagation factors based on delay registrations are mainly of lower values than those based on the proposed method. The distribution of propagation factors is illustrated by the green bars in Fig. 14, which shows that the propagation factors are typically between 1 and 2, with relatively few being over 2.

#### *4.4.3. Propagation factors based on the algorithm as a function of initial delay*

In Fig. 15, the propagation factors for the case sample are plotted against the initial delay. The delay propagations that consist of $n = 2, \ldots, 6$ and $n > 6$, where $n$ is the number of trains involved, are marked in the figure. Fig. 15 (a) shows that there is a high variability for smaller initial delays. The weight of propagation factors around 2 consists of delay propagations involving two trains, ($n = 2$), and is in accordance with the results in Figs 10 and 14. The mean initial delay in the sample is 16 min, while the median is 7 min.

To understand the behaviour of the propagation factor we take a closer look at Equation (1). If all the affected trains in a dynamic delay propagation receive the same delay as the initial delay, meaning $p_i = p_1$ for $i = 2, \ldots, n$, the propagation factor

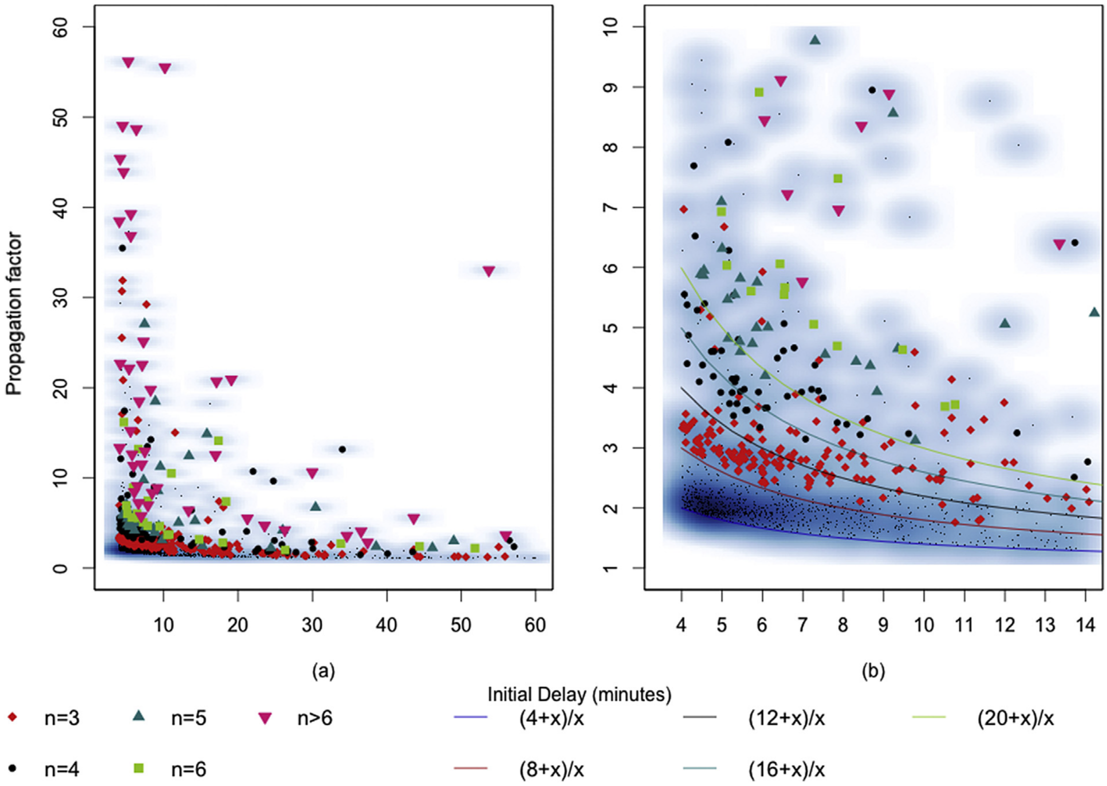


**Fig. 15.** Propagation factor plotted against initial delay for the five-month case. The points are separately emphasized for $n = 2$ (the point ".") and for $n = 3, 4, 5, 6$ and $n > 6$ as given in the figure, hence covering all possible dynamic delay propagations. The curves are given by Equation (3), for $n = 2, 3, 4, 5$ and 6.

will be $y(p_1) = n$. For small initial delays close to 4 min, we would therefore expect the propagation factor to be scattered right above $y(p_1) = 2$ for dynamic delay propagations consisting of $n = 2$ trains, and right above $y(p_1) = 3$ for $n = 3$, etc. Indeed, this can be observed in Fig. 15 for $n = 2$, 3 and 4, however, for larger numbers of involved trains the propagation factors have higher values.

When the initial delay, $p_1$, increases we see in Fig. 15 (b) that the propagation factor $y(p_1)$ decreases towards a value close to 1. Rearranging Equation (1), the propagation factor can be written as

$$y(p_1) = \frac{\sum_{i=1}^{n} p_i}{p_1} = 1 + \frac{\sum_{i=2}^{n} p_i}{p_1} = \frac{xn + \sum_{i=1}^{n} \Delta p_i}{x + \Delta p_1},$$

using $p_i = x + \Delta p_i$, where $x$ is the delay limit from the algorithm in Section 3.2.1.

For small secondary delays, we can write the propagation factor as

$$\lim_{\sum_{i=2}^{n} \Delta p_i \to 0} y(p_1) = \frac{x(n-1) + p_1}{p_1} = 1 + \frac{x(n-1)}{p_1}. \tag{3}$$

We see that $y(p_1)$ will never be 1, because $\sum_{i=2}^{n} p_i$ will always be larger than $x(n-1)$, or in this case $239(n-1)$ seconds (see Section 4), the delay limit of $x$ times the number of affected trains in the dynamic delay propagation, $n-1$. The lines in Equation (3) are drawn in Fig. 15 (b) for $n = 2, \ldots, 6$. Dots which are higher compared to the corresponding lines, indicate increased delays in the affected trains. For each increase in $n$, the dots are clustered higher above the corresponding line. This is because $\sum_{i=2}^{n} \Delta p_i$ gains one more element that is to some degree larger than zero.

To summarize the results, we see that when

- $y(p_1) \approx \frac{4(n-1)+p_1}{p_1}$, for $n \geq 2$: the affected train(s) is (are) minimally delayed above 4 min, with $n = 2$, when

- $\frac{4+p_1}{p_1} < y(p_1) < 2$: the affected train is less delayed than the delayer train,
- $y(p_1) > 2$: the affected train is more delayed than the delayer train,

with $n > 2$, when

- $\frac{4(n-1)+p_1}{p_1} < y(p_1) < n$: the affected trains are on average less delayed than the delayer train,
- $y(p_1) > n$: at least one of the affected trains is more delayed than the delayer train.

### 4.5. Propagation factor based on Potthoff's equations

In order to compare these results with the theoretical propagation factors, we calculated expected propagation factors from Equation (2), based on information about the railway line and the schedule, given in Table 3. The average line utilization is defined as the minimum technical headway divided by the average timetable headway. The calculations are performed on smaller sub-lines of the line, as the studied line is relatively long, and there are different traffic and infrastructure profiles in different parts of the line. Based on the assumptions in Potthoff's equation, this indicates that the line, from an operational perspective, functions as a set of separate lines.

Table 3 gives the calculated propagation factor as a function of initial delay, showing that the propagation factor increases with higher initial delays. The propagation factor for initial delays just above the buffer time is between 1.1 and 2.2. This is well in line with the values in Fig. 14.

Interestingly, the propagation factor in Fig. 15 decreases with increasing initial delay, as opposed to the expected calculated values in Table 3. Clearly, the propagation factor in Equation (2) is bound to increase with $p_1$, because both of the terms $\frac{p_1}{t_b}$ and $-\frac{(a-1)t_b}{p_1}$ ensure this. Nonetheless, the propagation factor in Equation (3), that we derived from Equation (1), is indeed decreasing with $p_1$. The difference lies in the value of $n$. In Equation (3), we investigate the behaviour of the propagation factor for each different value of $n$ separately. However, it is logical to assume that $n$ is dependent on $p_1$. In fact, in the derivation of Potthoff's equation, an approximation is made that the delay of the last affected train is negligible, which leads to the relation $n - 1 = \frac{p_1}{t_b}$ between the number of affected trains and the initial delay (Jernbaneverket, 2015). Therefore, larger initial delays should theoretically delay more trains. Furthermore, from Equation (3), and as illustrated by the lines in Fig. 15, it is clear that the propagation factor $y(p_1)$ has larger values with e.g., $n = 3$ than with $n = 2$.

Fig. 15 illustrates that large initial delays can delay one train, as well as several trains, while small initial delays can delay several trains, as well as just one train. Nevertheless, the propagation factors form rather dedicated clusters for different numbers $n$ of trains in the delay propagation. This is evident for $n = 2$ and $n = 3$, but less distinct for $n > 3$. In addition, the centres of the clusters increase in both initial delay and propagation factor with increasing $n$, as expected from the recent discussion. We calculated the median value of the propagation factor and the initial delay for each of $n = 2, \ldots, 9$. The median was chosen instead of the mean, because both initial delay and propagation factor have a few outliers that are much higher than the remaining data, which can skew the mean towards high values. The median values of propagation factor and initial delay are plotted in Fig. 16 for $n = 2, \ldots, 9$, and compared with the calculated expected propagation factor of four of the sub-lines in Table 3. Fig. 16 shows that the median values partially follow the lines of the calculated propagation factors.

### 4.6. Sum of delay and additional delay

In Fig. 17, the sum of delays, $\sum_{i=2}^{n} p_i$, of all affected trains in a dynamic delay propagation is plotted against initial delay. As in Fig. 15, the delay propagations that consist of $n = 2, \ldots, 6$ and $n > 6$ (number of involved trains) are marked in figure. As was shown in Section 4.4.3 (see Equation (3)), $\sum_{i=2}^{n} p_i$ will always be larger than $x(n-1)$ minutes. And as expected, in Fig. 17 (a) and (b) we see that the sum of delays lies above the horizontal lines $y \approx 4, 8, 12, \ldots$ for $n = 2, 3, 4, \ldots$, respectively. If we assume that all trains in a dynamic delay propagation receive the same delay as the initial delay, i.e. $p_i = p_1$ for $i = 2, \ldots, n$, then it would mean that the sum of delays follows the line $\sum_{i=2}^{n} p_i = (n-1)p_1$. These lines are drawn in Fig. 17 (b) for $n = 2, \ldots, 6$. For $n = 2$ it is clear that when

- $p_2 \approx p_1$, the affected train is delayed with about the same delay as the delayer train,
- $p_2 < p_1$, the affected train is less delayed than the delayer train,
- $p_2 > p_1$, the affected train is more delayed than the delayer train.

For delay propagations with $n > 2$, we see that when

- $\sum_{i=2}^{n} p_i \approx (n-1)p_1$: the affected trains are on average delayed by the same amount as the delayer train,
- $\sum_{i=2}^{n} p_i < (n-1)p_1$: the affected trains are on average less delayed than the delayer train,
- $\sum_{i=2}^{n} p_i > (n-1)p_1$: at least one of the affected trains is more delayed than the delayer train.

**Table 3**
Calculation of propagation factor based on Potthoff (1970). (*Initial delay is 25 min and **35 min, as the equation is not valid for lower initial delays on this particular sub-line).

| Part of line | Prop. factor (5 min initial delay) | Prop. factor (15 min initial delay) | Sections | Min. technical headway (min) | Av. time-table headway (min) | Buffer time (min) | Av. line utilization (%) |
|---|---|---|---|---|---|---|---|
| Trondheim-Søberg | 1.1 | 3.8 | 6 | 7.5 | 12.26 | 4.76 | 61 |
| Støren-Dombås | 1.4* | 3.1** | 12 | 13.75 | 37 | 23.25 | 37 |
| Lillehammer-Hamar | 1.2 | 4.5 | 8 | 11.6 | 16.32 | 4.72 | 71 |
| Hamar-Eidsvoll | 2.2 | 5.8 | 10 | 11.4 | 15.22 | 3.82 | 75 |
| Lillehammer-Eidsvoll | 1.5 | 7.9 | 18 | 11.6 | 16.32 | 4.72 | 71 |

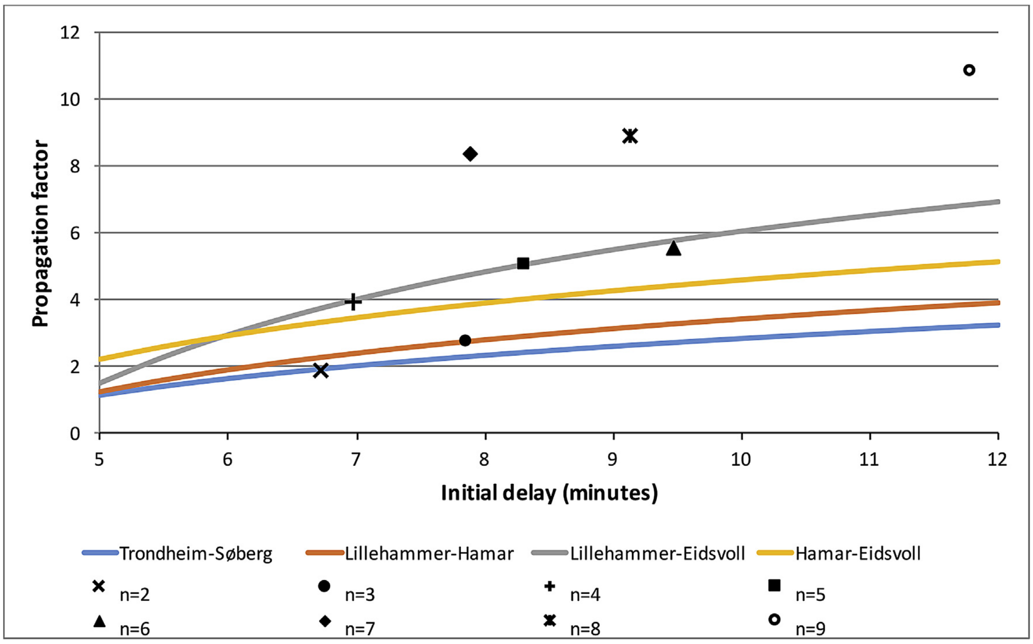


**Fig. 16.** Median propagation factor in the case sample compared to the expected calculated propagation factor from Equation (2). The lines represent the expected propagation factor of four of the sub-lines in Table 3, calculated based on Potthoff (1970). The marked points are the calculated median value of propagation factor and median initial delay for each of n = 2, …, 9 in the case sample.

Dots that are below the corresponding line indicate that the affected trains are on average delayed with less than the initial delay. When the dots are more than a few minutes lower than the corresponding line we may assume that the crossing was moved or in other ways optimized to minimize the influence from the delayer train. When the dots are above the corresponding line, the affected train(s) is (are) delayed with more than the initial delay. If this difference is unexpectedly large, we can assume that at least one of the affected trains in the dynamic delay propagation had a large delay prior to the crossing. We observe that when the initial delay increases past a certain value, fewer dots follow the drawn lines, and the majority of the dots scatter horizontally below the corresponding drawn lines.

To look closer at the occurrences where affected trains had a large delay prior to the crossing we plotted the sum of additional delays, $\sum_{i=2}^{n} p_i^{add}$, against the sum of delays of affected trains, $\sum_{i=2}^{n} p_i$, given in Fig. 18. The algorithm identified in total 1616 unique events of dynamic delay propagations in the case period. In 135 of the identified dynamic delay propagations, at least one of the trains did not have enough information to calculate the delay on the previous station, which was therefore excluded in the plot in Fig. 18.

As noted previously, the sum of delays of affected trains lies above the lines of value 4, 8, 12, … for $n = 2, 3, 4, \ldots$, respectively, which in Fig. 18 (a) are shown by the vertical lines. The dashed line in Fig. 18 is $\sum_{i=2}^{n} p_i^{add} = \sum_{i=2}^{n} p_i$, for simplicity denoted $g(z) = z$. We see that when

- $g(z) \approx z$ (meaning $\sum_{i=2}^{n} p_i^{add} \approx \sum_{i=2}^{n} p_i$): the affected train(s) was (were), on average, on time at the previous station and all delay was caused by the delay propagation,

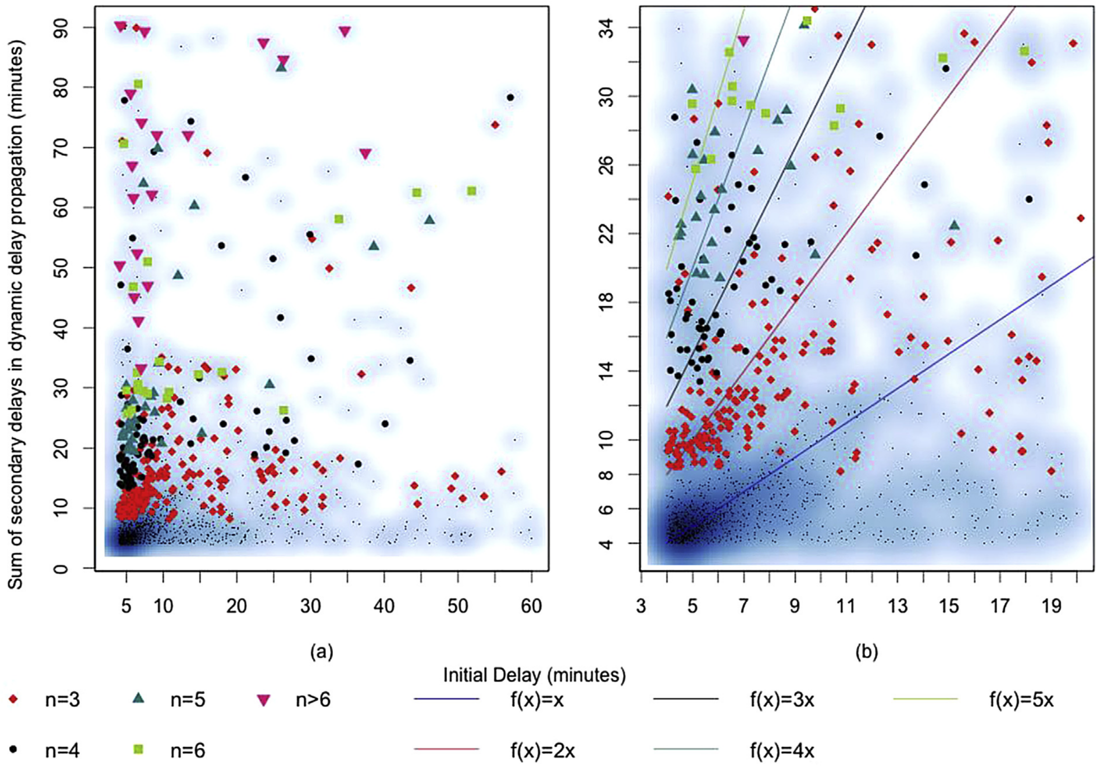


**Fig. 17.** The sum of delays of all affected trains in the dynamic delay propagations plotted against initial delay.

- $g(z) < z$ (meaning $\sum_{i=2}^{n} p_i^{add} < \sum_{i=2}^{n} p_i$): the affected train(s) was (were), on average, to some extent delayed prior to the crossing,
- $g(z) > z$ (meaning $\sum_{i=2}^{n} p_i^{add} > \sum_{i=2}^{n} p_i$): the affected train(s) was (were), on average, before schedule prior to the crossing.

Fig. 18 shows how delayed the affected trains were prior to the crossing. To emphasize the occurrences with large delays, we start by assuming that the affected train(s), on average, was (were) at least 239 s delayed on the previous station, and we rewrite the sum of additional delays as

$$\sum_{i=2}^{n} p_i^{add} = \sum_{i=2}^{n} p_i^j - \sum_{i=2}^{n} p_i^{j-1} = \sum_{i=2}^{n} p_i^j - (n-1)x - \sum_{i=2}^{n} \Delta p_i^{j-1},$$

where $x$ is the delay limit.

If the affected trains are on average delayed by 239 s at the station prior to the crossing, then $\sum_{i=2}^{n} \Delta p_i^{j-1} = 0$ and the sum of additional delays will be a value on the line $f(z) = z - (n-1)x$. The lines are plotted in Fig. 18 for $n = 2, \ldots, 6$. Consequently, we see that when

- $z > \sum_{i=2}^{n} p_i^{add} > z - (n-1)x$: the affected train(s) was (were), on average, not delayed at the station prior to the crossing, meaning, on average, less than 239 s,
- $z - (n-1)x > \sum_{i=2}^{n} p_i^{add}$: at least one of the affected trains was delayed prior to the crossing.

Hence, Fig. 18 (b) shows that the largest portion of delay propagations with $n = 2$ lies between the lines $f(z) = z$ and $f(z) = z - 3.98$. Similarly, the largest portion of delay propagations with $n = 3$ lies between the lines $f(z) = z$ and $f(z) = z - 7.97$, and so on. Fig. 18 (a) shows some clear outliers where at least one of the affected train(s) in the dynamic delay propagations had large delays prior to the crossing. For this reason, we can say that the points that lie below the corresponding line had at least one train that was delayed by something else prior to the crossing.

Unlike Fig. 17, the y-axis in Fig. 18 shows the difference in delays for affected trains before and after the interaction, i.e. the additional delay (see Fig. 5), while Fig. 17 shows the sum of delays of the affected trains at the crossing station. Negative y-

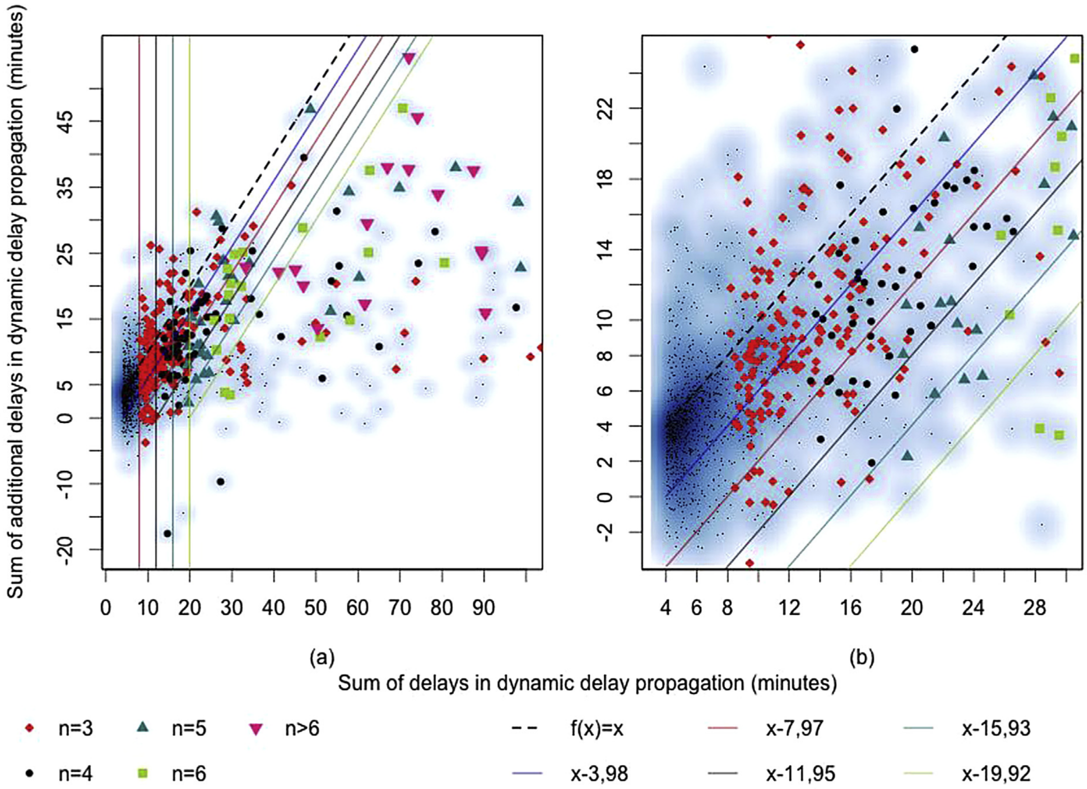


**Fig. 18.** Additional delays for affected trains involved in delay propagations, plotted against sum of delays. The y-axis shows the difference in delay between the last station before the interaction with another train, and the delay when leaving the station where it was an interaction.

values in Fig. 18 indicate that trains became less delayed, even though they did have an interaction with a delayed train. The mean value of this sum is 10.4 min, and the median is 5.8 min.

## 5. Concluding discussion

This paper has presented an algorithm for quantifying delay propagation. We applied the algorithm on railway traffic data and showed that the implementation can be successfully used for analysis. The algorithm and software allows for a nuanced approach to the analysis of "everyday" delays. Traditional methods of descriptive statistics and tallying uncover little about the correlation between trains and presumed knock-on effects. This tool allows for a first attempt at indicating the direction of knock-on effects in single-track traffic.

The algorithm and software developed in this study give knowledge of how delays and delay propagations behave on single-track railways. The program is not a simulation tool, and cannot directly be used to predict outcomes of future changes to schedules. Instead, it analyses real-world data gathered from actual traffic, finds delays and knock-on delays that have happened, and returns the networks of dynamic delay propagations. As such, the best return would be in strategic work to improve future schedules through analysis of current patterns.

### *5.1. Practical implications*

With regard to the practical application and the ability to extend to other lines, the proposed method is applicable on any single-track line. The theoretical propagation factors calculated from Potthoff's equation (Equation (2)) will depend on more detailed data concerning the information about the line and the schedule. The theoretical decreasing curves of propagation factors dependent on $n$, (Equation (3)), represent the expected development on any single-track line. However, how the actual propagation factors are distributed as a function of initial delay will vary between the railway lines, influenced by the schedules and other factors.

While the total number of delay minutes in the system might be low, imprecision in rapidly crossing traffic (as found on extensive single-tracked systems such as the Norwegian railway) causes lasting service degradation, and in effect, introduces a kind of "glass ceiling" for the overall degree of punctuality that is practically achievable.

The propagation of train delays in a railway network depends on several factors, including timetable and infrastructure characteristics. The presented method can be applied for analyses that are interesting in punctuality improvement work, including timetable analyses. The method can support identification of which parts of a line, and at what time of day, delay propagation is most severe. We can also identify individual trains that are subject to primary and secondary delays. An empirical approach to delay propagation on single-tracked railways may be beneficial in punctuality analysis to quantitatively identify single trains causing excessive problems for punctuality of the system. In Norway, such a tool can create the evidence basis for existing performance improvement measures, for instance, serving as a source of quantitative evidence in a system such as PIMS (Veiseth et al., 2011).

Performance measurement of railway traffic in several countries, including Norway, is based on two types of punctuality data. The first is time registrations; showing at what time a train arrived or departed a defined point in the railway system. These registrations are commonly based on data from the signalling system. These data are typically of relative high quality, especially if the registrations include seconds and not only minutes. The second data set is related to delay causes. Train dispatchers, or similar personnel, typically register these data manually. These data include a higher degree of subjectivity. Because the coding frequently includes a tagging of responsibility for a delay, it can cause heated discussions about root cause and responsibilities for delays. Reviews of the registrations have shown that some of the coding can be questioned. However, the performance measurement that is based on the delay coding typically receives more attention than the performance measurement based on pure time registrations. It is therefore a paradox that of the two available data sets, the best data are the least utilised, and the data of lower quality get much of the attention. Our proposed approach is to utilize the time-based data to support the delay causes data, for quality assurance and analytical purposes. From a longer time perspective, it would be interesting and possible to use time-based data in a structured way to support, or possibly even replace, manual registrations.

### 5.2. About delay propagation factors

In this paper, we calculated propagation factors of the case sample and compared the results to the delay cause registrations, and to the theoretically expected delay propagation. In the case sample, the mean propagation factor is about 2 (see Fig. 14). This means that an initial primary delay is expected to generate secondary delays of the same magnitude as the initial delay. This corresponds with the finding that over 70% of dynamic delay propagations consist of two trains (see Fig. 10), as the propagation factors for dynamic delay propagations consisting of two trains were clustered around 2 (see Fig. 15). The median propagation factor for each $n$ appears to be close to the value $n$ (Fig. 16), where $n$ is the number of trains in the dynamic delay propagation including the delayer train. Thus, based on Figs 15 and 16, we can say that the relationships between the initial delay and the propagation factor is in general $1:n$ for dynamic delay propagation with less than seven trains, and initial delays around 4–8 min.

The propagation factors for the delay cause registration were lower than the case sample, with a weight below 2. The delay cause registration data have some inadequacies that can have been a contributing factor to the low values. However, in light of the findings, the low values of propagation factors can in general be explained by either that mostly single knock-on delays can be extracted from the registrations, or that the initial delays are much higher than the secondary delay of each individual train they delay.

We calculated the theoretically expected propagation factor based on the equation by Potthoff (1970) on smaller sub-lines of the case line. The calculated propagation factors are between 1.1 and 2.2 for initial delays just above the buffer time of each sub-line. The propagation factors in the case sample decrease with increasing initial delay, while the theoretical propagation factors increase with initial delay. It was shown that this was due to the dependency between the initial delay and the propagation factor. Hence, larger initial delays will theoretically delay more trains. However, calculations of the median values of propagation factors for $n = 2, \ldots, 9$ in Fig. 16, showed that the median values do not lie in ascending order according to $n$. Thus, the practical propagation factor, according to our case sample, is not linearly dependent on the initial delay. A reasonable explanation is that Potthoff's equations, which we used in Fig. 16, strictly prioritize the trains travelling in one direction, adding all delays to the trains travelling in the other direction. However, in our case, such a strict prioritization of direction is not applied in the practical railway operation.

Furthermore, we examined the relationship between, and frequency of, how often trains were the delayer train and how often they were affected by dynamic delay propagations (see Table 2, Figs. 13 and 14). Some trains may be more exposed to delay propagations, for instance as a result of tighter schedules. The trains with high frequency as the delayer train and the trains that have ratio larger than one are candidates for special study when analyzing punctuality and timetable improvements. The scatterplots also identify trains for further study. In each of the scatterplots (Figs. 15, 17 and 18), we separated the values based on the number of trains in the dynamic delay propagations. The corresponding trend lines show how we expected the values to behave. However, there are some clear outliers in each of the plots. The trains with higher propagation factors or higher sum of secondary delays, or trains with lower sum of additional delay, are candidates for special study to further verify the algorithm when analyzing punctuality and timetable improvements.

### 5.3. Further work

Further development of the algorithm should include validation with competing methods of single-track delay propagation, including theoretical models for further calibration and validation of the conditions considered a knock-on delay. A vision for this work is to utilize time data of relatively high accuracy to improve data about delay causes, a type of data that is of lower accuracy. In particular, manual registration of secondary delays can be supported by this type of analysis.

Future iterations would also have to better account for and handle delays in the junctions between single- and multi-track infrastructure, as well as better handling of delays induced in first-/final-stop (where head codes change). For the punctuality improvement work, it would be informative to analyse the network of dynamic delay propagations, and the interactions at such junctions.

## Acknowledgement

This publication is a result of a cooperation between a student project at the Norwegian University of Science and Technology and the research project PRESIS at SINTEF Technology and Society. The authors would like to thank Sverre Wattum, Arild Almås Berg, Torkil Schjetlein and Ole Hartvik Skogstad for their contributions in the initial work on this topic, and Arild Almås Berg for programming the prototype. The authors are grateful for all data provided by Jernbaneverket (now Bane NOR).

## References

Conte, C., 2007. Identifying Dependencies Among Delays. Georg-August-Universität Göttingen.
Cule, B., Goethals, B., Tassenoy, S., Verboven, S., 2011. Mining train delays. In: Gama, J., Bradley, E., Hollmén, J. (Eds.), Advances in Intelligent Data Anlysis X, vol. 7014. Springer, Berlin, pp. 113–124.
D'Ariano, A., Pranzo, M., Hansen, I.A., 2007. Conflict resolution and train speed coordination for solving real-time timetable perturbations. IEEE Trans. Intelligent Transp. Syst. 8 (2), 208–222.
D'Ariano, A., Pranzo, M., 2009. An advanced real-time train dispatching system for minimizing the propagation of delays in a dispatching area under severe disturbances. Netw. Spatial Econ. 9, 63–84.
Flier, H., Gelashivili, R., Graffagnino, T., Nunkesser, M., 2009. Mining railway delay dependencies in large-scale real-world delay data. In: Ahuja, R.K., Möhring, R.H., Zaroliagis, C.D. (Eds.), Robust and Online Large-scale Optimization, vol. 5868. Springer, Berlin, pp. 354–368.
Goverde, R.M.P., 2005. Punctuality of Railway Operations and Timetable Stability Analysis TRAIL Thesis Series No. T2005/10. TRAIL Research School for Transport, Infrastructure and Logistics, Netherlands.
Goverde, R.M.P., 2010. A delay propagation algorithm for large-scale railway traffic networks. Transp. Res. Part C Emerg. Technol. 18 (3), 269–287.
Goverde, R.M., Hansen, I.A., 2001. Delay propagation and process management at railway stations. In: 5$^{th}$ World Conference on Railway Research (WCRR 2001). Köln, November 25-29, 2001. WCRR.
Handstanger, A.C.T., 2008. Scheduled waiting time from crossing on single track railway lines. PhD Thesis. NTNU Norwegian University of Science and technology.
Hansen, I.A., 2001. Improving Railway Punctuality by Automatic Piloting." Intelligent Transportation Systems. IEEE.
Harris, N.G., Mjøsund, C.S., Haugland, H., 2013. Improving railway performance in Norway. J. Rail Transp. Plan. Manag. 3 (4), 172–180.
Jernbaneverket, 2014. Jernbanestatistikk, vol. 2014. Viewed online March 23 2016. http://www.jernbaneverket.no/contentassets/8846eedda25f46afac36410aaa7053bf/jernbaneverket_statistikk_2014_web_enkeltsider.pdf.
Jernbaneverket, 2015. Kapasitet på jernbane. Accessed August 22, 2016. http://www.jernbanekompetanse.no/wiki/Kapasitet%5fp%b5%5fjernbane%23Enkeltspor.
Jernbaneverket Trafikkdivisjonen, 2016. Punktlighet i togtrafikken 2015. http://www.jernbaneverket.no/contentassets/c63c2f99e7cb46e98cc81962b2c5e7ab/punktlighet-persontog-pr-maned-2015-1.pdf. Viewed online October 2016.
Lindfeldt, A., 2010. A study of the performance and utilization of the Swedish railway network. In: First International Conference on Road and Rail Infrastructure. CETRA, Croatia.
Lindfeldt, A., 2015. Railway Capacity Analysis: Methods for Simulation and Evaluation of timetables, Delays and infrastructure. Doctoral thesis. KTH, Stockholm, Sweden.
Lindfeldt, O., 2007. Quality on Single-track Railway Lines with Passenger Traffic: Analytical Model for Evaluation of Crossing Stations and Partial Double-tracks. KTH, School of Architecture and the Built Environment (ABE), Transport and Economics (closed 20110301), Traffic and Logistics (closed 20110301). KTH.
Olsson, N.O.E., Haugland, H., 2004. Influencing factors on train punctuality – results from some Norwegian studies. Transp. Policy 11, 387–397.
Potthoff, G., 1970. Verkehrströmungslehre, Band1, Die Zugfolge auf Strecken und in Bahnhöfen. Transpress, Berlin.
Radtke, A., Bendfeldt, J.-P., 2001. Handling of railway operation problems with RailSys. In: Proceedings of the 5th World Congress on Rail Research (Cologne, Germany).
Samuel, H., 1961. Railway Operating Practice. Odhams Press Limited, Long Acre, London, UK.
Schwanhausser, W., 1974. Die Bemessung von Pufferzeiten im Fahrplangefuge der Eisenbahn. PhD thesis. Veroffentlichungen des Verkehrswissenschaftlichen Instituts der RWTH-Aachen. Heft 20.
UIC (International Union of Railways), 2013. UIC Code 406 Capacity, second ed.
Veiseth, M., Hegglund, P.M., Wien, I., Olsson, N.O.E., Stokland, Ø., 2011. Development of a punctuality improvement method. TQM J. 23 (3), 268–283.
Warg, J., 2013. Economic evaluation of capacity with simulations. In: Proceedings of the 10th World Congress of Railway Research (Sydney, Australia).
Yuan, J., Hansen, I.A., 2007. Optimizing capacity utilization of stations by estimating knock-on train delays. Transp. Res. Part B Methodol. 41 (2), 202–217.

# Publication II

Sørensen, A. Ø., Bjelland, J., Bull-Berg, H., Landmark, A. D., Akhtar, M. M. & Olsson, N. O. E. (2018). Use of mobile phone data for analysis of number of train travellers. *Journal of Rail Transport Planning & Management.* doi:10.1016/j.jrtpm. 2018.06.002.

Journal of Rail Transport Planning & Management 8 (2018) 123–144

Contents lists available at ScienceDirect

Journal of Rail Transport Planning & Management

journal homepage: www.elsevier.com/locate/jrtpm

# Use of mobile phone data for analysis of number of train travellers

Anette Østbø Sørensen[a,*], Johannes Bjelland[b], Heidi Bull-Berg[c], Andreas Dypvik Landmark[c], Muhammad Mohsin Akhtar[a], Nils O.E. Olsson[a]

[a] Norwegian University of Science and Technology, Dept. of Mechanical and Industrial Engineering, N-7491, Trondheim, Norway
[b] Telenor Research, Oslo, Norway
[c] SINTEF AS, Postboks 4760 Sluppen, N-7465, Trondheim, Norway

ARTICLE INFO

*Keywords:*
Railway
Ridership
Train
Mobile phone data

ABSTRACT

Several studies have pointed to the difficulties of obtaining good data on train ridership. There are at least two challenges regarding these data. First, train operators consider such data confidential business information, especially in high resolution. Second, the data that actually are available vary in quality and coverage. This paper studies mobile phone data as an alternative measure to obtain data about train ridership.

Handset counts were obtained from one telecom operator for selected mobile phone base stations and compared with timetable data and APC. The selected base stations are located so that it is likely that a large share of the mobile phone traffic is generated by train passengers. The number of units connected to a base station is found to correspond relatively well with the trains that pass close to the base stations. A ratio between the handset count and APC data appear as promising in utilizing handset count to calculate train ridership, with ratios around one in the rush hours. We discuss preliminary results as well as methodological and technical challenges.

To make sure that we do not violate privacy concerns, the data used in the study have been approved by personal privacy representatives.

## 1. Introduction

### 1.1. Research on train ridership is important

When analysing public transportation, including trains, ridership is an important factor. The number of travellers is a measure of demand for transportation services, which is important information for planning and evaluations. With updated ridership information, planners should be able to get a detailed, continuous and accurate vision of the travel behaviour of their customers. This is important in planning and improving the transportation service. Other uses of ridership numbers are calibration and validation of transport models. Boyle (1998) identifies four main reasons why ridership data are collected. Firstly, ridership is reported to external funding and oversight agencies. Secondly, it monitors trends over time. Thirdly, ridership is a key performance indicator at various levels of the transportation system. Finally, ridership data identifies locations with the greatest boarding and alighting activity, which is important not only for its own purpose, but because the safe management of the railway may depend upon it. Other issues that call for data on ridership on trains include fare equipment location optimization, fare policy change and train schedule (Li, 2000). In addition, revenue distribution in integrated public transportation systems can be based on ridership data.

According to Vuchic (2005), the purpose of obtaining data on passenger volume and load count is to monitor trends and travel behaviour over time. Such data show passenger volumes on different sections of a line, the maximum number of passengers on

* Corresponding author.
*E-mail address:* anette.o.sorensen@ntnu.no (A.Ø. Sørensen).

https://doi.org/10.1016/j.jrtpm.2018.06.002
Received 16 March 2018; Received in revised form 27 May 2018; Accepted 6 June 2018
Available online 20 June 2018

different lines and when the maximum is reached, along with information on variations in passenger volume. Ridership data are also a key performance indicator in public transport (Vuchic, 2005).

On the other hand, the process of obtaining data on ridership creates least two major challenges. Firstly, train operators consider such data confidential business information, especially in high resolution (Vigren, 2017). Secondly, the data that actually are available vary in quality and coverage. Several studies highlight the unreliability of ridership data (including Chu and Chapleau, 2008 and Fowkes et al., 1985).

### 1.2. Mobile phone data and alternative technologies

Doi and Allen (1986) studied a rapid transit line for a period of about six and a half years (from 1978 to 1984) based on ridership data provided by a transit authority. Even though some studies combine several data sets, most studies on ridership rely on one data source. Wang et al. (2011) explored the application of archived data from Automated Data Collection Systems (ADCS) to transport planning with a focus on bus passengers' travel behaviour. They claimed that it was the first known attempt to validate the results by comparing automated ridership data with manual passenger survey data. Passenger distribution in the urban Copenhagen rail network is, according to Nielsen et al. (2014), tracked based on a combination of Electronic Weighing Equipment (EWE) and Automatic Passenger Count (APC). The two systems provide complementary information, since the weight-based estimation provides information about the total traffic volume and automatic passenger counting provides information on passenger flow. The two systems can also be used to perform quality assurance of each other's measurements. Zhao et al. (2007) combine data from the automated fare collection system and the automated vehicle location system to examine the rail-to-bus trip sequence to obtain a clearer picture of ridership patterns. De Regt et al. (2017) combine smart card and Global System for Mobile Communications (GSM) data to examine spatial and temporal patterns of public transport usage versus overall travel demand. The methodology was applied to a case study in Netherlands, and was shown to be valuable in supporting tactical transit planning and decision making.

Sørensen et al. (2017) identify several technologies for measuring ridership on trains. The technologies and approaches include (1) manual counts and surveys, (2) on-board sensors, such as door passing, weight, CCTV and Wi-Fi-use, (3) ticketing systems, ticket sales or ticket validation, and (4) tracking of travellers for larger part of the journey, such as tracking of mobile phones and payments. Pelletier et al. (2011) presents an overview of the first developments of smart card. Smart cards are used to store individual data such as identification, biometrics, photos, banking data, transportation fares, etc. In transit, the main purpose of smart cards is to collect revenue, but they also produce detailed data on onboard transactions which can be useful to transit planners on both a strategic, tactical and operational level. Smart cards in public transit are usually issued by the operators to be used on their own system, and the cards are typically tapped over the reader when the user enters the vehicle.

Pelletier et al. (2011) summarise the pros and cons of smart card use in public transit which they revealed in their literature review. Some of the pros and cons are also valid for mobile phone data. Disadvantages are for instance that the data cannot provide information on trip purpose or on user assessment of service, and that development cost is high. Furthermore, the ultimate destination is not provided. Advantages include that the user role in data collection previously achieved by the survey process is minimized, as well as improved data quality and increased amount of statistics available.

Data from on-board sensors and ticketing systems are typically managed by the transportation providers. However, surveys, payments statistics and mobile phone data may be available to stakeholders outside the public transportation system, which can be an advantage because access to ridership data can be an issue for business reasons. Furthermore, mobile phone data appears to be an interesting option because they can track complete journeys.

Mobile phone data can be used to derive good estimates of dynamic quantities, such as travel times, train occupancy levels and origin-destination flows, for transportation studies (Aguiléra et al., 2014). The advantages of mobile phones as sources of data include:

- the potential to generate information about travels that combine different modes of travel (such as walking, bus, train),
- to track journeys that include transfer between trains,
- to estimate commuting patterns,
- and to derive estimates of travel times, train occupancy levels and origin-destination flows.

Several studies on mobile phone data in transportation research utilise Call Detail Records (CDR) data. This paper studies a different type of mobile phone network source that will be presented in Section 3.2.1. Three main types of mobile phone data are collected using passive collection: CDR data, Probes data and Wi-Fi data (Larijani et al., 2015). CDRs are generated by phone communication activities and contain relevant information about the activity (e.g., caller/callee, time, duration) and the location of the cell phone tower that handles the communication (Zhao et al., 2016). Studies have shown that CDR data can be used to study habits and mobility patterns of mobile users (Bianchi et al., 2016; Zhao et al., 2016), to study user movements (Leo et al., 2016), and to calculate commuting matrices with a very high level of accuracy (Frias-Martinez et al., 2012). Studies have also looked at utilizing mobile data to estimate intra-city travel time (Kujala et al., 2016) and have shown that mobile data could be employed as a real-time traffic monitoring tool (Järv et al., 2012).

Studies point out that CDR data are coarse in space and sparse in time (Becker et al., 2013) because people's phone communication activities are unevenly distributed in space and time. The bias of CDR data in human mobility research depends on what research question one wants to answer and how frequently, as well as when and where, one uses the mobile phone to contact others (Zhao et al., 2016). It has therefore been suggested that researchers should use CDR data with caution.

CDR data contain information about the caller/callee, so they are not anonymous. Consequently, studies that utilise CDR data are required to protect privacy through measures such as anonymizing the data (i.e., removing personal identification), only using the minimum of information needed for the studies, only presenting aggregated results and not focusing the analysis on individual phones (Becker et al., 2013). With pseudo-anonymized data (i.e., the ID is replaced with a code), the record must be pre-processed to reduce probability of re-identification. A common procedure is to decrease time resolution or increase space granularity (Bianchi et al., 2016). Norwegian Law states that collected personal information should only be used for the specific purpose for which it was originally collected (Drageide, 2009). As a consequence, any use of CDR data that goes beyond billing requires an active consent from the subscriber. Furthermore, the EU General Data Protection Regulation (GDPR) will be enforced in May 2018.

We have found no publications that combine data from mobile phone networks with comparable registrations of number of travellers based on on-board ridership data from train operators, even though others have combined mobile phone data with other types of public transport data such as Holleczek et la. (2014).

### 1.3. Research purpose

This paper studies how mobile phone data can be used to analyse the number of travellers on trains. This research has both long- and short-term perspectives. In a short-term perspective, we study how mobile phone data can be used to analyse the number of travellers on trains. In a long-term perspective, we can measure traffic flows in new ways to cover whole journeys. In addition to providing information about ridership, mobile phone data can provide information about journey flows that is not limited to each mode of transportation, but for complete journeys including travel modes such as walking, bus and train. Related to train travel, we can track train journeys that include transfer between trains, which is difficult to obtain using established techniques. Such tracking may raise personal privacy concerns, but it is not necessary to identify individual trips but to focus on flows and movements of large groups, and such data can be made anonymous (Olsson and Bull-Berg, 2015). We will then be able to see transport patterns and not only measure the volume of traffic at those points where there is a count. One can also seek explanations by combining ridership data with, for example, data on punctuality or weather. We are interested in this type of data to evaluate major transport infrastructure investments such as new double tracks of railway tunnels. Several such projects are ongoing in Norway. We investigate how mobile phone data can be used in future evaluations of these projects.

The purpose of this study is to test the use of mobile phone data to measure train ridership and to investigate the potential for using mobile phone data to describe travel patterns that include train travel. Our research questions (RQs) follow.

- RQ1. Is it possible to combine mobile phone data with railway infrastructure and train traffic data?
- RQ2. What are suitable formats for presenting and analysing train ridership based on mobile phone data?
- RQ3. To what extent is the format of available mobile phone data suitable for measuring the number of mobile units passing close to the railway line?

We will discuss our results as well as methodological and technical challenges with such an approach to estimate train ridership compared to other established methods.

## 2. Analysing ridership on trains

### 2.1. Use of information about number of travellers

The distribution of travel demand can be analysed and presented as a function of time or as a spatial distribution (Vuchic, 2005). Spatial distribution measures the volume of travellers in different parts of a transportation network. In a railway context, spatial distribution is used for different parts of the network. However, it would be interesting to track spatial distribution of a larger part of journeys, not just the railway ride, to ideally include the whole trip from origin to destination. The time distribution of number of travellers can be studied in different time perspectives, including variations during a day, during the week, yearly variations and long-term developments spanning several years. Daily variations in commuter transport are characterised by the morning and afternoon rush hour peaks.

Train ridership is influenced by a number of factors, including fares, transit time, transit comfort characteristics and feeder accessibility of transit, price and service characteristics of the competing modes, seasonal variations and monthly working day variations, as well as socioeconomic conditions of the service areas in the medium or long term (Doi and Allen, 1986). Demand for railway travel is typically expressed in number of travellers. Other more nuanced measures include the common format origin-destination matrices. Vuchic (2005) lists a set of relevant key performance indicators related to ridership:

- average passenger trip length, total passenger-km divided by number of passengers,
- average passenger volume, total passenger-km divided by line length,
- coefficient for flow variations to indicate the degree to which passenger volume peaks along a line,
- coefficient of passenger exchange, what proportion of passengers that are exchanged along a line,
- riding habit, how much of a population in an area that utilises the transport in question (such as commuter railway), and
- market share, use of a particular type of transportation in relationship to total travel volume in the same market.

Traditionally, there are multiple methods for calculating the demand between an origin and destination point. The most common is the O-D matrix that characterizes the transitions of a population between different geographical regions representing the origin (O) and destination (D) of a route (Frias-Martinez et al., 2012). The most commonly used method for populating these matrices is user surveys. Strengths of traditional surveys are that they include important information about the respondent, such as age and gender, and also include information about the purpose of the trip (Alexander et al., 2015). A major problem with user surveys is declining response rates (Schoeni et al., 2013), which may introduce bias into the samples. Such surveys are typically not done with a higher frequency than yearly, but may also be conducted less frequently and not necessarily on a regular basis. Consequently, this method may possess low frequency, high cost, varying data quality, low precision and susceptibility to errors. Alternatives to traditional transport surveys include an origin-only automatic fare collection system, as proposed by Zhao et al. (2007), and mobile phone data (e.g., Jiang et al., 2013). Mobile phone data can be used to describe people's movement patterns, as illustrated in the study by Calabrese et al. (2013) who analysed the mobile phone records of a million users in Boston to describe transportation needs.

Passenger counting is the key measuring parameter associated with ridership. Different measurement types and ridership estimation techniques are applied for different network levels. The selection of the appropriate network level is dependent on the particular use and issue being addressed (Gordillo, 2006; Boyle, 1998). Gordillo (2006) and Boyle (1998) identify uses of passenger counting and ridership calculation based on the way data are measured. The uses are different for each measurement type, and the operator normally uses multiple types to fulfil different purposes. The types and their relevant usages are as follow.

- System Level Use: Tracking system-wide ridership totals to assess changes in ridership.
- Route Level Use: Ridership by route, passenger loads at maximum load points, compiling ridership by day type and time period and monitoring schedule adherence. Performance measures are frequently calculated at the route level, and running time adjustments, route revisions, and ridership trends also rely on route-level data. Route level data are used primarily for planning and scheduling.
- Trip Level Use: Data on ridership by trip is used to add or delete trips and to adjust running times, schedule adherence and passenger loads. Link-loading between two adjacent points are important for rail capacity.
- Stop Level Use: Data on entries per station and boarding and alighting by stop are typically used in adjusting running times and in service planning in assessing route performance.
- Origin-destination level: Measures the number of passengers travelling between a pair of stations. OD data are used to help with revenue maximisation.

The Norwegian National Rail Administration (former Jernbaneverket, now Bane NOR) has published annual Official Railway Statistics for Norway. The railway statistics include aggregated data on number of travellers, passenger kilometres, and number of sold single tickets and monthly tickets. The practice for measuring ridership is manual counting at chosen stations on each railway line.

### 2.2. Mobile phone network data

Much research has focused on developing methods to extract meaningful information about human mobility from mobile phone traces and understanding its limitations (Alexander et al., 2015). Mobile phone data can be utilised in estimating commuting patterns and travel times for individuals. Chaudhary et al. (2016) discuss collecting information about occupancy levels of public transportation system using smartphones. They show that patterns observed can predict occupancy level in a bus, with accuracy up to 92 percent. Higuchi et al. (2015) identify a number of innovative uses based on mobile devices, including several technologies that typically are found in smartphones, such as GPS, Wi-Fi, and Bluetooth. Mobile phone data sets allow for a statistical analysis of human activities at a fine level of detail (Leo et al., 2016).

Various approaches can be utilised for calculating this information by analysing the exchange of information between the mobile base station and cellular network. Most studies perform some kind of trip extraction to extract the movements relevant for traffic analysis from the raw cellular network data (e.g., Calabrese et al., 2011; Doyle et al., 2011; Alexander et al., 2015; Iqbal et al., 2014). Because cellular network data can contain a lot of noise, there is no obvious definition of what a movement/trip is (Gundlegård et al., 2016). Hence, trip extraction algorithms vary a lot among different authors.

An origin-destination matrix can be computed based on the extracted trips (e.g., Calabrese et al., 2011; Larijani et al., 2015). For instance, Alexander et al.'s (2015) method estimates average daily origin-destination trips from triangulated mobile phone records of millions of anonymized users. The CDR records are converted into cluster locations and inferred to be home, work or other depending on observation frequency, day of week and time of day. The aggregation of OD flows gives an estimate of the number of cell phone users who are travelling, but only those of the operator that provided the data (Gundlegård et al., 2016). As a result, this can only give information about how the travel demand distributes relatively between different OD pairs. To estimate the total travel demand in terms of the number of people travelling, authors use different scaling factors (e.g., Alexander et al., 2015; Iqbal et al., 2014; Toole et al., 2015; Calabrese et al., 2011).

Several authors have also tried to reconstruct the specific travel mode and route that a user took for a trip, which is challenging. However, as Larijani et al. (2015) showed, detection of the trip segments in which people take the metro is promising, because underground tunnels are served by dedicated base stations (Gundlegård et al., 2016). Xu et al. (2016) used a large-scale mobile phone data set to estimate demand of bicycle trips in a city. Another approach is to use smartphone travel surveys based on smartphone applications to capture accurate details about individuals' travel behaviour, as presented by Assemi et al. (2016). However, extracting required information (e.g., travel mode and purpose) from the data captured by smartphone applications is relatively complex.

Holleczek et al. (2014) showed that urban mobility patterns and transport mode choices can be derived from mobile phone CDR coupled with public transport data. This public transport dataset consists of trips made by 4.4 million anonymized users of Singapore's public transport system. The advantage is that passengers in Singapore use smart cards when getting on and off trains and buses, hence the data include station and time of departure and arrival of each trip.

Calabrese et al. (2013) show that mobile phone data can be used to describe people's movement patterns, as an alternative to traditional transport surveys. They obtain mobile phone records of a million users in Boston during a three-month period to describe transportation needs. They discuss three challenges using mobile data. The first is that demographic information about individuals was not available due to privacy concerns. Secondly, mobile users were not necessarily representative of the whole population. Thirdly, the data were not formatted for this type of analysis. To address the first challenge, they used aggregated data in which users were collected into groups corresponding to the most detailed level of economic and demographic data that were available. The second issue introduces sample bias amongst the population. To validate the representativeness, Calabrese et al. (2013) calibrated the data based on information from security inspections of the vehicles, which included mileage condition, to check if the estimated mileages seemed realistic. Mobile phones are a growing data source through activated apps. Apps can track complete journeys, especially if the users have allowed apps to use GPS for tracking. Such apps can be supplied by private or public transportation entities, or they can be apps for navigation, health monitoring or other types. Data from these apps are typically managed by the organisation issuing the app, and not by the mobile network managers.

There are both advantages and disadvantages by using mobile phone data. For instance, CDR data contain approximate locations when the phone communicates with a cell phone tower, hence providing an inexact and incomplete picture of daily trips. Furthermore, the mobile phone data are not able to provide information about the traveller, like age, income or purpose of trip, as a survey would (Alexander et al., 2015). On the other hand, mobile phone data are automatically collected, which makes them more frequent and economical than, for instance, a survey. In addition, as mobile phone data can be gathered over a longer time period, it can capture information such as variations in the travellers' daily travel behaviour (Alexander et al., 2015).

Having established the need and uses of ridership data, the short-term vision of this project is to investigate if available mobile data and railway traffic data are a viable method for calculating number of travellers on a specific train route. This data can serve as a source of validation, quality assurance and triangulation for the currently available data in railway industry.

## 3. Approach

Three sets of data were obtained for analysis purposes: punctuality data for specific train stations, mobile phone data in a specific format from adjacent cell sites, and actual passenger counts from the trains. The data sets are described in Section 3.2. The purpose of this study was to investigate suitability of mobile data to find ridership on trains and to investigate if it is possible to combine mobile data with railway infrastructure and train traffic data.

The first part of the analysis focused on combining mobile telephone data with data describing the railway and railway traffic. As we will see, a key issue is to relate peaks in mobile telephone connections to the passing of trains. The final part of the analysis utilises counts of passengers on the trains and connects the number of travellers on the trains to the number of mobile phone connections.

The analysis was done in the following steps:

- identifying base stations,
- connecting trains to base stations,
- graphic inspection of handset counts in relationship to trains passing the base station,
- analysis of data resolution, comparing data sets of five- and one-minute collection time intervals,
- statistical analysis using a proposed algorithm,
- extracting the peaks in handset count, correlated with trains passing the base station, and
- comparison and validation with actual ridership data.

A key step of the approach was to locate suitable mobile phone base stations close to the studied railway line. Thus, the approach of selecting base stations is described subsequently, along with analysis methods. To begin with, the following section describes the use case.

### 3.1. Use case description

The railway line selected as case for this preliminary study goes into a city in Norway, where people commute daily to work from towns on the outskirts of the city. The analyses look at five base stations located near the railway tracks and in connection with five of the train stations on the selected railway line. To anonymize the data, we denote the train stations as U, V, W, X and Y, where station U is farthest away and station Y is closest to the city, as illustrated in Fig. 1. The base stations are denoted B1, B2, B3, B4 and B5.

Data were collected in three time periods:

1. eight consecutive days in the spring of 2016,
2. three consecutive days in the fall of 2016, and
3. nine consecutive days in the spring of 2017.

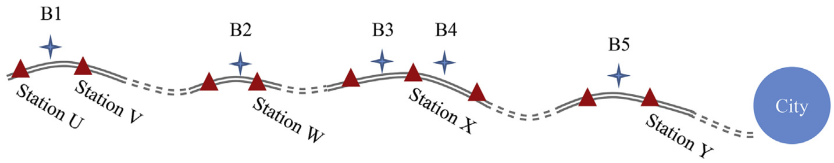


**Fig. 1.** Illustration of the studied railway line, with the five base stations in relationship to the train stations.

Punctuality data were made available for all three time periods. The mobile phone data in time period 1 were collected with five-minute time intervals between each collection time. In time period 2, the mobile data were collected with one-minute time intervals between each collection time. This data set is less complete than the data set with five-minute collection intervals. In time period 3, the mobile data were collected with one-minute time intervals between each collection time. This data set is complete. Automatic Passenger Count (APC) data were made available for time periods 1 and 3.

### 3.2. Available data and research material

#### 3.2.1. Mobile phone data

The mobile phone data sets are counts of the number of handsets recorded at the selected base stations. Since the base station has limited range and a mobile subscriber can be connected to one of six base stations based on the signal strength, due to a hexagonal symmetry for frequency reuse (Mac Donald, 1979), the handset counts for a certain base station serve as an indicator of the number of people using the mobile network in the coverage area of the base station.

The counts are the total number of connections to a base station cell, and the data set consists of *collect_time*, *cell_id*, and *count_handsets*, meaning the number of handsets connected to a base station cell at a given point in time. Example of data format is shown in Table 1. For the purpose of this study, a script was made that extracts these data with certain intervals. When a phone turns up on another base station, it will no longer be counted on the previous one. The handset count will serve as an indicator of the number of people using the mobile network in the coverage area of the base station, at the exact time the data are collected.

It is worth mentioning, though obvious, that passengers without mobile phones and those whose mobile phones are switched off or not working, will not be recorded by this approach. In the source system used to extract the count data, it is not possible to see each individual event. However, at any time, the mobile network operator can extract the number of mobile phones that were last seen on the base station.

The mobile phone data used in this study are pure counts by cell by time unit, which are anonymous data that are not covered by the privacy legislation and cannot be used to identify a person. Both from an ethical and legal point of view, it is important to protect personal information and respect people's privacy. Data that do not include personal information are basically unproblematic, both as individual data sources and the combination of several sources. Combination of different data sources in which persons are the link between the various data is more problematic. Data from different sources can be combined with personal data without revealing personal information, but this can be challenging. Anonymity in datasets is typically achieved by aggregation, in which each group includes so many persons that individuals cannot be identified. To make sure that we do not violate private privacy, the data used in the study have been approved by privacy representatives.

#### 3.2.2. Railway traffic and infrastructure data

The Norwegian National Rail Administration (Jernbaneverket at the time of the study, now Bane NOR) records punctuality data for individual trains at arrival and departure at stations. Data are recorded in a database (TIOS), of which we have obtained a copy. The data describe the movements of the trains through the network. These records include scheduled and actual arrival and departure times for each train at every station, train number and operating company, and class information (e.g., freight, running empty, or passenger train). Example of data format is shown in Table 2.

The physical layout of the Norwegian railway network is described in several formats, many of them being available on the internet (BaneNor, 2017). We have utilised this information to combine train data and mobile phone data. In particular, we calculated when trains passed close to the mobile phone base stations.

#### 3.2.3. Automatic passenger count (APC) data

The third data set is passenger counts from trains based on Automatic Passenger Counting (APC). The APC system is installed on a sample of the vehicles on the railway line studied in this work. The APC data that were made available were collected from the trains on two of the railway lines that pass station Y, which we denote Line 1 and Line 2. The APC system registers the number of people who

**Table 1**
Sample table showing the data format of mobile phone data.

| Collection time | Cell id | Cell easting | Cell northing | Handset count |
|---|---|---|---|---|
| 2017-06-01 06:19:00 | xxx | xxxx | xxxx | 159 |

**Table 2**
Sample table showing the data format of TIOS data.

| Date | Train no. | Station code | Scheduled arrival | Actual arrival | Scheduled departure | Actual departure |
|---|---|---|---|---|---|---|
| 2017-06-01 | 123 | XYZ | 17-06-01 11:59 | 17-06-01 11:59 | 17-06-01 12:00 | 17-06-01 12:00 |

**Table 3**
Sample table showing the data format of APC data.

| Train no. | Product name (start-stop) | Direction | Date | Day | Location (train station) | Sum boarding | Sum alighting | Sum passengers |
|---|---|---|---|---|---|---|---|---|
| 123 | Small town-Other town | Towards Other town | 17-06-01 | Thu | City C | 37 | 32 | 189 |

board and alight through each train door on every station by means of sensors in the doorways. Norsk Regnesentral (Norwegian Computing Centre) has developed a mathematical tool that, based on the APC data, uses a statistical model to calculate the total number of passengers on each train at different points in time (Teknisk Ukeblad, 2014). The data set that was made available to us is the calculated total number of passengers on each train when the train is leaving the station. Example of data format is shown in Table 3.

### *3.3. Approach of selecting base stations*

The base stations in close proximity to the railway tracks were located based on coordinates and description of coverage. The Quantum geographic information system (QGIS) was used to import the coordinates for the base station, together with a map of the area and the railway lines. The initial criteria used to select base stations were that the base station should be within 2 km from the tracks, but also excluding base stations more than 1 km from the end train station in the opposite direction. Generally, we also excluded cells with descriptions of indoor coverage. These restrictions resulted in approximately 600 cells divided between around 100 base stations. To arrive at a more appropriate number of cells for our preliminary analysis, we singled out the base stations that with greater certainty would be connected to train travellers, e.g., including base stations with descriptions containing railway, train stations or railway tunnel. The number was further narrowed down to about 10 by selecting base stations that (1) were located between train stations, (2) were not located near a main road, and (3) were located in more deserted areas. Mobile phone data were obtained from the selected base stations and compared with timetable data. The selected base stations are located so that it is likely that a large share of the mobile phone traffic is generated by train passengers.

### *3.4. Applied methods*

As mentioned in Section 3.3, the selected base stations were located between the train stations on the case line. A first step in the analysis was to compare the handset counts on the base stations with when trains are passing the location in proximity to the base station to see if it is possible to connect passing of trains with possible jumps in the handset count data. The train traffic data give the actual time of when the trains arrive and depart from the train stations. We therefore needed to find how long after the train leaves the train station the train passes the location in proximity to the base station.

We looked into two different approaches to find the approximate time of when the train passes the base station. The first approach is based on the distances and times between the stations, with the assumption of constant speed. The second approach is based on the allowed speed on the railway lines.

**Approach 1**: With the assumption that the train has an average constant speed from station $A$ to station $B$ (see Fig. 2), the time $s$ can be expressed as $s = \frac{y}{x}t$ minutes. The travel time from station $A$ to station $B$ can either be the scheduled travel time or the actual travel time for a specific train at a specific time. We used actual travel time when it was available in the train traffic data.

**Approach 2**: Based on the speed limits on each section of the line, which are rarely constant between two stations, the average speed $v$ from the nearest train station to the base station can be calculated, and the time is found to be $s = \frac{y}{v}60$ minutes.

Based on approach 1 or approach 2, we could then estimate the time for when trains passed the base stations, utilizing data that

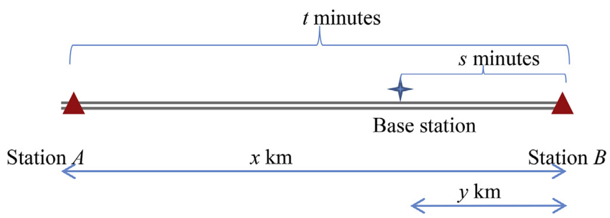


**Fig. 2.** Illustration of a base station located between two train stations and the notations used to calculate travel time of the train from train station B to the base station.

show actual train movements. If the train is travelling from $A$ to $B$, the point in time when the train passes the base station is then given by $t_{arrival}(B) - s$. If the train is travelling from station $B$ to station $A$, the point in time when the train passes the base station is given by $t_{departure}(B) + s$. Based on preliminary analyses, the approaches yielded quite similar results. We therefore chose to use approach 1 throughout this study.

#### 3.4.1. Algorithm

This section presents an algorithm that compares the collection times with the calculated approximate times the trains pass the base stations. The method aims at quantifying the impact of a train passing and its extent. The main point of interest is to categorize the collection times for mobile data to reflect whether a train has passed the base station or not. This basic categorization distinguishes the closest count collection points to the passing of a train and their adjacent values, with the objective of determining if a train has passed between two subsequent collection times. Further sub-categorization is made to analyse the effects of rush hours and direction.

The algorithm, as presented below, compares each single collection time with an array containing the calculated times of all trains that pass the base station. The collection time closest to the time of train passing is the minimum value of the time difference between the collection time and the train time. When the value is less than the collection time interval, that implies that the specific collection point is adjacent to a train passing. Otherwise, if the value is greater than the collection time interval, the specific collection time instant is not reflective of a train passing.

The input values of the algorithm are the three known variables: number of handsets (numerical value), denoted *count_handsets*, collection times (date and time value), denoted *collect_time,* and time of train passing at base station (date and time value), denoted *train_time*. In addition, the collection time interval is denoted $t_{CI}$. The output variables calculated in the algorithm are illustrated in Fig. 3:

- $D(i)$: time difference between collection time $i$ and the nearest train passing time before it (value in minutes);
- $T(i)$: string value "Yes" or "No", indicating if a collection time is adjacent to a train passing since the last collection time. 'Yes' if a train has passed within the collection time interval (i.e., $D(i) < t_{CI}$), 'No' otherwise, i.e., $D(i) > t_{CI}$; and
- $I(i)$: percentage increase in handset count, i.e., percentage increase from the previous count at a specific collection time, given by $I = \frac{d_c^i}{count(i-1)}100$, where $d_c{}^i$ is the difference in handset count given by $d_c^i = c_i - c_{i-1}$.

The steps of the algorithm are as follows.

1. Find the length of the arrays *collect_time* and *train_time,*
   $n$ = length(*collect_time*); $m$ = length(*train_time*).
2. **For** $i$ in 1 to $n$; **for** $j$ in 1 to $m$: compare the value *collect_time*($i$) to the value *train_time*($j$). If the value *collect_time*($i$) is greater than *train_time*($j$), subtract it from the collection time value and calculate the difference in minutes. The value is saved in a temporary array $P$. That is,
   **If** *collect_time* ($i$) > *train_time* ($j$)
   $P(j)$ = *collect_time* ($i$) - *train_time* ($j$).
3. Calculate the closest train passing in minutes, i.e., time difference $D(i)$, for a specific collection time $i$ as the minimum value in $P$.
   $D(i) = \min(P)$
4. Step 2 and 3 are repeated for each $i$, which will result in a difference value array $D$ of length $n$.
5. **For** each *collect_time*($i$),
   **If** the corresponding Difference value is less than $t_{CI}$ minutes: $D(i) < t_{CI}$
   then the $T$ value is Yes: $T(i)$ = 'Yes'
   **else** if it is greater
   $T(i)$ = 'No'.
6. **For** $i$ in 1 to $n$, the percentage increase $I(i)$ in each handset count is calculated from the previous count value as,
   $I(i) = \frac{c_i - c_{i-1}}{c_{i-1}} 100.$

The described algorithm was implemented in MATLAB R2017b (9.3), a numerical software environment and programming language developed by MathWorks. The values calculated in the algorithm were then analysed in the statistical software R 3.4.

#### 3.4.2. Graphic inspection

The graphic inspection entails plotting the handset counts against time. The handset counts were compared to the calculated approximate time of when trains are passing the base station, illustrated by vertical lines in Fig. 4. The assumption is that as the train pass the base station, the number of users connected to the specific base station increases, corresponding to the people travelling on the trains, because the passengers on the trains also get connected to the specific base station and thereby increase the count at the next measurement point, so that at the next collection time, a higher value is observed.

The directions the trains are travelling in are included in the analysis, where the trains are divided into two categories: going towards the city, and going away from the city. The direction of the train is illustrated in Fig. 4 by green lines for trains travelling towards the city and red lines for trains travelling away from the city. Morning rush hour trains going into an urban centre were

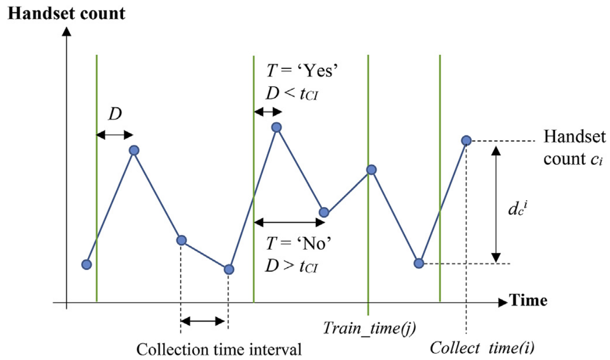


**Fig. 3.** Illustration of the output variables of the algorithm. The variable D is the difference in time between a collection time and the nearest train time. T is the string 'Yes' or 'No', whether D is less than the collection time interval or not. Handset count $c_i$ at collect_time(i) and difference in handset count $d_c^i$.

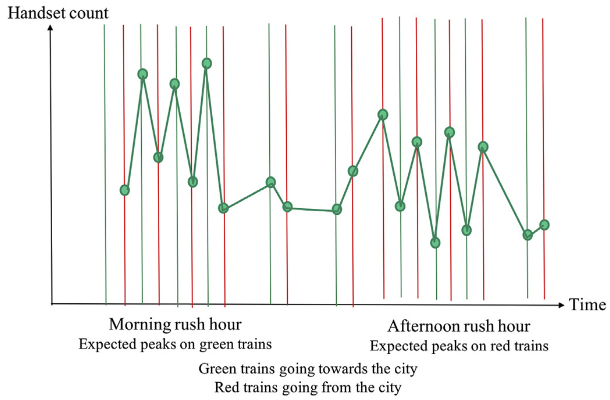


**Fig. 4.** Illustration of expected correlation between the handset counts and the morning and afternoon rush hours, where the vertical lines represent the calculated times the trains are passing the base station. Trains travelling towards the city are illustrated by green vertical lines, and trains travelling away from the city are illustrated by red lines. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

expected to have more passengers than the trains going in the other direction, especially in typical suburban and rural areas. We would therefore expect peaks in the diagrams when trains going into the city pass a mobile base station, as illustrated in Fig. 4. The opposite would be expected in afternoon rush hour. We would expect a lower variation in periods when there are no trains, compared to when trains pass.

### *3.4.3. Collection time interval*

The collection time interval of how frequently the handset counts are collected is expected to have an effect on the results of the analysis. The collection time interval was studied by aggregating a data set with one-minute collection time intervals down to a data set with five-minute intervals. Because the handset counts are snapshots of the number of handsets connected to the base station, the aggregation is done by extracting every fifth measure collected at the times 00, 05, 10, etc., past each hour to create a new data set with five-minute collection time intervals. These two data sets were compared by graphic inspection and compared to the calculated times the trains passed the base station.

### *3.4.4. Statistical analysis—Violin plot*

The statistical analysis was based on the categories presented in Table 4 and Table 5.

To validate and compare the results of the statistical analysis, certain assumptions need to be considered. The assumptions are based upon the general trends believed to exist based on previous general traffic patterns.

1. The train passing causes an increase in the number of handsets in addition to the natural variation of the mobile handset count, i.e., the 'Yes' category is higher than the 'No' category.

**Table 4**
General categories for statistical analysis and the values included in the analysis.

| Category | Condition | Values | Description |
|---|---|---|---|
| Yes | $D(i) < t_{CI}$ $T(i)$ = 'Yes' | *collect_time*($i$), *count_handset* ($i$), $I(i)$ | All collection points and adjacent handset counts, as well as the percentage increase in handset count, where a train has passed since the previous collection time. |
| No | $D(i) > t_{CI}$ $T(i)$ = 'No' | *collect_time*($i$), *count_handset* ($i$), $I(i)$ | All collection points and adjacent handset counts, as well as the percentage increase in handset count, where a train has not passed since the last collection time. |

**Table 5**
Categories for statistical analysis within the category of events where a train has passed since the previous collect time, i.e., D(i) < 1, T(i) = 'Yes'. Direction towards the city is denoted by true (TR) and false(FL).

| Category | Conditions | | Description |
|---|---|---|---|
| | Collection time (hrs) | Direction | |
| TR | All day | Towards the city | All trains travelling towards the city |
| FL | All day | Away from the city | All train travelling away from the city |
| MRTR | 06:00 to 09:59 | Towards the city | Trains travelling towards the city in the morning rush hours |
| MRFL | 06:00 to 09:59 | Away from the city | Trains travelling away from the city in the morning rush hours |
| ERTR | 15:00 to 18:59 | Towards the city | Trains travelling towards the city in the evening rush hours |
| ERFL | 15:00 to 18:59 | Away from the city | Trains travelling away from the city in the Evening rush hours |

2. In the morning (MR) and evening rush hours (ER), the number of travellers on the train is higher in count than at other hours of the day. This means the 'MRTR', 'MRFL', 'ERTR', and 'ERFL' categories possess higher statistical values than 'TR' and 'FL'.
3. In the morning rush hour, the number of passengers travelling on trains towards the city is higher than the number of passengers on trains travelling away from the city, i.e., 'MRTR' statistical values are higher than 'MRFL'.
4. In the evening rush hour, more passengers travel away from the city than travel towards the city, i.e., the 'ERFL' category has higher value than 'ERTR'.

Information about the distribution of the counts in each of the categories can be visualised with violin plots. The violin plot is a combination of boxplot and a kernel density plot to reveal structures within the data (Hintze and Nelson, 1998). The box plot shows four main features of a variable, centre, spread, asymmetry and outlier, which are also included in the violin plot. Similar to box plots, violin plots allow us to compare and visualize the relationship between numerical and categorical variables. In addition, the violin plots have a rotated density distribution on each side, showing the distributional characteristics of batches of data.

### *3.4.5. Extracting the peaks in handset count*

Apart from the peaks, we expect that the handset count data will show a varying trend throughout a day and a week. We investigated therefore methods to extract the peaks. This would remove the variation contributed, for instance, from the people within the range of the base station cell who are not on the train. Two approaches were tested. The first extracts the daily variations to find the extent of the peaks. This was done with a simple moving average (SMA) in R. This method will not give the exact value of the peaks, so a more accurate method will be preferred in later analyses, but for this preliminary study, we show results of the SMA analysis. The second approach to extract the peaks is to calculate the difference in handset count for each collection time as the difference between the current and the previous handset count value, that is, $d_c^i = c_i - c_{i-1}$.

### *3.4.6. Comparison to actual ridership*

The number of travellers from the APC data was compared to the handset counts by calculating the ratio for each passing train. The algorithm was used to connect the collection times with the trains. The ratio was calculated as handset count divided by APC count. The results of this comparison are shown in Section 4.5.

## 4. Results

The following subsections present the results from each analysis in the study. The number of handsets and the timings of the trains passing adjacent to the base station were analysed systematically. Having calculated the approximate times for trains passing at the base station, the trends are plotted, and general analysis based on the correlation of trains passing and counts variation is undertaken. The trends at different times of the day are studied separately. Moreover, the train direction is added, and analysis for trains travelling in different directions is considered. In the statistical analysis, categorization of data is based on the passing of trains and calculation of different statistical values to compare and contrast the results. The availability of the actual passenger count from the train operator served as a mechanism for validation of the ratio between the number of passengers and number of handsets on the adjacent cell sites. The handset counts are indexed in the presented analyses to anonymize the data.

The analysis is divided into a number of steps:

1. graphic inspection of peaks in handset counts in relationship to trains passing the base station,
2. analysis of data resolution, comparing data sets of five- and one-minute collection time intervals,
3. statistical analysis of the output values of the proposed algorithm,
4. extracting the peaks in handset count, analysing the extracted peaks using the proposed algorithm, correlated with trains passing the base station, and
5. comparison and validation with APC data.

### 4.1. Graphic inspection

Fig. 5 illustrates the actual handset counts with five-minute collection time intervals plotted against time on base station B1, located between station U and station V. The time period is one day from six in the morning to midnight. The handset counts are compared to the approximate time when trains are passing the base station, represented by the vertical lines. The handset count data have some distinct peaks that seem to coincide with some of the trains passing the base station.

Fig. 6 shows an example of morning rush hours on a Thursday and afternoon rush hours on a Friday on base station B1 between station U and station V. In this example, the expected pattern is not visible (see assumption 3 and 4 in Section 3.4.4). For this base station, the peaks mostly coincide with the trains travelling towards the city, regardless of time of the day.

### 4.2. Collection time interval—five minutes versus one minute

In Fig. 7, the handset counts with one-minute collection time intervals is compared to the aggregated data set with five-minute intervals. Fig. 7 (a) shows the data set of one-minute collection time intervals on base station B5 before station Y, Cell 2, from the second data set. Fig. 7 (b) shows the aggregated data of the handset count with five-minute intervals for the same data set on the same day and base station. The handset counts with one-minute collection time intervals in Fig. 7 (a) are closely related to the passing of trains. The five-minute aggregated handset counts in Fig. 7 (b) show fewer and less distinct peaks, with lower values, and are less obviously relatable to the trains passing the base station. With one-minute collection time intervals as shown in Fig. 7 (a), the peaks are more frequent and show that a large share of the peaks is in connection with trains travelling away from the city in the afternoon.

### 4.3. Statistical analysis

This section analyses the output of the algorithm presented in Section 3.4.1. The analysis is done on base station B5 that is located before train station Y, which is the station with the most frequent trains passing on the studied railway section. The one-minute collection time interval data set is used for a time period of five days, i.e., Monday-Friday. The output is studied based on the categories described in Tables 4 and 5, which are direction (trains travelling towards or away from the city), time of day (morning or evening rush hour), and whether or not a train passed between the collection times (category Yes or No). The assumption is that the count values will be higher in the collection time intervals when trains are passing the base station (assumption 1 in Section 3.4.4).

Fig. 8 shows violin plots of the categories described in Table 5. The density distribution of handset counts over five days is given on the y-axis. The time of day on the x-axis shows the categories of morning rush hour, evening rush hour, and the rest of the day. Which direction the trains are travelling in, i.e., towards the city or away from the city, is given by the colours as described in the legend. These are handset counts categorised in the 'Yes' category, identified by the proposed algorithm as collection times when a train has passed the base station since the previous collection time. In addition, the 'No' category when no trains had passed the base station since the previous collection time is included and categorised by the time of day. The median values in handset count for each

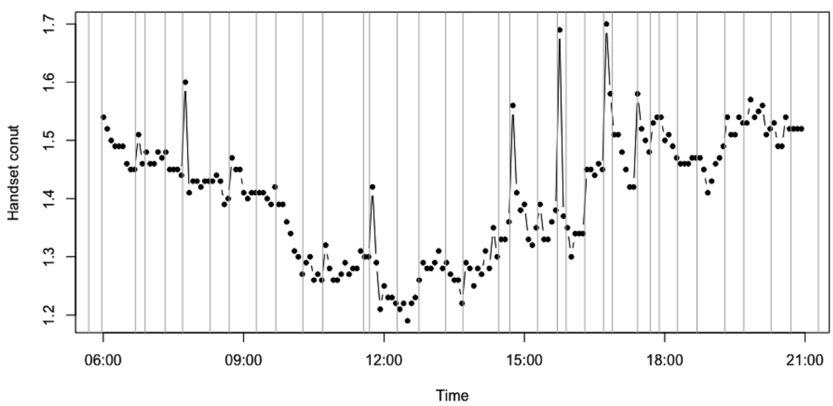


**Fig. 5.** Handset count collected on a Friday on base station B1 located between station U and station V, from the first data set with five-minute collection time intervals. The vertical lines represent the calculated times when trains passed the base station.

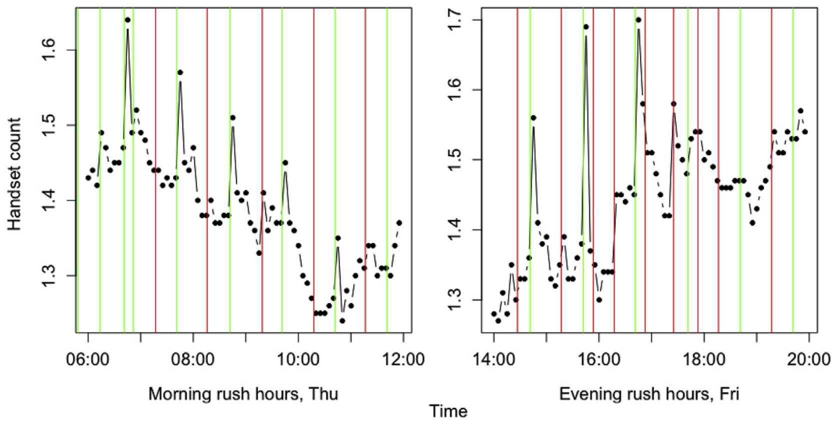


**Fig. 6.** Handset count from the morning and afternoon rush hours on base station B1 located between station U and station V, from the first data set with five-minute collection time intervals.

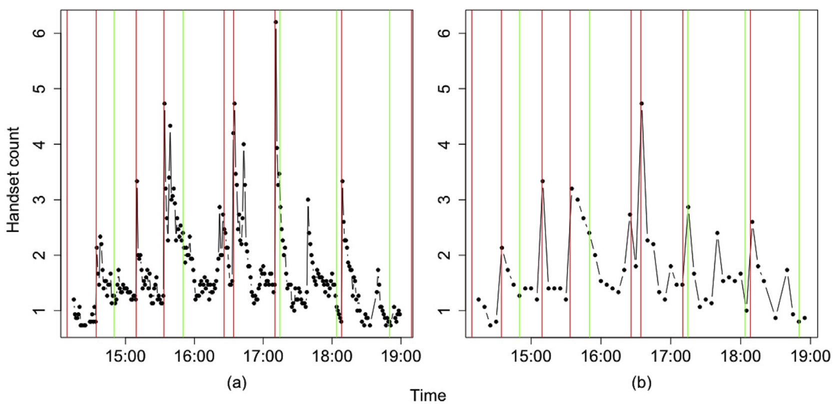


**Fig. 7.** Handset count with one-minute collection time intervals in (a) and the one-minute interval handset count aggregated to five-minute intervals in (b) on base station B5 before station Y, Cell 2, from the second data set. Times when trains are passing are represented by the vertical green lines (towards the city) and red lines (from the city) for Line 1.

of the categories are shown by the white circle. The black lines give the standard box plot for each of the categories. Violin plots with higher or lower values than the box plot whiskers, for instance, as the maximum value in the evening rush hour towards the city, indicate outlier(s).

The evening rush hour in Fig. 8 is the category with the highest median handset counts, compared to the other hours of the day. The direction away from the city in the evening rush hour has a median that is slightly higher than the other subcategories. In the morning rush hour, it is the direction towards the city that has the highest median, but just barely. The density distributions for the evening rush hour show one local maxima for each of the three categories, in which the category with no trains contains the highest collected handset count. On the other hand, the 'No' category shows a distribution with higher density at the lower part of the violin plot. Both the morning rush hour and the rest of the day show bimodal distributions, with two local maxima in each of the density functions. This means that within these categories there are high densities of both low handset counts and high handset counts. The morning rush hour towards the city shows a distribution with higher density at the upper part of the violin plot, i.e., the largest portion of the handset counts in this category has high values. All three categories in the rest of the day have distributions with the highest densities at the lower parts of the violin plots.

Summarising the handset counts over five days, Fig. 9 shows that the average peaks in handset count is higher in the evening than in the morning. It also appears that the peaks are more frequent in the evening rush hour. Fig. 9 also shows that the average daily variation is quite distinct between rush hours and the middle of the day.

### *4.4. Extracting the peaks in handset count*

The handset counts fluctuate throughout a day, with smaller fluctuations on the weekends, as shown in Fig. 9 and the count part

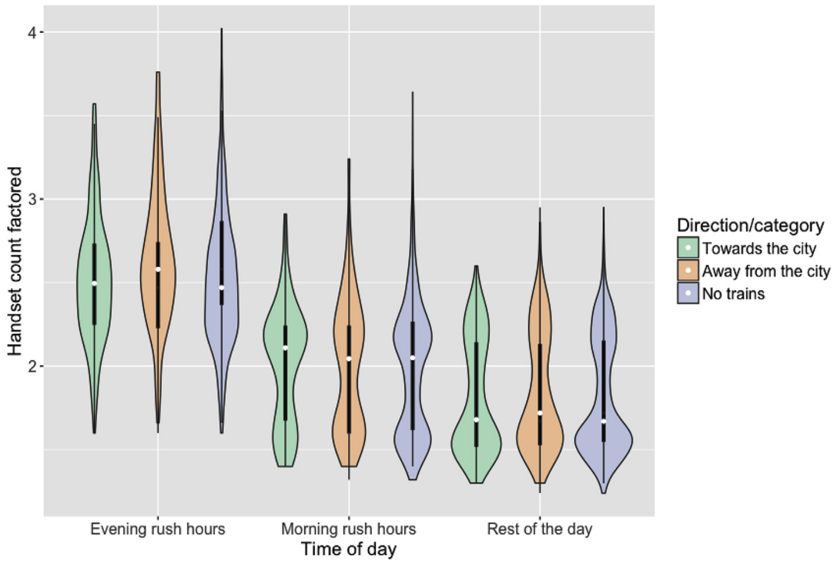


**Fig. 8.** Violin plots showing the density distribution of handset counts over five days (y-axis), with median value (white circle) of the categories of (see Table 5) direction (x-axis) and rush hour (colour in legend) when trains had passed within the collection time interval prior to the collection time and time of day for category 'No' when no trains had passed since the previous collection time (see Table 4). The handset counts, which are represented in the figure as factors, are from the third data set collected at base station B5, with one-minute collection time intervals. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

of Fig. 10. Since there is a variation through the day, we investigated methods to extract the peaks, presented in Section 3.4.5. The approaches are analysed with the handset counts collected on base station B1 for the first time period of collecting data, with five-minute collection time intervals. The result of the simple moving average is given in Fig. 10, showing the handset counts, the trend line and the irregular values, which are simply handset counts minus trend values. The irregular values show jumps in the counts, presumably when trains are passing the base station.

The second method of extracting the peaks as the difference in handset count, $d_c^i = c_i - c_{i-1}$, is shown in Fig. 11. The difference in count, $d_c^i$, is the same value used to calculate the percentage increase in handset count $I$ in the algorithm in Section 3.4.1. The percentage increase is analysed with violin plots in Fig. 12 for data collected at base station B5 from the third data set with one-minute collection time intervals for a time period of five days, i.e., Monday-Friday. The violin plots in Fig. 12 show the density distribution of the percentage increase on the y-axis for each of the categories described in Table 5. The time of day on the x-axis shows the categories of morning rush hour, evening rush hour, and the rest of the day. Which direction the trains are travelling in, i.e., towards the city or away from the city, is given by the colours as described in the legend. These are percentage increases

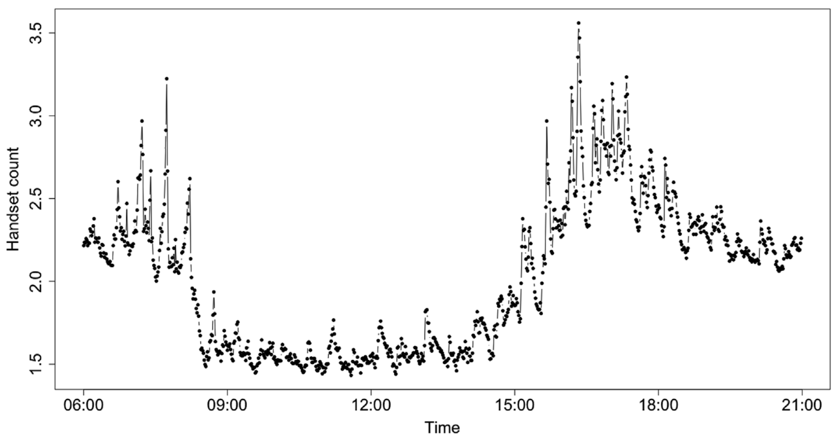


**Fig. 9.** The sum of handset counts over five days (Monday-Friday) on base station B5 located before station Y from the third data set with one-minute collection time intervals.

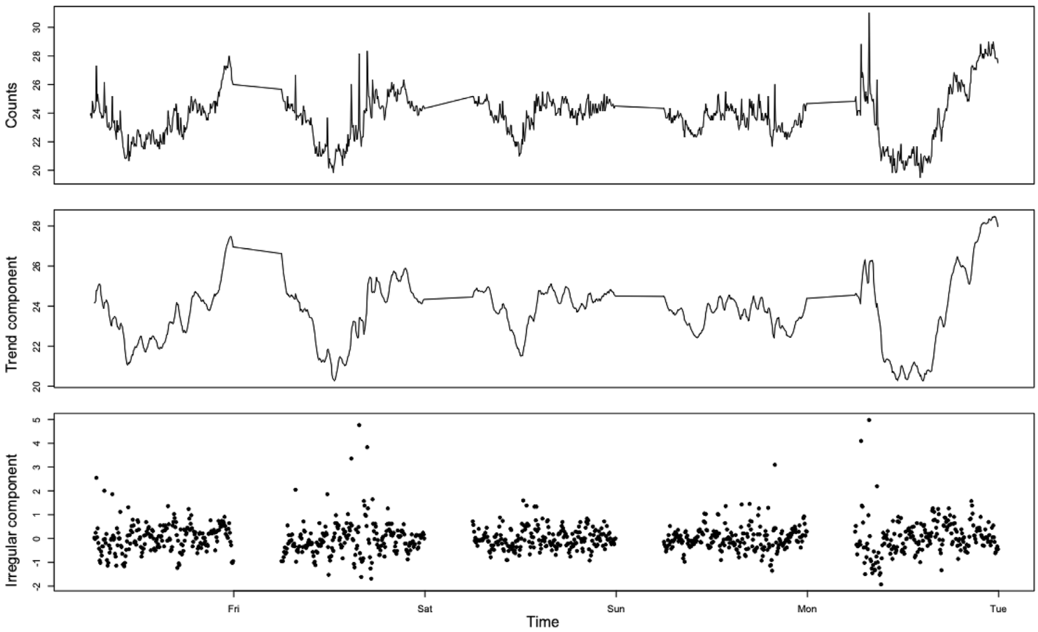


**Fig. 10.** Counts, trend component and irregular component at base station B1 between station U and station V for six days in 2016, with five minutes as collection time intervals.

categorised in the 'Yes' category, identified by the proposed algorithm as collection times when a train has passed the base station since the previous collection time. In addition, the 'No' category when no trains have passed the base station since the previous collection time is included, categorised by the time of day. The median values in percentage increase for each of the categories are shown by the white circles. The black lines give the standard box plot for each of the categories.

The percentage increases categorised as direction away from the city in Fig. 12 have highest median values for all times of the day. The other categories have medians that are negative or close to zero. Most of the violin plots have a density distribution around zero with one local maxima, except the evening rush hour towards the city, which has a large portion of negative values. However, this category also has the highest registered percentage increase. The category when no trains passed shows density distributions around zero with both positive and negative values and a few high percentage increases.

### 4.5. Comparison to actual ridership

This section presents the comparison and validation of the handset counts with APC data. The passenger counts from the APC data are plotted in Fig. 13 for two of the train lines as the trains passes base station B5, towards the city of Line 1 shown by the green line and away from the city of Line 1 and Line 2 shown by the red lines. The y-axis gives the APC count, however the values are indexed to anonymize the data. Fig. 13 shows a clear daily variation in which most passengers are travelling towards the city in the morning and away from the city in the evening, with the exception of a few trains away from the city in the morning on Monday and Thursday, and a few trains towards the city on Thursday evening. Fig. 14 show the handset counts on base station B5 from the third data set with one-minute collection time intervals. Line 1 and Line 2 are shown in Fig. 14 by the vertical lines of same colours as in Fig. 13. The handset counts show a daily variation in Fig. 14 similar to Fig. 9 with distinct variation between the morning, the middle of the day and the evening. Fig. 14 also show that the frequency of peaks in handset counts are higher in the evening rush hours compared to the morning rush hours. This result matches the observation from the violin plot of the percentage increase, in which there were most positive increases for trains travelling away from the city.

The ratios of the number of handsets to the number of travellers (handset count/APC count) are given in Fig. 15 for base station B5 from the third data set with one-minute collection time intervals. Fig. 15 gives the calculated ratio for trains travelling towards the city and away from the city for five days. A smooth curve was fitted by loess regression as shown by the black line, with confidence bands. For the trains travelling towards the city (see Fig. 15 (b)) the ratio centres around one from 6 a.m. in the morning to 3 p.m. in the afternoon. From 3 p.m. when the evening rush hours commence the ratio increases towards around three and four. Ratio above 1 means that the handset counts are higher than the automatic passenger counts. For the trains travelling away from the city (see Fig. 15 (a)) the ratio has a close to inverted shape, with a decrease from around four in the morning from 6 a.m. to 10 a.m. during the morning rush hours. From 10 a.m. the ratio has a small curve centred around one, and with a slight increase towards the evening. The ratio for trains travelling towards the city on Thursday and Friday in Fig. 15 (b) has three higher ratio-values on, respectively, one and

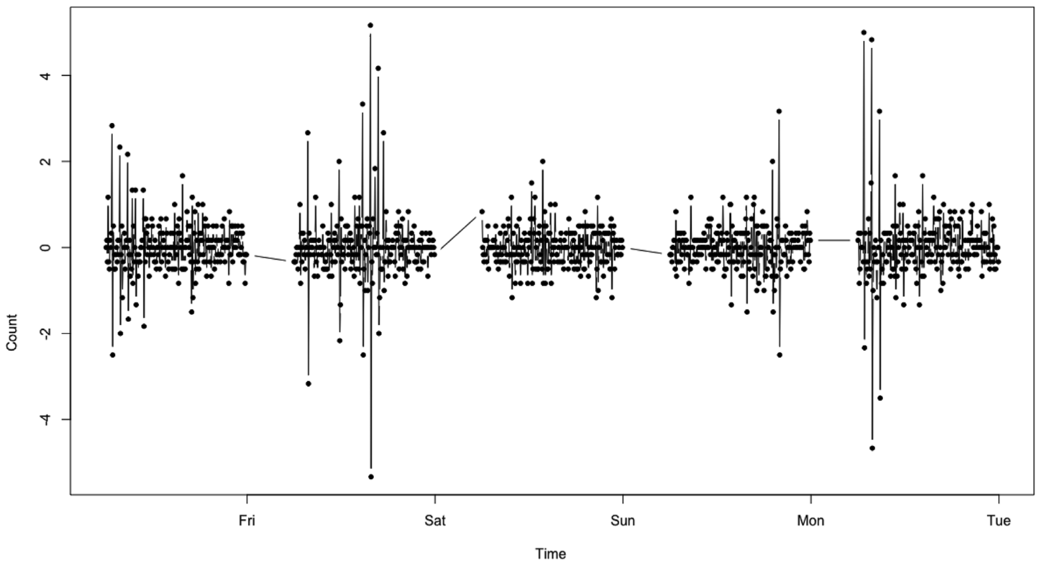


**Fig. 11.** Difference in count, $d_c^i = c_i - c_{i-1}$, for the basestation located between station U and station V for six days in 2016, with five minutes as collection time intervals.

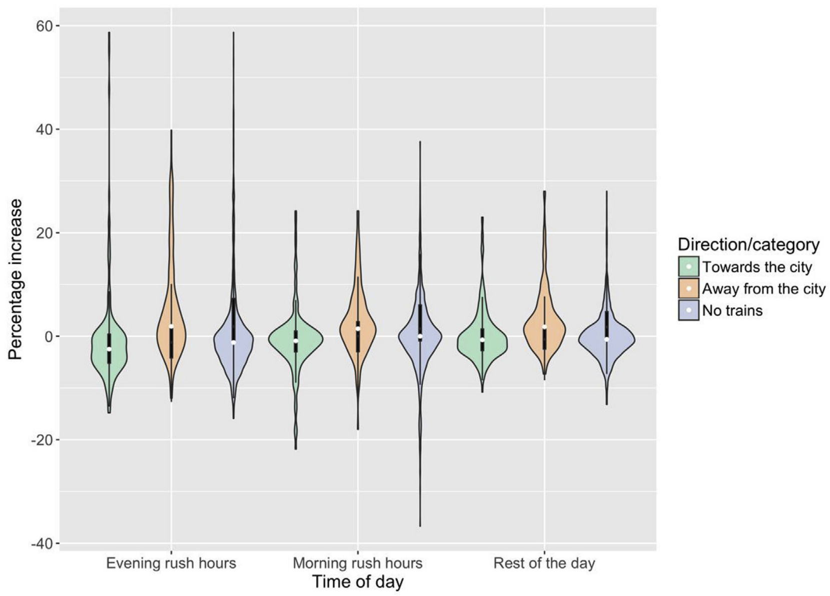


**Fig. 12.** Violin plots showing the density distribution of percentage increase I over five days (y-axis), with median (white dot) of the categories (see Table 5) direction (x-axis) and time of day (colours in legend) when trains had passed within the collection time interval prior to the collection time and time of day for category 'No' (see Table 4). The percentage increases $I$ in handset counts are from the third data set, collected at base station B5 with one-minute collection time intervals. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

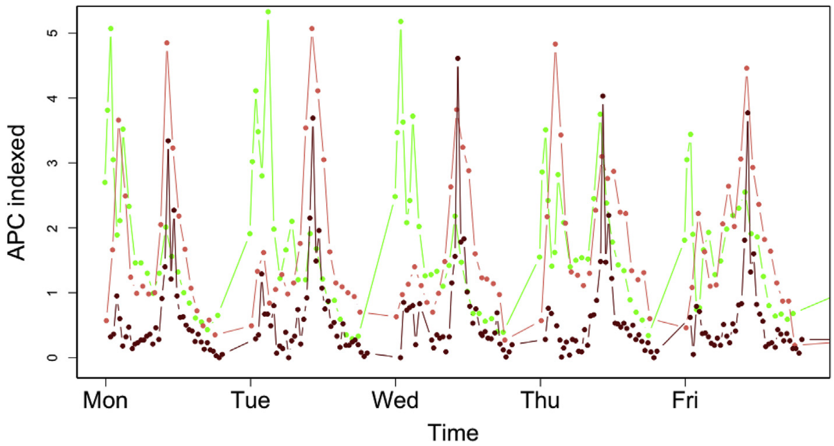


**Fig. 13.** Automatic passenger count, represented as indexed values on the y-axis, for two different train lines as the trains pass base station B5 located before station Y, where green line is Line 1 direction towards the city, light red line is Line 1 away from the city and dark red is Line 2 away from the city. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

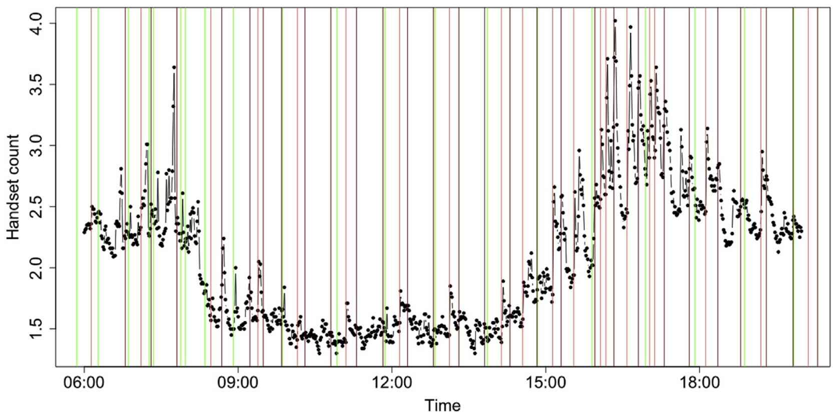


**Fig. 14.** Handset counts from the third data set with one-minute collection time intervals collected on a Thursday at base station B5 located before station Y. The vertical lines represent the same train lines as in Fig. 13, where green lines are Line 1 direction towards the city, light red lines are Line 1 direction away from the city, and dark red lines are Line 2 direction away from the city. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

two of the morning trains. For trains travelling away from the city in Fig. 15 (a), four of the morning trains on Thursday has lower ratio-values than the other days.

## 5. Discussion

This section discusses some of the results presented in Section 4.

### 5.1. Graphic inspection and collection time interval

The expected relationship between the rush hours and the direction (see assumption 3 and 4 in Section 3.4.4) was not confirmed in Fig. 6. However, in Fig. 7 (a), we noticed that for base station B5 with one-minute collection time intervals, a large share of the peaks is in connection with trains travelling away from the city in the afternoon rush hours, which is what we were expecting. One-minute collection time intervals appear as more favourable than the intervals of five minutes both in capturing the frequency in peaks and peaks that are connectable to trains passing the base station, as seen in Figs. 5–7.

Station Y has higher train frequency than the previously studied stations, i.e., station U and station V, as seen in Fig. 7. Several of the peaks in the handset counts corresponds nicely with trains that pass the base station. Some of the peaks have collection times between two different train passing times. In these cases, the train passing related to the peak is most likely the time prior to the peak,

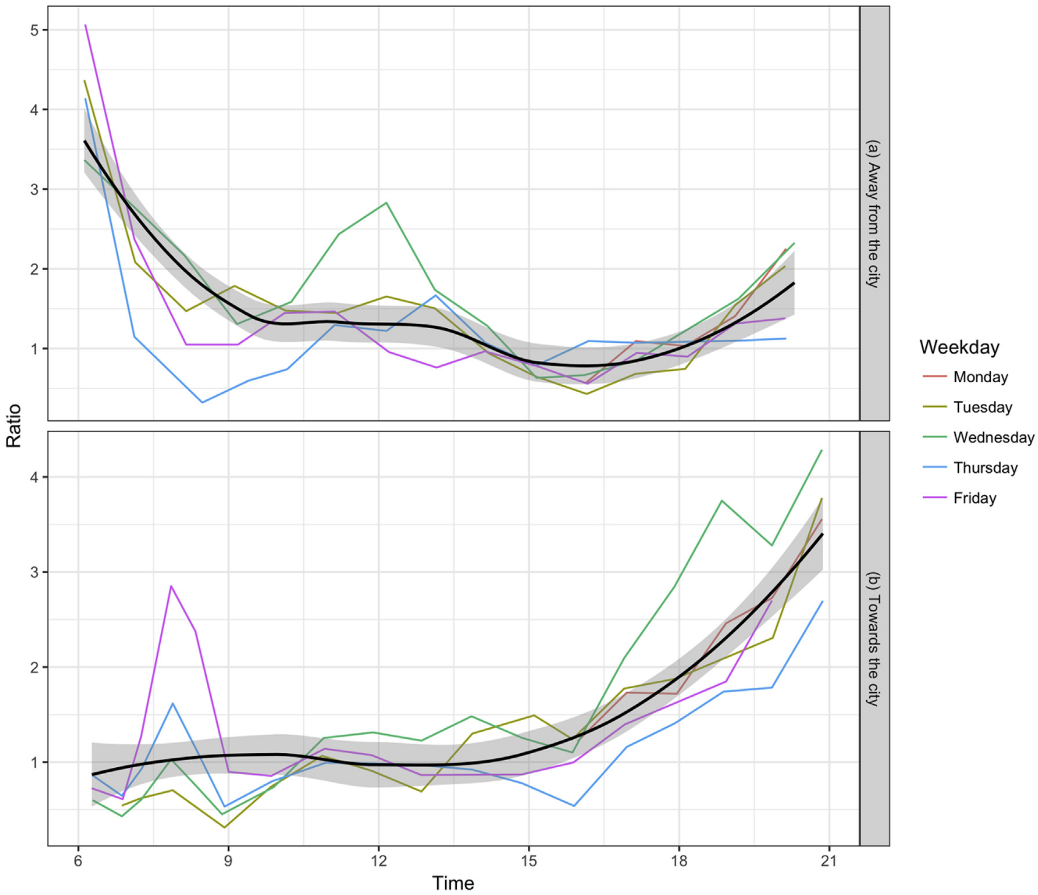


**Fig. 15.** Ratio of handset counts to the number of passengers (APC) for trains going towards the city (b) and away from the city (a) on Line 1 for five days at base station B5 located before station Y. The black lines represent a smooth curve fitted by loess regression with confidence bands.

because the mobiles are registered on the base station until they are connected to a new base station, so the mobiles are most likely registered on the selected base station also for a small amount of time after the train has passed. This is the logic behind the suggested algorithm. However, the question is if this also applies when the collection time is closer to the time of the train that passes right after the collection time. When the trains pass the base station with a high frequency and the time interval between trains are small some of the peaks seem to be placed in such a way that we cannot be certain which peak corresponds with which train.

### 5.2. Statistical analysis

The violin plots were used in Section 4.3 to analyse the output of the algorithm. The median value was slightly higher for the 'Yes' categories (see Fig. 8). For the morning rush hours, the median value was highest for trains travelling towards the city, supporting the assumption made in Section 3.4.4 (assumption 3). For the evening rush hours, the median value was highest for trains travelling away from the city, in agreement with the assumption (assumption 4). Fig. 8 showed bimodal distributions for both 'Yes' and 'No' categories with two local maxima in each of the density functions (i.e., for the categories morning rush hours and rest of the day). The low count values in the 'Yes' category was expected because the graphic inspection has shown that there are not peaks every time a train passes (see for instance Figs. 7 and 14). What is more concerning is that the 'No' category has so many high values. Fig. 9 showed an average daily variation with distinct variation between the morning, the middle of the day and the evening. Thus, this may be a reason why the 'No' category in Fig. 8 has a high density of handset counts in both high and low count values. Another explanation could be that some peaks are caused by trains that are about to pass, but are a few seconds before the collection time. Hence, the high value in handset count would be categorised as a 'No' category by the algorithm.

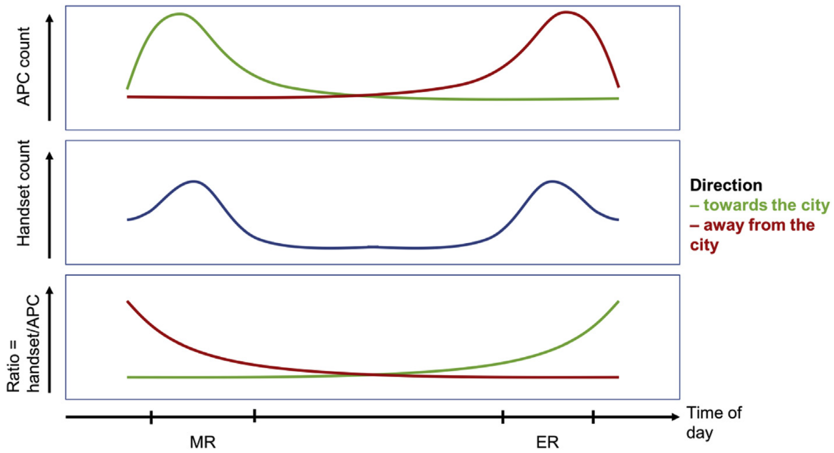


**Fig. 16.** Illustration of the logical connection between the handset counts, the APC data and the ratio. The drawn curves are the supposed daily variation in APC count, handset count and ratio, respectively. The assumed daily variations are based on the results presented in Fig. 9 and 14 (handset count), Fig. 13 (APC), and Fig. 15 (ratio handset count/APC).

### 5.3. Extracting the peaks in handset count

The violin plots of the percentage increase (Fig. 12) show that the 'No' category has values with positive increases, and also increases of more than 10%. These peaks are probably deviations that are not caught by the algorithm, as discussed above.

### 5.4. Comparison to actual ridership

The shapes of the ratio curves in Fig. 15 was expected because of the daily variation of the handset counts and the APC data. This is illustrated in Fig. 16, showing the assumed daily variations based on the previously presented results of handset counts, APC counts and ratio. The daily variation in the handset count (see Figs. 9 and 14) show low values in the middle of the day and higher values before and during the morning rush hours and during and after the evening rush hours. First inspecting the trains travelling towards the city, we see that the ratio increases in the evening (see Fig. 15 (b)) because the handset counts increases towards the evening (Fig. 14) while the APC data remain low (Fig. 13). Likewise, for the trains travelling away from the city we see that the ratio starts at a high value and decreases because the handset counts have high values in the morning and then decrease (Fig. 14) while the APC data is low from the morning until the evening rush hours commence (Fig. 13).

For trains travelling towards the city on Thursday and Friday, the higher ratio-value on three of the morning trains (see Fig. 15) are possibly because the APC data show fewer passengers on these days in the morning compared to the other days (see Fig. 13). In the same way, for trains travelling away from the city on Thursday, the lower ratio-values on four of the morning trains (see Fig. 15) are probably because the APC data show more passengers on this day in the morning compared to the other days (Fig. 13).

## 6. Concluding discussion

This study has investigated the potential for using mobile phone data to describe travel patterns that include train travel. We have tested the use of mobile phone data to measure train ridership. We find that there is a connection between the train passing and changes in the handset counts. Although the variation is different for different base stations, there is a significant positive increase when trains pass.

We have shown that it is possible to combine mobile data with railway infrastructure and train traffic data. These preliminary results show that there is a connection between the train passing and changes in handset counts. However, it is also evident that a lot of the trains passing the base station are not detectible in the handset counts. Furthermore, some peaks seem to occur even though there were no trains passing the base station. We also showed that one-minute collection time intervals are needed, especially on the railway sections that have high frequency of trains passing.

The main implication of the findings is that mobile phone data can potentially be used for ridership analyses.

### 6.1. Combining mobile phone data with railway infrastructure and train traffic data (RQ1)

This study show that it is possible to combine mobile phone data with railway infrastructure and train traffic data, thus answering the first research question. We successfully combined data on rail traffic, timetables and infrastructure with handset counts from mobile phone base stations. Provided that the mobile phone base stations are chosen carefully and the collection time intervals for handset counts are suitable, the handset counts correspond well with the passing of trains. We tested collection time intervals of five minutes and one minute. Especially for high frequency parts of the railway network, a one-minute collection time is needed. Even for lines with less traffic, a one-minute data set appears to have advantages over a five-minute data set.

However, the findings also show that there is potential to improve the method of connecting the collection times with the times of trains passing the base station. As Fig. 8 showed, the distribution of handset counts that are not connected to a train passing shows a larger portion of high values than we would prefer. This could indicate that the algorithm is not able to connect all the peaks in handset counts to trains passing. For instance, the base station has a certain range, so considering the time frames in which the trains are passing within the range of the base station and how early or late in that time frame, it is likely that the mobile phones are connected to the base station in question. And we could consider the extension of the trains from the first to the last carriage, which can vary.

To consider the range of the base station, a time margin can be included in the method to connect collection times with when trains pass the base station, which can be done in two ways. Either time can be added on the *train_time*, some before and some after the calculated point in time when the train is passing the base station, to include when one thinks the train is within range of the base station. Depending on how large this time interval is, a result can be that the train will be connected to more than one collection time. This does not necessarily have to be a disadvantage. Alternatively, one can add a time interval in the collection time, for instance, by saying that if the train's time falls within the interval created as a half minute before and a half minute after the collection time, then the train is connected to the collection time. A third way could be to both add a time interval around the collection time and a time interval around the train time. In this approach, the result would most likely be that the trains will be connected to more than one collection time.

### 6.2. Suitable formats for presentation (RQ2)

One research question was what suitable formats are for presenting and analysing train ridership based on mobile data. We have tested and described different formats for analysing and presenting train ridership based on handset counts. In particular, we have used both absolute numbers of handsets and changes in the handset counts. We have also tried different approaches for estimating the probability that a count peak is related to the passing of a train. Although it is possible to draw conclusions and validate assumptions from graphic analysis, there is a need to develop a uniform mechanism for categorizing the available data and assigning as well as calculating variables for validating the assumptions. Hence, an algorithm was developed yielding the required output variables so that tangible results in relationship to the hypothesis can be extracted for the categories both in preliminary and time and direction analysis.

This section discusses the strengths and weaknesses of each of the formats used in this paper for presenting and analysing the results.

#### 6.2.1. Graphic inspection and collection time interval

Graphic inspection was used in Section 4.1 to visualize the handset count and compare the collection times to the calculated approximate times of when trains are passing the base station. The strength of graphic inspection is that it is good to visualize how the collected handsets and trains are connected. The weakness of graphic inspection is that it is less suitable to comparing several days or base stations.

#### 6.2.2. Violin plot

The strengths of the violin plot are that it allows us to compare and visualize the relationship between numerical and categorical variables, as well as showing the distributional characteristics of the data. The violin plots would be suitable for comparing distributions of handset counts (and percentage increases) in a study to compare several methods, for instance, to test different methods of calculating the times the trains are passing the base station or for testing other methods of connecting train times with collection times, as discussed above. What is desired is that the category that contains collection times when no trains passed has no high peaks in handset counts. As Fig. 9 showed, it is an advantage to separate the handset counts in rush hours and the rest of the day with no rush hour. As the APC data show, more people are travelling towards the city in the morning and away from the city in the evening than during the rest of the day. Thus, even though the violin plots do not show a clear difference between these categories, it will be favourable to make this separation between directions in further testing.

#### 6.2.3. Extracting the peaks

As learned from the results on collection time intervals, one-minute collection intervals give better precision in detecting peaks that occur in close time proximity to trains passing the base station. The five-minute collection time interval was less accurate. Extracting peaks on a data set collected at one-minute collection time intervals could be more useful than with five-minute intervals as we saw from the results in Section 4.4.

The violin plots of the percentage increase on the data set with one-minute collection time interval showed that both high and low values were connected through the algorithm to trains passing the base station. And plots of the handset count like Fig. 14 show that handset count seems to increase in stages, or at least that as a train passes the handset count increase gradually in more than one collection time. The difference in handset count may therefore be questioned as a good approach to extract the peaks. A method similar to SMA could perform better when the collection time intervals are less than five minutes.

#### *6.2.4. Comparison to APC ratio*

A useful way to utilise the handset count data would be to find an average ratio between the handset count and the automatic passenger count. The comparison of the actual passenger counts with the number of handsets needs to take specific direction and timeframe into consideration to gauge the exact number of travellers on the passing train. Thus, to calculate the actual number of passengers on the train, the ratio of numbers of local people connected to the base station is one of the factors that needs to be taken into account. The ratio is almost constant during specific times. Further testing of how to connect the train times to the collection times will also affect the ratio between the APC data and handset count. A good connection between the train times and collection times will improve the results of the ratio and will facilitate the ratio serving as a practical validation indicator to observe the effect of passengers on the train with the number of counts.

### *6.3. Measuring number of passengers using mobile phone data (RQ3)*

The last research question was to what extent the format of available mobile phone data is suitable for measuring the number of mobile units passing close to the railway line. The mobile phone data available for this study was handset counts. The handset counts are compared to the number of travellers as measured by on-board APC equipment. The results indicate that handset counts and changes in handset counts can give an indication of the number of travellers on a passing train. The results are, however, not as conclusive as for train detection. The ratio of handsets over passengers varies. It is likely that a large-scale calibration is needed, using more data than we had available, to increase the accuracy of handset counts as indicators of the number of travellers.

To find a good approximation of the number of passengers on the train, we should consider factors that can affect the precision of the measurement. There are some issues that may introduce bias into the handset counts. For instance, there are most likely some people without phones on the trains, or possibly business people with more than one phone. Another issue is that these data are only from one telecom operator, and the number of travelling passengers can be unpredictably divided among different mobile operators. Other factors that can affect the precision of the measurement include,

- the daily variation (morning-afternoon-evening);
- the need to gather data for longer time periods. For evaluation purposes, there is a need for data covering a relatively long period, typically a few years. We recommend that data are stored with the highest possible resolution and that how the data are collected and processed is clearly described. In addition, it would help to find better average values, maybe over years, for monthly variations. Mobile data for longer time durations will account for seasonal effects and smooth the effect of irregular fluctuations in the data and can provide a better bias for developing more precise mathematical model;
- if available, base stations located in railway tunnels would be preferred; and
- taking into account the normal variations of mobile data and specific factors and conditions associated with the specific cell sites can provide a more linear relationship with the ridership and closeness of cell sites to the train station (indoor cell sites for the stations or located exactly at the train station). Variations between the base stations include different areas, different prerequisites for the base stations to pick up phone signals, or numbers of people passing or being in the area.

Other uncertainties with this approach of measuring train ridership are that we cannot say for sure that the time stamps for the collection times and the train times are taken by synchronized watches. There may be a small time-lag that we do not know about.

When discussing the suitability of the available mobile phone data it is also reasonable to discuss whether it is worthwhile to invest and to apply mobile phone data for ridership analysis. If universal and scalable methods are developed, then the costs should be significantly less than with methods used now.The data is already there, so a lot of the expenses lies in developing systems and testing good algorithms.

Table 6 summarises some pros and cons of handset counts and mobile phone data.

**Table 6**
Pros and cons of handset counts and mobile phone data.

| Pros | Cons |
|---|---|
| Need no direct access to train or railway network – the source of data and data owner is the mobile network operators | Little experience |
| Preserves personal privacy (for handset counts) | Personal privacy (not for handset counts, but an issue for other mobile phone data) |
| Data is available | Separate logs for each mobile network (this study is with mobile phone data from one network) |
| The data gives units connected to base station, (and potentially movement in network) | The ultimate destination is not provided – at this point (which is true for the data source utilised in this paper, however other mobile phone data may provide destination information) |
| The user role in data collection previously achieved by the survey process is minimized | The data cannot provide information on trip purpose or on user assessment of service |
| Improved data quality and increased amount of statistics available | Costs regarding developing systems, testing good algorithms and safeguarding personal privacy. |

### 6.4. Further research

The statistical and graphic analyses and the evaluation of hypothesis developed provide a clear picture of different aspects that can be analysed and looked into combining mobile data and train data. The important fact is that this is only one format of data that has been looked into in detail. The format provides a good starting point for looking into the concept of utilizing mobile data in ridership evaluation and that longer data duration and higher resolution can point out the exact quantifiable figures for the number of travellers. Other formats of anonymous data formats also need to be explored separately to determine if they can be combined with the format used in this report. The availability of longer time interval data and other formats can provide the opportunity of developing systemic models and frameworks for analysing the ridership from different angles. Hence, mobile data can be declared a viable source for calculating the number of travellers on trains.

Further steps needed to make mobile phone data usable tools for measuring ridership on trains:

- Calibrate the ratio of handset counts to the number of passengers using data sets for longer time periods
- Test other methods to determine if a train is connected to the base station at the moment of a collection time.

The easiest approach to measure the number of passengers would be to get direct access to APC and other ridership data. However, mobile phone data opens up a wider set of options for analysis, provided that personal privacy issues are managed and the telecom operators can develop business models for supplying such data.

## Acknowledgements

We thank the Concept Research Program at NTNU for support and constructive feedback on previous versions of parts of this text. We would also like to thank Ida Kristoffersson at the Swedish National Road and Transport Research Institute (VTI) for valuable feedback.

## References

Aguiléra, V., Allio, S., Benezech, V., Combes, F., Milion, C., 2014. Using cell phone data to measure quality of service and passenger flows of Paris transit system. Transport. Res. C Emerg. Technol. 43, 198–211.
Alexander, L., Jiang, S., Murga, M., González, M.C., 2015. Origin-destination trips by purpose and time of day inferred from mobile phone data. Transport. Res. C Emerg. Technol. 58 (Part B), 240–250.
Assemi, B., Safi, H., Mesbah, M., Ferreira, L., 2016. Developing and validating a statistical model for travel mode identification on smartphones. IEEE Trans. Intell. Transport. Syst. 17 (7).
BaneNor, 2017. Network Statement. BaneNor. http://www.banenor.no/en/startpage1/Market1/Network-Statement/, Accessed date: 6 December 2017.
Becker, R., Cáceres, R., Hanson, K., Isaacman, S., Loh, J.M., Martonosi, M., Rowland, J., Urbanek, S., Varshavsky, A., Volinsky, C., 2013. Human mobility characterization from cellular network data. Commun. ACM 56, 74–82.
Bianchi, F.M., Rizzi, A., Sadeghian, A., Moiso, C., 2016. Identifying user habits through data mining on call data records. Eng. Appl. Artif. Intell. 54, 49–61.
Boyle, D.B., 1998. Passenger Counting Technologies and Procedures. TCRP Synthesis of Transit Practice 29. Transportation Research Board, Washington, DC.
Calabrese, F., Diao, M., Di Lorenzo, G., Ferreira, J., Ratti, C., 2013. Understanding individual mobility patterns from urban sensing data: a mobile phone trace example. Transport. Res. Part C 26, 301–313.
Calabrese, F., Di Lorenzo, G., Liu, L., Ratti, C., 2011. Estimating origin-destination flows using mobile phone location data. IEEE Pervasive Computing 10 (4), 36–44. https://doi.org/10.1109/mprv.2011.41.
Chaudhary, M., Bansal, A., Bansal, D., Raman, B., Ramakrishnan, K.K., Aggarwal, N., 2016. Finding occupancy in buses using crowdsourced data from smartphones. In: Proceedings of the 17th International Conference on Distributed Computing and Networking, vol. 35 ACM. http://dx.doi.org/10.1145/2833312.2833460.
Chu, K.A., Chapleau, R., 2008. Enriching archived smart card transaction data for transit demand modeling. Transport. Res. Rec.: J. Transport Res. Board 2063, 63–72. http://dx.doi.org/10.3141/2063-08.
De Regt, K., Cats, O., Van Oort, N., Van Lint, H., 2017. Investigating potential transit ridership by fusing smartcard data and global system for mobile communications data. Transport. Res. Rec. 2652, 50–58.
Doi, M., Allen, W.B., 1986. A time series analysis of monthly ridership for an urban rail rapid transit line. Transportation 13 (3), 257–269. https://doi.org/10.1007/BF00148619.
Doyle, J., Hung, P., Kelly, D., McLoone, S., Farrell, R., 2011. Utilising Mobile Phone Billing Records for Travel Mode Discovery. Trinity College Dublin ISSC 2011, June 23-24 2011.
Drageide, V., 2009. Towards Privacy Management of Information Systems. MS thesis. The University of Bergen.
Fowkes, A.S., Nash, C.A., Whitening, A.E., 1985. Understanding trends in inter-city rail traffic in Great Britain. Transport. Plann. Technol. 10, 65–80.
Frias-Martinez, V., Soguero, C., Frias-Martinez, E., 2012. Estimation of urban commuting patterns using cellphone network data. In: Proceedings of the ACM SIGKDD International Workshop on Urban Computing, pp. 9–16.
Gordillo, F., 2006. The Value of Automated Fare Collection Data for Transit Planning: an Example of Rail Transit Od Matrix Estimation. Thesis, S. M.. Massachusetts Institute of Technology. http://hdl.handle.net/1721.1/38570.
Gundlegård, D., Rydergren, C., Breyer, N., Rajna, B., 2016. Travel demand estimation and network assignment based on cellular network data. Comput. Commun. 95, 29–42.
Higuchi, T., Yamaguchi, H., Higashino, T., 2015. Mobile devices as an infrastructure: a survey of opportunistic sensing technology. J. Inf. Process. 23 (2), 94–104.
Hintze, J.L., Nelson, R.D., 1998. Violin plots: a box plot-density trace synergism. Am. Statistician 52, 181–184.
Holleczek, T., Yu, L., Lee, J.K., Senn, O., Ratti, C., Jaillet, P., 2014. Detecting weak public transport connections from cellphone and public transport data. In: Proceedings of the 2014 International Conference on Big Data Science and Computing, 2014. ACM, pp. 9.
Iqbal, M.S., Choudhury, C.F., Wang, P., González, M.C., 2014. Develoment of origin-destination matrices using mobile phone call data. Transport. Res. C Emerg. Technol. 40, 63–74.
Järv, O., Ahas, R., Saluveer, E., Derudder, B., Witlox, F., 2012. Mobile phones in a traffic flow: a geographical perspective to evening rush hour traffic analysis using call detail records. PLoS One 7 (11), e49171. http://dx.doi.org/10.1371/journal.pone.0049171.
Jiang, S., Yang, Y., Fiore Jr., G., J, F., Frazzoli, E., González, M., 2013. A review of urban computing for mobile phone traces: current methods, challenges and opportunities. In: Proceedings of the ACM SIGKDD International Workshop on Urban Computing.
Kujala, R., Aledavood, T., Saramäki, J., 2016. Estimation and monitoring of city-to-city travel times using call detail records. EPJ Data Science 5 (1). http://dx.doi.org/

10.1140/epjds/s13688-016-0067-3.
Larijani, A.N., Olteanu-Raimond, A.-M., Perret, J., Brédif, M., Ziemlicki, C., 2015. Investigating the mobile phone data to estimate the origin destination flow and analysis; case study: Paris region. Transport. Res. Procedia 6, 64–78.
Leo, Y., Busson, A., Sarraute, C., Fleury, E., 2016. Call detail records to characterize usages and mobility events of phone users. Comput. Commun. 95, 43–53. http://dx.doi.org/10.1016/j.comcom.2016.05.003.
Li, J.P., 2000. Train station passenger flow study. In: Proceedings of the 2000 Winter Simulation Conference.
Mac Donald, V.H., 1979. Advanced mobile phone service: the cellular concept. Bell System Technical Journal 58 (1), 15–41. http://dx.doi.org/10.1002/j.1538-7305.1979.tb02209.x.
Nielsen, B.F., Frølich, L., Nielsen, O.A., Filges, D., 2014. Estimating passenger numbers in trains using existing weighing capabilities. Transportmetrica: Transport. Sci. 10 (6), 502–517.
Olsson, N.O., Bull-Berg, H., 2015. Use of big data in project evaluations. Int. J. Manag. Proj. Bus. 8 (3), 491–512.
Pelletier, M.-P., Trépanier, M., Morency, C., 2011. Smart card data use in public transit: a literature review. Transport. Res. C Emerg. Technol. 19 (4), 557–568. http://dx.doi.org/10.1016/j.trc.2010.12.003.
Schoeni, R.F., Stafford, F., Mcgonagle, K.A., Andreski, P., 2013. Response rates in National panel surveys. Ann. Am. Acad. Polit. Soc. Sci. 645 (1), 60–87. http://dx.doi.org/10.1177/0002716212456363.
Sørensen, A.Ø., Olsson, N., Akhtar, M., Bull-Berg, H., 2017. Approaches, Technologies and Importance of Analysis of Number of Train Travellers. (Manuscript submitted for publication).
Teknisk Ukeblad, 2014. Matematiske Verktøy Skal Gi Kortere Stop På Togstasjonene, January 29 2014. http://www.tu.no/artikler/matematisk-verktoy-skal-gi-kortere-stopp-pa-togstasjonene/224974.
Toole, J.L., Colak, S., Sturt, B., Alexander, L.P., Evsukoff, A., González, M.C., 2015. The path most traveled: travel demand estimation using big data resources. Transport. Res. C Emerg. Technol. 58, 162–177.
Vigren, A., 2017. Competition in Public Transport. Essays on Competitive Tendering and Open-access Competition in Sweden. Doctoral Thesis in Transport Science. KTH Royal Institute of Technology, Stockholm, Sweden.
Vuchic, 2005. Urban Transit. Operations, Planning and Economics. John Wiley and Sons, Hoboken, NJ.
Wang, W., Attanucci, J.P., Wilson, N.H.M., 2011. Bus passenger origin-destination estimation and related analyses using automated data collection systems. J. Public Transp. 14 (4), 131–150.
Xu, Y., Shaw, S.-L., Fang, Z., Yin, L., 2016. Estimating potential demand of bicycle trips from mobile phone data – an anchor-point based approach. ISPRS Int. J. Geo-Inf. 5 (8), 131. http://dx.doi.org/10.3390/ijgi5080131.
Zhao, J., Rahbee, A., Wilson, N.H.M., 2007. Estimating a rail passenger trip origin-destination matrix using automatic data collection systems. Comput. Aided Civ. Infrastruct. Eng. 22 (5), 376–387.
Zhao, Z., Shaw, S.L., Xu, Y., Lu, F., Chen, J., Yin, L., 2016. Understanding the bias of call detail records in human mobility research. Int. J. Geogr. Inf. Sci. 30 (9), 1738–1762.

# Publication III

Sørensen, A. Ø., Olsson, N. O. E., Akhtar, M. M. & Bull-Berg, H. (2019). Approaches, technologies and importance of analysis of number of travellers. *Urban, Planning and Transport Research*, 7(1), 1-18. doi:10.1080/21650020.2019.1566022.

Urban, Planning and Transport Research
An Open Access Journal

ISSN: (Print) 2165-0020 (Online) Journal homepage: https://www.tandfonline.com/loi/rupt20

# Approaches, technologies and importance of analysis of the number of train travellers

Anette Ø. Sørensen, Nils O. E. Olsson, Muhammad Mohsin Akhtar & Heidi Bull-Berg

**To cite this article:** Anette Ø. Sørensen, Nils O. E. Olsson, Muhammad Mohsin Akhtar & Heidi Bull-Berg (2019) Approaches, technologies and importance of analysis of the number of train travellers, Urban, Planning and Transport Research, 7:1, 1-18, DOI: 10.1080/21650020.2019.1566022

**To link to this article:** https://doi.org/10.1080/21650020.2019.1566022



Published online: 21 Jan 2019.

URBAN, PLANNING AND TRANSPORT RESEARCH
2019, VOL. 7, NO. 1, 1–18
https://doi.org/10.1080/21650020.2019.1566022

Routledge
Taylor & Francis Group

ARTICLE

OPEN ACCESS Check for updates

# Approaches, technologies and importance of analysis of the number of train travellers

Anette Ø. Sørensen[a], Nils O. E. Olsson[a], Muhammad Mohsin Akhtar[a] and Heidi Bull-Berg[b]

[a]Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology (NTNU), Trondheim, Norway; [b]Department of Applied Economics, SINTEF Technology and Society, Trondheim, Norway

**ABSTRACT**

Several studies have pointed to the difficulties of obtaining good data on train ridership. This paper is a literature review on how the number of travellers on trains are measured, including technologies and practices for measuring actual ridership. There a number of publications and practical work done on estimating ridership. We find there are several technologies that can be applied for measuring ridership on trains. The technologies and approaches include (1) Manual counts and surveys, (2) On-board sensors such as door passing, weight, CCTV and Wi-Fi-use, (3) Ticketing systems, ticket sales or ticket validation, and (4) Tracking of travellers for larger part of the journey, e.g. by mobile phones and payments. Data from on-board sensors and ticketing systems are both managed by public transportation providers. By contrast, surveys, payments statistics and mobile phone data may be available to stakeholders outside the public transportation system, which can be an advantage, as access to ridership data can be an issue for business reasons. Furthermore, mobile phone data appears as an interesting option, as they can track complete journeys. New technologies, and especially mobile phone data, are therefore of special interest in future uses of ridership data for evaluations and quality assessments.

**ARTICLE HISTORY**
Received 2 March 2017
Accepted 3 January 2019

**KEYWORDS**
Railway; ridership; number of passengers; train; evaluation

## 1. Introduction

Access to good relevant data can be a challenge when evaluating railway investments (Frumin, 2010; Volden & Samset, 2013). This may seem like a paradox, when the volume of data and different sensoring technologies are maturing. In this paper, we discuss to what extent different technologies have the potential to provide information on train ridership.

Information about ridership in public transportation has long been an issue (Vuchic & Newell, 1968). In a long-term perspective, we are interested in this type of data for use in the evaluation of major transport infrastructure investments, such as new double tracks, railway tunnels, new timetables or fare changes. You may measure the change in travel time from home to work (and vice versa) and changes in travel

**CONTACT** Anette Ø. Sørensen anette.o.sorensen@ntnu.no Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

patterns before and after a major investment in new railway lines for a larger population, which is a suitable way to evaluate the effect of the investment project on.

According to OECD (2002), evaluation can be defined as a systematic and objective assessment of an ongoing or completed project, program or policy, its design, implementation and results. Scriven (1991) use another definition, stating that evaluation is the process of determining the merit, worth or value of something. Data quality and availability are key issues in evaluations (Olsson & Bull-Berg, 2015; Parthasarathi & Levinson, 2010; Small, 1999), and access to good relevant data is a common challenge in project evaluations. This may seem like a paradox, as the volume of data generally increases. The Norwegian Ministry of Communication requests ex-post cost-benefit analyses to be carried out five years after major transportation infrastructure projects are finished. These evaluations are typically made based on the situation at one point in time. New types of data can add both precision and new perspectives to evaluations (Tanaka, 2015). There is, for instance, a potential to generate information about travels that involve different modes of travel.

We review how the number of travellers on trains are measured, including technologies and practices for measuring actual ridership. New technologies are of special interest and especially mobile phone data. There are a number of publications and practical work done on estimating ridership, including estimating changes in ridership due to different changes such as new infrastructure. While predicting a future change naturally is more difficult than to measure an existing situation, obtaining good data on the numbers of travellers on trains is not a trivial task. In addition, experience shows that detailed information on ridership is frequently considered to be business confidential information from the train operator's point of view.

## 2. Analysing ridership on trains, quality assessments and Big Data applications

Two major challenges arise in the process of obtaining data on ridership. One issue is that train operators often treat such data as confidential business information, especially high-resolution data (Vigren, 2017). A second issue is that the data that is available tend to vary in both quality and coverage. The unreliability of ridership data has been addressed in the previous research, such as Fowkes, Nash, and Whitening (1985), Chu and Chapleau (2008), and Kežić and Durango-Cohen (2018).

There is currently a focus on digital transformation in railways (Pierigud, 2018), for instance, within safety (Parkinson & Bamford, 2016), freight transport (Green, 2017) and maintenance (Tute, 2018). Several undertakings and programs focus on this transformation, including Shift2Rail Joint Undertaking, Britain's Digital Railway, and Network Rail's Offering Rail Better Information Services (ORBIS) program. Ongoing digitalization of railways open opportunities for new and expanding measurements of ridership. As an example of the possibilities following digitalization and technological development, Mabrouk and Zagrouba (2018) review how video surveillance system can be used to map behaviour, including travellers on train stations. In the following, we briefly review some of the background for Big Data in transportation and implications for ridership measurements, including quality assessments.

### 2.1. Big Data in telecommunication and transportation

The term 'Big Data' is increasingly becoming mainstream and its usage is no longer limited to a finite set of industries. The term has evolved over the years and as defined by De Mauro, Greco, and Grimaldi (2016): Big Data represents the information assets characterized by such a high volume, velocity and variety to require specific technology and analytical methods for its transformation into value. While accessibility to Big Data is on the rise, however access to such large chunks of information provides challenges of estimation, integration, and validation (Toole et al., 2015). On one hand, Big Data is attractive in the sense that it can be extracted instantaneously at a low cost and the available samples are long-running and cover multiple aspects simultaneously. On the other hand, the sheer size of the data means that it lacks the contextual demographic information pertaining to privacy, resolution of data and inherent noise (Toole et al., 2015).

The type of Big Data that is extracted from cellular data is labelled as 'digital breadcrumbs' by Pentland (2012). This type of data consists of numbers and facts and is structured as it is comprised of datasets of variables which can be easily tagged and categorized. The data are employed itself by the Telecommunication companies and the trend to utilize Big Data in such organizations is gradually increasing. According to Bughin (2016), the five functional domains in which the telecom industry uses Big Data is; sales and marketing, customer care, competitive intelligence, network load optimization, and supply chain optimization. Zheng et al. (2016) suggest application areas of Big Data in transportation, that includes transportation analytics and forecasting with social signals, and crowdsourcing mechanisms for transportation through social media. They also classify the analysis approaches into three main types; i.e. Statistical Analysis, Data mining and Visual Analytics.

In the transportation sector, owing to the inability of the current methods to predict the real-time traffic growth and trends, there is a growing focus in developing smart, efficient, as well as sustainable, transportation systems. The reason for this is twofold. On one hand, it shapes the strategies and operational models, planning and service quality of transport companies. On the other hand, it is imperative in detecting the needs of individual users in terms of time efficient, adaptive travel planning and support of mutually beneficial social behaviour (Fiosina, Fiosins, & Müller, 2013). Utilizing Big Data can prove to be vital towards developing sustainable and environment friendly transport systems.

### 2.2. Use of information about the number of travellers

The number of travellers is a measure of demand for transportation services. Ridership is an important factor when analysing public transportation, including trains. Information on ridership is also important for business planning and evaluations. However, train operators frequently consider such data as business confidential information, especially in high resolutions. In addition, the data that actually is available is of varying quality and coverage (Fowkes et al., 1985).

Boyle (1998) summarize four main reasons for why ridership data is collected:

- ridership is reported to external funding and oversight agencies.
- it monitors trends over time.

- ridership is a key performance indicator at various levels of the transportation system.
- ridership data helps to identify locations with the greatest boarding and alighting activity.

Train ridership is influenced by a number of factors, including fares, transit time, transit comfort characteristics and feeder accessibility of transit, price and service characteristics of the competing modes, seasonal variations and monthly working day variations, as well as socio-economic conditions of the service areas in the medium or long-term (Doi & Allen, 1986).

There has been significant discussion about actual ridership on trains and urban rail. Several authors claim that ridership is systematically overstated by traffic demand models and project promoters (Altshuler & Luberoff, 2003; Flyvbjerg, Bruzelius, & Rothengatter, 2003; Parthasarathi & Levinson, 2010). These studies typically focus on factors influencing the inaccuracy in forecasts. However, we dare to claim that few researchers have focused on inaccuracies in ridership data.

### 2.3. Quality assessment and improvement of passenger transportation by train

Multiple criteria decision-making (MCDM) are methods applied in quality assessments and improvement of passenger transportation. Such methods can be applied to evaluate railway quality from both a passenger perspective and seen from the railway organizations. To begin with, we provide examples of MCDM in a passenger perspective, which is an area where ridership information is highly relevant.

Maskeliūnaitė and Sivilevičius (2012) examine the criteria describing the quality of passenger transportation and determine their significance by applying the method of comparative analysis, i.e. the Analytic Hierarchy Process (AHP) method. Ranks are assigned to the criteria by train passengers, service staff and administrative staff. These are criteria for train elements and technical state of railway tracks, for railway trip planning and technology, for a price of a trip ticket, and for a safe railway trip. They offer an algorithm for displaying the quality of train travel.

Chen, Leng, Mao, and Liu (2014) study passenger transfer capacity, which is one of the fundamental requirements to validate designs and ensure efficient operations at large transport terminals. The study focuses on the transport terminals at major cities where a number of transport services connecting local and inter-city traffic convert. They propose an integrated weight-based multi-criteria evaluation on transfer performance at large integrated transport terminal, as a robust and flexible method. The evaluation is developed by the formulation of criteria systems and integrated performance factors. The methodology is verified through a case study on the Beijing South Railway Station in China, verifying the feasibility and effectiveness of the method.

Paha, Rompf, and Warnecke (2013) explore determinants for customer choice behaviour in passenger rail competition on two cross-border routes. The dataset was collected by performing almost 700 interviews on-board trains on the two cross-border routes. They estimate a multinomial Logit discrete choice model for demand, to analyse the preferences of the interviewees for different long-distance transport services. The results imply that entry into the commercial passenger rail market may be more difficult than often thought.

De Oña, Eboli, and Mazzulla (2014) analyse transit service quality on the basis of the perceptions directly expressed by the passenger of the services. The experimental data were collected in a survey conducted with a sample of more than 16.000 passengers. The survey covered service characteristics such as safety, cleanliness, comfort, information, and personnel. The tool used for evaluating service quality is a Classification and Regression Tree Approach (CART). An important advantage of the CART model is that is does not need to establish a functional relationship among variables as ordinary statistical modelling techniques, such as regression models. Also, it can effectively handle multi-collinearity problems. The existence of multicollinearity is very frequent in these satisfaction surveys. Disadvantages are that the classification tree models are generally 'unstable' because the building of the trees is based on their seed number, which is random, and therefore different trees could be obtained and the results might vary.

Batley, Dagray, and Wardman (2011) develop a dynamic model of rail demand at the market level, yielding short and long-run elasticities with respect to lateness. The short-run market-level model could possibly encompass a range of responses including changes to departure time, ticket type, route, operator and/or mode. In the long-run, the market-level could extend to an even wider range of responses, such as residential location, car ownership and employment. Their primary contribution is reporting estimates of the elasticity of demand for rail in response to changes in service performance. They find that whilst rail travellers show considerable distain for experiences of lateness, such experiences will not necessarily dissuade them from travelling by train.

MCDM methods can be applied to evaluate railway quality seen from the railway infrastructure managers' and train operators' perspective. Train ridership information is a relevant parameter for such assessments. Maskeliūnaitė and Sivilevičius (2014) suggest a mathematical model for the evaluation of the train elements and technical state of railway tracks. The influence of all the criteria on the trip may be evaluated by comprehensive quality index (CQI) or quality index K. The model is used for calculating the values of each of the criteria based on the normalized weight of the criterion multiplied by its variable. They claim that the model and the techniques may be applied to determine the quality or effectiveness of other objects or processes, which can be described by sets of criteria.

The control of train movement in large passenger railway stations is connected with a number of specific problems, including delays to incoming trains. Solving these problems has an impact on the quality of service provided to passengers at a station and in the surrounding railway network. Jánošíková, Kavička, and Bažant (2014) study the dispatcher's decision-making process in the situations of delays to incoming trains, which require the dispatcher to flexibly solve problems related to potential train routing conflicts. The article models the dispatcher's decision-making process using a mathematical programming approach, which includes the operation scheduling and platform track assignment in a large passenger railway station. The model can be classified as a multi-criteria MIP model. The approach has two goals: to minimize the deviations from the valid timetable, and to minimize the inconvenience caused to passengers. The inputs to the model contain potential delays of incoming trains. The outputs are assignments of platform tracks to arriving trains, and adjusted arrival and departure times.

Christogiannis and Pyrgidis (2014) investigate the impact of traffic composition on the economic profitability of a new railway corridor, using a mathematical model. Traffic composition denotes the percentile distribution of passenger and freight trains in circulation on the railway infrastructure, where the possible compositions are mixed, dedicated passenger trains and dedicated freight trains. The economic profitability is characterized by the financial indicators net present value (NPV) and internal rate of return (IRR) of the investment. Their results show that the basic criterion for the selection of the optimum scenario concerns the characteristic of transportation demand. Here demand of transportation is the type of goods being transported (passengers and/or freight) and volume transported, meaning the number of passengers or the number of tonnes transported per day.

## 3. Approaches for obtaining information about the number of travellers

There are several ways to obtain data on train ridership. Boyle (1998) found that most US agencies use more than one method. Manual technologies have been in practice from the start and were still surprisingly the dominant method of gauging passenger count. Vuchic (2005) mentions that surveys and manual passenger counts are well established in railway transit services to obtain information on passenger volume and load counts. Railway agencies have relied on traveller surveys and manual counts to collect data on train ridership.

Tanaka (2015) looks at the prospects of using Big Data in the railways. He suggests technologies for measuring actual ridership, such as data collected at automatic ticket gates, data of vehicle weight, positional data of individuals and data of numbers of passengers in each station.

### 3.1. Traditional methods

According to Boyle (1998) more than two-thirds of the studied US transport agencies used paper and pencil to collect ridership data. Almost as many used Electronic Registering Fareboxes (EFR). Other applied methods were On-Board Surveys, Vehicle Operator Trip Cards, Estimate from Passenger Revenue, Checkers and Hand-Held Units, Automatic Passenger Counting (APC) and Smart Cards. They found that large transport systems were more likely to rely on manual procedures. Smaller systems were more likely to use ERFs. More advanced digital technologies such as APCs and smart cards were used independently of system size.

The practicalities for conducting manual observations, storing and summarizing them are well established (described in Vuchic (2005)). This has been the main data source for the Norwegian railways. In addition, different travel behaviour surveys have been carried out. Using surveys and manual counts provide transport organizations with a reasonable snapshot of existing demand on their transport system. Increasing the resolution of manually obtained information is typically costly and requires personnel resources. A number of APC systems are therefore developed. According to Boyle (1998), accuracy was a major concern in any data collection effort. The use of manual techniques can result in errors both in the collection and registration phases. Such errors tend to be random in nature.

### 3.2. Technologies and methods for automatic passenger counting

According to Vuchic (2005), the most complete passenger counts can be obtained from fully controlled stations. Fare collection systems are platforms for collecting passenger fares and controlling access to the transportation service. There is a trend towards automating the fare collection process, introducing Automated Fare Collection (AFC) Systems. In addition to collecting fares, such systems can be used to track not only the number of passengers, but also the entry and exit points for travels. Frumin (2010) use such entry and exit data and develops a methodology for building an unbiased estimate of existing travel patterns on the London Underground.

Automatic Passenger Counting (APC) is gaining popularity. An APC is an electronic device, which accurately records boarding and alighting data on transit vehicles such as trains. Sensors are located in the doorways to a vehicle. When a person passes, the sensors count movements and determines if they are entering or exiting the vehicle. Barabino, Di Francesco, and Mozzoni (2014) addresses the challenges with using APC to measure ridership on busses. They mention some challenges, such as matching of data to the bus stops, tackling anomalies, and building intelligible performance reports.

APC units are not needed on every vehicle in the fleet. Boyle (1998) found that the eight agencies that make regular use of APCs equipped about 10% of their fleet with APC units, and rotated these units throughout the routes in the system. APCs can be based on different technologies. One way is using infrared lights above the doorways to a vehicle (Chu, 2010). These beams cross the stairwells and are spaced so that the order in which the beam is broken by a person determines if they are entering or exiting the vehicle. Boyle (1998) mentions that another technology for APC is treadle mats. Mats can be applied to the vehicle steps. They contain switches that close when the mat is stepped on. The transitions of closing and opening switches and the times between them determine passenger flows. They mention that in certain climates, treadle mats can be difficult to maintain. Another technology for APC is the use of closed-circuit television (CCTV) and intelligent people counters to log numbers of travellers getting on and off a vehicle (Saponara, Pilato, & Fanucci, 2016). On-board CCTV is frequently installed on trains for surveillance and safety. The technology can also be applied for detection for people counting. Other methods include magnetic loop detectors and automatic register plate recognition systems (Kujala, Aledavood, & Saramäki, 2016). All these methods require physical equipment installation.

Dlamini (2011) quantifies the environmental impacts of the different rail ticketing options available in Japan. For the different ticketing options, the author quantified the energy requirement, environmental impact through an established indicator and the associated $CO_2$ emissions. In terms of energy requirements, the environmental impacts of the various ticketing systems from highest to lowest are; the paper ticket, the limited use plastic ticket, the smart card with gate doors, and the smart card without gate doors. Most energy is consumed by ticket-related machinery during standby time that contributes 50–76% of the system energy requirement.

Nielsen, Frølich, Nielsen, and Filges (2014) present a method for estimating passenger numbers based on electronic weighing equipment (EWE). EWE is installed in many modern trains because it supplies data for the braking system. This information can be used to estimate the number of passengers in the trains, as the weight of a train is a function of the number of passengers in the train at any particular time. Provided that the weighing equipment is installed on a train, this can be a cost-efficient way of estimating ridership. Nielsen et al. (2014) show that EWE-based monitoring can provide estimates with higher accuracy than infrared sensor technology. The weighing system also has the potential to provide a complete sample of weight, and thus ridership. According to Nielsen et al. (2014), passenger distribution in the urban Copenhagen rail network is tracked based on a combination of EWE and APC. The two systems provide complementary information, as the weight-based estimation provides information about the total traffic volume and automatic passenger counting provides information on passenger flow. The two systems can also be used to do quality assurance of each other's measurements.

### 3.3. Existing ridership measurements in Norway

The Norwegian National Rail Administration (Jernbaneverket, now Bane NOR) publish annual reports on the Official Railway Statistics in Norway. The railway statistics include aggregated data on the number of travellers, as well as passenger kilometres, and the number of sold single tickets and monthly tickets. The practices for measuring ridership is manual counting at chosen stations on each railway line.

In 2013, testing of APC system from the German Dilax began in Norway (Zachariassen, 2014). The APC registers the number of people that embark and disembark through each train door on every station, by means of sensors in the doorways. Norsk Regnesentral (Norwegian Computing Center) has developed a mathematical tool that will, based on the APC data, use a statistical model to calculate the total number of passengers and to develop a model for generalizing the data from the vehicles with APCs to be applicable for the complete train traffic system in the greater Oslo area (Teknisk Ukeblad, 2014).

Several evaluations have mentioned the lack of ridership data as a problem. Commissioned from Concept, Nilsson, Nyström, and Pyddoke (2012) evaluated the railway investment on a third and fourth railway track west of Oslo a few years after the commissioning of the project. They commented that there was lack of information about the number of travellers before and after the investment. The ex-ante evaluation had made a forecast on how the number of travellers would change, which was based on conditions of population increase and increase in employment and income. In the ex post evaluation, they had aggregated data on the percentage increase of the number of travellers. However, they comment on the lack of numbers that are consistent over time on the railway line.

### 3.4. New types of data

Hilbert (2013) proposes a classification of different types of data and data sources based on tracking words, locations, nature, behaviour, economic activity, and finally tracking other data. Related to the evaluation of buildings we suggest a division into the following categories according to how data is collected or generated:

- Internet traffic, including activity on social media and data from search engines
- Movement-related data, and different visualizations, including pictures, video, BIM models and maps

- Physical environment, typically from different types of sensors
- Commercial activity, the use of payment services and consumption patterns

New types of data open many possibilities for the analysis of transport measures (Barabino et al., 2014; Bianchi, Rizzi, Sadeghian, & Moiso, 2016). A significant portion of Big Data is geospatial data, generated from sources such as mobile devices and RFID sensors. Geospatial Big Data gives both opportunities and challenges, as discussed by Lee and Kang (2015). In 2012, Virginia Railway Express looked into the latest payment technologies to pilot at key stations, as well as technology to verify ticket purchase and use along with ticket history (Henry & Grant, 2012). The technology options included mobile ticketing, radio frequency tags and Near Field Communication (NFC). The ticket database would be used for analytical purposes such as ridership, travel patterns, boarding at particular stops, client use of facilities by time of day, and other information that enables the providers to better plan their services. They highlight that such solutions should allow for secure mobile phone ticketing with both electronic and visual ticket verification with handheld devices. A key component is the ability to support smartphone technology. This would include mobile devices utilizing the prevalent versions of the mobile operating systems.

Travel time and ridership can be detected using GPS traces. These approaches are indeed innovative and capture in detail individual travel behaviour, but are limited by their sample sizes (e.g. number of volunteers) and currently face scaling difficulties (Holleczek et al., 2014). Also, a low penetration of smartphones on a global scale and limited access to GPS related information from Telecom Operators because of user privacy policies also hinders this to be an effective mode for calculating travel times going forward. Chaudhary et al. (2016) discuss collecting information about occupancy level of public transportation system using the potential of smartphones. Smartphones have inbuilt sensors like GPS which can be used to extract locational intelligence of the commuters. They describe that collected information can be stored in a database for analysis to obtain occupancy level patterns for different routes on different days. They show that patterns observed are used to make predictions of occupancy level in a bus, with an accuracy up to 92%.

Higuchi, Yamaguchi, and Higashino (2015) identifies a number of innovative use forms based on mobile devices, including several technologies which typically are found in smartphones. They mention, among other options, GPS, Wi-Fi, Bluetooth, FM radios and sound recognition. Moreover, it has been showed that it is possible to detect transportation mode based on the GPS sensor on mobile devices and knowledge of the underlying transportation network (Stenneth, Wolfson, Yu, & Xu, 2011). Proximity sensing can be based on logging of Bluetooth units. This is applied by the Norwegian Road Authority (Olsson and Bull Berg, 2015).

Some studies using mobile phone data have been done. For instance, Xu, Shaw, Fang, and Yin (2016), Kujala et al. (2016), Holleczek et al. (2014), Calabrese, Di Lorenzo, Lui, and Ratti (2011), and Jarv, Ahas, Saluveer, Derudder, and Witlox (2012) all show that cell phone data can be used to describe people's movement pattern. However, most studies have not focused on the rail in particular, but typically addressing travel in general. Mobile phone datasets allow deriving a statistical analysis of human activities at a fine level of details (Leo, Busson, Sarraute, & Fleury, 2016). It has been shown that cell phone data can be used to derive good estimates of dynamic quantities, such as travel times, train occupancy levels and origin-destination flows, for transportation studies (Aguilera et al., 2014). Due to this reason, mobile phone data can be utilized in estimating the commuting patterns and travel times for individuals.

There is a variety of approaches that can be utilized for calculating this information by analysing the exchange of information between the mobile base station and cellular network. Most studies perform some kind of trip extraction in order to extract the movements relevant for traffic analysis from the raw cellular network data (e.g. Alexander, Jiang, Murga, & González, 2015; Calabrese et al., 2011; Doyle, Hung, Kelly, McLoone, & Farrell, 2011; Iqbal, Choudhury, Wang, & González, 2014). Because cellular network data can contain a lot of noise, there is no obvious definition of what a movement/trip is (Gundlegård, Rydergren, Breyer, & Rajna, 2016). Hence, trip extraction algorithms vary a lot among different authors. An origin-destination matrix can be computed based on the extracted trips (e.g. Calabrese et al., 2011; Larijani, Olteanu-Raimond, Perret, Brédif, & Ziemlicki, 2015). The aggregation of OD-flows gives an estimation of the number of cell phone users that are travelling, however, only those of the operator that provided the data (Gundlegård et al., 2016). As a result, this can only give information about how the travel demand distributes relatively between different OD-pairs. To estimate the total travel demand in terms of the number of people travelling, authors use different scaling factors (e.g. Alexander et al., 2015; Calabrese et al., 2011; Iqbal et al., 2014; Toole et al., 2015). Several authors have also tried to reconstruct the specific travel mode and route that a user took for a trip, which is challenging to classify. However, as Larijani et al. (2015) showed, detection of the trip segments in which people take the metro is promising, because underground tunnels are being served by dedicated base stations (Gundlegård et al., 2016).

There are three main types of mobile phone data collected using the passive collection: Call Detail Records (CDR) data, Probes data and Wi-Fi data (Larijani et al., 2015). CDRs contain anonymized traces of a user at approximate locations when the phone communicates with a cell phone tower. Gundlegård et al. (2016) present new algorithms for dynamic demand and route choice estimations, that enables efficient use of CDR data for understanding mobility from a transportation planning perspective. Alexander et al. (2015) present methods to estimate average daily origin-destination trips from triangulated mobile phone records of millions of anonymized users. The records are converted into cluster locations and inferred to be home, work or other depending on observation frequency, day of week and time of day. They compared their method to find origin-destination trips with travel surveys.

The advantage with mobile phone data is that it is automatically collected, and hence more frequently and economic than travel survey data (Alexander et al., 2015). In contrast, the disadvantages with mobile phone data are the lack of information about the traveller, like age, income or purpose of a trip.

### *3.5. Summary of approaches for ridership measurements*

Cottrill and Derrible (2015) describe the current methods as sporadic and inefficient, small sample sizes in relation to the target and a poor level of accuracy. Table 1 lists a number of identified technologies for identification of the number of travellers on trains and other transport vehicles. There are many technologies for automatic data collection of ridership, including measuring on-train loads from automatic measurements of train payload weight, and data based on passengers passing through doors. We find a potential to use mobile phones and Wi-Fi-based data. Furthermore, we find that the majority of approaches are depending on access to a train station or rolling stock equipment. In addition, we find that most technologies measure either number of travellers continuously, or at entry- and exit points at train stations. Mobile-based measurements can be applied independent of the traffic operators, and they have the potential to measure complete trips, including several modes of transportation. These features are two arguments in favour of investigating the measurement of ridership using mobile phone technologies, even though there already exist several technologies for ridership measures.

All of the data sources in Table 1, except manual counts and surveys, comprise of data sets that can be characterised as Big Data, as they have the possibility to automatically generate large amounts of data of high resolution. However, till now it has been common to use arrays of information collected for instance by surveys, which is typically not regarded as Big Data.

## 4. Future uses of ridership measurements for evaluations and quality assessments

The term most usually used to define ridership is linked trips. Linked trips are associated with total riders and gauges the actual number of complete trips from origin to destination, which includes transfers. The fundamental advantage associated with transportation investment is dramatically dependent on the extent of mobility it contributes to. Energy savings, air quality contributions, congestion relief, offsetting roadway infrastructure needs, etc., all require the transit services to be utilized by travellers for these benefits to be captured (Polzin & Page, 2003)

Having established the fact that the travel time and its possible optimization are based upon the modes of travel, whose individual usage patterns and density needs to be scrutinized. The foremost reason to evaluate ridership is to determine that there are multiple alternatives available for travellers and the particular mode of transport is having the desired usage and impact as the investment was targeting. Transportation mode inference is a tool to determine the transportation mode of an individual traveller or a group of travellers, based on the speed, travel time or other information that can be collected from their trips. This tool has been used to provide traveling services, manage transportation and plan cities (Wang, Calabrese, Di Lorenzo, & Ratti, 2010). Analysing ridership is imperative for policymakers, understanding how various factors affect transit ridership is crucial to the development of robust policies that will match the goals of sustainable transport (Choi, Lee, Kim, & Sohn, 2011), business planning revenue estimation and taking measures for the future are also reasons for looking into ridership analysis. Stasko, Levine, and Reddy (2016) summarize the application areas of ridership data. These areas are: setting regular scheduled frequency, designing supplemental service during events and construction, planning system expansions, and determining the need for additional station exit capacity, assessing impacts of service

Table 1. Summary of studied technologies and approaches for ridership measurement.

| Indicator on ridership | Source | Part of trip | Data owner | Strength | Weakness |
|---|---|---|---|---|---|
| Manual counts | On-board personnel | Single point measure | Transport service operators | Established | Requires manual resources varying accuracy |
| Surveys | Questionnaires | Tailored | Client conducting survey | Can be tailored to collect an array of information | Requires manual resources. Cannot measure continuously |
| Passing of doors | Automatic logging of door passing | Entry and exit numbers | Transport service operators | Accurate | Cost of installation. Only a part of vehicle fleet typically has equipment |
| Vehicle weight | Rolling stock control system | Number of people continuously | Transport service operators | No additional equipment (on new trains). Need not identify individuals | Requires that weight measure is already installed. |
| Video counting | On-board CCTV | Number of people continuously | Transport service operators | No additional equipment (if already installed). | Requires that CCTV is already installed. |
| Ticket sales | Sales data | Origin- destination on single tickets. Main entry and exit for monthly tickets | Transport service sales organisation | Need no additional equipment | Business sensitive. Difficult to identify individual trips using periodic tickets |
| Ticket use | Ticket validation equipment | Entry and exit stations | Transport service sales organisation | Identifies individual trips, possibly both entry and exit | Requires ticket validation equipment |
| Wi-Fi use | No. of units logged on to train Wi-Fi | Number of units connected, continuously | Transport service operators | No additional equipment (if already installed). | Measures only users of Wi-Fi |
| Proximity sensing | Bluetooth recognition | Number of units passing | One who has access and authority to place equipment | Flexible location | Identifies only mobile units with enabled Bluetooth |
| Train operator smartphone app | Travel detection (GPS etc.) | Location of unit | Transport service operators | Need no additional equipment | Measures only app users |
| Mobile base station data | Mobile network operator | Units connected to base station, and movement in network | Mobile telephone network operator | Need no direct access to train or railway network | Little experience. Personal privacy. Separate logs for each network |
| Payment statistics | Payment transactions in connection to travel | Movements between points of commercial activity | Payment transaction manager | Need no direct access to train or railway network | Indirect measurement of ridership |

management decisions on travel times, computing passenger-based performance metrics (e.g. waiting time and travel time) and reporting to outside agencies.

There can be countless travelling scenarios for a person travelling between two points. The travel scenarios and travel times between an origin and destination point are typically systemized as origin-destination estimations and matrices (Frias-Martinez, Soguero, & Frias-Martinez, 2012; Iqbal et al., 2014; Vuchic, 2005; Wang, Attanucci, & Wilson, 2011). Two common cases of travel scenarios will be

- Work to Home
- Home to Work

The path taken in either direction involves a combination of travelling modes. Public transport will be taken into consideration for people travelling to work and vice versa, and the focus will be on the railway. This can be illustrated by the path described in Figure 1. The journey from Work to Home for an individual may consist of:

- Getting from Home to station C
- Train journey from C to A
- Interchange at station A
- Train journey from A to B
- Getting from B to Work

Mobile phone data has the potential to reveal transport patterns and not only measure the volume of traffic at those points where there is a count. One can also seek explanations by combining ridership data with, for example, data on punctuality or weather data.

Bull-Berg, Olsson, and Sørensen (2017) have done some analyses of train ridership based on mobile phone data. With reference to Figure 1, they looked at the number of handset counts that are likely to be related to train passengers at selected points, like A, B, C and D. Next step will be to obtain data showing flow between base stations, such as from A to B in Figure 1. This is interesting in general, and in particular, if there is a transfer point between

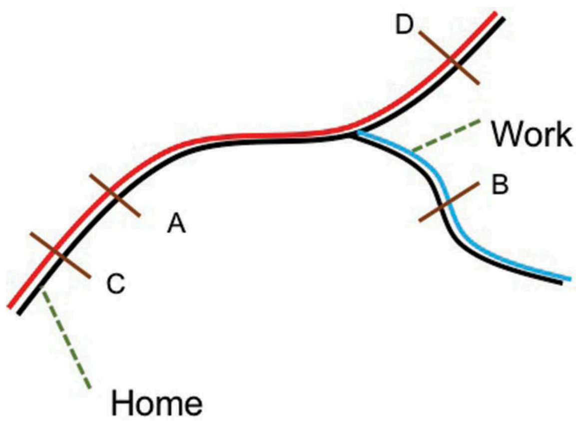


Figure 1. Illustration of travel patterns.

A and B. Even more interesting would be to look at flows between like 'home' areas and likely 'work' areas. This is interesting in itself, but particularly in an evaluation perspective to compare total travel time before and after an investment.

The long-term vision is to utilize mobile data in estimating the benefits of large transport investments. The possibilities to look at different parameters associated with transport is huge, such as traffic density, commuting patterns, utilization percentages of specific train lines, capacity analysis to name a few. Moreover, the usage of mobile data may not only be limited to railways but expanded to other modes of transport as well. Multiple large infrastructure projects are ongoing in Oslo and its neighbouring areas, the vision is to investigate how mobile phone data can be used in future evaluations of these projects.

Mobile phone data can serve as a reliable data source to estimate benefits and analyse the real-time impact of such projects. It can be utilized for calculating stats and facts in alternative analysis and justifying the opportunity space as well as benefit calculation and providing more information to counter the uncertainties. Post evaluation or ex ante evaluation is also a mandatory requirement for Norwegian Government which includes cost benefit analysis. Mobile data thus have the potential to replace the traditional survey data in these evaluations to provide real time information at the point of evaluation. The total time for the complete journey needs to be estimated, the analysis for travel time during different phases of the journey can also be looked into if needed and the causes for increased or factors for optimal travel times can be looked into once these statistics are available for different regions. In accordance with the factors mentioned above the travel time can be a vital input in project evaluations and can serve as a baseline for planning, scheduling, reporting, finance as well as public affair teams for the train operator.

## 5. Conclusion

The paper has reviewed how the number of travellers on trains is measured, including technologies and practices for measuring actual ridership. Furthermore, our study has investigated the potential for using mobile phone data to describe travel patterns that include train travel. We find that there is a set of technologies that can be applied for measuring ridership on trains. The technologies and approaches include:

- Manual approaches, such as manual counts and surveys (based on interviews, questionnaires, etc.)
- On-board sensors, such as door passing, weight, CCTV and Wi-Fi-use
- Ticketing systems, ticket sales or ticket validation
- Tracking of travellers for a larger part of the journey, such as tracking of mobile phone and payments

Use of on-board sensors is established. However, it mainly measures ridership on individual rolling stock units. It is less suitable to measure multi-mode journeys, or even transfers between lines in a train system. Ticketing systems can measure journeys within the ticketing system, which can cover both transfers between lines and between travel modes within the same public transport system. However, it requires ticket validation at entry and preferably also at the exit. Both on-board sensors and ticketing are data that are managed by the public transportation providers. They may not necessarily be available to others, such as evaluators.

Tracking of travellers for larger parts of journeys can be done by traditional surveys. It is costly and is depending on response rates. It also measures people's stated preferences and impressions of their travel time. Payment tracking requires access to payment statistics. It only records electronic payments, and only records visit to commercial areas. Surveys, payments statistics and mobile phone data may be available to stakeholders outside the public transportation system, such as evaluators and others. Finally, mobile phone data appears as an interesting option. It can track complete journeys, with accuracy on the level of coverage of base stations, or even more accurate if apps allowing for GPS tracking are used.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

This study was partially supported by the CONCEPT program at the Norwegian University of Science and Technology (NTNU).

## References

Aguiléra, V., Allio, S., Benezech, V., Combes, F., & Milion, C. (2014). Using cell phone data to measure quality of service and passenger flows of Paris transit system. *Transportation Research Part C: Emerging Technologies*, *43*, 198–211.

Alexander, L., Jiang, S., Murga, M., & González, M. C. (2015). Origin-destination trips by purpose and time of day inferred from mobile phone data. *Transportation Research Part C: Emerging Technologies*, *58*(Part B), 240–250.

Altshuler, A., & Luberoff, D. (2003). *Mega-projects*. Washington, DC: Brookings Institution.

Barabino, B., Di Francesco, M., & Mozzoni, S. (2014). An offline framework for handling automatic passenger counting raw data. *IEEE Transactions on Intelligent Transportation Systems*, *15*(6), 2443–2456.

Batley, R., Dagray, J., & Wardman, M. (2011). The impact of lateness and reliability on passenger rail demand. *Transportation Research Part E: Logistics and Transportation Review*, *47*, 61–72.

Bianchi, F. M., Rizzi, A., Sadeghian, A., & Moiso, C. (2016). Identifying user habits through data mining on call data records. *Engineering Applications of Artificial Intelligence*, *54*, 49–61.

Boyle, D. B. (1998). *Passenger counting technologies and procedures. TCRP synthesis of transit practice 29*. Washington, DC: Transportation Research Board National Academy Press.

Bughin, J. (2016). Reaping the benefits of Big Data in telecom. *Journal of Big Data*, *3*(1).

Bull-Berg, H., Olsson, N., & Sørensen, A. (2017). *Mobiltelefondata – Potensiell datakilde i evaluering av jernbaneutbygginger?* (Progress report). Trondheim: Concept Research Program, NTNU. Retrieved from https://www.ntnu.no/documents/1261860271/1262021752/Rapport+pilot+big+data_+ENDELIG.pdf/81ab8e79-04e0-4c18-8639-87dfe1ad6970?version=1.0

Calabrese, F., Di Lorenzo, G., Lui, L., & Ratti, C. (2011). Estimating origin-destination flows using mobile phone location data. *IEEE Pervasive Computing*, *10*(4), 36–44.

Chaudhary, M., Bansal, A., Bansal, D., Raman, B., Ramakrishnan, K. K., & Aggarwal, N. (2016, January). *Finding occupancy in buses using crowdsourced data from smartphones. Proceedings*

*of the 17th international conference on distributed computing and networking* (p. 35). Article No. 35 Singapore, Singapore: ACM.
Chen, S., Leng, Y., Mao, B., & Liu, S. (2014). Integrated wight-based multi-criteria evaluation on transfer in large transport terminals: A case study of the Beijing south railway station. *Transportation Research Part A: Policy and Practice*, *66*, 13–26.
Choi, J., Lee, Y. J., Kim, T., & Sohn, K. (2011). An analysis of metro ridership at the station-to-station level in Seoul. *Transportation*, *39*(3), 705–722.
Chu, K. A., & Chapleau, R. (2008). Enriching archived smart card transaction data for transit demand modeling. *Transportation Research Record: Journal of the Transportation Research Board*, *2063*, 63–72.
Chu, X. (2010, December). *A guidebook for using automatic passenger counter data for national transit database (NTD) reporting*. Tampa: National Center for Transit Research at CUTR. University of South Florida. http://www.nctr.usf.edu/wp-content/uploads/2011/03/77803.pdf
Cottrill, C. D., & Derrible, S. (2015). Leveraging Big Data for the development of transport sustainability indicators. *Journal of Urban Technology*, *22*(1), 45–64.
Cristogiannis, E., & Pyrgidis, C. (2014). Investigation of the impact of traffic composition on the economic profitability of a new railway corridor. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *228*(4), 389–401.
De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of big data based on its essential features. *Library Review*, *65*(3), 122–135.
De Oña, R., Eboli, L., & Mazzulla, G. (2014). Key factors affecting rail service quality in the Northern Italy: A decision tree approach. *Transport*, *29*, 75–83.
Dlamini, N. G. (2011). The environmental impacts of various rail passenger ticketing options in Japan. *Transportation Research Part D: Transport and Environment*, *16*, 365–371.
Doi, M., & Allen, W. B. (1986). A time series analysis of monthly ridership for an urban rail rapid transit line. *Transportation*, *13*(3), 257–269.
Doyle, J., Hung, P., Kelly, D., McLoone, S., & Farrell, R. (2011, June 23-24). Utilising mobile phone billing records for travel mode discovery. In *ISSC 2011*. Dublin: Trinity College.
Fiosina, J., Fiosins, M., & Müller, J. P. (2013). Big data processing and mining for next generation intelligent transportation systems. *Journal Teknologi*, *63*(3), 21–38.
Flyvbjerg, B., Bruzelius, N., & Rothengatter, W. (2003). *Megaprojects and risk: An anatomy of ambition*. Cambridge, UK: Cambridge University Press.
Fowkes, A. S., Nash, C. A., & Whitening, A. E. (1985). Understanding trends in inter-city rail traffic in Great Britain. *Transport Planning and Technology*, *10*, 65–80.
Frias-Martinez, V., Soguero, C., & Frias-Martinez, E. (2012). *Estimation of urban commuting patterns using cellphone network data. Proceedings of the ACM SIGKDD international workshop on urban computing* (pp. 9–16). Beijing, China: ACM
Frumin, M. S. (2010). *Automatic data for applied railway management: Passengerdemand, service quality measurement, and tactical planning onthe London overground network* (Doctoral Dissertation). Department of Civil and Environmental Engineering and the Operations Research Center. Cambridge: Massachusetts Institute of Technology.
Green, A. (2017). Technology and digitalisation at forefront of freight development. *International Railway Journal*, *57*(7), 50. https://urn.nb.no/URN:NBN:no-nb_digibok_2008071400110
Gundlegård, D., Rydergren, C., Breyer, N., & Rajna, B. (2016). Travel demand estimation and network assignment based on cellular network data. *Computer Communications*, *95*, 29–42.
Henry, C., & Grant, S. (2012, April 17–19). Implementing new automated ticketing technology at Virginia railway express. Paper No. JRC2012-74054. In *2012 joint rail conference* (pp. 449–454). Philadelphia, PA: American Society of Mechanical Engineers.
Higuchi, T., Yamaguchi, H., & Higashino, T. (2015). Mobile devices as an infrastructure: A survey of opportunistic sensing technology. *Journal of Information Processing*, *23*, 94–104.
Hilbert, M. (2013, September 17). Big Data for development: From information to knowledge societies. Retrieved from http://ssrn.com/abstract=2205145

Holleczek, T., Yu, L., Lee, J. K., Senn, O., Ratti, C., & Jaillet, P. (2014). *Detecting weak public transport connections from cellphone and public transport data. Proceedings of the 2014 international conference on big data science and computin* (p. 9). Beijing, China: ACM.
Iqbal, M. S., Choudhury, C. F., Wang, P., & González, M. C. (2014). Development of origin–Destination matrices using mobile phone call data. *Transportation Research Part C: Emerging Technologies*, *40*, 63–74.
Jánošíková, Ĺ., Kavička, A., & Bažant, M. (2014). Optimal operation scheduling and platform track assignment in a passenger railway station. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *228*(3), 271–284.
Jarv, O., Ahas, R., Saluveer, E., Derudder, B., & Witlox, F. (2012). Mobile phones in a traffic flow: A geographical perspective to evening rush hour traffic analysis using call detail records. *PloS one*, *7*, 11.
Character Kežić, M. E. L., & Durango-Cohen, P. L. (2018, July 4). New ridership for old rail: An analysis of changes in the utilization of Chicago's urban rail system, 1990–2008. *Research in Transportation Economics*, *71*, 17–26.
Kujala, R., Aledavood, T., & Saramäki, J. (2016). Estimation and monitoring of city-to-city travel times using call detail records. *EPJ Data Science*, *5*(1), 6.
Larijani, A. N. M., Olteanu-Raimond, A.-M., Perret, J., Brédif, M., & Ziemlicki, C. (2015). Investigating the mobile phone data to estimate the origin destination flow and analysis; case study: Paris region. *Transportation Research Procedia*, *6*, 64–78.
Lee, J. G., & Kang, M. (2015). Geospatial big data: Challenges and opportunities. *Big Data Research*, *2*(2), 74–81.
Leo, Y., Busson, A., Sarraute, C., & Fleury, E. (2016). Call detail records to characterize usages and mobility events of phone users. *Computer Communications*, *95*, 43–53.
Mabrouk, A. B., & Zagrouba, Z. (2018). Abnormal behavior recognition for intelligent video surveillance systems: A review. *Expert Systems with Applications*, *91*, 480–491.
Maskeliūnaitė, L., & Sivilevičius, 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*(3), 544–566.
Maskeliūnaitė, L., & Sivilevičius, H. (2014). The model for evaluating the criteria describing the quality of the trip by international train. *Technological and Economic Development of Economy*, *20*(3), 484–506.
Nielsen, B. F., Frølich, L., Nielsen, O. A., & Filges, D. (2014). Estimating passenger numbers in trains using existing weighing capabilities. *Transportmetrica A: Transport Science*, *10*(6), 502–517.
Nilsson, J. E., Nyström, J., & Pyddoke, R. (2012). *Före och efter - uppföljning av en jäernvägsinvestering i Oslo* (VTI noteat 38-2012). Stockholm: The Swedish National Road and Transport Research Institute (VTI). Retrieved from https://www.vti.se/sv/Publikationer/Publikation/fore-och-efter_669292
OECD. (2002). *No. 6: Glossary of key terms in evaluation and results based management*. Paris: Organisation for Economic Co-operation and Development.
Olsson, N. O., & Bull-Berg, H. (2015). Use of big data in project evaluations. *International Journal of Managing Projects in Business*, *8*(3), 491–512.
Paha, J., Rompf, D., & Warnecke, C. (2013). Customer choice patterns in passenger rail competition. *Transportation Research Part A: Policy and Practice*, *50*, 209–227.
Parkinson, H. J., & Bamford, G. (2016). The potential for using big data analytics to predict safety risks by analyzing rail accidents. In *3rd international conference on railway technology: research, development and maintenance* (pp. 5–8). Cagliari, Sardinia, Italy.
Parthasarathi, P., & Levinson, D. (2010). Post-construction evaluation of traffic forecast accuracy. *Transport Policy*, *17*(6), 428–443.
Pentland, A. (2012). Society's nervous system: Building effective government, energy, and public health systems. *IEEE Computer*, *45*, 31–38.
Pieriegud, J. (2018). Digital transformation of railways. Retrieved from https://shift2rail.org/wp-content/uploads/2018/04/DIGITAL_TRANSFORMATION_RAILWAYS_2018_web.pdf

Polzin, S. E., & Page, O. A. (2003, November). Ridership trends of new start rail projects. In *Transportation research circular E-C058: Experience, economics, and evolution – from starter lines to growing systems. 9th national light rail transit conference* (pp. 319–337). Portland, OR.

Saponara, S., Pilato, L., & Fanucci, L. (2016). Exploiting CCTV camera system for advanced passenger services on-board trains. In *Smart cities conference (ISC2), 2016 IEEE international* (pp. 1–6). Trento: IEEE.

Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Thousand Oaks, CA: Sage.

Small, K. A. (1999). *Project evaluation. Essays in transportation economics and policy: A handbook in honor of John R. Meyer, eds. by Gómez-Ibáñez, J. A., Tye, W. B. and Winston, C.: VI, 577 s.* Washington, DC: Brookings Institution Press.

Stasko, T., Levine, B., & Reddy, A. (2016). Time-expanded network model of train-level subway ridership flows using actual train movement data. *Transportation Research Record: Journal of the Transportation Research Board*, (2540), 92–101. doi:10.3141/2540-11

Stenneth, L., Wolfson, O., Yu, P. S., & Xu, B. (2011). *Transportation mode detection using mobile phones and GIS information. Proceedings of the 19th ACM SIGSPATIAL international conference on advances in geographic information systems* (pp. 54–63). Chicago, IL: ACM.

Tanaka, M. (2015). Prospective study on the potential of Big Data. *Quarterly Report of RTRI*, *56* (1), 5–9.

Teknisk Ukeblad (2014, January 29). Matematiske verktøy skal gi kortere stop på togstasjonene. http://www.tu.no/artikler/matematisk-verktoy-skal-gi-kortere-stopp-pa-togstasjonene/224974

Toole, J. L., Colak, S., Sturt, B., Alexander, L. P., Evsukoff, A., & González, M. C. (2015). The path most traveled: Travel demand estimation using big data resources. *Transportation Research Part C: Emerging Technologies*, *58*, 162–177.

Tute, R. (2018). How advancing technology is allowing network rail to view rail like never before. Retrieved from http://www.infrastructure-intelligence.com/article/jan-2018/how-advancing-technology-allowing-network-rail-view-rail-never

Vigren, A. (2017). *Competition in public transport: Essays on competitive tendering and open-access competition in Sweden* (Doctoral thesis, comprehensive summary). Stockholm: KTH Royal Institute of Technology.

Volden, G. H., & Samset, K. (2013). *Etterevaluering av statlige investeringsprosjekter. Konklusjoner, erfaringer og råd basert på pilotevaluering av fire prosjekter* (Concept Report No. 30). Norwegian University of Science and Technology University. Retrieved from http://www.concept.ntnu.no/Publikasjoner/Rapportserie/Nr.%2030_webutgave30_norsk.pdf

Vuchic, V. R. (2005). *Urban transit. Operations, planning and economics*. Hoboken, NJ: John Wiley and Sons.

Vuchic, V. R., & Newell, G. F. (1968). Rapid transit interstation spacings for minimum travel time. *Transportation Science*, *2*(4), 303–339.

Wang, H., Calabrese, F., Di Lorenzo, G., & Ratti, C. (2010, September). Transportation mode inference from anonymized and aggregated mobile phone call detail records. *2010 13th International IEEE Conference on Intelligent Transportation Systems (ITSC)* (pp. 318–323). Funchal: IEEE.

Wang, W., Attanucci, J. P., & Wilson, N. H. (2011). Bus passenger origin-destination estimation and related analyses using automated data collection systems. *Journal of Public Transportation*, *14*(4), 131–150.

Xu, Y., Shaw, S.-L., Fang, Z., & Yin, L. (2016). Estimating potential demand of bicycle trips from mobile phone data – an anchor-point based approach. *ISPRS International Journal of Geo-Information*, *5*(8), 131.

Zachariassen, E. (2014, January 29). Matematiske verktøy skal gi kortere stopp på togstasjonene. *Teknisk Ukeblad*. Retrieved from https://www.tu.no/artikler/matematisk-verktoy-skal-gi-kortere-stopp-pa-togstasjonene/224974

Zheng, X., Chen, W., Wang, P., Shen, D., Chen, S., Wang, X., … Yang, L. (2016). Big Data for social transportation. *IEEE Transactions on Intelligent Transportation Systems*, *17*(3), 620–630.

# Publication IV

Sørensen, A. Ø., Olsson, N. O. E., & Ekambaram, A. (2015). Evaluation and learning—Experiences from a construction project in Norway. *Procedia Economics and Finance*, 21, 510-517. doi:10.1016/S2212-5671(15)00206-3.

Available online at www.sciencedirect.com

ScienceDirect

Procedia Economics and Finance 21 (2015) 510 – 517

Procedia
Economics and Finance

www.elsevier.com/locate/procedia

8th Nordic Conference on Construction Economics and Organization

# Evaluation and learning – Experiences from a construction project in Norway

Anette Østbø Sørensen[a], Nils O. E. Olsson[a] and Anandasivakumar Ekambaram[b,*]

[a]Norwegian University of Science and Technology, 7491, Trondheim, Norway
[b]SINTEF, 7465, Trondheim, Norway

**Abstract**

The purpose of this paper is to study the relation between evaluation and learning. This paper will look at a construction project in Norway – considering evaluation of the project in connection with learning and knowledge sharing. In this regard, the paper describes different approaches to evaluation and learning, and proposes a model. The model distinguishes between an internal and external perspective when it comes to evaluation of projects, and between a structured and an informal perspective when it comes to learning. With the focus on this model, the paper also presents enablers and barriers of learning and knowledge sharing. The model provides a structured illustration of the connection between project evaluation and learning. And thus, the model would be useful for, for instance, determining and applying learning mechanisms for both internal and external evaluation of projects. Based on our model, traditional project evaluations can be categorized as external-structured. This type of evaluations appears to not necessarily be an important tool for learning. Internal structured approaches, such as an experience report, have been in high demand. We found that external informal learning was of importance. One example was when consultants shared experiences in their home organization. This paper is based on qualitative case study approach.


Selection and/ peer-review under responsibility of Tampere University of Technology, Department of Civil Engineering

*Keywords:* construction projects; evaluation; learning; knowledge sharing.

* Corresponding author. Tel.: +47 40 47 81 74
*E-mail address:* siva@sintef.no


Selection and/ peer-review under responsibility of Tampere University of Technology, Department of Civil Engineering
doi:10.1016/S2212-5671(15)00206-3

## 1. Background

This paper discusses the role of evaluation in relation to experience transfer. The topic of sharing of knowledge and experience in projects and project-based organizations has been discussed for several years. It is not a new topic. However, there is an increasing focus on knowledge sharing and learning in project-based organizations. A wider perspective of management of projects has gained significant attention in recent years. Topics such as project governance (Müller et al., 2014; Biesenthal and Wilden, 2014), project owner (Johansen et al., 2012) and governance of knowledge (Pemsel and Müller, 2012; Pemsel et al., 2014) provide and / or encourage a wider perspective of managing projects and thus emphasize the importance of knowledge sharing and learning in project environments.

Two definitions of evaluation are "A systematic and objective assessment of an ongoing or completed project, program or policy, its design, implementation and results" (OECD, 2000), and "The process of determining the merit, worth or value of something" (Scriven 1991). This paper is about ex-post evaluation. Ex-post evaluation can be described as evaluation of an intervention after it has been completed (OECD 2000). The purpose of this paper is to study the relation between evaluation and learning. We will study different approaches to learning, and propose a model to distinguish between an internal and external perspective, and between a structured and an informal perspective. We also intend to use the model as a framework to study experience transfer in one case project.

## 2. Methodology

In the empirical part of the paper, we have used a qualitative case study research approach, as described by Yin (2008). Information relating to the site was obtained from three main sources: literature related to the sites, other relevant documents, interviews and on-site inspection.

Case study data are collected in a case-study protocol. The protocol includes collected documentation, transcribed notes from interviews and codification of results to fit the applied evaluation framework. We have mainly studied learning in one project, using multiple sources. However, the research also addressed how this project interacted with other projects.

There are two ways to improve reliability in this type of qualitative research, according to Moisander and Valtonen (2006). The first is to make the research process transparent, and the second is to pay attention to theoretical transparency. Both ways to improve reliability have been applied in the research. We have described the theoretical basis for the proposed model, the research process. We have involved all three authors in the analysis of empirical data and results.

## 3. Approaches to evaluation

Evaluators who aim at including a user perspective often prefer holistic evaluations based on a diverse set of approaches and indicators typically combinations of quantitative and qualitative evaluations (OECD 2000). According to Cracknell (1989), the logical framework was developed in the United States during the 1960s. It was adopted by several foreign aid agencies. It has later been adopted for use in project management in general, and proved particularly useful for analyzing public investments. As described by Samset (2003), the logical framework includes a number of different dimensions to be addressed in an evaluation, including efficiency, effectiveness, sustainability, relevance and impact of a project.

We will now look somewhat deeper into the experience transfer approach of evaluations.

## 4. Knowledge sharing and learning

The term knowledge has several definitions. One definition of knowledge given by Davenport and Prusak (1998) highlights the role of knowledge in interpreting and evaluating information. We acknowledge that we use our knowledge to interpret or evaluate information. We believe that the interpretation of the information – the understanding of the information after the interpretation – may also change our knowledge or add new elements to it. This may happen consciously or unconsciously. In this regard, it can also be said that we use information to develop knowledge. Karlsen et al. (2004) view knowledge as the capacity, embodied in the brains of people and embedded in social practices, to interpret information, transforming it into fresh knowledge.

In this paper, we use the term knowledge sharing to include sharing of information, as well as reflection and sensemaking by the individuals who involve in the sharing process, and possible interaction between the individuals who involve in the sharing process. In this regard, our definition incorporates 3 categories of knowledge that Spender (2008) described:

- Knowledge-as-data: The category tends to suggest that knowledge is viewed as an object, and to point out the explicit and objective characteristics of knowledge.
- Knowledge-as-meaning: This category deals with reflection and sensemaking.
- Knowledge-as-practice: This category views knowledge beyond the cognitive spectrum – beyond the sense-making aspect. It incorporates tacit characteristics of knowledge.

Reflecting on the description given by Karlsen et al. (2003), we consider knowledge sharing as a process through which knowledge that has been acquired in one situation is utilized in another situation.

Sharing knowledge can be seen as a tricky issue in projects. There can be a dilemma, specially for project managers, to determine to what extent that they can focus on learning and knowledge sharing in projects, since the major focus is on the iron triangle; namely time, cost and quality that are associated with the project. In other words, in a time-constraint, temporary work-environment, how can one allocate time and energy to sharing knowledge?

A study on sharing lessons learned in construction companies in UK carried out by Carrillo et al. (2013) suggests that sharing lessons learned would lead to learning for similar projects in the future, to avoid making mistakes and repeat success, to provide a competitive edge over other companies, to learn lessons for consecutive stages of on-going projects. One can distinguish between a structured (or hard) and an informal (or soft) approach to knowledge sharing. The structured approach typically focuses on knowledge-as-data. A common approach is to create knowledge repositories of knowledge items. Knowledge repositories are electronic databases that are created for access by users. The content resides on distributed heterogeneous computing systems that use different database management software. The repository is searchable and knowledge items can be tagged with metadata and keywords to facilitate searches. The repositories can be filled through knowledge harvesting. This can serve as input into the knowledge repositories. In the other end of a structured-to-informal scale is a knowledge management approach that focus on human interaction as the repository and communication media for knowledge. This informal approach includes storytelling and ad-hoc experience transfer. Storytelling is the use of stories in organizations as a communication tool to share knowledge. Storytelling uses a range of techniques to engage, involve and inspire people, using common language and a narrative form that people find interesting and even fun. The informal approach incorporates tacit elements of knowledge, and it does not typically focus on knowledge-as-data.

In this paper, we consider the term learning as a process in which knowledge sharing plays a significant role.

## 5. Learning and project evaluation

So far, we have seen knowledge sharing in projects in general. Now, we will look at the connection between knowledge sharing (along with learning) and project evaluation.

Project success is often measured in terms of efficiency and effectiveness. In general, efficiency is related to producing direct outputs, and effectiveness is related to added value for owners and users. A project's ability to produce its immediate outcome can be measured in terms of efficiency. It is a question of doing things right and producing project outputs in terms of the agreed scope, quality, cost, and time. It is a measure internal to the project and restricted to the project or contractor's perspective. The longer-term effects of the project can be measured in terms of effectiveness – or, in other words, doing the right things. It is an external measure. Eikeland (2001) relates effectiveness to how the results of a project contribute to added value for owners and users. In OECD terms, effectiveness measures the realization of the project's objectives (OECD 2000). This is the perspective of the project owner or financing party, who in many types of projects might have a perspective similar to that of the users.

Learning and evaluation can be related to each other in different ways, and on different levels. Begin with learning from evaluations. A key purpose of project evaluations is to provide learning to the project organization.

The aim is typically to provide learning that can be utilized in later phases of the project, or in coming projects. Learning can be related to the evaluation itself – how can we do the evaluation better next time (in order to, for example, obtain better learning). The evaluators can achieve learning not only of conducting evaluations, but also about the particular type of projects.

We can now briefly look at evaluations of learning. This may address issues such as; how did the project achieve learning during the project, and from previous experiences? How well did they spread this learning within the project team, and to relevant stakeholders such as suppliers and users?

The relation between learning and evaluation is addressed in the present trend towards meta-evaluations (Patton 2013), or evaluations of evaluations of similar characteristics, and typically learning related to the project.

Pemsel and Weiwiora (2013, page 32) say that "the risk of knowledge loss at the projects end is a serious problem for PBOs [Project-Based Organizations]". There are several ways to deal with this situation. One of the ways is to look at what has happened in the project and evaluate it – at the end and / or at different points of time during the project. Williams (2007) claim that learning lessons from project reviews is important and an integral part of the learning organization. Making an organization a learning organization is one of the important steps to obtain competitive advantage (Senge, 2006). The project evaluation process incorporates reflection that would lead to learning. We will focus on the role of reflection in learning now.

The process of evaluation of a project includes the project members' reflection on what has happened in the project. Boud et al. (1996) suggest that reflective skills are needed in order to turn an experience into learning. When reflection is considered in connection with a professional action that an organizational member participates, then it can be viewed as reflection-on-action and reflection-in-action (Schön, 1998). Reflection-on-action is a process in which the individual reflects on his or her past experience or on a future act deliberately or unintentionally. Reflection-in-action is a process in which the individual reflects on what he or she is experiencing while he or she is engaging in the activity. Evaluation of a project at its end is hence a process of reflection-on-action. It is interesting and important to look at this process with respect to learning. How can evaluation of a project lead to learning and knowledge sharing?

Argyris and Schön (1996) discuss learning as understanding and eliminating the gap between the expected result and the actual result of an action. This gap can be eliminated either by making changes (taking corrective measures) within the existing values and norms, or by changing the existing values and norms. The former is called as single-loop learning and the latter is called as double-loop learning. Single-loop learning is connected to maintaining efficiency – doing things right according to the existing values and norms. But, the double-loop learning is about doing the right things, by questioning the existing values and norms. This is important especially in a dynamic work environment; because, in order to be effective in such an environment, one probably has to think out-of-box at least now and then.

Single and double loop learning can be accomplished during the evaluation process where collective reflection, sensemaking, discussion and knowledge sharing take place.

The above discussion primarily focuses on internal evaluation. The evaluation process can be carried out both internally and / or externally. Though internal evaluation will lead to learning as it was mentioned above, external evaluation too has its benefits. External evaluators are outsiders. They would think differently and look at the project that they have to evaluate from a kind of an open and wider perspective – free from the value system of the people who belong to the organization (insiders). Senge (2006) discusses mental models as deeply ingrained assumptions that influence how we understand the world. The external evaluators (outsiders) are likely to have a different set of mental models than that of the insiders.

When outsiders look at "the insider's reality", then the outsiders are more likely to ask questions or commenting on assumptions, expectations, judgments and interpretation based on which the project was planned and carried out. These assumptions and interpretations might be implicit for the insiders (including those who carried out the project), or the insiders might not be aware of their assumptions. Outsiders can ask fundamental questions in such a way to challenge the mental models of the insiders, initiate a critical reflection process among the project participants that can lead to double-loop learning; the project participants will then be able to find new, better solutions beyond their existing norms and values. Eriksson (2013) discusses how the process of evaluation – specially, external evaluation – can lead to exploration of new knowledge. We thus note that evaluation and learning can include internal and external stakeholders.

Another perspective on evaluations is to distinguish between a structured (or hard) and an informal (or soft) approach to knowledge sharing. The structured approaches are focused on reports, databases and other formal and tangible means for learning. The informal approaches are concerned with human interaction as the main tool for learning. These two perspectives on learning, internal and external, as well as structured and informal, can be combined in a matrix, as shown in Fig. 1. The distinction between internal and external in Figure 1 is done

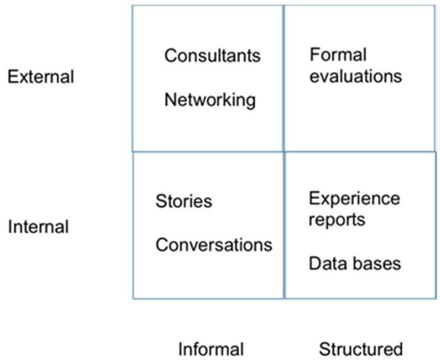


Fig. 1. Illustration of approaches to knowledge sharing in different perspectives

between the organization in question, such as the Railway authority, and stakeholders outside the organization. The distinction can be questioned. One example of a grey area is consultants who are hired to serve as members of the project team. We classified them as external. External evaluations are examples of structured and external learning. Structured approaches can also be internal, including databases, knowledge repositories and internal reports. Informal learning approaches can be internal. This includes common conversations at coffee machines, stories and other informal activities, which can be of different types. Informal approaches can also be external. This can take place at different networking arenas, such as conferences, courses, professional social media, but also at the base organization of consultants and researchers.

The model can be used to illustrate to organizations that there are different alternatives for learning, and that different combinations of external, internal, structured and informal approaches can be used. In addition, the model can be used to study what learning approaches that have been used in a project, as we will do in this paper.

## 6. The case – Gevingåsen tunnel

The railway project was a major public investment executed by the Norwegian National Rail Administration (Jernbaneverket; JBV). The project was the construction of the railway tunnel through Gevingåsen and the connecting railway between Hommelvik and Hell on the Nordland line. The building was started the spring of 2009, and finished the fall of 2011. The railway tunnel through Gevingåsen is the first step in the modernisation of the southern part of the Nordland lane. The Nordland line is the railway section connecting Trondheim and Bodø with 726 km of tracks. The southern part of the Nordland lane is the section between Trondheim and Steinkjer, called "Trønderbanen".

The Gevingåsen tunnel superseded a railway track vulnerable of landslide and with a major need of maintenance. The main goal of the project was to improve the operating economy, through reducing the travel time and increasing the capacity on the railway section in question. In addition, the project would result in increased security level on the railway section.

The railway section Trondheim-Stjørdal carries a great deal of traffic, and before the project was carried out the capacity was almost fully utilized. The railway line between Hommelvik and Hell was the bottleneck of the named section, because distance in travel time between Hommelvik station and Hell station was twice as long as the travel time between the other stations on the section.

The chosen concept of the project was to build a single tracked tunnel between Hommelvik and Hell. The tunnel is 4.4 km in length and has four escape routes. The project included 5.7 km of new railway tracks, and reduced the railway section with 16 km of tracks. Two of the escape routes are in connection with the road tunnel that runs parallel with the railway tunnel. The old railway tracks were demolished.

The tunnel resulted in increased capacity on the railway section Trondheim-Værnes, from 5.4 trains per hour to 8 trains per hour.

## 7. Learning and evaluation in the case

### *7.1. Summary of evaluation*

The project is evaluated to have a high relevance and sustainability compared to the strategic objectives, and the tactical objectives are evaluated to be achieved. The project seems to have had a positive impact on the punctuality and number of passengers on the railway section in question. The implementation process is evaluated to be successful. Smaller overruns in time, cost and SHE is what gives a lower score on efficiency than the other

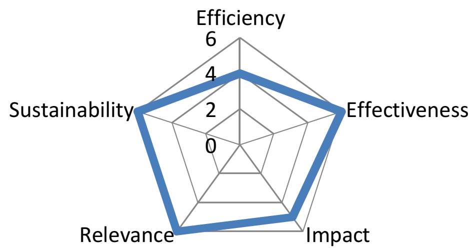


Fig. 2. Radar-plot illustrating the successfulness of each evaluation criterion

evaluation criteria. Fig. 2 shows the successfulness of each evaluation criteria. The overall evaluation is that the project was successful.

The project was a step in the process to modernize a larger railway section; this limits the evaluation since the result of the project is dependent on the subsequent projects.

### *7.2. Experience of learning and knowledge sharing*

In the following section, we present knowledge gained in the project and examples of knowledge sharing related to the project. The information is collected through interviews, and includes aspects of knowledge sharing in the project, barriers and enablers of knowledge sharing, and lessons learned in the project.

*Internal-structured:* In the early phase of the project, information about similar projects is gathered to obtain knowledge on how these projects where done and what can be learned from them. The information is gathered through documents concerning previous projects, end-reports, on-site inspections and contact with the project managers on the previous projects. (The two latter are informal approaches, which is described more below.) The documentations of projects are saved in searchable electronic databases, such as Saksrom and ProArc. The searchable database makes it easy to access the documentation. However, as projects are complex and of different sizes, it can be challenging to find exactly what you are looking for and be able to see the link to other projects.

During the project, deviations are registered in a register that is available for the entire organization. Also, empirical data concerning costs in the project are gathered. These are useful for calculations of rough estimates for later projects. During and/or in the end phase of the project, audits, inquiry of financial management, are implemented.

In the end phase of the project, an internal evaluation is implemented and an end-report is written. The report should document the experiences from the project, and should include mistakes and negative experiences as well as the good experiences. However, experiences that could contribute to better decision making in future projects are often lost when a project is finished. Documentation of such knowledge could increase the learning from knowledge sharing and make it possible to make faster decisions.

In general, project managers obtain useful and different experiences when working with either small or large projects. Smaller projects mean additional responsibilities in several areas, such as economical responsibility. It is useful to have experience from both. A key is to find the balance between top competences and being all-rounders. The training of new project managers is formally arranged through trainee positions, where the candidates follow every phase of a project and have a mentor and teaching supervisor.

It can be useful to give feedback on requirements and regulations if any of these do not work as anticipated. For instance, during the project the project group experienced challenges with the new requirements on the railway signaling plant (FATC). The requirements were therefore worked through and changed after the project.

*Internal informal:* Two other projects in the Norwegian National Rail Administration were carried out in the same time period in other regions; the tunnel in Bærum (just outside Oslo) on the railway section between Sandvika and Lysaker, and new double tracks between Barkåker and Tønsberg. The projects had limited contact in the lifespan of the Gevingåsen project, and the contact was mostly in the early phases. The projects could have gained knowledge through more contact, for instance through exchanging experiences from writing formal applications or in implementing emergency exercise. Also, sharing grounds for decision-making could contribute to more rapid decisions.

The project consists of several phases, research work, planning and constructing. The responsibility of the phases is divided on smaller groups. To ease the transition between the phases people from other project phases are invited to discuss and give suggestions to ideas to make an optimal result. The project also cooperated with the operations managers to ease the transition from the finalization of the project to managing the installation. By thoroughly preparing the plans before project start, the project avoids a lot of changes during the process, which costs both time and money

The project group gains a lot of experiences through the work. It could be of advantage to use the same project group in a following project, in order to gain benefit from the experience that the group has built up.

Other internal and informal ways to gain knowledge is through self-evaluation during the project.

In general, professional networks make it possible to share knowledge through informal (and formal) meetings with coworkers inside the organization (and external contacts). This is an important platform for informal sharing of knowledge to gain useful information to a project. Assemblies, conferences and courses are examples of arenas where such connections can be made. For the employees in the organization, assemblies are arranged regularly for project managers, in trade unions, between different departments, as well as nationwide assemblies. Factors that lower the threshold to contact connections are the availability of videoconferences, open plan office, and co-localization of the departments located in the same city. This availability can contribute to encourage obtaining and sharing knowledge in an ad-hoc manner. This ad-hoc, informal setting provided the Gevingåsen project opportunities to gain knowledge from other projects. The above examples of informal knowledge sharing can also be a source of disturbances, if the contacting is too frequent and not well prepared.

*External informal:* In the project, information on the road-tunnel project through Gevingåsen was also gathered. They had several meetings with the Norwegian Public Road Administration and the project manager about the capacity and security of emergency exits, weaknesses in the mountain, supporting, etc.

External consultants and competence were hired to work on the project. This was useful in several ways. For the organization, the external personnel obtain specific knowledge concerning the organization and discipline from working on the project, which can then be shared in their home organization. In the specific examples given to us the knowledge was informally shared. The bringing of knowledge back to external companies is useful when the organization uses the same company in a future project. For the project, external resources are useful since each has a network of contacts that can be used, if needed, into the project. People who worked on the project were in retrospect hired to speak about the knowledge they gathered through the project.

*External-structured:* Formal evaluation implemented by external evaluators is an example of external and structured knowledge sharing. Neither the studied project, nor the projects they were in contact with had been through formal evaluations. Cooperation with other organizations on specific education and training is another example. The Norwegian National Rail Administration cooperates for instance with the Norwegian Public Road Administration on the tunnel school.

## 8. Reflections on learning and evaluation

The purpose of this paper was to study the relation between evaluation and learning. We have studied different approaches to learning, and proposed a model to distinguish between an internal and external perspective, and between a structured and an informal perspective.

Traditional project evaluations can be categorized as external-structured. Our study indicates that this type of evaluations is not necessarily an important tool for learning. Internal structured approaches, such as an experience report, have been in high demand. Somewhat surprisingly, external informal learning proved to be of importance, for example when consultants share experiences in their respective home organizations.

The model that we suggest provides a structured illustration of the connection between project evaluation and learning. The model would be useful for, for instance, determining and applying learning mechanisms for both internal and external evaluation of projects.

## References


Argyris, C., Schön, D. A., 1996. *Organizational learning II : theory, method, and practice*. Reading, Mass.: Addison-Wesley.

Biesenthal, C., Wilden, R., 2014. Multi-level project governance: Trends and opportunities, *International Journal of Project Management*, Vol. 32.

Boud, D., Keogh, R., Walker D., 1996: *Reflection: Turning Experience into Learning*, Kogan Page, London.

Carrillo, P., Ruikar, K., Fuller, P., 2013. When will we learn? Improving lessons learned practice in construction, *International Journal of Project Management*, Vol. 31

Cracknell, B., 1989. Evaluating the effectiveness of the Logical Framework system in practice. *Project Appraisal*, Vol 4, no 3, pp 163-167,

Davenport, T. H., Prusak, L., 1998. *Working Knowledge - How organizations manage what they know*. Harvard Business School Press.

Eriksson, P. E., 2013. Exploration and exploitation in project-based organizations: Development and diffusion of knowledge at different organizational levels in construction companies, *International Journal of Project Management*, Vol. 31.

Eikeland, P. T., 2001. Teoretisk analyse av byggeprosesser, Samspill i byggeprosessen, prosjektnr. 10602, Trondheim : SiB. ("*Theoretical analysis of construction processes, Cooperation in the construction process*, Project number 10602").

Johansen, A., Steiro, T., Ekambaram, A., 2012. Knowledge management – what do uncertainty management and the project owner perspective have to do with it?, *Proceedings of the European Conference on Knowledge Management*, Spain.

Karlsen, J. T., Gottschalk, P., 2003. An empirical evaluation of knowledge transfer mechanisms for IT projects, *Journal of Computer Information Systems*, Fall.

Karlsen, J. T., Gottschalk, P.,2004. Factors affecting knowledge transfer in IT projects, *Engineering Management Journal*, March, 16-1.

Moisander, J., Valtonen, A., 2006. *Qualitative marketing research: a cultural approach*. Sage Publications, London

Müller, R., Pemsel, S., Shao, J., 2014. Organizational enablers for governance and governmentality of projects: A literature review, *International Journal of Project Management*, Vol. 32.

OECD, 2000. *OECD glossary of evaluation and results based management terms*. Organization for Economic Co-operation and Development, Paris.

Patton, M. Q., 2013. Meta-evaluation: Evaluating the evaluations of the Paris declaration, *The Canadian Journal of Program Evaluation,* Vol. 27 No. 3 Pages 147–171

Pemsel, S., Müller, R., 2012. The governance of knowledge in project-based organizations, *International Journal of Project Management,* Vol. 30.

Pemsel, S., Wiewiora, A., 2013. Project management office a knowledge broker in project-based organisations, *International Journal of Project Management,* Vol. 31.

Pemsel, S., Wiewiora, A., Müller, R., Aubry, M., Brown, K., 2014. A conceptualization of knowledge governance in project-based organizations, *International Journal of Project Management*, Vol. 32.

Samset, K., 2003. Project Evaluation: *Making Investments Succeed*. Tapir Academic Press, Trondheim.

Schön, D. A., 1998. *The Reflective Practitioner, How Proffessionals Think in Action*: Ashgate.

Scriven, M., 1991. *Evaluation thesaurus* (4th ed.). Thousand Oaks, CA: Sage.

Senge, P. M., 2006. *The Fifth Discipline – The art and practice of the learning organization*. Random House Business Books.

Spender, J. C., 2008. Organizational learning and knowledge management: Whence and whither, *Management Learning*, Vol 39 (2).

Yin R. K., 2008. *Case Study Research Design and Methods*. Thousand Oaks, California, SAGE Inc.

# Publication V

Sørensen, A. Ø., Olsson, N. O. E. & Landmark, A. D. (2016). Big Data in Construction Management Research. I: *Proceedings of the CIB World Building Congress 2016, Volume III. Building up business operations and their logic. Shaping materials and technologies*. Tampere University of Technology, ISBN 978-952-15-3743-1, 405-416.

# Big Data in Construction Management Research

Anette Ø. Sørensen,
Department of Production and Quality Engineering, Norwegian University of Science and Technology
anette.o.sorensen@ntnu.no
Nils O.E. Olsson
Department of Production and Quality Engineering, Norwegian University of Science and Technology
nils.olsson@ntnu.no
Andreas D. Landmark
SINTEF Technology and Society
andreas.dypvik.landmark@sintef.no

## Abstract

The paper is a literature review on Big Data in project management of construction projects. The literature shows practical examples of use and potential use of Big Data in construction management research. Big Data has become common as a business term in most businesses. However there is little published management scholarship that tackles the challenges of using Big Data, or even that explores the opportunities for new theories and practices that Big Data might bring about. There is a need for further discussion of the possible implications Big Data can have for construction management research.

The construction process can be studied in a number of dimensions. We structure documented and potential use of Big Data related to construction projects in the time perspective of a typical construction project, from concept preparation and brief, through design and construction to use.

We have identified studies describing Big Data applications and theory. Thematically they fall into three broad categories: 1) New construction equipment; generating, sharing and storing data about use. 2) Data from internal IT systems; such as planning, procurement and Building Information Modelling (BIM) can be utilised. Lastly, 3) people generate an increasing flow of information, which can be useful if handled with care. In combination this addresses the life-cycle from concept to decommissioning. We find that the construction phase appear to have received most attention from researches. We also find that several studies are applicable to more than one phase of a construction project. We find a potential for increased use of Big Data methods and applications within construction. While some data and applications have been analysed in isolation previously, there is a potential to combine different types of data.

**Keywords:** Big Data, Construction management, Construction projects.

# 1. Introduction

Big Data has become common as a business term in many industries, like manufacturing, transportation, retail, finance and IT. However, according to George et al. (2014), there is little published literature in management scholarship that deals with the challenges of using big data or even explores the opportunities for new theories and practices that Big Data might bring about. Olsson et al. (2015) discuss how Big Data can be applied in project evaluation and in project management research. But there is a need for further discussion of the possible implications Big Data can have for construction management research.

## 1.1 More quantitative data available

On a worldwide basis, the total amount of digital data created and replicated each year is expected to increase exponentially up to 2020 (Tien, 2013). This is also the case for data that can be applied in construction projects and management. The principles used in Big Data can also be applied on smaller quantitative data sets. This include data stored in company internal IT-systems.

The definition of Big Data is shifting as software tools become more powerful. Big Data was first defined as data sets whose sizes are too large for commonly used software tools to capture, manage and process within a tolerable elapsed time (Manyika et al., 2011; Tien, 2013; Waller & Fawcett, 2013). However, other definitions will probably be needed, as Big Data is becoming a part of commonly used software tools. The uniqueness of Big Data is the volume, velocity and variety, the three V's (Courtney, 2012; Russom, 2011). The volume refers to the size of data sets, containing a few terabytes to many petabytes. But it is the variety and velocity of the generated data that makes the data sets so big. Variety refers to the variety of sources. In addition, the data are measured and captured in more detail, such a location, time and metadata, giving both structured and unstructured data sets (Russom, 2011; Waller & Fawcett, 2013). The velocity of data refers to the speed at which the data is generated, from being recorded, updated or measured monthly and weekly to more frequent updates such as daily, hourly or continuously (Courtney, 2012). The access to real time or almost real time information makes it possible for a company to be much more agile than its competitors (McAfee & Brynjolfsson, 2012).

Courtney (2012) mentions veracity and value when describing Big Data. Veracity refers to a quality of the data sets, while value is reference to the goal from using the data sets. Veracity is a description of how the measures are reliable, referring to the accuracy and the quality of the generated data. Value refers to turning the Big Data-sets into value for the business. Size does not need to be the only defining part of Big Data, and data can be discussed along more than the mentioned three dimensions. George et al. (2014) points out that there are discussions among practitioners that "big" is no longer the defining parameter, but rather "smart", including a fine-grained nature of the data. The data available to companies are often unstructured (Davenport et al., 2012). The sources of digital data can include retail transactions, security cameras, internally registered data in the organisation, time-stamps, GPS-tracking, sensor-data from instrumented machinery and metavalues of documents.

The main reason to carry out data analysis is to derive information from data, knowledge from information, and wisdom from knowledge (Tien, 2013). And this is the purpose of Big Data. Big Data can give new information and knowledge for decision-making. For instance, Big Data can be used to make more precise predictions, and it follows that better predictions yields better decisions (Jagadish, 2015). McAfee and Brynjolfsson (2012) found that the more companies characterized themselves as data-driven, the better they performed on objective measures of financial and operational results. More and more business activity is digitized, and new sources of information are available (McAfee & Brynjolfsson, 2012). This also applies to the construction industry.

## 1.2 Use of new quantitative data in construction management

Technology has always been a part of construction management, both in research and practice, but there are new technological demands. This means that new data can support the trend towards an intelligent built environment, covering the whole lifecycle of facilities. Big Data has a potential to generate new insights into the costs, designs and processes of construction management. The aim is to develop tools and approaches for intelligent, efficient and sustainable construction management. Such strategies need to be sufficiently flexible to meet requirements resulting from changes in user-demands, technology and other framework conditions, while at the same time increase efficiency. It is a potential to integrate areas such as Computer Aided Facility Management (CAFM), Building Information Modelling (BIM), and Integrated Building Control and Monitoring Systems (BCMS). Such knowledge can later be utilized in decision making support, innovation of technical systems and in the education and training of project managers.

Monitoring activity across a large, complex construction site is particularly difficult because there are so many moving parts, and because the jobs being performed change frequently. In contrast to production industry, most construction sites are also temporary by nature, often challenging the investment in production infrastructure. Several reports document (including Egan, 1998; Lo et al., 2006; Durdyev & Mbachu, 2011) that construction lags behind other industries such as manufacturing in terms of productivity, and blamed the situation on problems with planning, coordination, and communication.

Modern construction equipment also generates data through usage. Producers of equipment such as trains have for some time utilised equipment life-cycle management data, which are generated in large scales through the period of production, operation, and maintenance. There is broad recognition of value of data and information obtained through analysing it. This type of data is also possible to generate from construction equipment. The exponential growth in this type of data means that new measures are needed for data management, analysis and accessibility. MIT Technology Review (2015) reports on the use of drones to monitor construction progress. Once per day, drones automatically patrol the work site, collecting video footage. The footage is then converted into a three-dimensional picture of the site, which is fed into software that compares it to computerized architectural plans as well as the construction work plan showing when each element should be finished. The software can show managers how the project is progressing, and can automatically highlight parts that may be falling behind schedule.

A discussion has emerged about use of Big Data and performance measurement for micro management and continuous monitoring of employees. Using drones to monitor activity continually can be controversial. Such controversies have occurred related to monitoring of employees in other sectors. As reported by the New York Times (2015), the company Amazon monitor employees in warehouses using sophisticated electronic systems to ensure they are packing enough boxes every hour. In a similar way, the Amazon also uses a self-reinforcing set of management, data and psychological tools to spur its tens of thousands of white-collar employees to do more and more.

### 1.3 Purpose and research questions

There is a need for broader discussions of Big Data in society and its implications for management research in general, including construction and project management. George et al. (2014) points out that even though "Big Data" has become commonplace as a business term, there is little published management scholarship that tackles the challenges of using such tools, or that explores the opportunities for new theories and practices that Big Data might bring about. The purpose of this paper is to investigate different use and potential use of Big Data in construction management research. The construction process can be studied in a number of dimensions. We structure documented and potential use of Big Data related to construction project along a time axis of a typical construction project, from concept preparation and brief, through design and construction to use.

Our research questions addressed in this paper are:

- Which applications of Big Data in construction project management have been published in recent years, based on the defined literature search criteria?
- Which time phases of a construction project are recent Big Data research relevant to?

## 2. Method

We review previous research and structure the papers based on the time perspective of a typical project. Several approaches were used to identify relevant literature. Searches in the Norwegian library database (Bibsys) were conducted, covering both books and academic journals. Searches were made using several search engines on the Internet, such as Emerald, Science Direct, Wiley Online Library, and Google Scholar. During the database searches, both titles and keywords were examined. In the searches in Google Scholar the entire texts are searched, consequently providing more search results of both relevant and irrelevant papers. Exclusion keywords were provided in the search, to exclude papers from non-relevant research fields. The collected material was subsequently examined in more detail. The main keywords used in the search were «big data», «construction project/industry», and (construction/project) management. Relevant literature was also found in the journal for automation in construction, as this is relevant technology and data source even though the papers do not use the phrase "big data".

Searches were also made with capital projects and engineering projects as search keywords, but with limited additional search results. Exclusion keywords used were psychology, ecology, biology, constructionism, health, and network construction. The overview of search keywords, databases and number of results is given in Table 1.

In the search results, there are several papers that are not available or do not deal with construction industry. Most of these were excluded by the excluding keywords. However, several papers are within the areas of software development, data management, construction of data centre facilities, Big Data project etc. but are not directly related to the construction industry. The papers where excluded by the search keywords, based on headings, abstracts, or a quick search for construction, management, and Big Data within the main text. Exclusion criteria were used to exclude studies that are not relevant to answer the research questions. Only papers that give a clear contribution or are of clear relevance to Big Data in construction projects or management are of interest. The papers are limited to 2014 and 2015, and Table 1 show that most research contribution has been published these last couple of years.

*Table 1: Overview of search keywords, databases and results. (*Year 2005-2015, **Year 2011-2015, ***Year 2014-2015)*

| *Keyword* | *Search engine /data-base* | *No. of papers* |
|---|---|---|
| *«big data» AND construction AND management* | *Oria* | *121* (95)**** |
| *«big data» AND «construction project OR industry» AND management* | *Google Scholar* | *(591)** (462)**** |
| *«big data» AND «construction project OR industry» AND management –psychology –ecology -constructionism* | *Google Scholar* | *(418)** (319)**** |
| *«big data» AND construction AND «project management» -health, -ecology, -constructionism, -"network construction", -psychology* | *Google Scholar* | *838 (581)* (553)*** |
| *«big data» AND «construction project OR industry» AND «project management» –psychology –ecology –constructionism -biology* | *Google Scholar* | *168 (142)** (103)**** |

The literature that we found to be relevant to our research questions, where structured into the time perspective of a typical construction project inspired by the phases in the RIBA (The Royal Institute of British Architects) model. The RIBA Plan of Work is the definite UK model for the building design and construction process (RIBA Plan of Work 2013). The RIBA model has served as an inspiration to a recently presented Norwegian phase model for construction (bygg.no, 2015). The Norwegian model uses a time frame similar to the RIBA model, but highlights the perspectives of owner, user, supplier and society. The Norwegian model has the following phases: Strategic definition; Concept development; Concept design; Detailed design; Construction; Handover; Use and facility management and finally, Disposal. The selected literature is categorised based on this life cycle perspective. In addition to the time perspective, we found need for adding a second dimension, and discuss a number of alternatives.

# 3. Literature

As mentioned in the previous section, the construction phases that we structured the literature based on, are; Strategic definition; Concept development; Concept design; Detailed design; Construction; Handover; Use and facility management; and Disposal. The result of the structuring of the literature that we found to be relevant is given in Table 1.

A significant portion of Big Data is geospatial data, generated from sources such as mobile devices and RFID sensors. Geospatial Big Data gives both opportunities and challenges, as discussed by Lee and Kang (2015). Related to the construction industry, location aware data can give useful information into urban planning, by providing information in the early project phases on how people use public spaces and infrastructure.

Caron (2015) focus on the importance of an early engagement of stakeholders to improve the forecasting/planning process and manage the project. Engagement of stakeholders from the stage of the project life cycle, in which they may be involved in or be impacted by the project, can significantly increase the available amount of both explicit and tacit knowledge. Managing stakeholders is critical in large construction projects, and in infrastructure major projects this is especially difficult due to the diversity of stakeholders and interests. A study by Bakht et al. (2014) evaluate stakeholders' impact on infrastructure megaprojects through analysis of Big Data captured from online social media.

During the project phases, from early preparation to handover and the liability period, a project produces a lot of documents. Kähkönen et al. (2015) study content and characteristics of current practice of Electronic data/document management systems (EDMS) in building construction projects. This includes project cost, number of files and file accessing actions (open, view, download). They looked at 15 building development projects, which used the same data system.

According to Naderpajouh et al. (2015), effective frameworks to facilitate data-driven decision-making are noticeably lacking in the construction industry. They developed a hierarchical definition for health of the construction industry that was used to propose a framework to benchmark, interpret and analyse data associated with the status of the health of the industry. The dimensions of the framework include; positive financial performance; stability and resilience to shocks, pleasant working atmosphere, applying the best expertise, science, and technology in the production process, and producing high quality products.

Collection of new data and analytical approaches has potential to develop new insights in project management maturity, as examined by Williams et al. (2014). In particular two Big Data analytical techniques is highlighted to have potential to develop understanding of maturity in organisations. These are social network analysis and text analysis. The study by Whyte (2015) analysed change management practices in three separate organizations, including a large construction project, who all deliver complex projects, rely on digital technologies to manage large data-sets, and use configuration management to establish and maintain integrity. The organisations use for instance configuration management in the concept, product identification and definition stages, and as control at the "as-built" stage before handover. The authors conclude that the unstructured, uncontrolled nature of Big Data presents challenges to complex projects that deliver assets. Martínez-Rojas et al. (2015) demonstrate that suitable data handling facilitates and improves the decision-making process and helps to carry out successful project management. They analysed what the main information and communication technologies are, and reviewed the proposals that exist in the literature that focused on the management of information and knowledge from a general point of view in the field of project management.

Barista (2014) presents some early successful applications of data-driven design and planning applications, and how Building Teams can benefit from this. For instance, designers can capture and analyse data from key building performance metrics, such as energy use intensity, to optimize early prototypes. Based on feedback data from building occupants, firms can evaluate design concepts against the real world, and help the building team to understand how people interact with spaces. Redmond et al. (2015) studied how social network analysis and energy usage analyses can be a source to create integrated models for green building design. The main objective was to highlight green building technologies, while at the same time engaging end-users and harnessing their collective knowledge in building design. Several papers also treat opportunities and challenges in combining BIM and Big Data (e.g. Chen et al., 2015).

Several of the papers cover the construction phase. Akhavian et al. (2015) investigate the prospect of using built-in smartphone sensors as data collection and transmission nodes in order to detect detailed construction equipment activities. The method demonstrated a perfect success in recognizing the engine off, idle, and busy states of construction equipment. The work by Teizer (2015) outlines early results for vision-based sensing technology for tracking of temporary assets on infrastructure construction sites. Research and practical industry applications demonstrate promising work towards automated visual recording and progressing of temporary construction resources. Guo et al. (2015) show that Big Data can be used for behaviour observation in China metro construction. The suggested framework was verified in an example to be able to analyse semantic information contained in images effectively, extract worker's unsafe behaviour knowledge automatically.

Automation is a field in construction that uses new data sources and technologies. Machine-to-machine (M2M) installed on construction machines could be used to recommend overhauls to end users at the optimum timing according to Vanzulli et al. (2014). To track the progress of earthwork processes at underground construction sites, Bügler et al. (2014) suggests a novel method that combines two technologies based on computer vision, photogrammetry and video analysis. Combining these data sources allows exact measurement of the productivity of the machinery and determining site-specific performance factors. The construction quality of the material roller-compacted concrete, used in construction of storehouse surfaces, is affected by factors such as the roller compaction, concrete temperature and construction climate. Liu et al. (2015) proposes a real-time construction quality monitoring method, to provide the construction operations on site with timely collection and comprehensive analysis of construction data from the construction process. Wang et al. (2015) presents a method for automatic object recognition and rapid 3D surface modelling, including point cloud data collected from a construction jobsite. Yang et al. (2015) reviews state-of-the-art vision-based construction performance monitoring methods. According to Skibniewski et al. (2015), the construction and operation of infrastructure systems have opportunities for improvement through research on robotics and automation. Automated equipment could cut waste, improve job safety and the overall quality of construction projects. Performance monitoring can be made more effective with tools that better characterize the extent to which construction plan are being followed and the extent to which workers and equipment are being fully utilized.

*Table 2: Categorisation of literature in the time perspective of a construction project.*

<table>
<tr><th>Strategic definition</th><th>Concept dev.</th><th>Concept design</th><th>Detailed design</th><th>Construction</th><th>Hand -over</th><th>Use</th><th>Dis-posal</th></tr>
<tr><td colspan="3">Lee et al. (2015)</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="4">Caron (2015), Bakht et al. (2014)</td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="6">Kähkönen et al. (2015), Naderpajouh et al. (2015)</td><td></td><td></td></tr>
<tr><td></td><td colspan="3">Redmond et al. (2015)</td><td colspan="4"></td></tr>
<tr><td></td><td colspan="5">Williams et al. (2014), Whyte (2015)</td><td></td><td></td></tr>
<tr><td></td><td></td><td colspan="5">Martinez-Rojas et al (2015)</td><td></td></tr>
<tr><td></td><td></td><td colspan="2">Barista (2014)</td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td colspan="2">Chen et al. (2015)</td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td>Akhavian et al. (2015), Teizer (2015), Guo et al (2105), Vanzulli et al, 2014, Bügler et al. (2014), Liu et al. (2015), Yang et al. (2015), Wang et al. (2015)</td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td colspan="3">Skibniewski et al. (2015)</td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td colspan="4">Lu et al. (2015)</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td>Hong et al. (2015), Ioannidis et al (2015), Isikdag (2015)</td><td></td></tr>
</table>

Lu et al. (2015) study the performance of construction waste management for the construction categories building, civil engineering, demolition, foundation, and maintenance and renovation. The study develops a set of key performance indicators (KPIs)/waste generation rates using an available Big Data set on construction waste management in Hong Kong. Demolition was found to be the most wasteful works that generate both non-inert and inert construction waste.

There have been published several examples of evaluation of buildings in use, using quantitative methods based on Big Data approaches. The study by Hong et al. (2015) reviews the state-of-the-art in the major phases for a building's dynamic energy performance, focusing on the operation and maintenance phase. Ioannidis (2015) presents Big Data and visual analytics techniques for comparing building performance under different scenarios and design. Data that provide useful information is energy consumption, building geometry, and space occupancy. Isikdag (2015) provides a method for facilitating the geographic information system (GIS) based fusion of information residing in digital building "Models" and information acquired from the city objects. The virtual BIM sensors in the proposed design pattern will provide geometric and semantic information together with information related to the state of the building elements, and the information can be used to represent the building within a GIS environment, and city monitoring/management.

# 4. Concluding discussion

Our first research question was related to which applications of Big Data in construction project management that have been published in recent years, based on the defined literature search criteria. We have identified studies describing Big Data applications and theory. Thematically they fall into three broad categories: 1) New construction equipment; generating, sharing and storing data about use. 2) Data from internal IT systems; such as planning, procurement and BIM can be utilised. Lastly, 3) people generate an increasing flow of information, which can be useful if handled with care. In combination this addresses the life-cycle from concept to decommissioning. We find that Big Data have been used in energy management. Data from internal IT systems, such as planning, procurement and BIM can be utilised. People generate an information flow, which can be useful but also treated with care to safeguard personal integrity. Heating, ventilation, and air conditioning (HVAC) and electricity management systems generate large volumes of data that can be applied for life-cycle management of the equipment, but also for describing the use of the building. None of the capabilities described in this study are entirely novel in inception nor unique to the construction industry, producers of major equipment such as train rolling stock have for some time utilised equipment life-cycle management data, which are generated in large scales through the period of production, operation, and maintenance.

Other categorisations of Big Data could have been applied. One alternative is to use a hierarchy of increasing aggregation; raw sensor data, databases of sensor data, reports from such systems, and different forms of presentations and reports based on input from several systems. Other alternative dimensions include different stakeholders involved in the project, type of data sources or different types of use of the data.

Our second research question was which time phases of a construction project that recent Big Data research is relevant to. The literature is presented based on a project phase perspective. We find that the construction phase appear to have received most attention from researches. We also find that several studies are applicable to more than one phase of a construction project.

We find a potential for increased use of Big Data methods and applications within construction. Typically, the power of data integration is hard to demonstrate in limited and small pilots, but requires critical mass before providing return. Big Data can be used to make more precise predictions, which can lead to better decisions. While data and applications related to different engineering disciplines, such as energy, have been analysed in isolation previously, there is a potential to combine different types of data. This creates opportunities, but also challenges related to personal privacy. Big Data appear to have a potential to generate new insights into the costs, designs and processes in project management. The aim is to develop tools and approaches for intelligent, efficient and sustainable construction project management. This research addressed Big Data. It is not necessarily the "big" part of Big Data that is key in construction applications. A possible development is that concepts from Big Data are applied on smaller datasets. Focus will then be on new data, or expanded use of existing data.

## References


Akhavian, R & Behzadan, A (2015). Construction equipment activity recognition for simulation input modeling using mobile sensors and machine learning classifiers. *Adv. Eng. Inform*.

Barista, D (2014). The Big Data revolution: How data-driven design is transforming project planning. *Building Design & Construction*.

Bügler, M, Ogunmakin, G, Teizer, J, Vela, PA & Borrmann, A (2014). A comprehensive methodology for vision-based progress and activity estimation of excavation processes for productivity assessment. *Proceedings of the 21st International Workshop: Intelligent Computing in Engineering (EG-ICE)*, Cardiff, Wales.

Bygg.no (2015). Bygg21 lanserer nytt rammeverk for stegene i en byggeprosess, (available online http://www.bygg.no/article/1256664 [accessed 27.11.2015])

Caron F (2015). Data management in project planning and control. *Int. J. Data Science*,1,42-57.

Chen, K, Lu, W, Peng, Y, Rowlinson, S & Huang, GQ (2015). Bridging BIM and building: From a literature review to an integrated conceptual framework. *Int. J. Proj. Manag.*

Courtney, M (2012). Puzzling out big data. *Engineering and Technology*, 7, 56-60.

Davenport, TH et al. (2012). How 'big data' is different. *MIT Sloan Management Review*, 54.

Durdyev, S et al. (2011). On-site labour productivity of New Zealand construction industry: Key constraints and improvement measures. *Construction Economics and Building*, 11(3), 18-33.

Egan, J (1998). *Rethinking construction*. Report of the Sir John Egan's Construction Task Force (July), London: HMSO.

George, G, Haas, MR & Pentland, A (2014) Big Data and management. *Acad. Manage. J.,* 57, 321-326.

Guo, S, Luo, H & Yong, L (2015). A Big Data-based Workers Behavior Observation in China Metro Construction, *Procedia Engineering*, 123, 190-197.

Hong, T et al. (2015). A review on sustainable construction management strategies for monitoring, diagnosing, and retrofitting the building's dynamic energy performance: Focused on the operation and maintenance phase. *Applied Energy*, 155, 671-707.

Ioannidis, D et al. (2015). Big Data and Visual Analytics for Building Performance Comparison, *Artificial Intelligence Applications and Innovations*. Springer International Publishing, 421-430.

Isikdag, U (2015). BIM and IoT: A Synopsis from GIS Perspective. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 40.

Jagadish, HV (2015). Big Data and Science: Myths and Reality. *Big Data Research*, 2, 49-52.

Kähkönen, K & Rannisto, J (2015). Understanding fundamental and practical ingredients of construction project data management. *Construction Innovation*, 15, 7-23.

Lee, J et al. (2015) Geospatial Big Data: Challenges and Opportunities. *Big Data Res.*, 2, 74-81.

Liu, Y et al. (2015). Study on real-time construction quality monitoring of storehouse surfaces for RCC dams. *Automation in Construction* 49: 100-112.

Lo, T, Fung, I, Tung, K (2006). Construction delays in Hong Kong civil engineering projects. *J. Constr. Eng. Manage*, 132 (6), pp. 636–649.

Lu, W et al. (2015). Benchmarking construction waste management performance using big data. *Resources, Conservation and Recycling* 105: 49-58.

Manyika, J, Chui, M, Brown, B, Bughin, J, Dobbs, R, Roxburgh, C & Byers A (2011). Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute.

Martínez-Rojas, M, et al. (2015). The Role of Information Technologies to Address Data Handling in Construction Project Management. *J. Comput. Civ. Eng.*: 04015064.

Mcafee, A & Brynjolfsson, E (2012). Big data: the management revolution. *Harvard business review*, 60-6, 68, 128.

MIT Technology Review (2015). New Boss on Construction Sites Is a Drone, (available online http://www.technologyreview.com/news/540836/new-boss-on-construction-sites-is-a-drone/)

New York Times (2015). Inside Amazon: Wrestling Big Ideas in a Bruising Workplace, (available online http://www.nytimes.com/2015/08/16/technology/inside-amazon-wrestling-big-ideas-in-a-bruising-workplace.html?_r=0)

Olsson, NO, Bull-Berg, H and Walker, D (2015). Use of big data in project evaluations. *International Journal of Managing Projects in Business,* 8.

Redmond, A, Papagelis, M & El-Diraby, T (2015). Employing an exploratory research stage to evaluate green building technologies for sustainable systems. *Proceeding of the International Conference on Civil, Structual and Transportation Engineering*, Ottawa, Canada, May4-5 2015.

RIBA Plan of Work 2013, (available online http://www.ribaplanofwork.com [viewed 11.2015])

Russom, P (2011). Big Data analytics. *TDWI Research, TDWI Best Practices Report.*

Skibniewski, M & Golparvar-Fard, M (2015). *Toward a Science of Autonomy for Physical Systems: Construction*. A white paper prepared for the Computing Community Consortium committee of the Computing Research Association, (available online http://cra.org/ccc/resources/ccc-led-whitepapers/)

Teizer, J (2015). Status quo and open challenges in vision-based sensing and tracking of temporary resources on infrastructure construction sites. *Adv. Eng. Informatics,* 29**,** 225-238.

Tien J (2013). Big Data: Unleashing information. *J. Systems Sc. and Systems Eng.*, 22, 127-151.

Vanzulli, B, Kosaka, M & Matsuda, F (2014). Servitization in a construction machinery industry by using M2M and cloud computing systems. *Service Systems and Service Management (ICSSSM)*, 2014 11th International Conference on. IEEE, 1-6.

Wang, C and Yong, KC (2015). Smart scanning and near real-time 3D surface modeling of dynamic construction equipment from a point cloud. *Automation in Construction* 49: 239-249.

Whyte, J et al. (2015). Managing change in the delivery of complex projects: Configuration management, asset information and 'big data'. *Int. J. Proj. Manag*, (available online http://dx.doi.org/10.1016/j.ijproman.2015.02.006)

Williams, N, Ferdinand, NP & Croft, R (2014). Project management maturity in the age of big data. *International Journal of Managing Projects in Business*, 7, 311-317.

Waller M & Fawcett S (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics*, 34, 77-84.

Yang, J et al. (2015). Construction performance monitoring via still images, time-lapse photos, and video streams: Now, tomorrow, and the future. *Adv. Eng. Inform*, 29.2: 211-224.