Weather and Passenger Train Delay Interactions in Southern Sweden Between 2008-2019

AALBORG UNIVERSITY
STUDENT REPORT

Master Thesis

Weather and Passenger Train Delay Interactions in Southern Sweden Between 2008-2019

Aalborg University Copenhagen

Michelle Ochsner
(20191434)
MSc. Eng. Sustainable Cities

Supervisor: Susse Georg
Co-examiner: Lief Gjesing
Hand in Date: June 4th 2021

Pages: 103
Figures: 38
Tables: 14

Abstract

Railways play an important role in sustainable transportation and in the mitigation of climate change. However, in order to fulfil this role railways must be resilient to disturbances such as those related to weather and climate. The aim of this thesis is to identify the relationship between precipitation, temperature, and dwell and run delays in Skåne, Southern Sweden between 2008-2019, and to determine how this relationship will change in the future in light of climate change. Currently, the punctuality of railways in Sweden is below desired targets, which can hinder current ridership numbers and future growth. Weather is one factor that has an impact on the punctuality of railways in Skåne, Southern Sweden.

The main theoretical framework behind this research is based on resilience thinking. Resilience is defined and used in many different fields, however for this thesis it was narrowed down to the scope of resilience within the transportation system and is defined as the system's ability to resist, absorb, and recover from a disruption related to precipitation or temperature. The main methods include a literature review, graphical evaluations, and both multiple linear and multiple logistic regression modelling to determine the relationship between the sum of precipitation, minimum temperature, maximum temperature, and run and dwell delays over 1, 7, 14, 21, and 28 days.

This thesis identified that the most statistically significant variables for dwell delays are the sum of precipitation over 7 days, the minimum temperature over 1, 7, 14, and 28 days, and the maximum temperature over 1 and 28 days. The most statistically significant variables for run delays are the sum of precipitation over 1, 7, 14, and 28 days, the minimum temperature over 1,7, and 28 days, and the maximum temperature over 1, 14, and 28 days. In other words, most delays occur when temperatures are below freezing and extremely warm, and with intense precipitation. This relationship is expected to become more significant as temperatures and the amount of precipitation increases due to climate change. Therefore, railways need to be more resilient to the future effects of climate change in order to ensure that railways remain on time, and act as passengers first-choice mode of transportation.

Kew words: temperature, precipitation, delays, railways, climate change, resilience, vulnerability

Preface

This thesis focuses on the relationship between weather, specifically temperature and precipitation, and passenger rail run and dwell delays in the region of Skåne between 2008-2019. It is written in collaboration with the Traffic and Roads Department of the Faculty of Engineering at Lund University and is intended to act as the first part of a greater five-year project that explores the effects of weather on train delays in Sweden in the past, present, and future. By identifying how precipitation and temperature impact railways and by increasing railways resilience to future climate change, railways can become more punctual and reliable both today and in the future.

This thesis would have been possible without all the people behind me. A huge thank you to my friends and family for their continued support and to my roommates for putting up with my crazy antics during another full semester of working from home. Although our master’s studies has played out differently than we imagined I would like to thank my fellow classmates for a great two years and thank them for their support as we navigated studying during the COVID-19 pandemic. Thank you to my supervisor Susse Georg for navigating me throughout the writing process, and for providing much support, insight, and advice on my research and structure of my thesis. Lastly, a big thank you to everyone at LTH for their support and advice throughout my thesis. I have learnt so much from you all and for that I am truly grateful. To Carl-William Palmqvist a huge thank you for all your help throughout this thesis. Thank you for the opportunity to work with you, for all your advice and insight, and for your patience and support throughout this process. I look forward to working with you again in the future. The list of thanks could go on and on, I am really grateful to have so many people to thank! So once again thank you and I hope you enjoy reading this thesis.

Table of Contents

List of Figures

List of Tables

INTRODUCTION

1.0 Introduction

Historically, there has been an enormous social and technological change in the succession of the transportation sector (Armstrong & Preston, 2011). Railways originally dominated and quickly expanded in the 19th and early 20th centuries, but by the mid-20th century motorised road transport and commercial aviation quickly dominated the sector for short-haul trips and long-haul trips respectively (Armstrong & Preston, 2011). This progression has also led to the reduction of rail services across the whole world (Armstrong & Preston, 2011). Today, it is becoming more evident what implications motorised road and air transportation have on global greenhouse gas emissions. In 2017, the transportation sector was responsible for 28% of total CO_2 emissions (Stéphan & Blayac, 2021). In Europe, the transportation sector accounted for a larger share of emissions, and specifically 40% in Sweden (Stéphan & Blayac, 2017). Therefore, reducing CO_2 emissions within the transportation sector should be a priority for global policymakers (Stéphan & Blayac, 2021). Often, a modal shift towards more public transportation or cleaner modes has been considered as one of the main ways to significantly reduce emissions (Stéphan & Blayac, 2021); especially as more people are moving to urban areas, it is important to ensure that these people choose sustainable modes of transportation (United Nations, 2018). Now faced with global climate change, we must rethink the entire transportation system (Armstrong & Preston, 2011).

Our society is becoming more dependent on critical infrastructures. Infrastructure is defined as critical when its absence has a serious impact on our well-being, safety, and health; and transportation is considered to be such an infrastructure (Adjetey-Bahun, et al., 2016). Today, weather conditions play a role in the punctuality of railways (Zakeri & Olsson, 2018). As global climate change progresses the effects of weather on the punctuality of railways is expected to become more severe (Zakeri & Olsson, 2018). Therefore, it is important to ensure that railway systems and infrastructure is resilient. Ensuring that trains are able to bounce back from disruptions can assist with their ability to provide reliable and sustainable transportation for all and increase sustainable transportation ridership.

1.1 Railways as a Sustainable Form of Transportation

Railways are often considered to be a form of sustainable transportation (Brons & Rietveld, 2008). They offer an efficient transportation system built on low carbon emissions, low environmental impacts, social equity, and positive economic growth (Brons & Rietveld, 2008). Compared to motorised vehicles and aviation, railways are responsible for significantly less amount of greenhouse gas emissions. In Europe, rail is responsible for less than 0.5% of greenhouse gas emissions in the transportation sector, making it one of the most sustainable modes of freight and passenger transportation (European Commission, n.d.). In Sweden in particular, The Swedish Transport Administration, Trafikverket, recorded in 2017 that the total Swedish railway network includes 12,000 km of electrified rail lines out of a total of 14,600 km (Trafikverket, 2017). This further supports the idea of railways as a sustainable mode of transportation.

In 2020, the EU launched the "*Sustainable and Smart Mobility Strategy*" an initiative aimed at creating a sustainable transportation system within the European Union (European Union, 2020). The main goal of the initiative is to cut 90% of greenhouse gas emissions in the transportation sector by 2050 (European Union, 2020). Additional goals for the railway sector specifically include doubling high-speed rail traffic across Europe by 2030, creating healthy and sustainable interurban and urban mobility by 2050, and to create synergies between multiple modes of transportation (European Union, 2020). To put a spotlight on the "*Sustainable and Smart Mobility Strategy*," The European Commission declared 2021 as the "*European Year of the Rail*" (European Commission, 2020). This initiative was created with the intention to highlight railways as a smart, safe, and sustainable mode of transportation (European Commission, 2020). Today only about 11% of goods and 7% of passengers travel via rail, therefore the initiative aims to increase this ridership by creating momentum, highlighting how an increase in rail ridership can continuously decrease greenhouse gas emissions (European Commission, 2020).

Since the relative decline of railway use during the 20th century, in the past few decades, there has been a bit of a resurgence, particularly in high-speed passenger rail and long-distance freight services (Armstrong & Preston, 2011). Railways have also proven to play a valuable and significant role in medium-distance passenger transport, and commuter transportation within larger urban areas, to, and from (Armstrong & Preston, 2011).

1.2 Punctuality and Train Delays

Railway transportation in Sweden has been growing by 3% annually since the 1990s, with the primary growth found in passenger trains for regional and local trips (Palmqvist, 2019). This implies that if this growth rate continues, more people will be taking trains and therefore the demand for reliable services will increase. Therefore, it is important to ensure that trains arrive on time to their destination in order to highlight rail as the best choice of transportation for passengers and to keep the increasing trend in ridership numbers. In order for trains to be an attractive mode of sustainable transportation, they must be reliable, predictable, and punctual (Brons & Rietveld, 2008).

Punctuality is a key indicator of the performance of railway services (Palmqvist, Olsson, & Hiselius, 2017a) and the national goal for punctuality in Sweden states that 95% of trains should be punctual (Palmqvist, 2019). Punctuality can be measured in different ways and in Sweden, a train is considered to be punctual if it arrives with no more than five minutes delay at the final stop (Palmqvist, 2019). Although the goal for punctuality is 95%, realistically it has been close to 90% which is deemed as too low by the industry (Palmqvist, 2019). Related to punctuality is the concept of delays, which are measured using a time unit, typically minutes (Palmqvist, Olsson, & Hiselius, 2017a). There are various types of delays but, this thesis looks at two: dwell delays and run delays, as they are the two main and prevalent types of delays. Dwell delays refer to when a train is stationary at a station for a longer than scheduled amount of time, while run delays occur when a train is moving between one station to another. Dwell delays and run delays are the delays looked at in this thesis because they are considered to be the main types of delays. For instance, multiple run delays along the same track can lead to knock-on delays, and therefore it is important to consider how to decrease run delays in order to prevent knock-on delays. Additionally, dwell delays are closely related to run delays as a run delay on a single line track can also lead to a longer dwell time at a station.

In order to achieve a high level of punctuality, delays need to be reduced. Rail delays in Sweden can be a result of many different factors such as maintenance, infrastructure failures, timetable planning errors or weather (Palmqvist, Olsson, & Hiselius, 2017b). In this thesis, the focus is on the impact of weather on rail delays.

1.3 Climate/Weather and Railways

Railways are required to operate under high operational standards, ensuring that rail is safe and reliable with uninterrupted services (Misnevs, Melikyan, & Bazards, 2015). Extreme weather events and changing climate conditions are putting transportation infrastructure investments at risk due to the long-life span and expenses of this type of infrastructure (European Commission, 2013). Additionally, the resilience and preparedness to the impacts of climate today and in the future is essential to keeping operational standards of railways high (European Commission, 2013; Misnevs, Melikyan, & Bazards, 2015). Railways are typically more vulnerable than road infrastructure because of the lack of excess capacity, limited rerouting options, and single-line tracks (Mattsson & Jenelius, 2015). There is less research focus dedicated to both waterborne and rail transportation compared to road, air, or urban (Aparicio, et al., 2013), due to priority and investments.

As previously mentioned, since rail is a mode of transportation with a relatively small environmental impact, it has the potential to help mitigate climate change (Armstrong, Preston, & Hood, 2017). However, this potential can only be achieved if railways are resilient and adaptable to the increasing extreme weather phenomenon that is associated with climate change predictions (Armstrong, Preston, & Hood, 2017). Currently, recent extreme weather events have demonstrated the railway industry's vulnerability to the current climate and most likely this sensitivity will increase (Armstrong, Preston, & Hood, 2017). Therefore, in order to better understand how climate change will impact rail delays and help rails reach their potential of mitigating climate change, it is important to gain a solid understanding of past weather conditions.

1.4 Problem Formulation and Research Questions

Based on the above mentioned, weather is one factor that can affect punctuality and therefore cause disruptions to the system. Therefore, it is important to understand the effects weather has on rail delays in order to ensure train transportation is reliable both today and in the future. In this thesis, the effects of precipitation and temperature on passenger train delays in Skåne, Sweden between 2008-2019 is explored. This time period is chosen as climate has already been changing significantly since the early 2000s (SMHI, 2015). Therefore, this thesis looks over these 11 years in order to better understand the effects of weather during the 2000s on delays. This has led to the following problem formulation:

What are the implications of the past effects of precipitation and temperature on past, present, and future passenger railway delays in Skåne, Sweden?

In order to elaborate on the problem formulation, the following sub-questions are formulated as followed:

Research Question 1: What is the relationship between precipitation and rail delays?

Research Question 2: What is the relationship between minimum temperature and rail delays?

Research Question 3: What is the relationship between maximum temperature and rail delays?

Research Question 4: What are the future implications of resilient railways in Skåne given the past precipitation and temperatures conditions and their effect on train delays?

The purpose of research questions 1-3 is to examine each weather variable individually but also to see the connections of their impact on rail delays. This thesis provides a quantitative study which uses graphical evaluations and descriptive statistics in order to determine the relationship between weather and delays. Research question 4 serves as a reflection question to take the information learned about the weather conditions between 2008-2019 and apply the knowledge to the concept of resilience to highlight what Skåne railways can learn and do differently in the future.

1.5 Thesis Structure

The aim of this thesis is to demonstrate the Skåne railway industry's current vulnerability to precipitation and temperature and to discuss how this vulnerability is expected to increase with climate change. This thesis contributes to the research already done on the effects of weather on railway delays by highlighting the importance of resilience in light of future climate change projections. The ability to carry out such an empirical analysis has greatly improved compared to previous studies due to the amount of data accessible. Additionally, it adds something new to this field of research by investigating the weather over a range of different time-periods, specifically 1, 7, 14, 21, and 28 days, instead of only investigating the instantaneous daily effects which has been mainly studied so far. Analysing the accumulative effects allows us to study the impact of events such as heatwaves, cold snaps, and periods of intensive rainfall. Furthermore, this thesis is focused on Southern Sweden, an area where fewer studies have been conducted.

Figure 1 below highlights the overall structure of the project with the overall problem formulation and the four research questions and how they will be addressed using various methods and theory. The methods used are discussed in Chapter 2.0. The theoretical framework behind this thesis is resilience thinking with a focus within the transportation sector which is elaborated upon in Chapter 3.0 and encompasses all research questions.

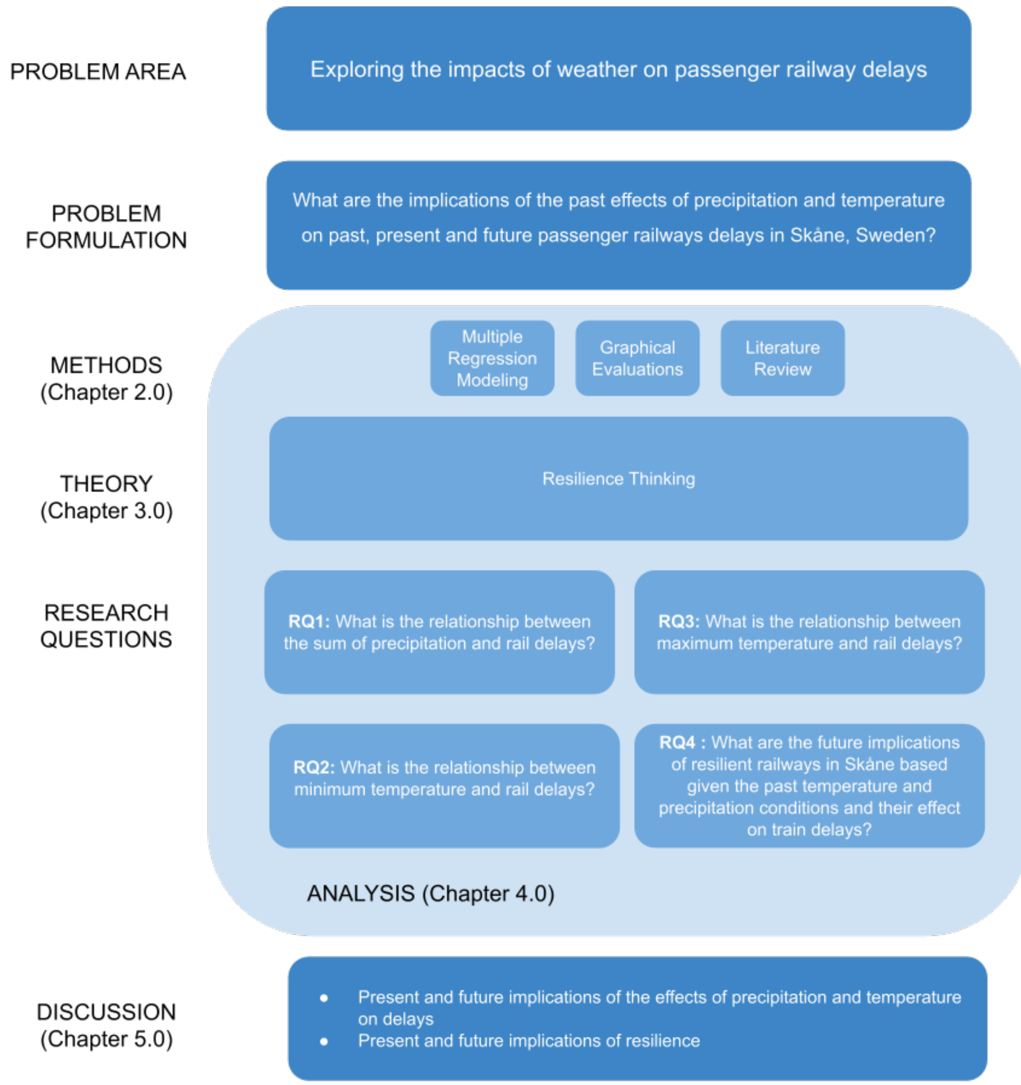
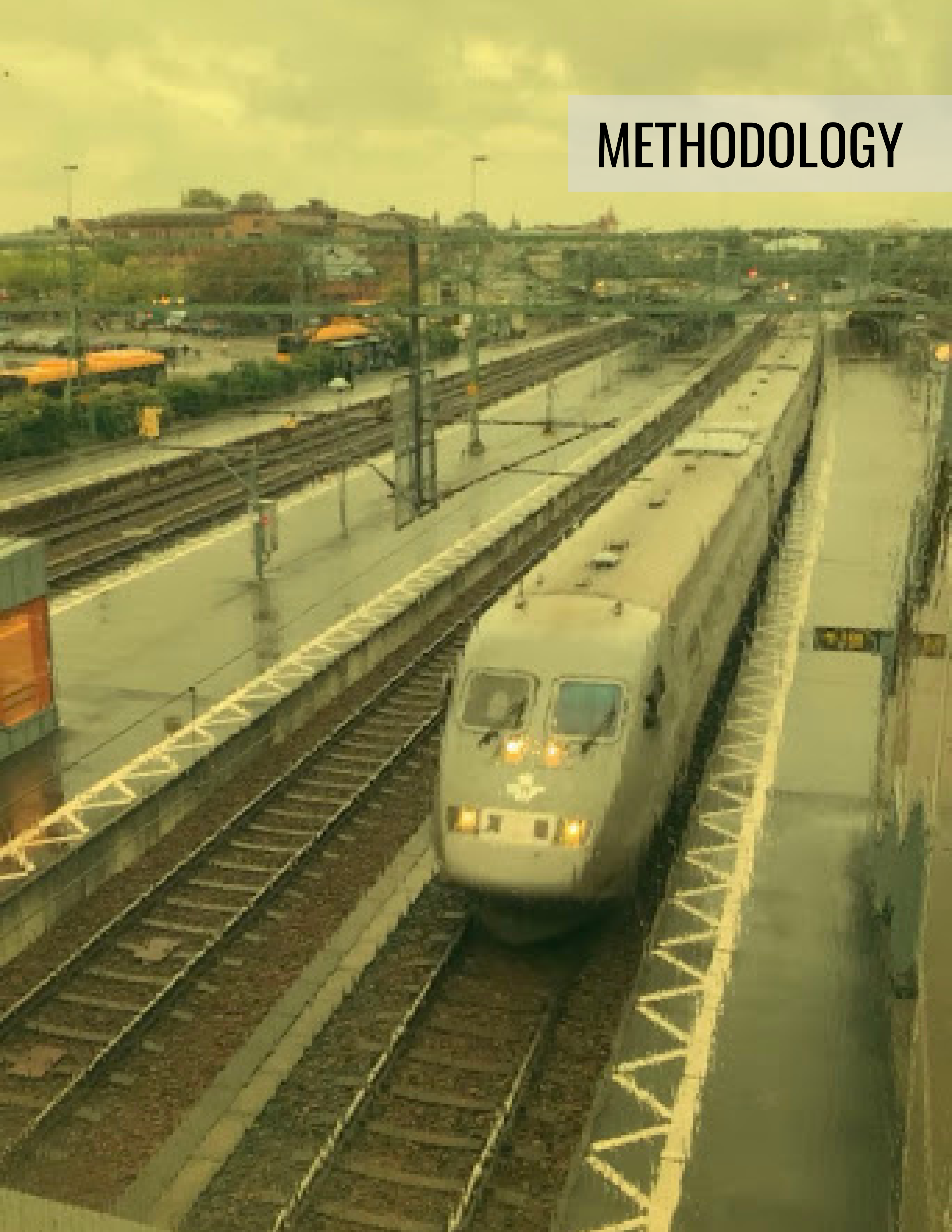


Figure 1. Structure of Thesis Demonstrating the Methods and Theories Used to Answer the Problem Formulation and Subsequent Research Questions

Research questions 1-3 are answered in Chapter 4.0. They are asked in order to determine the statistical relationship between precipitation, temperature, and dwell and run delays. First, graphical evaluations are used in order to visually interpret how these weather variables impact dwell and run delays. Next both multiple linear regression and multiple logistic regression models were used to determine which variables are the most statistically significant to dwell and run delays and the models were compared. These questions are asked in order to determine how the sum of precipitation, minimum temperature, and maximum temperature over 1, 7, 14, 21, and 28 days shaped both dwell and run delays in Skåne between 2008-2019. Understanding the past effects of weather helps to better determine the future implications of climate change on railways.

Question 4 is answered in Chapter 5.0 and is used as a more reflective question to understand how future climate change projections will shape passenger train delays in Skåne.

METHODOLOGY



2.0 Methodology

The following chapter will expand on the quantitative methods used to answer the research questions of this thesis. It also describes the reasoning behind the choice of methods. It describes the datasets used in this thesis, the variables which were analysed, and the methods for the analysis. Method choices were influenced by past experiences and lessons learnt throughout the course of the master's studies at AAU as well as by collaborating partners at Lund University.

2.1 Research design: Case study

The knowledge for this thesis is produced through a quantitative case study which is focused on one specific geographical location and examines the presumed causal relationship between weather and rail delays. Here the case is focused on the railway operations in Skåne, the southern region of Sweden. A case study was chosen as it pertains to detailed, in-depth, context-based knowledge about a specific place or setting used to answer the research questions in a report (Flyvbjerg, 2006). Here a case study method enables the researcher to examine the data within a specific context and explore relationships within the data (Yin, 2009). Yin (2009) describes case studies as being either exploratory, descriptive, or explanatory. It can be argued that this case study is descriptive as it sets out to describe the natural phenomena that occurs within the dataset. This study is interested in quantitatively describing the relationship between the studied weather variables and rail delays that occur in a specific geographical region, Skåne, Sweden. Chapter 3.0 will go into more detail about other studies that have been conducted within this topic in different areas and describe how weather can affect rail delays.

This descriptive case study can be further defined as a longitudinal study. With a longitudinal study, research is conducted over a longer period of time (Yin, 2009; Bryman, 2012). This study uses a dataset of precipitation, minimum temperature, maximum temperature, dwell delays, and run delays over 11 years from 2008-2019. This time period was chosen in order to understand the effects of weather on rail delays during the 2000s. Having data over a longer period of time allows for identifying trends in the data and the past results can be used in order to infer what may happen in the future (Bryman, 2012). In this case, understanding the past relationships between weather and delays can assist in understanding how climate change may future shape this relationship.

2.1.1 The Case of Skåne

Today, more than half of the world's population live in cities, and this number is expected to continuously grow (United Nations, n.d.). This means that more people will be moving to urban areas and the demand for railways is likely to increase. For instance, Skåne is Sweden's most southern county, and the population accounts for about 13% of Sweden's total population (European Union, n.d.). Malmö is the largest urban area in Skåne, and Sweden's third-largest city (European Union, n.d., Figure 2). Southern Sweden is also often considered as part of Greater Copenhagen (Greater Copenhagen, n.d.). This area comprises of about 4.3 million people living in both Southern Sweden and Eastern Denmark (Greater Copenhagen, n.d.). Approximately 15,000 people commute daily between Skåne and Eastern Denmark (Region Skåne och Helsingborg Stad, 2017), making it a dense transportation area. Denmark is not considered in this thesis; however, it is important to note how densely populated the region is and to highlight the number of commuters from Denmark to Sweden that are affected when rails are delayed.

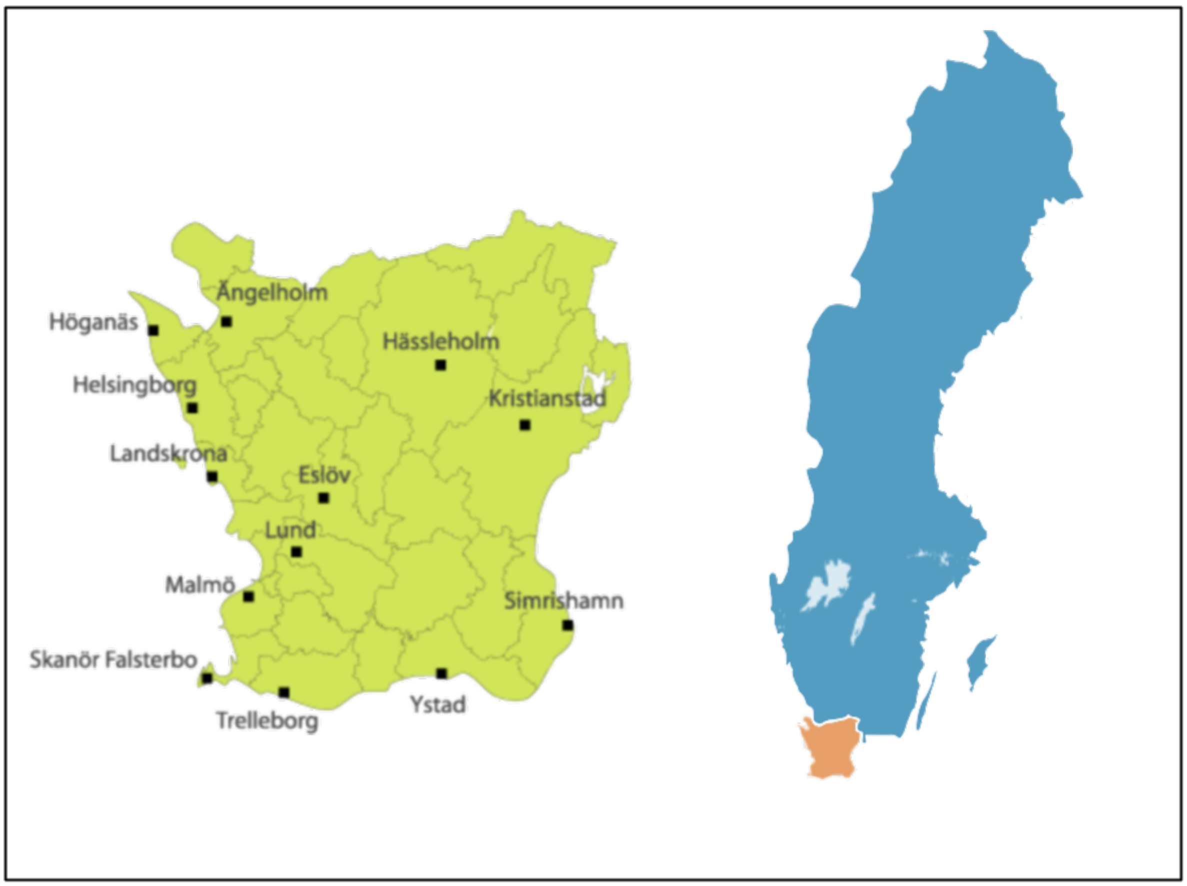


Figure 2. Map of Skåne zoomed in and in comparison, to all of Sweden (Regionfakta, 2018)

The largest commuter flows are between Lund and Malmö and vice versa, as well as between other suburbs of Malmö (Region Skåne och Helsingborg Stad, 2017). Commuter flows in Skåne are continuously increasing (Region Skåne och Helsingborg Stad, 2017). In Lund, Helsingborg, and the surrounding municipalities of Malmö between 63-77% of people commute outside of their municipality for work (Region Skåne och Helsingborg Stad, 2017). In 16 of 33 municipalities in Skåne, more than 50% of the population commute to another municipality for work (Region Skåne och Helsingborg Stad, 2017). For these reasons Skåne is chosen for the case study. It is one of Sweden's most densely populated regions and its close proximity to Denmark leads to heavy commuter flows. As the lines are mostly electrified, rails provide a sustainable choice of transportation for the people living in this area. Therefore, it is important to research the effects weather can have on train delays to ensure people are arriving at their destinations on time and will continue to use railways in the future. It is important to keep trains punctual and reliable in order for ridership to continue to increase.

Currently, there are about 71 active train stations in Skåne (Region Skåne och Helsingborg, 2017). Pågatågen is the main railway system running around Skåne and is run by the public transport administration Skånetrafiken, while Öresundståg and Krösatågen connect Skåne to other regions of Sweden and Denmark. SJ is the main train operator within the rest of Sweden and a few high-speed lines connect Skåne to other major Swedish cities (Stockholm and Göteborg), and to Copenhagen, Denmark (Figures 3-4).

Figure 3. Map of the Railway System in Skåne, Highlighting Påga-, Öresunds, and Krösatågen Stations and Lines (Skånetrafiken, 2020)

Figure 4. SJ High-speed Rail Lines in Sweden and Denmark (EU Rail, n.d.)

Skåne has a milder climate compared to the rest of Sweden due to its geographic location (Figure 5). The figure below highlights the climate normals data between 1961-1990. Climate Normals are known as the average value of monthly climate variables, such as temperature or precipitation, that are calculated over a period of 30-years (SMHI, 2020). They are used to describe the current climate in order to help gain an understanding of how our climate is changing (SMHI, 2020). Additionally, they can also be used to describe the occurrence of extreme values in a month, for example, the minimum and maximum temperature (SMHI, 2020).

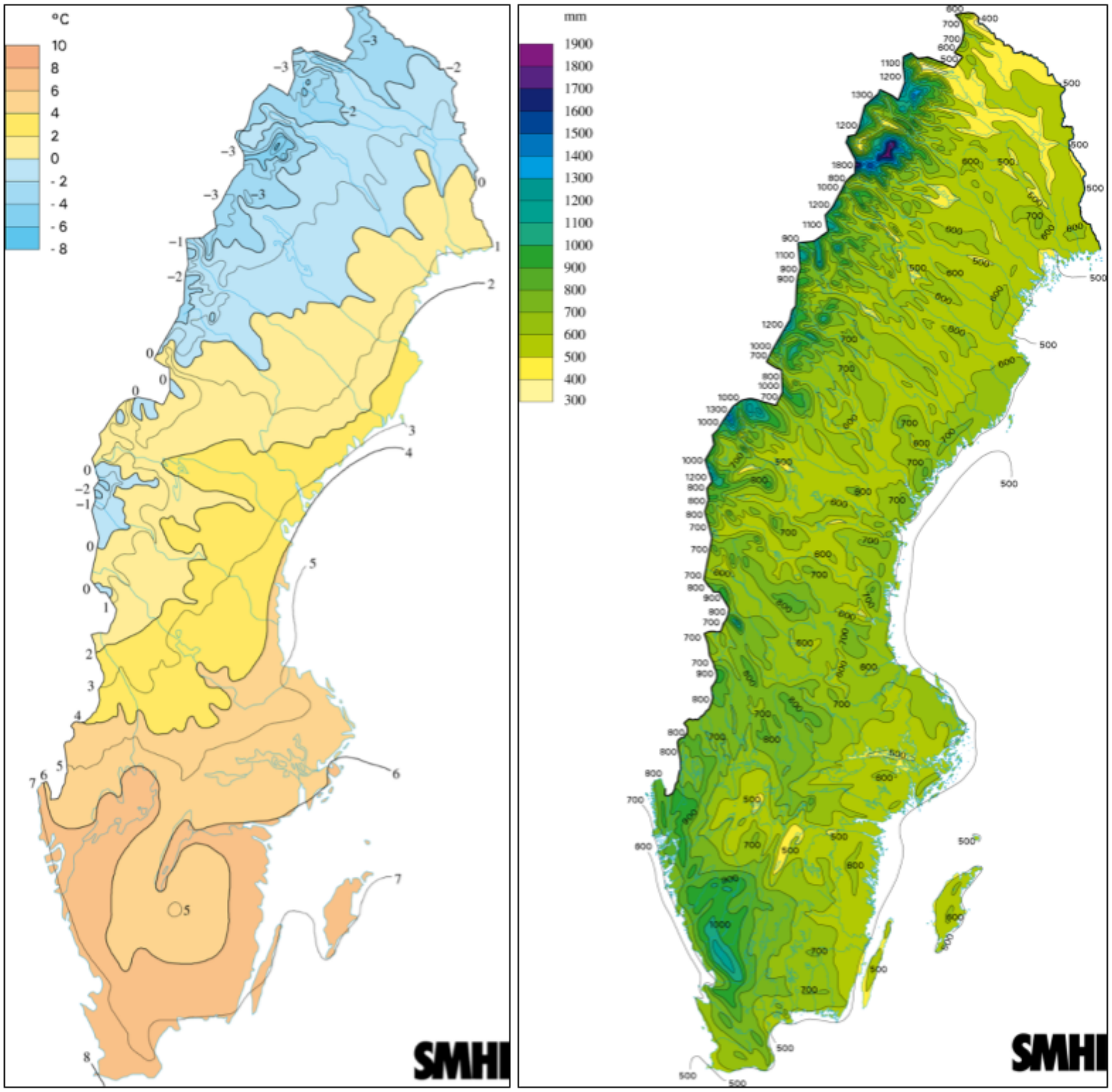


Figure 5. Climate Normals Data (1961-1990) for temperature and precipitation for Sweden. (SMHI, 2015)

Based on the information above, the average annual temperature in Skåne is around 7.2˚C. The average annual precipitation measured is roughly between 500-700mm.

Climate Change Projections in Skåne, Sweden

Average temperatures in Europe are continuing to increase, with warming rates found to be most prevalent in the high latitudes of Northern Europe (IPCC, 2014a). Scandinavia has been experiencing some of the strongest warming since the 1980s, especially during the winter months (IPCC, 2014a). Precipitation rates have also been increasing in Northern Europe (IPCC, 2014a).

It is important to acknowledge that climate is changing and to understand how it will change in order for railways to be more resilient.

The Swedish Meteorological and Hydrological Institute (SMHI) provides climate data and projections for Sweden. In 2015, they released a report describing the current and future climate scenarios in Skåne based on observations and modelling (SMHI, 2015). The climate projections are based on the Representative Concentration Pathway (RCP) 4.5 and RCP8.5 scenarios. An RCP scenario is a greenhouse gas concentration trajectory that was adopted by the Intergovernmental Panel on Climate Change (IPCC) (IPCC, 2014b). RCP8.5 represents a high emissions scenario, while RCP4.5 represents an intermediate effort to curb greenhouse gas emissions. Historically, between 1961-1990, the annual average temperature in Skåne was around 7.2˚C. At the beginning of the 2000s, there was 1-2 degrees increase experienced throughout Skåne, and projections show an additional 2-4 degrees increase towards the end of the century under the RCP4.5 scenario or 4-6 degrees under the RCP8.5 scenario. Winters have become milder over the past 23 years and this trend is predicted to continue in the future. Summer temperatures are also expected to continue to increase in the future. The report also discusses heat waves, where the number of consecutive days above 20˚C is expected to increase, with areas in the north of Skåne being affected the most (SMHI, 2015).

Under the RCP4.5 scenario, precipitation is expected to increase around 15% by the end of the century compared to the average annual rainfall of 748mm in Skåne during 1961-1990 (SMHI, 2015). Currently, during the winter months, precipitation falls mostly as rain but also occasionally as snow. Precipitation amounts are expected to increase, and due to the expected temperature increase, it is predicted that less precipitation will fall as snow. In the summer, precipitation is also expected to increase slightly, but not as much as winter months. In comparison, the RCP8.5 scenario predicts a 25% increase in precipitation (SMHI, 2015). With the expected increase in precipitation due to climate change across Skåne, it is expected that the number of days with more than 10mm of precipitation, which can lead to flooding, will also increase, with the greatest increase found in the north of Skåne (SMHI, 2015).

Based on the 2015 SMHI report, it is evident that the climate is continuously changing, and therefore it is important to understand how this will shape the operating conditions for the railway industry. Although the report is focused on future climate change, it also mentions how climate has already been significantly changing since the year 2000, further reiterating the reasoning for choosing 2008-2019 as the study period.

2.2 Literature Review

Bryman (2012) highlights the importance of searching through existing literature and writing a literature review once the research questions have been identified. The main tasks of a literature review include reviewing studies and theories in order to gain a foundation and understanding of the studies and literature that have already been conducted in the specific field of interest. The main intent of reading through the existing literature is to be able to recognise what has already been established in this area, what methods have been used, and what theories and concepts are relevant to the individual study. Additionally, a literature review allows for distinguishing any inconsistencies, controversies, and to identify any gaps in the research.

Accordingly, the literature is used to understand what past research has been done and what is already known within this field of research. Additionally, the literature review gave inspiration for what quantitative methods could be used and how the research could be conducted in order to answer the research questions. Lastly, the literature review served as an important piece of understanding climate change projections in Skåne, in order to infer how the results will change in the future.

Literature for this thesis was found mainly using the Aalborg University (AAU) and Lund University (LU) Library databases. Within these databases the Journal of Transport Geography, Natural Hazards, Transport Policy, and Transportation Research A, D, E, and Procedia were the most frequently used journals to look for articles. Furthermore, some papers were already found during the author's internship at Lund University from September-November 2020, and some additional resources were provided by the collaborating partner. Keywords that were used when searching through articles in the database included: railways, delays, weather, climate, temperature, precipitation, flooding, climate change, and adaptation. Once the initial list of relevant literature was established, the reference list for each paper was also searched, and if a reference was recognised in multiple papers' reference lists and literature reviews then that paper was flagged as one of importance to read.

2.3 Datasets

The analysis used in this thesis was based on a database containing two main datasets. The first dataset included all train movements in Skåne between 2008 and 2019. It includes the date, train mission, station name, station signature, the type of delay, scheduled arrival time, actual arrival time, and the duration of the delay. The two types of delays studied in this thesis are dwell and run delay. This dataset was obtained by Trafikverket and covers over 40,637,498 passenger train passages (Figure 6).

The second dataset includes historical meteorological observations of precipitation, minimum temperature, maximum temperature in Skåne which was downloaded from SMHI on February 5, 2021 (SMHI, 2021). Stations decommissioned before 2008 were not included (Table 1). It is also important to note that there were some discrepancies in the weather data when during some hours or days data was not collected, which resulted in a NULL value. NULL values were filtered out from the main dataset worked with. In addition, for unknown reasons there are more precipitation stations commissioned compared to air temperature stations. Only passenger trains were included in this analysis due to the differences in timetable planning, and traffic control between passenger trains and service and freight trains (Palmqvist, Olsson, & Hiselius, 2017b).

Table 1. Overview of the Weather Data Downloaded

Weather Variable	Unit	Number of Observations	Observation Frequency	Number of Weather Stations
Minimum Temperature	˚C	1,810,677	Hourly	20
Maximum Temperature	˚C	1,810,677	Hourly	20
Precipitation	mm	311,051	Daily	51

Since trains are often travelling long distances, they travel through various weather conditions (Palmqvist, Olsson, & Hilelius, 2017b). In order to account for this and since the locations of meteorological observation stations typically differ from the train stations, an algorithm was created which matches each train station to the nearest meteorological station (Palmqvist, Olsson, & Hileius, 2017b).

To manage these datasets, the Azure Data Studio database tool was used. Azure Data Studio uses Structured Query Language (SQL) to manage the databases. SQL is a domain-specific language which is used in programming and to manage data in database management systems. In this thesis Azure Data Studio was used to combine the original 40+ million passenger train observations with weather data to create one dataset containing 13,576,723 rows of information (Figure 6), including date, train mission, station signature, the type of delay, the size of delay, a binomial system to indicate if there was a delay, and all weather variables over 1, 7, 14, 21, and 28 days. 3,621,851 of the observations were dwell delays and 9,954,872 were run delays (Figure 6). These time periods were chosen as they are multiples of 7 which also coincide with one day, week, two weeks, and one month. These time periods help understand the effects of events such as intensive precipitation, heatwaves, and cold snaps on delays by analysing what was the sum of precipitation, minimum temperature, and maximum temperature 1, 7, 14, 21, and 28 days before a delay. Longer time periods allow for more understanding of the effects of extreme precipitation, heat, and cold on rail delays.

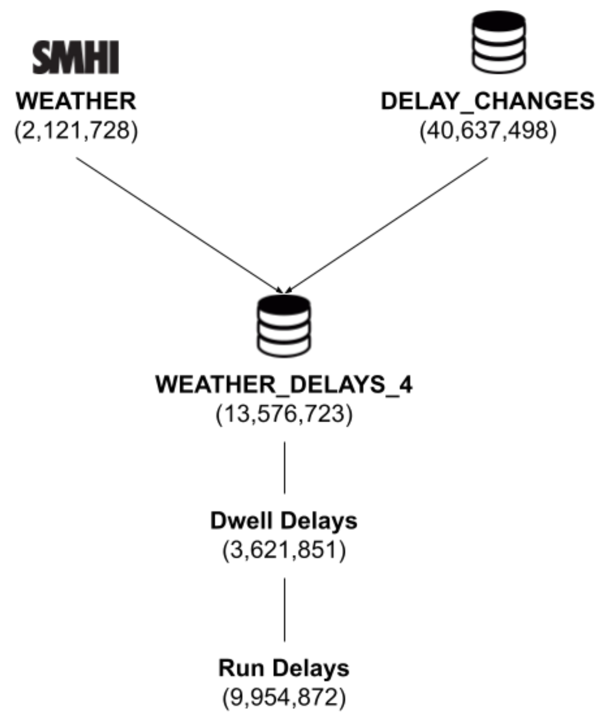


Figure 6. Data Collection Process. From original weather and delay observations to one table entitled weather_delays_4 which was processed in SQL and used for regression modelling in R Studio

2.4 Graphical Evaluations

The probability of both a run delay and dwell delay under each recorded weather variable was calculated using Azure Data Studio. Using the weather_delays_4 table in Azure Data Studio the probability of a dwell and run delay was calculated for each sum of precipitation, minimum temperature, and maximum temperature value observed over 1, 7, 14, 21, 28 days. Next, these tables were exported into Microsoft Excel. The number of observations differed for each value, for example there were more observations of 10˚C than for -19˚C. Therefore, in Excel the probability was converted to cumulative probability. Lastly, cumulative probabilities were plotted against each weather variable for each time period in order to visually analyse the trends and therefore relationship between weather and delays.

2.5 Regression

Both multiple linear and multiple logistic regressions were used in order to determine the statistical relationship between weather and delays. A regression in simple terms is a statistical model that analyses the relationship between a response variable and one or more predictor variables and their interactions (James, et al., 2013). The independent, or predictor variables are the weather variables, while the delays are the dependent or response variables. R Studio was used to run both the multiple linear regressions and multiple logistic regressions. R Studio is an open-source software which uses the R programming language to perform descriptive statistical analyses.

2.5.1 Multiple Linear Regression

A multiple linear regression model generally explains the relationship between multiple predictor or independent variables and one dependent or response variable (James, et al., 2013). A dependent variable is modeled as a function of multiple independent variables with corresponding coefficients (James, et al., 2013). With a linear regression modelling several important questions are being asked: what is the relationship between weather and passenger train delays? How strong is this relationship? And how accurately can the weather variables predict train delays? The standard equation for a multiple linear regression is:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon \quad (1)$$

Where β_0 and β_n represent two unknown constants which indicate the intercept and slope terms in a linear model. Additionally, x_n represents the predictor variables and Y is the response variable which is predicted on the basis of x (James, et al., 2013).

In a multiple linear regression, the null hypothesis indicates that there is no relationship between the independent and dependent variables (James, et al., 2013):

$$H_0 : \beta_1 = 0 \quad (2)$$

Versus the alternative hypothesis, which indicates that there is a relationship between the independent and dependent variables (James, et al., 2013):

$$H_0 : \beta_1 \neq 0 \quad (3)$$

The variables used in this multiple linear regression model are summarised in Table 2 below.

Table 2. Summary of the Variables in the Multiple Linear Regression Model (4)

Variables	Values	Description
Dwell or Run Delay	Y	Size of the delay in minutes
Sum of Precipitation_1	x_1	The sum of precipitation over 1 day in mm
Sum of Precipitation_7	x_2	The sum of precipitation over 7 days in mm
Sum of Precipitation_14	x_3	The sum of precipitation over 14 days in mm
Sum of Precipitation_21	x_4	The sum of precipitation over 21 days in mm
Sum of Precipitation_28	x_5	The sum of precipitation over 28 days in mm
Minimum Temperature_1	x_6	Minimum temperature over 1 day in ˚C
Minimum Temperature_7	x_7	Minimum temperature over 7 days in ˚C
Minimum Temperature_14	x_8	Minimum temperature over 14 days in ˚C
Minimum Temperature_21	x_9	Minimum temperature over 21 days in ˚C
Minimum Temperature_28	x_{10}	Minimum temperature over 28 days in ˚C
Maximum Temperature_1	x_{11}	Maximum temperature over 1 day in ˚C
Maximum Temperature_7	x_{12}	Maximum temperature over 7 days in ˚C
Maximum Temperature_14	x_{13}	Minimum Temperature over 14 days in ˚C
Maximum Temperature_21	x_{14}	Minimum Temperature over 21 days in ˚C
Maximum Temperature_28	x_{15}	Minimum Temperature over 28 days in ˚C

Two multiple linear regression models were run in this thesis. One for dwell delay and one for run delay with the predictor variables of the sum of precipitation, minimum temperature, and maximum temperature over 1, 7, 14, 21, and 28 days. The summary of this model is presented in the following equation

$$Y = \beta_0+\beta_1 x_1+\beta_2 x_2+\beta_3 x_3+\beta_4 x_4+\beta_5 x_5+\beta_6 x_6+\beta_7 x_7+\beta_8 x_8+\beta_9 x_9+\beta_{10} x_{10}+\beta_{11} x_{11}+\beta_{12} x_{12}+\beta_{13} x_{13}+\beta_{14} x_{14}+ \beta_{15} x_{15}+\varepsilon \quad (4)$$

Where Y is the size of dwell or run delay in minutes, β_i is the slope associated with each predictor variable (x_i), and β_0 is an intercept in the evaluated model.

2.5.2. Multiple Logistic Regression

In addition to a multiple linear regression model, a multiple logistic regression model was also performed in order to compare against graphical evaluations and determine the best suited model for this thesis. Logistic regressions are used to model the probability that the response variable (Y) belongs to a particular category rather than modeling the response of the response variable (Y) directly (James, et al., 2013). The response variable is dichotomous, coded as 1 in the presence of an outcome of interest, or 0 in the absence of the outcome of interest. The general logistic regression model which predicts the log of outcome Y is described in the formula below

$$\log(Y) = \left(\frac{(p(X)}{1-p(X)}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n \quad (5)$$

Here the coefficients β_i for the two category predictors is estimated by the maximum likelihood and are defined by the odds $\left(\frac{(p(X)}{1-p(X)}\right)$ The odds describe the likelihood of the occurrence of an event and is expressed as the probability of an occurrence versus the event not occurring. The odds were used to calculate the odds ratio, which is the ratio of one odds against the other. Equation 1 from the linear regression model is very similar to equation 5 here with the added log. Compared to the linear regression model where β_i indicates the average change in Y associated with a one-unit increase in X; a logistic regression infers that increasing X by one unit will change the log odds by β_i (James, et al., 2013). In simpler terms it multiplies the odds by $e^{\beta i}$.

Similarly, to the multiple linear regression the null hypothesis indicates there is no relationship between the predictor variables and response variables (James, et al., 2013). While the alternative hypothesis suggests there is a relationship between the two types of variables (James, et al., 2013). The variables used in this multiple logistic regression model are summarised in Table 3 below.

Table 3. Summary of the Variables in the Multiple Logistic Regression Model (7)

Variables	Values	Description
Dwell or Run Delay	$Y\begin{cases}0\\1\end{cases}$	No dwell or run delay Dwell or run delay present
Sum of Precipitation_1	x_1	The sum of precipitation over 1 day in mm
Sum of Precipitation_7	x_2	The sum of precipitation over 7 days in mm
Sum of Precipitation_14	x_3	The sum of precipitation over 14 days in mm
Sum of Precipitation_21	x_4	The sum of precipitation over 21 days in mm
Sum of Precipitation_28	x_5	The sum of precipitation over 28 days in mm
Minimum Temperature_1	x_6	Minimum temperature over 1 day in ˚C
Minimum Temperature_7	x_7	Minimum temperature over 7 days in ˚C
Minimum Temperature_14	x_8	Minimum temperature over 14 days in ˚C
Minimum Temperature_21	x_9	Minimum temperature over 21 days in ˚C
Minimum Temperature_28	x_{10}	Minimum temperature over 28 days in ˚C
Maximum Temperature_1	x_{11}	Maximum temperature over 1 day in ˚C
Maximum Temperature_7	x_{12}	Maximum temperature over 7 days in ˚C
Maximum Temperature_14	x_{13}	Maximum temperature over 14 days in ˚C
Maximum Temperature_21	x_{14}	Maximum temperature over 21 days in ˚C
Maximum Temperature_28	x_{15}	Maximum temperature over 28 days in ˚C

Two multiple logistic regression models were run in this thesis. One for dwell delay and one for run delay with the predictor variables of the sum of precipitation, minimum temperature, and maximum temperature over 1, 7, 14, 21, and 28 days. The summary of this model is presented in the following equation

$$\log(Y)=\left(\frac{(p(X)}{1-p(X)}\right)=\beta_0+\beta_1x_1+\beta_2x_2+\beta_3x_3+\beta_4x_4+\beta_5x_5+\beta_6x_6+\beta_7x_7+\beta_8x_8+\beta_9x_9+\beta_{10}x_{10}+\beta_{11}x_{11}+\beta_{12}x_{12}+\beta_{13}x_{13}+\beta_{14}x_{14}+\beta_{15}x_{15} \tag{7}$$

Where Y is represented as the probability of a delay to occur, β_i is the change in the log odds associated with each predictor variable x_i, and β_0 is an intercept in the evaluated model. In this model Y is coded as either 1, the presence of a dwell or run delay, or 0, the absence of a dwell or run delay.

SCOPE OF THE THESIS

3.0 Scope of the Thesis

The climate in Skåne has be changing since the 2000s and is expected to continuously change in the future (SMHI, 2015). Although railways play a great role in supporting a modal shift towards more sustainable transportation, their infrastructure and systems are already vulnerable to changes in weather and are becoming more vulnerable due to climate change. This vulnerability in turn impacts the ability of trains in Skåne to arrive to stations on time and calls for railways to increase their resilience. In addition, there is less of a comprehensive research focus on the impacts of weather and climate on railways compared to other sectors, such as air or road (Aparicio, et al., 2013). Therefore, it is important to understand the relationship between weather and rail delays to highlight the current challenges Skåne railways face in maintaining punctual operations, and therefore understand how to ensure railways are more resilient to disruptions in the future.

3.1 Resilience Thinking

For the past 40 years, resilience has been a topic of research and discussion used in many different fields with many different definitions (Fleming & Ledogar, 2008). Even within a field, there is variation in how resilience is used and defined (Fleming & Ledogar, 2008). There are over 70 different definitions of resilience found in the literature, and they mostly revolve around a cycle of, "*disruption, response, absorption, recovery, and learning*" (Chan, & Schofer, 2016, p. 05015004-1). Table 4 below tries to highlight some of the variations amongst and within the different fields.

Table 4. Various Fields of Research and How They Define/Use Resilience

Field of Research	**Definition**
Psychology and Psychiatry	• Having a positive adaptation to something despite enduring adversity (Fleming & Ledogar, 2008).
Engineering	• A resilient system is one that has an elastic response to external forces (Chan & Schofer, 2016). • A resilient engineered system is one that can bounce back from a disruption (Chan & Schofer, 2016).
Governments responsible for infrastructure resilience	• A resilient system is one that can bounce back stronger after a disruption with redundant and diverse designs in an effort to reduce the exposure to risk (Chan & Schofer, 2016). • Resilience includes mitigation of the impact of disruption through absorption, adaptation, and recovery (Chan & Schofer, 2016).
National Security	• Resilience is the ability to be prepared and adapt to changes and the ability to withstand and recover from deliberate accidents, attacks, threats, or incidents quickly (Chan & Schofer, 2016).
The IPCC	• "*A resilient system is one that can cope with hazardous events through response, adaptation, and learning*" (Chan & Schofer, p. 05015004-2).
Transportation	• The ability of a system to bounce back from a disruption (Chan & Schofer, 2016). • Coping preservation and capacity of tactical options after a disruption (Chan & Schofer, 2016). • The UK Department of Transportation defines resilience as "*the ability of the transport network to withstand the impacts of extreme weather, to operate in the face of such weather, and to recover promptly from its effects*" (Diab & Shalaby, 2020, p. 658).
The United Nations	• "*The ability of a system, community or society exposed to hazards to resist, absorb, accommodate to, and recover from the effects of a hazard in a timely and efficient manner, including through the preservation and restoration of its essential basic structures and functions"* (Bešinović, 2020, p.460).

Although there are many different definitions and applications of resilience, it is evident that a resilient system can return to its original state to some extent, after a disruption in a positive way. The focus of this thesis is on resilience thinking within the transportation sector.

3.1.1 Resilience in the Transportation Sector

Transportation is a type of critical infrastructure that is essential for the functioning of an economy and society (Bešinović, 2020). As transportation demands increase, the amount of congestion in railway networks is becoming more complicated to operate (Bešinović, 2020). As a result, urban mobility is becoming more fragile to unexpected changes in the network (Bešinović, 2020). Such changes may include disruptions due to weather, engineering works, infrastructure faults, disturbances due to daily variations in the operations, or disasters such as earthquakes or floods (Bešinović, 2020). Although it is recognised that transportation management and planning should improve their ability to bounce back from disruptions, disturbances, and disasters it is still challenging to determine how to address and identify the appropriate measures (Bešinović, 2020). This is mostly due to the lack of quantitative research to understand the effect these factors have on transportation (Bešinović, 2020). Weather is an important factor to look at because it is closely related to other factors such as infrastructure faults. For example, if it is extremely hot the track may buckle leading to a delay caused by weather which then impacts infrastructure. In countries such as the Netherlands, an increase in transportation demand has been leading to an increase in both the number of disruptions and the total duration of those disruptions (Bešinović, 2020), and it can be expected that other countries see a similar trend; therefore, highlighting the importance of considering resilience. The transportation system also consists of many subsystems; and sometimes one subsystem is able to cover for another subsystem and to some extent even reduce vulnerability (Mattsson & Jenelius, 2015). For instance, when the ashes of the 2010 volcanic eruption, Eyjafjallajökull, brought air traffic in Northern Europe to a halt, the road and rail transport systems provided an alternative option for many travellers (Mattsson & Jenelius, 2015).

There are fewer studies with a focus on the resilience of railways compared to other modes of transportation (Bešinović, 2020), which is something this thesis attempts to address and aims to fill this gap. For the purpose of this thesis, resilience thinking is narrowed down to uses related to railways. Additionally, this thesis uses Bešinović (2020)'s definition of the resilience of a railway transportation system, which is defined as "*the ability of a railway system to provide effective services in normal conditions, as well as to resist, absorb, accommodate, and recover quickly* *from disruptions or disasters*" (Bešinović, 2020, p. 461). The disruptions that are considered in this thesis are those related to weather, specifically temperature and precipitation. Figure 7 highlights the resilience of a railway transportation system.

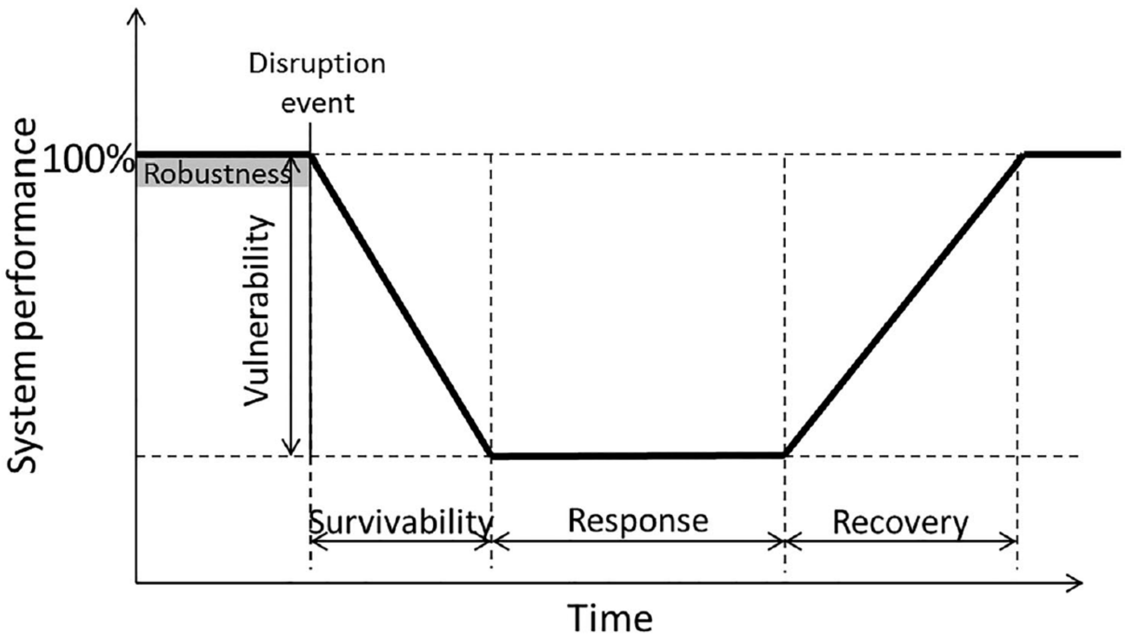


Figure 7. The Resilience of a Railway Transportation System (Bešinović, 2020)

The figure above indicates how a disruption event moves through a resilient railway transportation system as described by Bešinović (2020). **Vulnerability** is often referred to by how susceptible a system is to a disruption. Related to vulnerability is **robustness**, which has a more specific definition in railway infrastructure compared to other transportation modes. It can be defined as the "*ability to mitigate from various everyday delays caused by disturbances"* (Bešinović, 2020, p. 461). **Survivability** is the ability of a system to move from a normal state to a disrupted one in a steady manner. The system can fail either completely all at once (for example a power outage resulting in all electrified trains to come to a halt), or the system can fail gradually, slowly reaching its disrupted state. The **response** is typically known as the set of actions that are taken immediately after a disruption in order to provide the best possible service during the disruption, it may be ensuring public safety, providing alternative travel routes, etc. Lastly, **recovery** refers to the ability and the amount of time it takes the system to return to its original state. Depending on the type of disruption some stages may be completely omitted, or some may be prolonged.

Bešinović (2020) also describes two different methods to measure resilience: topological and system-based. Topological metrics emerge in complex network theory. Commonly, this is done by looking into the topological structure of a network and assessing its structural characteristics and assuming a failure of one component in the disrupted network while ignoring all the dynamic features within the system. On the other hand, system-based metrics have been receiving more attention lately as they overcome the limitations of graph methods by representing the supply and demand of a system and the system's response to a disruption and recovery from it. There are also four approaches to quantify resilience: data-driven, topological, simulation, and optimisation. Data-driven approaches use statistical methods to quantify the effects of a disruption that occurs on a transportation system. An example is quantifying the weather-related disruptions in a railway network. Topological approaches use complex network theory to assess different nodes individually in a system, for example when studying the vulnerability of metro networks in various cities. A simulation approach uses similar metrics as in topological approaches but also uses performance indicators in order to evaluate a network's performance in a stochastic environment. Methods used for identifying link vulnerability from the perspective of passengers is an example. Lastly, optimisation approaches use mathematical optimisation models in order to assess the resilience of railway networks, for example when determining the most critical elements in a network.

3.1.2. How Resilience is Used in This Thesis

It is evident that there are many different definitions of resilience and that it is used differently depending on the field or even within a specific field. Chapter 3.1.1 highlighted the ways resilience can be quantified and measured within the transportation sector and outlined the definition of resilience that will be used for this thesis. The disruptions that are studied in this research are those related to temperature and precipitation. First, it is important to quantify the past effects of weather on the railway system to get a better understanding of the baseline. Then we can learn from the past to better prepare for the future back on the climate change projections in Skåne.

This thesis uses the system-based metrics and quantifies resilience using the data-driven approach as described in Chapter 3.1.1 to frame the analysis in Chapter 4.0. Fewer studies focus on the resilience of railways compared to other modes of transportation (Bešinović, 2020). Therefore, the analysis of this thesis aims to evaluate resilience based on the past precipitation and temperature conditions using system-based metrics and a data-driven approach. In passenger railway networks the performance is typically evaluated based on the trains service adaptations and/or passenger discomfort or changes (Bešinović, 2020). To accomplish this many studies in the past have focused on measuring transport capacity, or cancellations and delays that were imposed on both rail operators and passengers (Bešinović, 2020). In this thesis, the dwell delay and run delay times due to the disruptions caused by temperature and precipitation are measured. Bešinović (2020) highlighted the importance of using more system-based metrics to capture the effects on transportation in order to obtain a more accurate assessment of resilience. Moreover, in order to quantify the resilience, a data-driven approach is taken. This is achieved by using regression modelling to derive the statistical relationship between the weather variables studied and the delays. Here the aim is to quantify the effects of weather disruptions on the railway network in Skåne, Sweden.

3.2 Previous Research on Punctuality and Climate

As mentioned, this thesis studies the effects of weather on rail delays in Skåne as weather-related research should gain more attention especially as the effects of climate change are becoming more understood (Bešinović, 2020). The following sub-chapter will highlight what is already known about this study area and how it applies to this research. First this sub-chapter will discuss the overall issue of punctuality and delays and then move into how weather can affect delays.

3.2.1 Punctuality and Train Delays

As previously mentioned, multiple factors impact the punctuality of railways. The focus of this thesis is the weather, but it is important to take into consideration that other aspects may influence delays such as infrastructure faults, passenger behaviour, disturbances due to operational issues. Additionally, various aspects may influence delays together. For instance, extreme low temperatures may disturb the rail infrastructure by causing track separation and brittle tracks (Xia, et al., 2013). This in turn leads to a daily problem caused by both weather and maintenance issues.

The two types of delays analysed in this thesis are dwell delays and run delays, which refer to the delays that occur when a train is waiting at a station and when it is moving from one station to another, respectively. However, it is important to note some other types of delays that may occur when there is a run or dwell delay. For example, a knock-on delay which occurs when trains are close to one another and so if one train is delayed it can easily spread to others (Palmqvist, 2019).

This can have serious implications on passengers commuting to work or trying to catch a connecting train.

Palmqvist, Olsson, & Hiselius (2017b) studied various influencing factors that affect passenger train punctuality in Sweden. They looked at weather, timetable planning, operational, and infrastructure factors. Weather is dependent on the geography of the area being studied, and therefore the weather variables which have the greatest impact on delays can vary between study areas; additionally, findings regarding weather will be addressed later in this sub-chapter. Timetable errors resulted when planning was insufficient; for instance, scheduling run and dwell times that are too short. On an operational level, the boarding and alighting of passengers during dwell times in congested areas can play a role in delays. Infrastructure failures also play a role in the delays. A study by Økland and Olsson (2020) concluded that low temperature, snowfall, reduced train lengths, and an increased volume of train services were the most influential factors on punctuality on Norwegian railways between 2005-2014. These factors differed from the original study conducted by the authors between 2005-2009 which highlighted an increased error in rolling stock and infrastructure, extensive work close to tracks due to maintenance, and the inability to display consistent punctuality even under normal weather conditions. Additionally, the paper considered factors from other studies that affect punctuality such as passenger behaviour, speed restrictions, and development close to tracks (Økland & Olsson, 2020). Overall, based on the literature review to explore which factors influence rail delays it is evident that there are many different factors; and they can all play a role in delays together but really independently (Olsson & Haugland, 2004).

In order for railways to be more punctual Palmqvist (2019) offers some suggestions, including removing switches so the chance of failure is lower, and the remaining switches can be maintained more easily, adding marking on platforms to indicate where passengers to wait in order to speed up the boarding progress, therefore decreasing dwell times, and to adapt railways to cope with today's weather variations and future climate change. Therefore, based on these suggestions, this thesis chooses to tackle the impact of weather. The above displays how complex the field of punctuality is, and that many studies are revolving around the issues related to punctuality. Many factors influence punctuality, and many of these factors are related. As mentioned, this thesis focuses on weather alone, however it is important to note the complexity and interrelatedness between other factors that impact punctuality (Økland & Olsson, 2020; Olsson & Haugland, 2004). This thesis now moves forward with the focus on the effects of weather on railways. Weather is chosen as railways are sensitive to disruptions and if they are unable to operate under the weather conditions of today it is likely that this will worsen as climate change progresses to more extreme weather events and conditions.

3.2.2. The Effects of Weather on Railways

There are many ways that weather can influence railways which ultimately lead to delays. It is important to note that geography has an impact on which weather phenomena are predominantly present in an area and how that influences the railway network in that area. The main quality dimensions for railway transportation include accuracy, reliability, and safety (Leviäkangas, et al., 2011). Extreme weather phenomena may threaten the high level of service railway operators aim for because of the impacts and consequences that affect these three dimensions (Leviäkangas, et al., 2011). The figure below highlights the impacts, consequences to infrastructure/operations, and traffic implications that various weather phenomena have (Ochsner, 2021). The figure also includes how weather events are related to the safety and maintenance of railways; however, the focus of this thesis is the relationship between weather and delays which can ultimately influence customer satisfaction and ridership rates. In addition, the figure shows how connected some weather phenomena are to others, for instance, how a blizzard results in both snowfall and wind gusts.

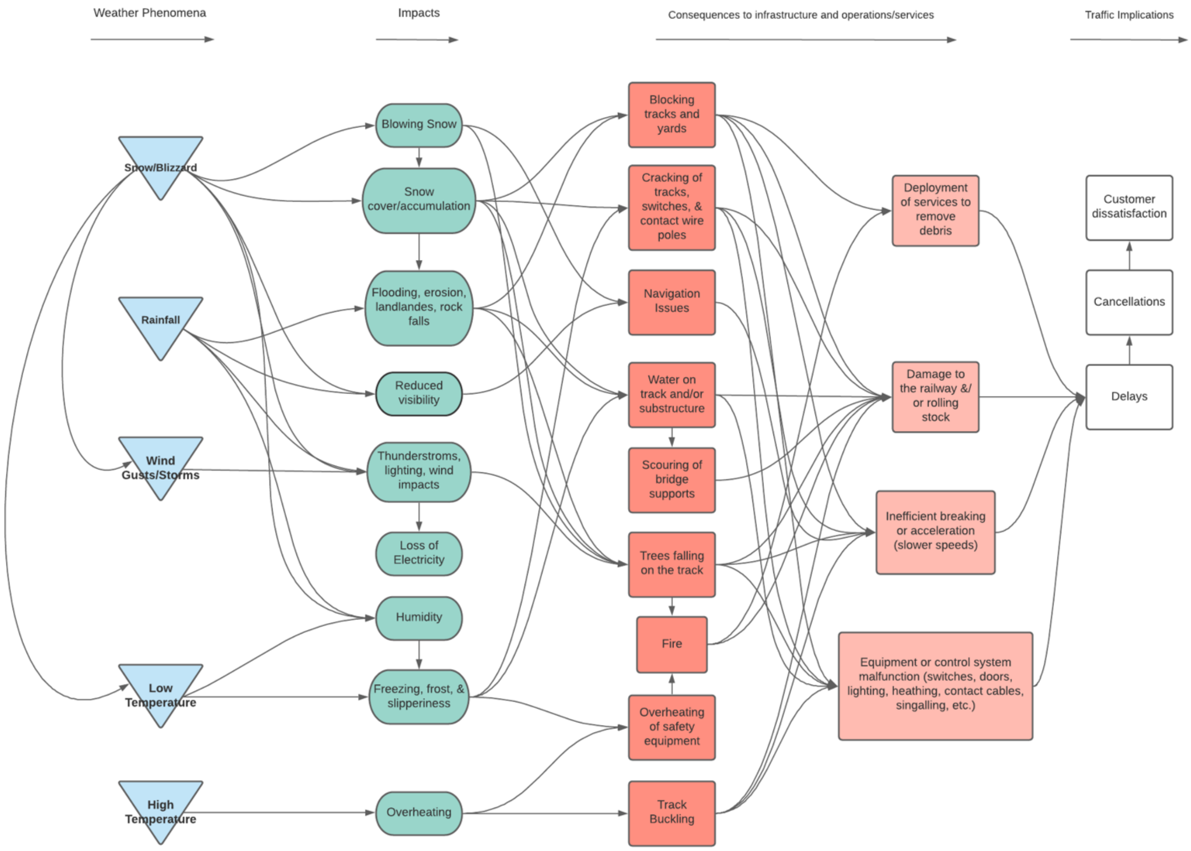


Figure 8. "Different weather phenomena and their impacts, consequences, and traffic implications for railways. Figure adapted from: Kreuz, et al., 2012" and further adapted to fit this region (Ochsner, 2021).

Snowfall can lead to various cascading triggers. On its own, it can accumulate on the track and make it difficult for rails to pass through (Ochsner, 2021). A blizzard is defined as a combination of snowfall, low temperatures, and wind gusts which can lead to snowdrifts, icy conditions, and low visibility (Leviäkangas, et al., 2011). This combination can lead to severe winter storms which can shut down rail traffic for a significant amount of time (Leviäkangas, et al., 2011). Intensive rainfall can lead to flooding of rail tracks (Leviäkangas, et al., 2011). It can also lead to poor visibility resulting in trains having to move at slower speeds (Leviäkangas, et al. 2011). Wind gusts are mainly associated with storms, heavy rainfall, blizzards, etc. (Leviäkangas, et al., 2011). Wind can result in falling trees or other debris which can block rail tracks or can even result in electricity failures during major storms which can cut off telecommunication services (Leviäkangas, et al., 2011; Lindgren, Jonsson, & Carlsson-Kanyama, 2009). Freezing temperatures can result in the formation of ice on rail tracks leading to slippery conditions and trains having to operate at slower speeds, or even track breakages (Oslakovic, et al., 2013). Ice formation can also occur on equipment such as overhead lines, contact wires, switch boxes, etc. resulting in malfunction

(Oslakovic, et al., 2013). Regarding extreme high temperatures, the main consequences of railway infrastructure are overheating (Dobney, et al., 2009; Ochsner, 2021). When the track becomes too hot it runs the risk of buckling; a track buckle is defined as a misalignment in the track which is serious enough to cause a derailment (Dobney, et al., 2009). High temperatures may also lead to the overheating of equipment which can lead to malfunctions and consequently delays, or even fire (Lindgren, Jonsson, & Carlsson-Kanyama, 2009). Extreme high temperatures can also spark a fire in dense vegetation or fallen trees around the railway tracks (Lindgren, Jonsson, & Carlsson-Kanyama, 2009).

In the past, there has been a greater focus on the effects of weather on the road sector, but more attention should be given to the railway sector (Zakeri & Olsson, 2018). As highlighted in Figure 8 above, there are many ways in which weather can lead to a railway delay. With climate change leading to predictions of warmer temperatures and more precipitation the railway's sensitivity to weather is expected to become worse (Armstrong, Preston, & Hood, 2017). Although various weather variables were discussed in this chapter this thesis is focusing on precipitation and temperature alone. Precipitation includes snowfall and rainfall but not the amount of precipitation accumulated on the ground. Wind and snow depth were originally considered but given the physical geography of Skåne, the focus shifted towards precipitation and temperature only. Average temperature was also first considered but the focus shifted to minimum and maximum temperatures in order to better understand the effects of extreme temperatures on railways.

Zakeri & Olsson (2018) studied the effects of severe weather conditions on delays and punctuality in Norway between 2007 and 2016 and concluded that harsher winters resulted in more delays compared to more mild winters, and that snow depth is the main weather variable that best explains daily and weekly punctuality variations. Ling et al. (2018) found that rail delay times in China are strongly correlated with bad weather and that these delays are most likely to occur during rainfall or snowfall. Chen, Wang, & Zhou (2021) concluded that the increasing amount of severe weather in more recent years has resulted in more challenges for transportation systems when researching the impact of severe weather conditions on high-speed rail and aviation in China. Palmqvist, Olsson, & Hiselius (2017b) concluded that in Sweden when temperatures fall below 0˚C, punctuality decreases exponentially. Brazil et al. (2017) studied the effects of weather conditions in Dublin and concluded that rain is the main delay factor. Zakeri & Olsson (2017) investigated the role weather conditions play on the performance of railways in Norway and highlighted the strong correlation between harsh winter conditions with extreme cold temperatures and punctuality. Diab & Shalaby (2020) studied the impact of outdoor track segments of the metro system in Toronto, Canada and concluded that the amount of snow on the ground and rainfall lead to more service interruptions. A study by Forzieri, et al. (2018) found that damage from heat waves, droughts, flooding, windstorms, and forest fires is expected to increase across Europe and with that requires higher costs of adaptation.

Some studies focused on the impact of one specific weather variable. For instance, Bubeck, et al., 2019 focused on the impacts of climate change on flood risk around Europe and concluded that currently the annual damage of flooding on railways in Europe is currently €581 million per year. This is expected to increase by 310%. Lastly, Dobney et al. (2009) quantified the effects of high summer temperatures related to climate change on rail buckling and consequently delays in the UK. They concluded that the number of buckles and therefore rail delays per year will increase if tracks are maintained to their current standard.

As seen in this literature review, numerous studies have occurred globally on the effects of weather on rail delays. They emphasise that weather today already plays a role in increasing the number of delays and that with climate change this is expected to worsen. Few have focused on Southern Sweden, and few investigate weather over a range of time-periods, as this thesis attempts to do.

ANALYSIS

4.0 Analysis Results

The analysis results are presented in the following chapter. First, the probability of a run or dwell delay is plotted against each weather variable for each of the five time periods. This is done in order to visualise the relationships and trends first that will help interpret the regression results in the next step of the analysis. A total of four regression models with 15 different independent or predictor variables were run against two different dependent or response variables. The predictor variables correspond to the minimum and maximum temperature, and the sum of precipitation over 1, 7, 14, 21, and 28 days while the response variables are dwell delays or run delays. Two multiple linear regression models and two multiple logistic regressions models were run. The results from these graphical evaluations and the four regressions are presented in this chapter.

4.1 Graphical Evaluation Results

The probability of a run delay and of a dwell delay was plotted against the sum of precipitation, minimum temperature, and maximum temperature over 1, 7, 14, 21, and 28 days in order to visualise the trends and where problems may occur. The probabilities were calculated in Azure Data Studio and then exported into excel files where they were plotted.

4.1.1 Sum of Precipitation

The following figures indicate the relationship between the sum of precipitation over 1, 7, 14, 21, and 28 days and the probability of a dwell or run delay.

Dwell Delay Probabilities for Precipitation

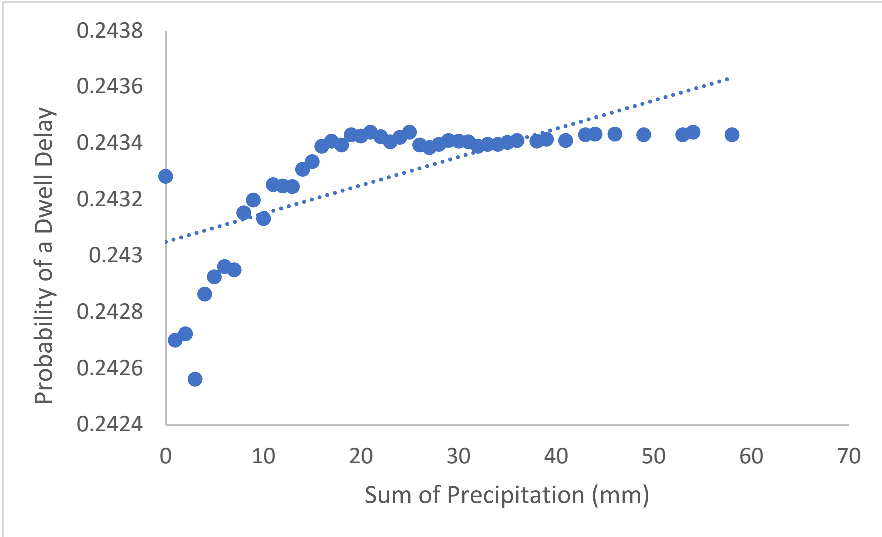


Figure 9. The Probability of a Dwell Delay for a Given Sum of Precipitation over 1 Day

Based on Figure 9 above, the of precipitation over one day ranges from 0mm-60mm. First the probability at 0mm shows a higher probability and then immediately dips down. Then the trend shows an increase of the probability of a delay as precipitation increases and eventually levels off. However, as it can been in figure 9 above that the increase in probability of dwell delays is very minimal.

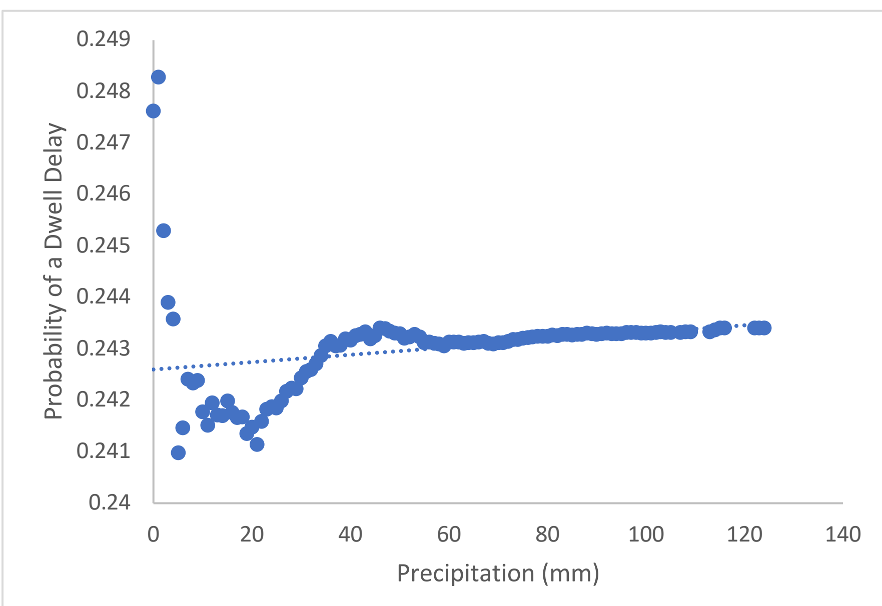


Figure 10. The Probability of a Dwell Delay for a Given Sum of Precipitation over 7 Days

Figure 10 above shows the probability of dwell delays when precipitation is summed over 7 days. Figure 2 shows a similar trend compared to Figure 9 with similar probabilities. Again, the trend shows a decrease first and then a gradual increase in the probability with precipitation increase. The range of precipitation values extends up to 125mm as 7 more days are added to the dataset. Between 0-20mm the trend appears to be more disordered, but this is minimal as the variation in the probability is minimal. The overall trends indicates that the probability of a dwell delay increases with the amount of precipitation cumulatively recorded over a 7-day period.

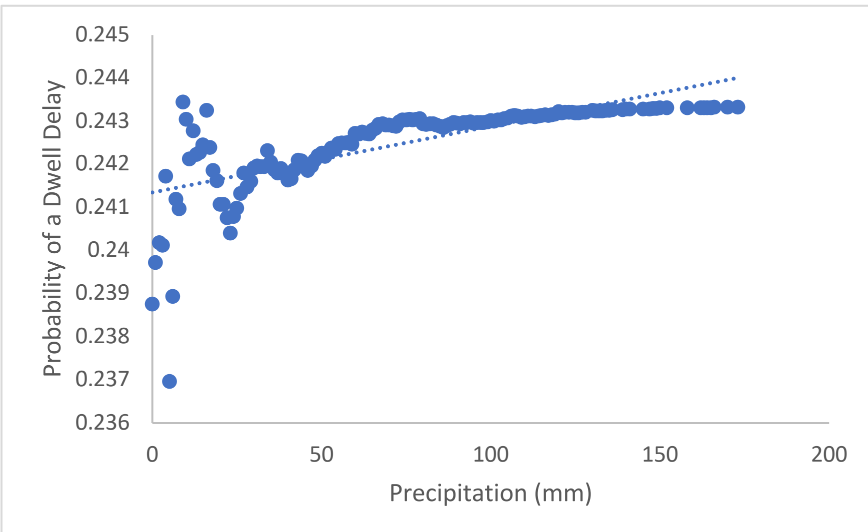


Figure 11. The Probability of a Dwell Delay for a Given Sum of Precipitation over 14 Days

In Figure 11 above, the first major difference compared to Figures 9 and 10 is that the probability axis shows lower values. The overall trend, however, indicates that dwell delay probability increases with the amount of cumulative precipitation, then decreases, but after increases again before slowly levelling out. As 7 more days are added, the precipitation values extend up to approximately 175mm.

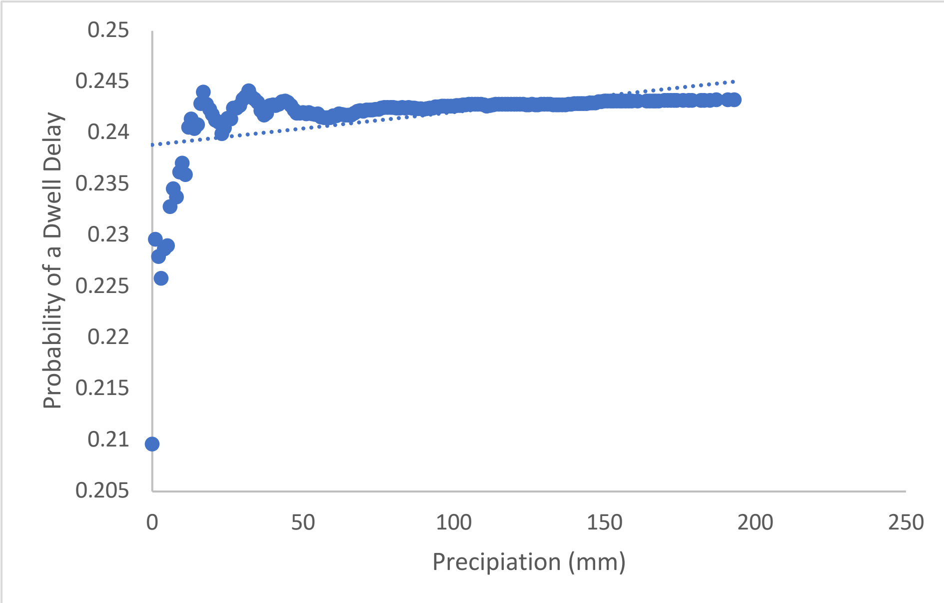


Figure 12. The Probability of a Dwell Delay for a Given Sum of Precipitation over 21 Days

Figure 12 above indicates a very similar trend to Figures 9-11, in that the probability of a dwell delay sharply increases with precipitation until around 50mm when it levels out. Here the precipitation ranges from 0-195 mm and the probability values include a wider range as well compared to Figures 9-11.

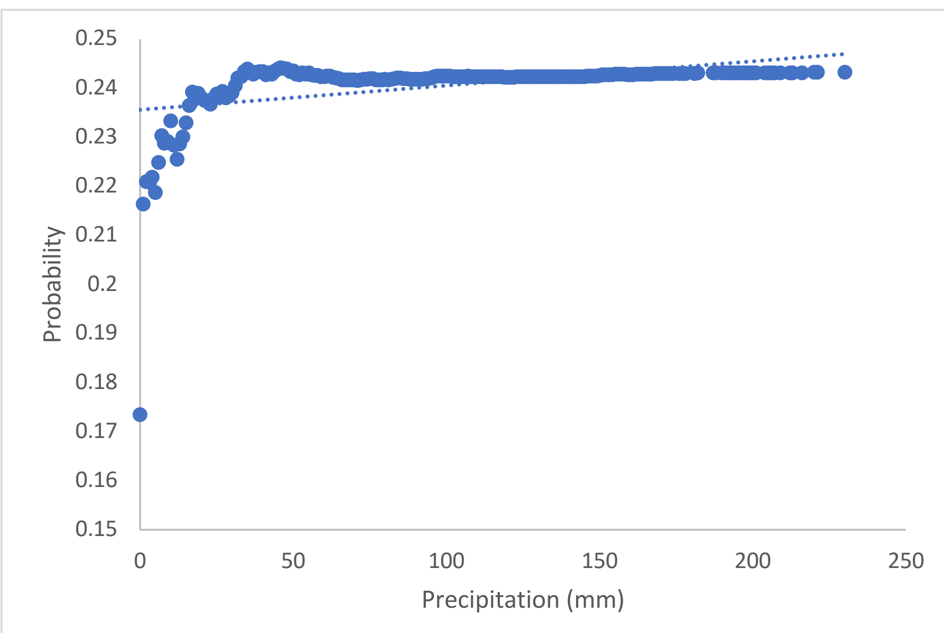


Figure 13. The Probability of a Dwell Delay for a Given Sum of Precipitation over 28 Days

Figure 13 highlights that the probability of a dwell delay over 28 days has arguably the least steep slope. Additionally, this figure includes the most precipitation values ranging from 0-230mm. The probability of a delay at 0mm is roughly 18% before the probability sharply increases to just under 22% with the next unit increase. The overall trendline matches Figures 9-12 in that precipitation has a positive relationship with dwell delays. This means that the more cumulative precipitation there is the higher probability there is for a dwell delay.

Run Delay Probabilities for Precipitation

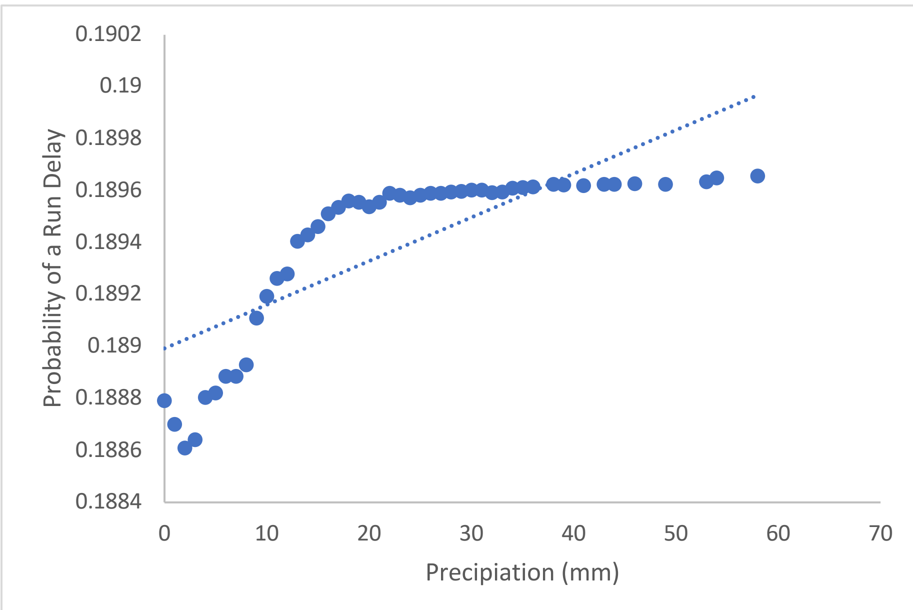


Figure 14. The Probability of a Run Delay for a Given Sum of Precipitation over 1 Day

Figure 14 above indicates first a decrease in probability with an increase in precipitation until about 3mm when the trendline then shows that the probability of a run delay increases with precipitation. Compared to the figures from dwell delay it can also be noted that the probabilities are lower. This is because, there are more scheduled run times compared to dwell times. This means that trains in Skåne are spending more time traveling between stations compared to the time spent at a station. Therefore, the chance of a dwell delay is higher as there are less values in the dataset. Again, the range of probability remains fairly small which indicates that the probability of a run delay does not significantly increase as precipitation increases.

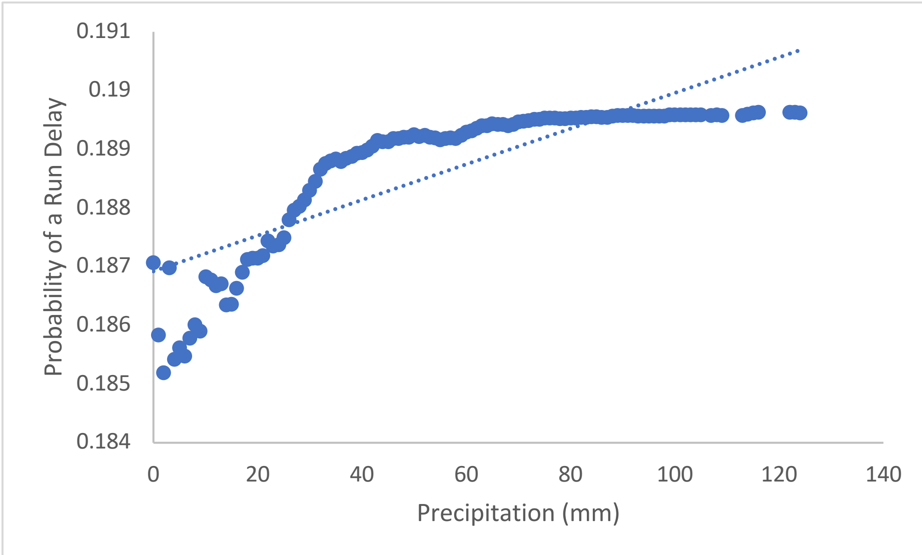


Figure 15 The Probability of a Run Delay for a Given Sum of Precipitation over 7 Days

The trend in Figure 15 above highlights that the probability of a run delay increases with precipitation. Similarly, to the other figures there are some fluctuations in the beginning of the plot where the probability decreases and increases again, however it eventually levels out to highlight the positive relationship between precipitation and the probability of a run delay.

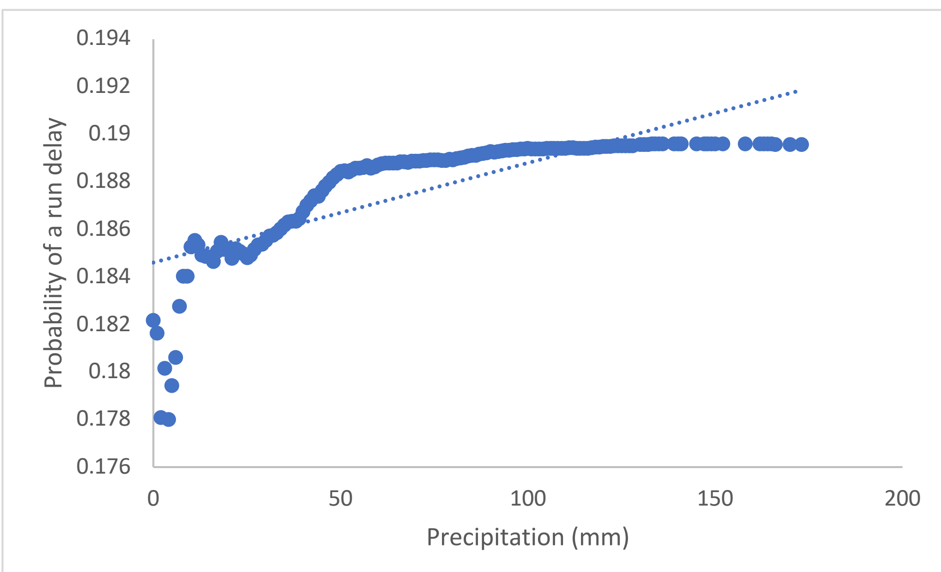


Figure 16 The Probability of a Run Delay for a Given Sum of Precipitation over 14 Days

Once again, Figure 16 above highlights that the probability of a run delay decreases sharply, then sharply increases again, and finally levels out to indicate a slight increase in the probability of a run delay with an increase in precipitation.

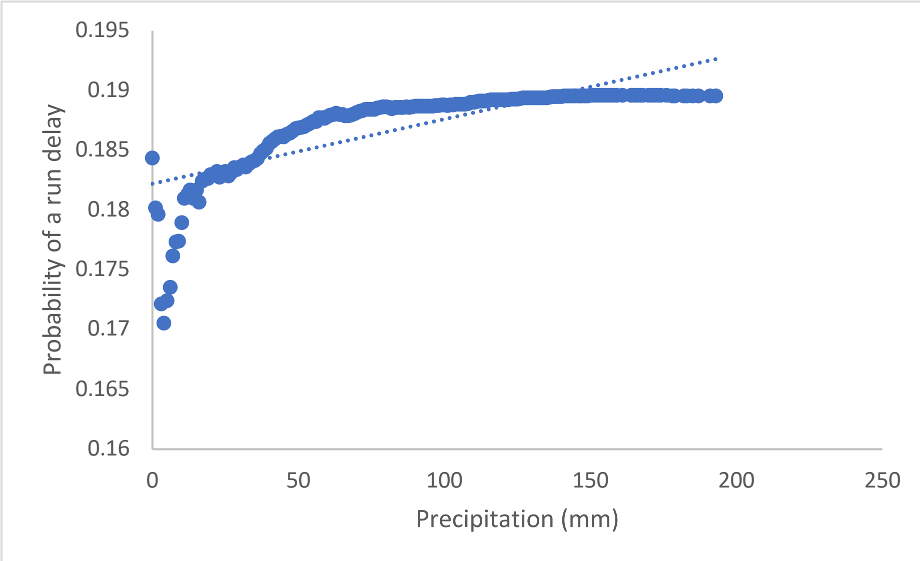


Figure 17 The Probability of a Run Delay for a Given Sum of Precipitation over 21 Days

Figure 17 above highlights that the probability decreases first before increasing similarly to Figures 14-16. Moreover, it shows that the more precipitation there is over 21 days the higher the probability of a run delay. At some point the probability levels out, only very slightly increasing.

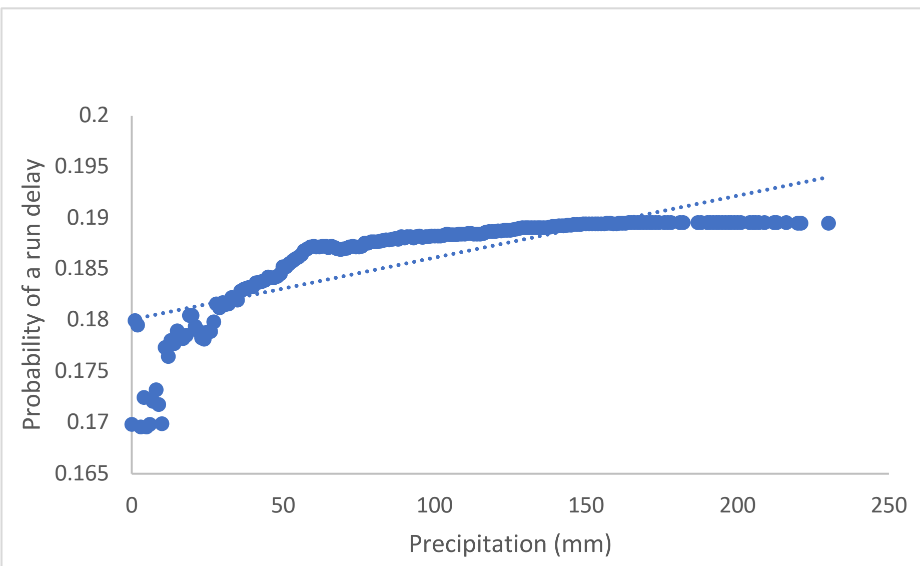


Figure 18 The Probability of a Run Delay for a Given Sum of Precipitation over 28 Days

Figure 18 above first indicates that the precipitation over 28 days has one of the greater ranges of probability. This infers that there is most likely a greater impact of 28 days compared to other time ranges due to the greater amounts of precipitation accumulated over 28 days. The trend again shows a decrease first in the probability of a run delay as precipitation increases before an increase, and eventually almost levelling out.

4.1.2 Minimum Temperature

The minimum temperature over 1, 7, 14, 21, and 28 days plotted against the probability of a dwell or run delay is highlighted in the following figures below.

Dwell Delay Probabilities for Minimum Temperature

Temperature is more complex to quantify the relationship with the probability of delays due to the negative and positive integers in the dataset. The minimum temperature is used to determine when temperatures are at their coldest. It is expected that there will be more delays when the temperatures are at their lowest. Additionally, it is expected that the longer period of time it is cold the more delays will occur.

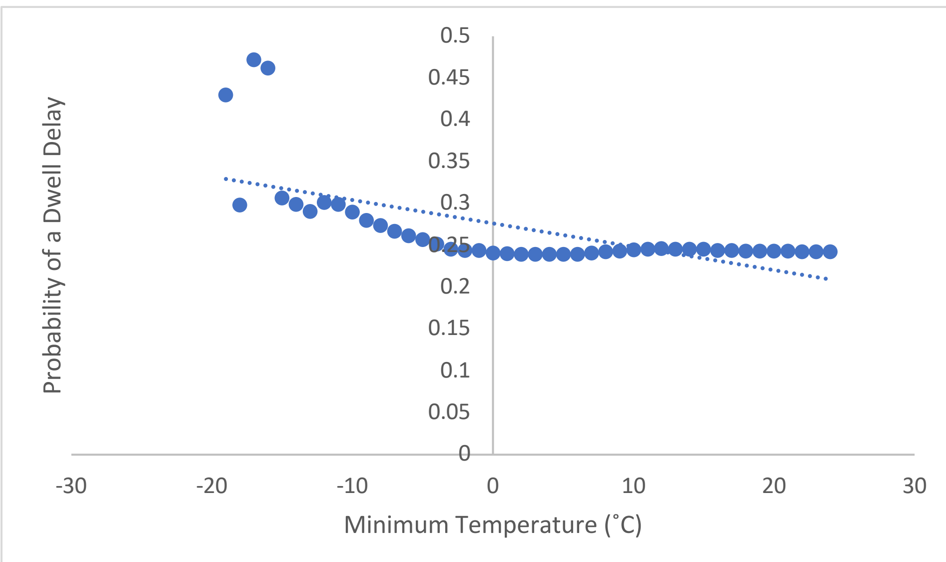


Figure 19 The Probability of a Dwell Delay for a Given Minimum Temperature over 1 Day

Figure 19 above highlights the expected results to hold true for minimum temperature over 1 day. The lower temperatures are associated with higher probabilities. As temperature increases the probability decreases with it. However, the probability increases again slightly when temperatures get quite high. It can be expected that if the minimum temperature of one day is high, for instance above 20˚C, then that day was particularly hot and therefore heat can also cause issues with delays as discussed in Chapter 3.2.2.

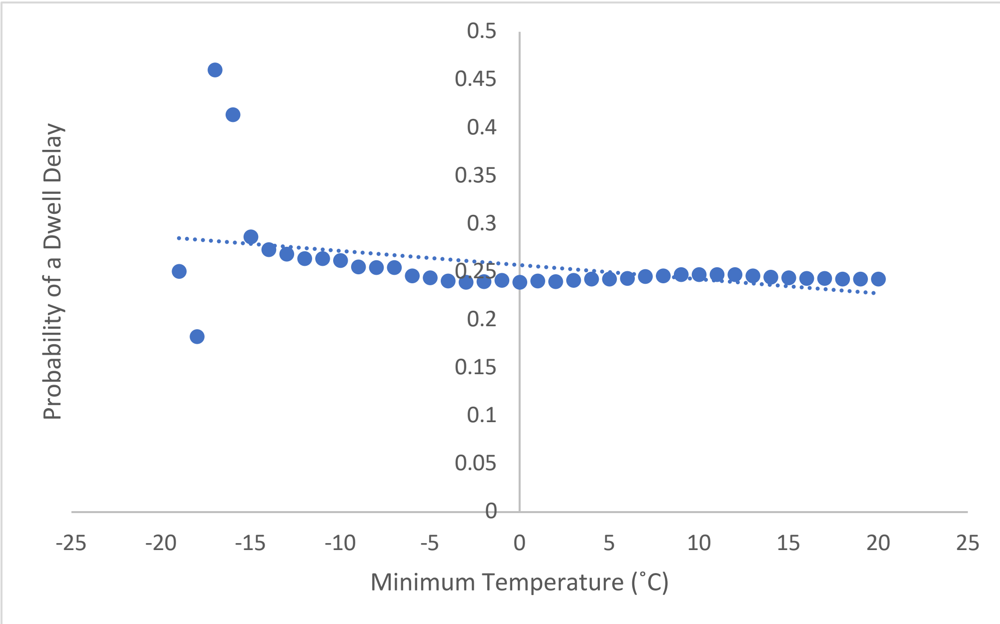


Figure 20 The Probability of a Dwell Delay for a Given Minimum Temperature over 7 Days

Figure 20 above indicates a negative relationship between the minimum temperature over 7 days and the probability of a dwell delay. The probability appears to be highest when the temperature is around -17˚C or -16˚C as opposed to -19˚C or -18˚C. However, overall, the figure highlights that the biggest problems with delays occur when it is cold, and the probability decreases as temperature increases.

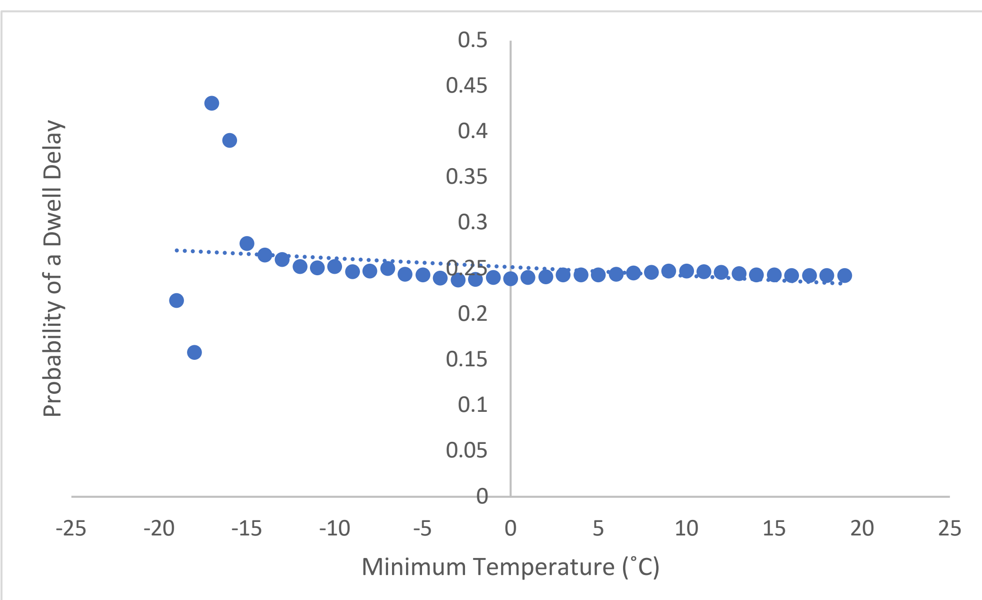


Figure 21 The Probability of a Dwell Delay for a Given Minimum Temperature over 14 Days

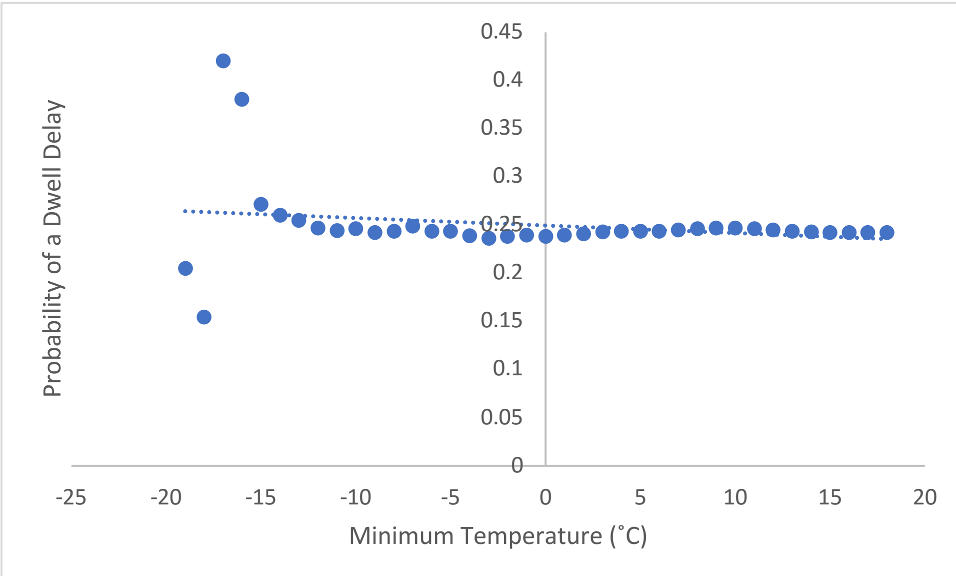


Figure 22 The Probability of a Dwell Delay for a Given Minimum Temperature over 21 Days

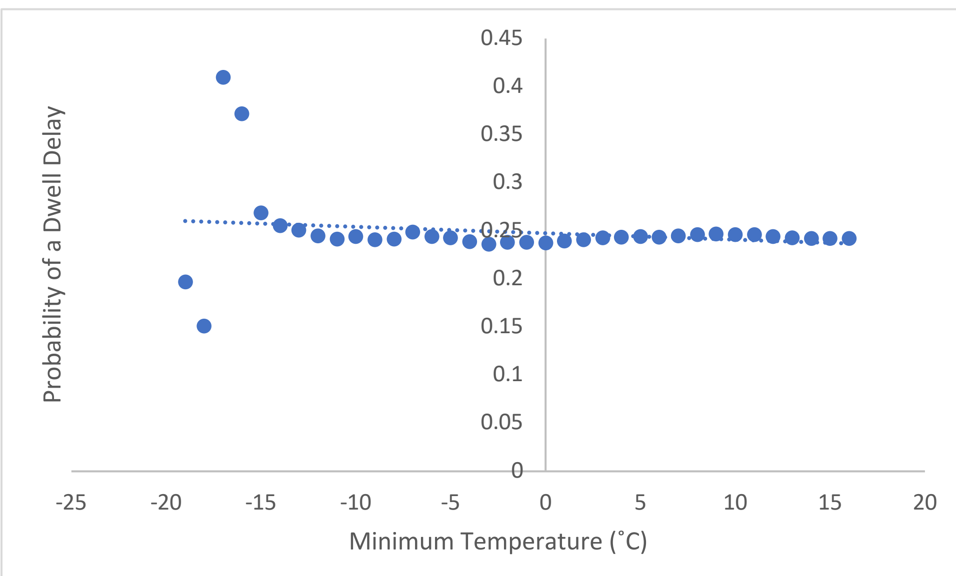


Figure 23 The Probability of a Dwell Delay for a Given Minimum Temperature over 28 Days

Figures 21-23 above show an almost identical trend as Figure 20 but with slightly different dwell delay probabilities. The overall trend indicates that there are more delays when it is extremely cold and as the temperature increases the probability of a dwell delay decreases.

Run Delay Probabilities for Minimum Temperature

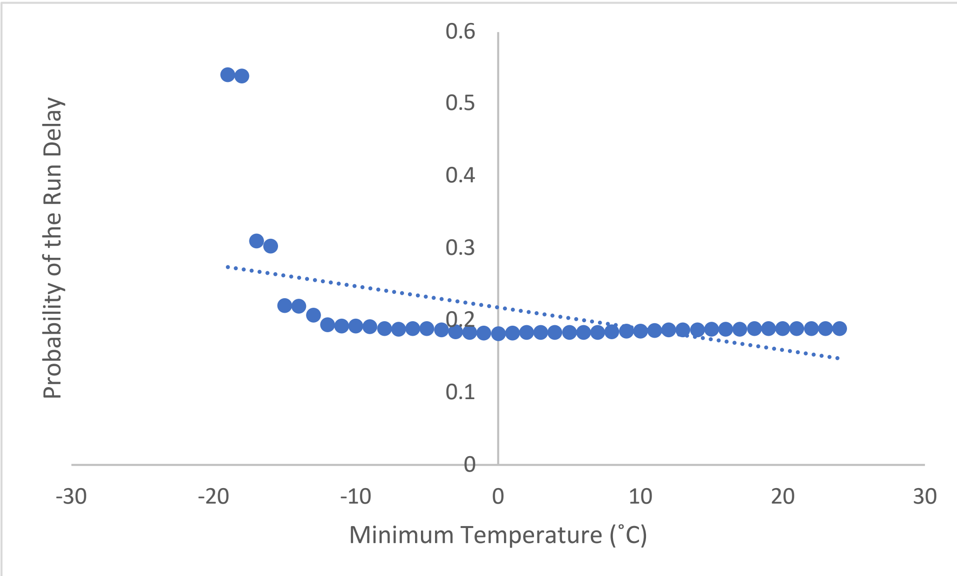


Figure 24 The Probability of a Run Delay for a Given Minimum Temperature over 1 Day

Figure 24 above indicates that the probability of a run delay decreases as the minimum temperature increases. On the right hand of the figure, it shows that the trend is starting to slightly increase again as temperatures get extremely warm. This means that the most delays occur when it is extremely cold and extremely hot. Overall, the figure shows that temperatures below freezing cause the most issues with run delays.

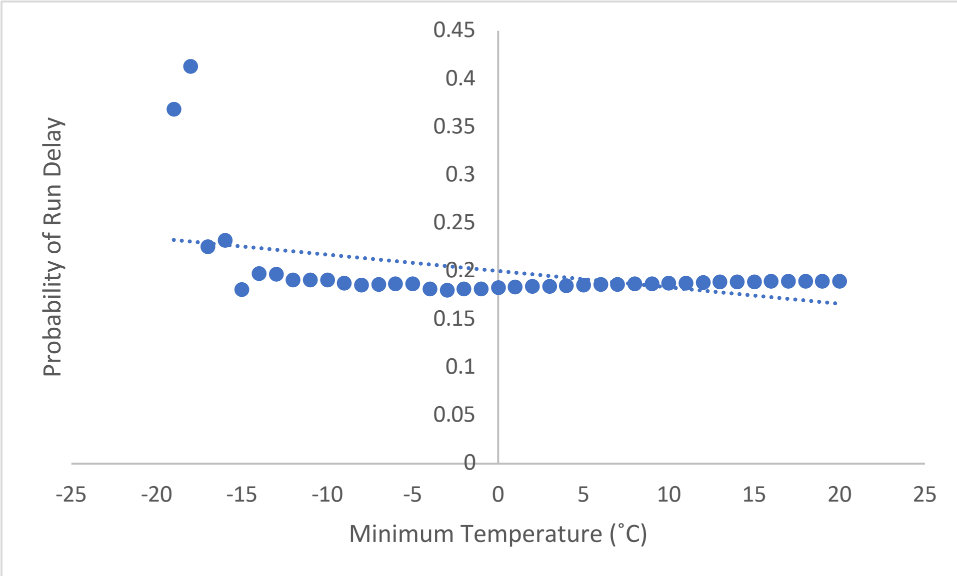


Figure 25 The Probability of a Run Delay for a Given Minimum Temperature over 7 Days

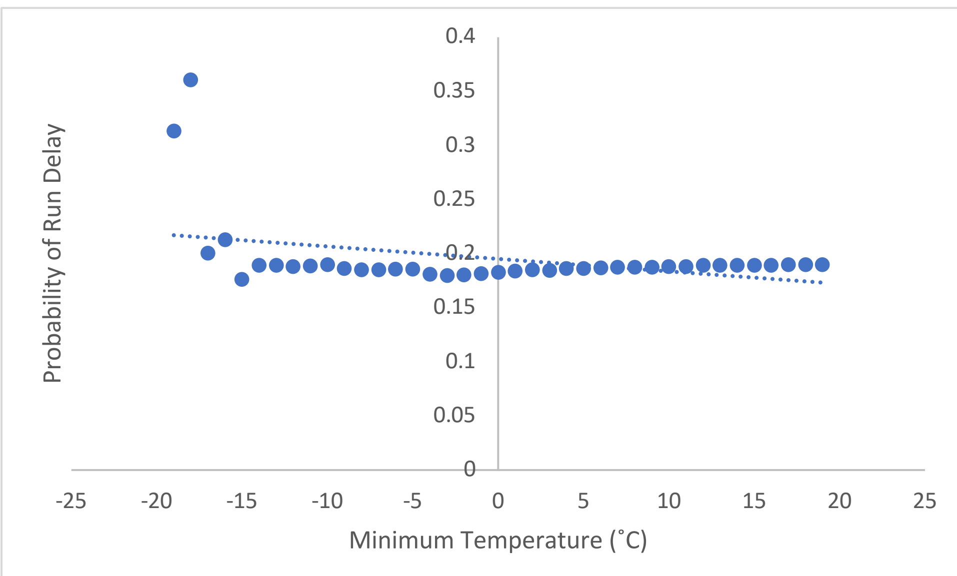


Figure 26 The Probability of a Run Delay for a Given Minimum Temperature over 14 Days

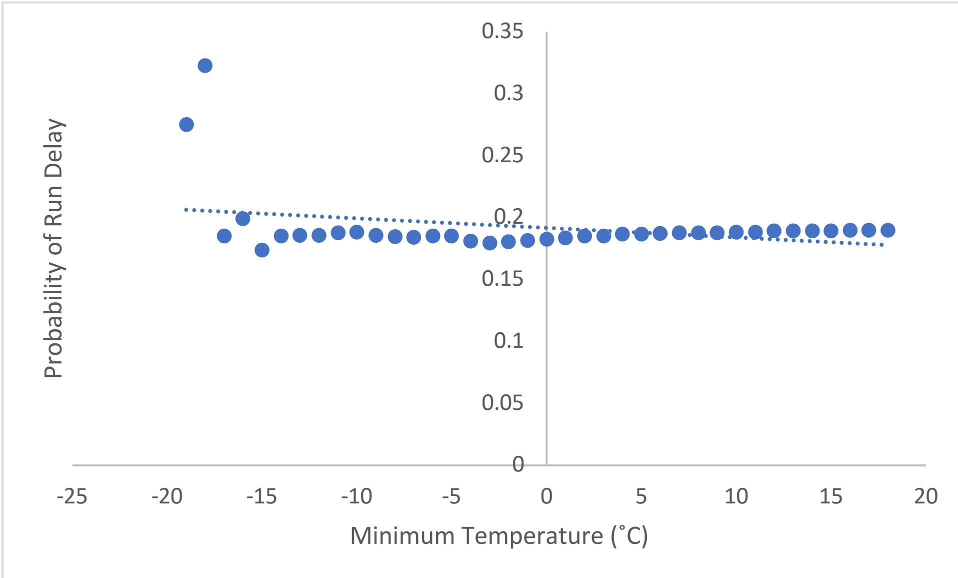


Figure 27 The Probability of a Run Delay for a Given Minimum Temperature over 21 Days

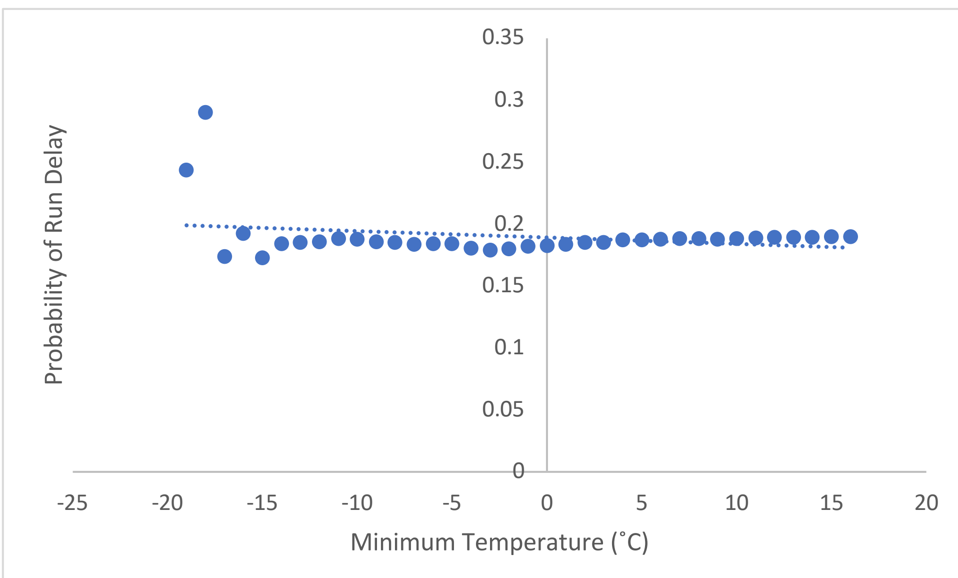


Figure 28 The Probability of a Run Delay for a Given Minimum Temperature over 28 Days

The trends over 7, 14, 21, and 28 days are very similar with the lowest temperatures having the largest probability of a run delay (Figures 25-28). Then the next three lowest temperatures recorded show some slight variation in probability but then the trend flattens out, slightly decreasing before starting to increase when temperatures are extremely hot. Overall, the trend appears to be negative; indicating that as temperature increases the probability of a run delay decreases.

4.1.3 Maximum Temperature

The maximum temperature over 1, 7, 14, 21, and 28 days plotted against the probability of a dwell or run delay is highlighted in the following figures below.

Dwell Delay Probabilities for Maximum Temperature

Temperature is more complex to quantify the relationship with probability of delays due to the negative and positive integers in the dataset. The maximum temperature is used to determine when temperatures are the warmest.

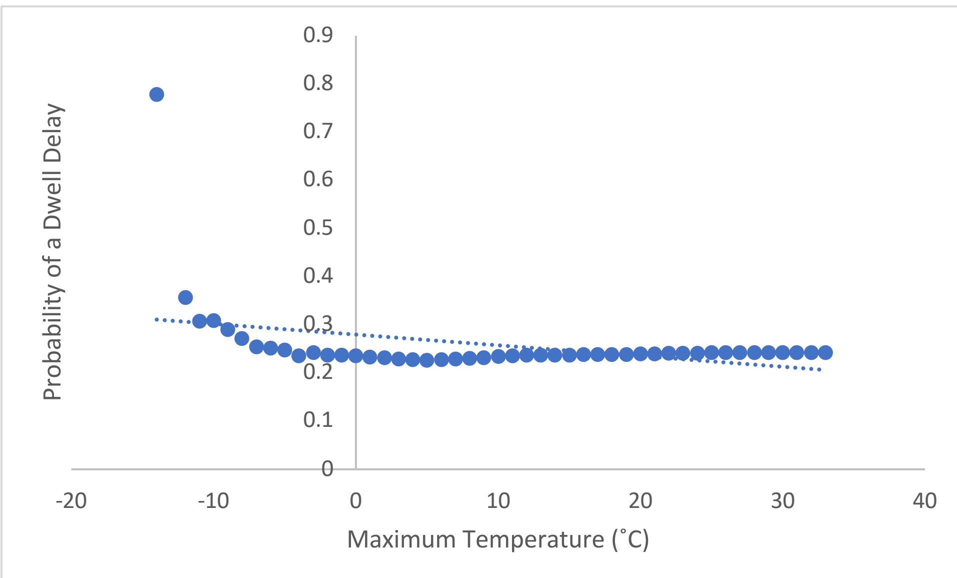


Figure 29 The Probability of a Dwell Delay for a Given Maximum Temperature over 1 Day

Figure 29 above highlights the probability of a dwell delay over one day. Here the greatest probabilities are below freezing. Since the time period of this figure is just over 1 day it is very likely that the maximum temperature experienced over one day is below freezing on extremely cold days. The trend shows a big decrease in the probability between the two lowest temperatures and then shows that as temperature decreases so does the probability of a dwell delay. Eventually, the trend shows that the probability starts to increase as temperatures get above around 20˚C.

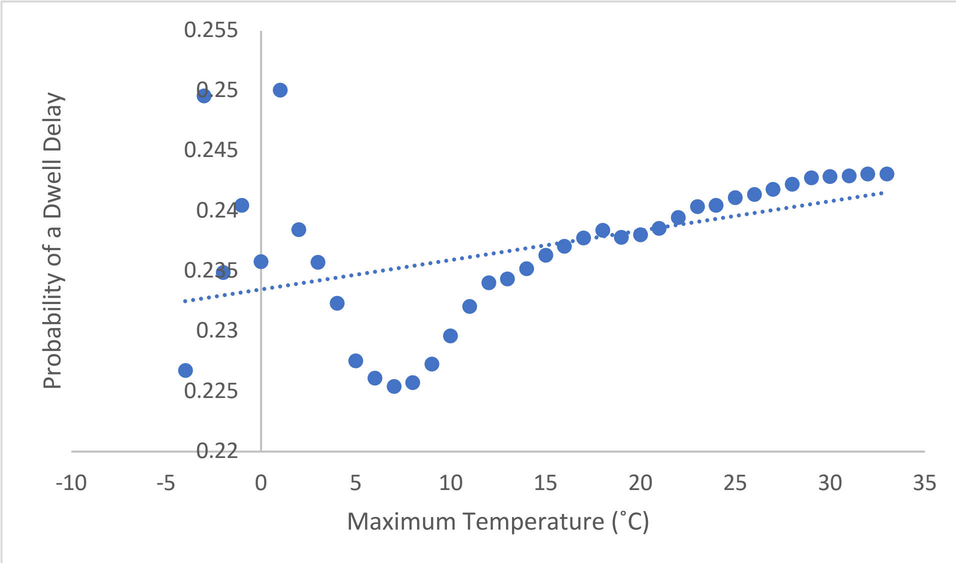


Figure 30 The Probability of a Dwell Delay for a Given Maximum Temperature over 7 Days

Figure 30 above highlights that the probability of a dwell delay is highest when temperatures are below freezing and when they are approaching extreme heat. This trend shows more variation compared to Figure 29. This most likely indicates that a period of extreme cold or extreme heat over one week are more impactful on dwell delays.

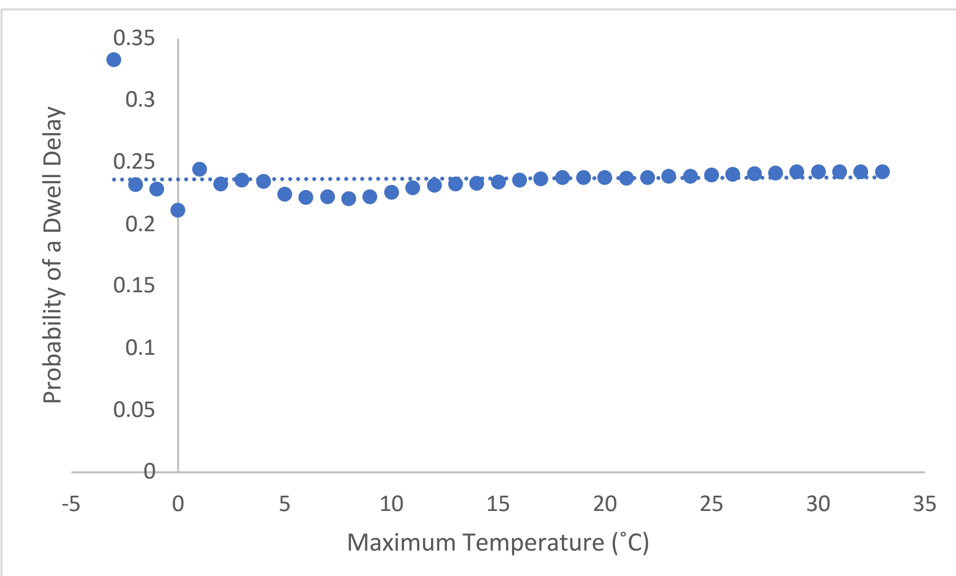


Figure 31 The Probability of a Dwell Delay for a Given Maximum Temperature over 14 Days

The probability of a dwell delay after 14 days of a given maximum temperature is shown in Figure 31 above. It can be seen that the highest probability of a delay occurs when the maximum temperature is below freezing. A maximum temperature over 2 weeks being below freezing indicates it was a particularly cold period of time and therefore it can be expected that there are more delays. Although, the trend decreases until temperatures hit around 0˚C, the trendline slowly increases again indicating that extreme hot temperatures are also related to higher probabilities of dwell delays. Overall, the figure indicates that temperature increases along with the probability of a dwell delay.

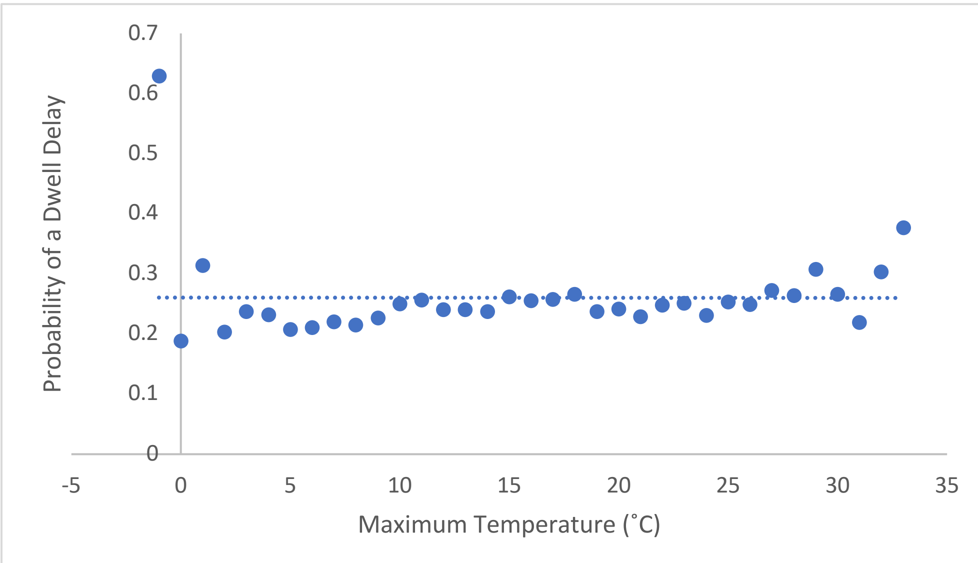


Figure 32 The Probability of a Dwell Delay for a Given Maximum Temperature over 21 Days

Figure 32 above shows a trend a bit more prominent compared to Figures 29-30. Over 21 days the lowest maximum temperature is -1˚C, which also corresponds to the highest probability of a dwell delay. At 0˚C the probability decreases by around 30% and then stays fairly constant. Once temperatures reach around 30˚C there starts to be a sharp increase in dwell delay probability. This indicates that the most problems with dwell delays occur when temperatures are below freezing and above 30˚C.

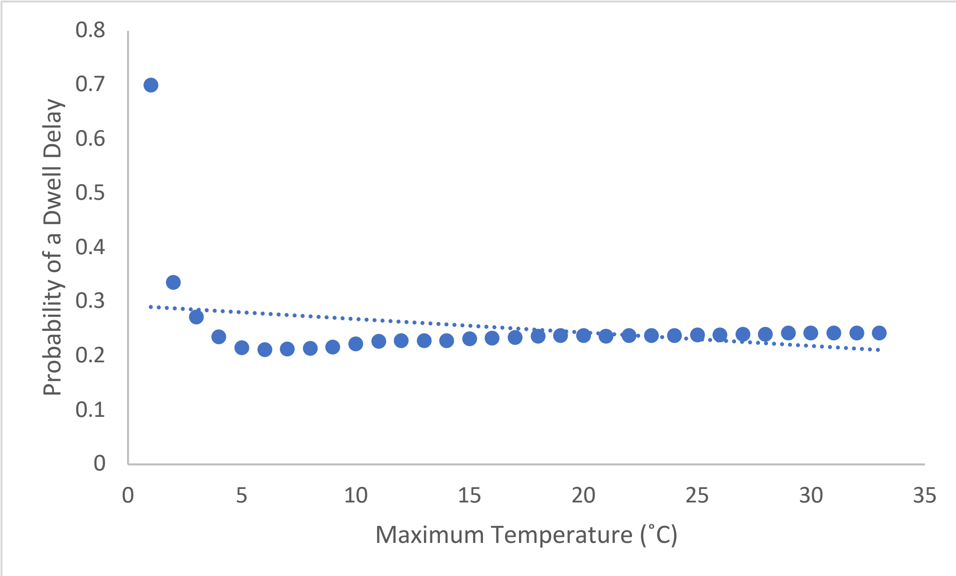


Figure 33 The Probability of a Dwell Delay for a Given Maximum Temperature over 28 Days

The maximum temperature over 28 days does not include any values below freezing as seen in Figure 33. However, the highest probability of a dwell delay still corresponds with the lowest temperature in the dataset. After, the probability decreases as temperature increases, however, around 10˚C the trend slowly starts increasing again. Once again this indicates that the most dwell delays occur on either tail-end of the temperature scale.

Run Delay Probabilities for Maximum Temperature

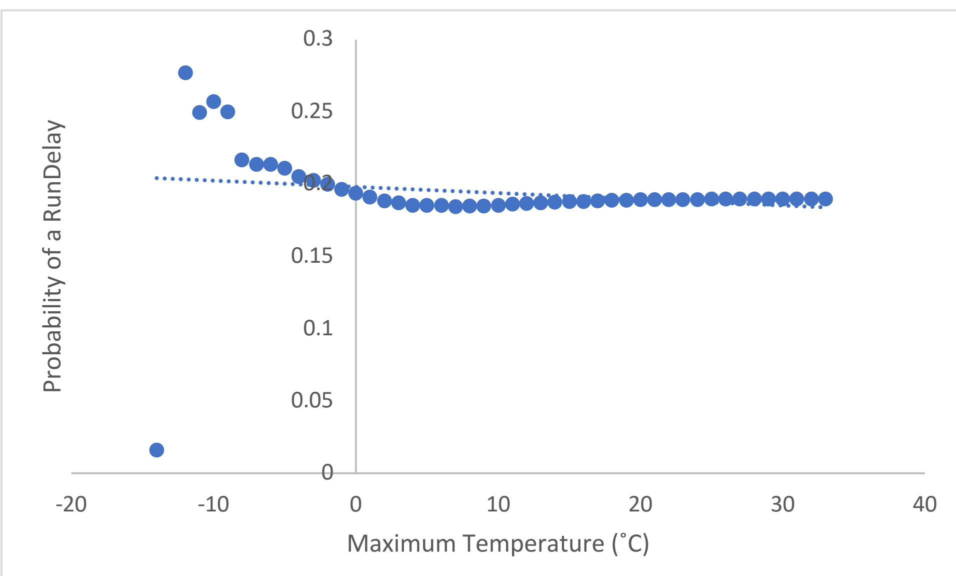


Figure 34 The Probability of a Run Delay for a Given Maximum Temperature over 1 Day

Figure 34 above highlights the probability of a run delay after a given maximum temperature over 1 day. Similarly, to Figure 29 there are more values below freezing since only 1 day is considered in this figure. Additionally, the lowest maximum temperature has the lowest probability before the probabilities sharply increase. From the figure it is evident that the higher chance of a run delay occurs when temperatures are below freezing. This indicates that as temperature decrease the probability of a run delay increases. However, this only occurs a certain point. Once temperatures begin to get more extremely high, the probabilities begin to increase again.

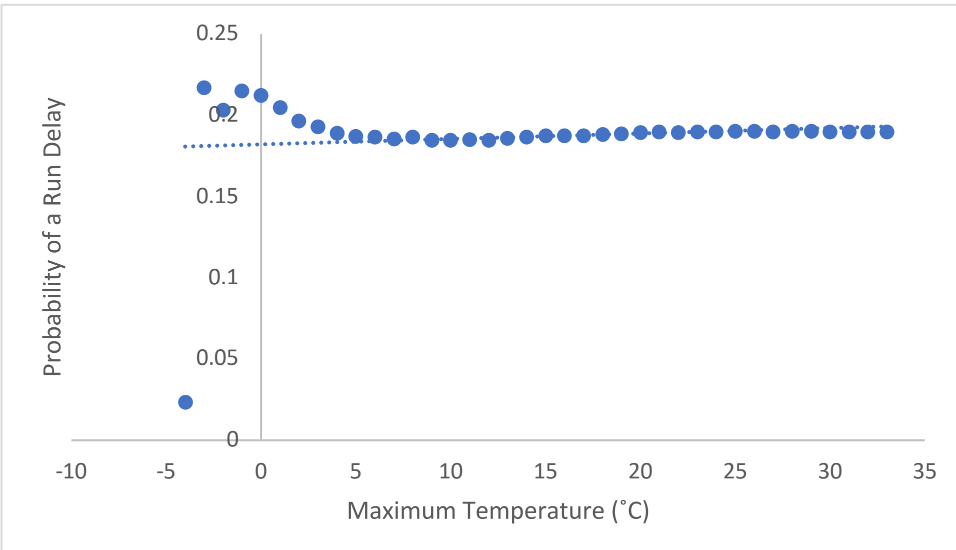


Figure 35 The Probability of a Run Delay for a Given Maximum Temperature over 7 Days

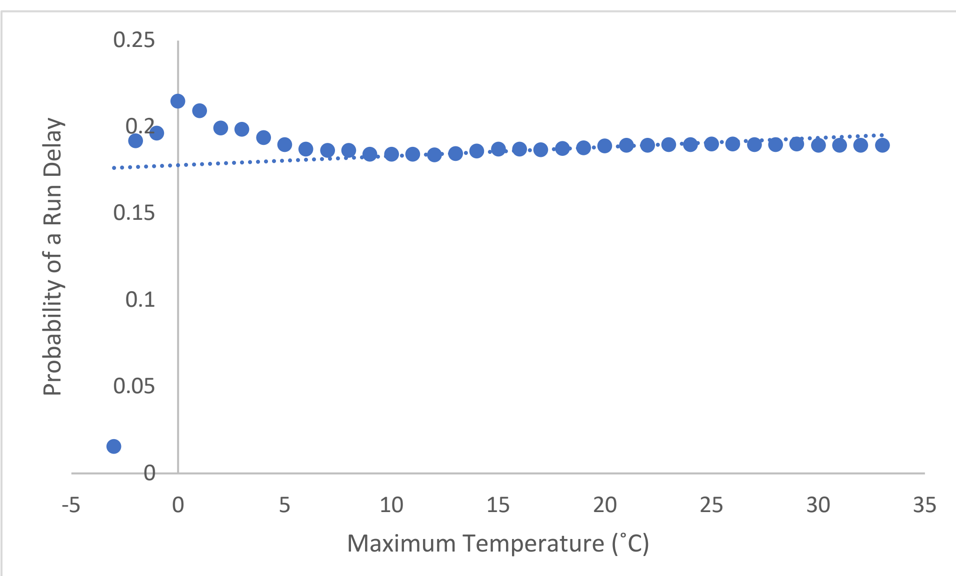


Figure 36 The Probability of a Run Delay for a Given Maximum Temperature over 14 Days

Figures 35 and 36 above display similar trends to Figure 34. The lowest maximum temperatures correspond to the lowest probability of a run delay. However, the next values below freezing correspond with the highest probabilities. When temperatures reach above freezing, they decrease slightly before starting to level out. These figures highlight that freezing temperatures over longer periods of time result in the highest chance of a run delay.

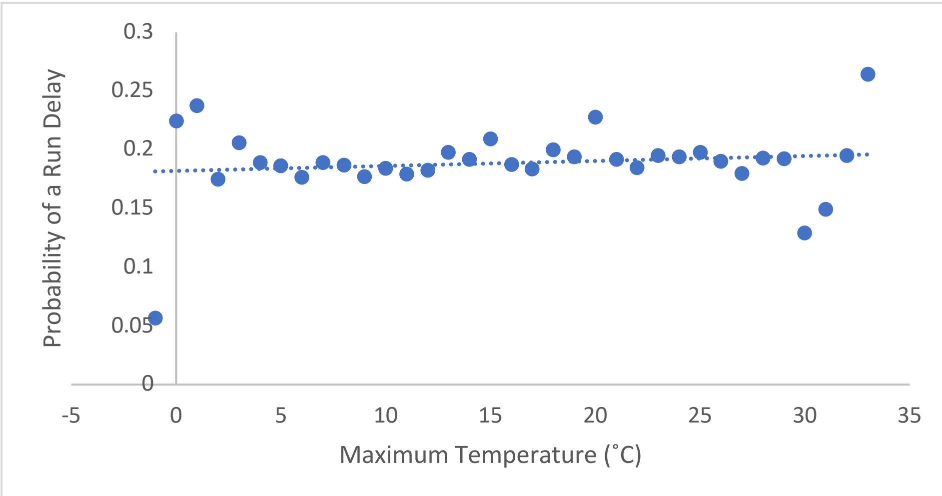


Figure 37 The Probability of a Run Delay for a Given Maximum Temperature over 21 Days

Figure 37 above indicates a similar trend to dwell delay probabilities shown in Figure 32. 21 days seems to highlight the effects of very high temperatures the most. The trend shows that the probability of a run delay increases with temperature. Here the highest probabilities corresponded with the warmest days.

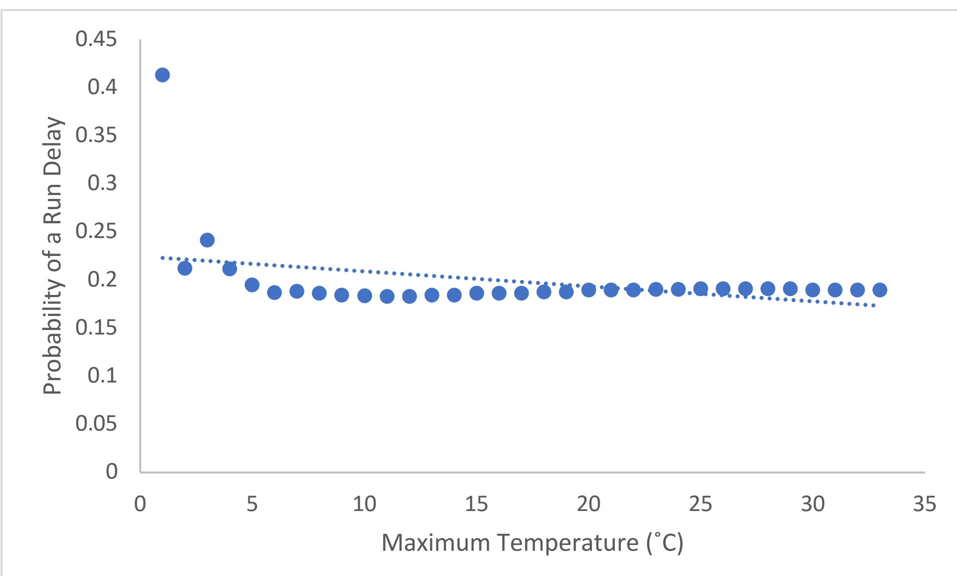


Figure 38 The Probability of a Run Delay for a Given Maximum Temperature over 28 Days

Figure 38 indicates that the maximum temperature over 28 days does not include temperatures below freezing. However, the highest probability of a run delay still occurs when the maximum temperature is the lowest. Then the trend sharply decreases before levelling out and eventually starting to increase as temperatures get very hot.

4.1.4 Graphical Evaluation Sub-Conclusion

Figures 9-38 show the probability of a dwell delay or run delay over 1, 7, 14, 21, and 28 days of a given sum of precipitation, minimum temperature, or maximum temperature. First, it can be seen that the probabilities of a dwell delay are higher than of run delay which is due to there being more and longer run times scheduled compared to dwell times. In regard to precipitation the figures highlight that the chance of both dwell and run delays increase with precipitation. When looking at both minimum and maximum temperature it is evident that highest chance of a delay occurs when temperatures are below freezing. Most figures indicate the trend slightly increasing when temperatures reach extremely high temperatures. With climate change projections it is expected that these temperatures will get higher. Therefore, the figures only showing a slight increase can be justified by the fact that Skåne over the past 11 years most likely has yet to experience the extreme high temperatures projected.

The next part of the analysis presents the results from the regression models. These models are used to quantify the relationships between weather and delays further by determining which variables are the most statistically significant, and therefore have the greatest impact on dwell and run delays.

4.2 Regression Analysis

The regression analysis results are presented in the following chapter. A total of four regression models with 15 different independent or predictor variables were run against two different dependent or response variables. The predictor variables correspond to the sum of precipitation, minimum temperature, and maximum temperature over 1, 7, 14, 21, and 28 days while the response variables are dwell delay or run delay. Two linear regression models and two logistic regression models were run. The results from these four regressions are presented in this sub-chapter.

4.2.1 Dwell Delay Multiple Linear Regression Model

The regression coefficients were estimated using the Linear Model (lm) in R. In this case the response variable is the size of the dwell delay. A linear regression model assumes there is a linear relationship between the predictor and response variables. The results for the linear regression model for a dwell delay are indicated in Table 5 below. The regression is based off equation 4 from Chapter 2.5.1.

Table 5. Summary of the Multiple Linear Regression Model for Dwell Delay (4)

Coefficients:		**Estimate**		**St. E.**	**Pr(>|t|)**	
	(Intercept)	β_0	0.265	0.003	< 2e-16	***
x_1	prec_sum_1	β_1	0.007	0.002	0.001	***
x_2	prec_sum_7	β_2	0.004	0.001	0.000	***
x_3	prec_sum_14	β_3	0.002	0.001	0.012	*
x_4	prec_sum_21	β_4	-0.002	0.001	0.053	.
x_5	prec_sum_28	β_5	0.001	0.001	0.038	*
x_6	temp_min_1	β_6	-0.004	0.000	< 2e-16	***
x_7	temp_min_7	β_7	-0.001	0.000	0.000	***
x_8	temp_min_14	β_8	-0.002	0.000	0.001	***
x_9	temp_min_21	β_9	0.000	0.001	0.842	
x_{10}	temp_min_28	β_{10}	-0.002	0.000	0.000	***
x_{11}	temp_max_1	β_{11}	0.005	0.000	< 2e-16	***
x_{12}	temp_max_7	β_{12}	0.001	0.000	0.193	
x_{13}	temp_max_14	β_{13}	-0.001	0.001	0.012	*
x_{14}	temp_max_21	β_{14}	-0.002	0.001	0.015	*
x_{15}	temp_max_28	β_{15}	0.007	0.001	< 2e-16	***
R^2= 0.001						

Note: Variables statistically significant at the 0 level are marked by a triple asterisk (***), variables statistically significant at the 0.001 level are marked by a double asterisk (**), variables statistically significant at the 0.01 level are marked by a single asterisk (*), and variables statistically significant at the 0.05 level are marked by a dot (.)

From the table above only the sum of precipitation over 1 and 7 days; the minimum temperature over 1, 7, 14, and 28 days; and the maximum temperature over 1 and 28 days are statistically significant. Therefore, in order to better understand the statistical significance, the regression model was run twice more until all variables were deemed statistically significant. In the first rerun variables 3, 4, 5, 9, 12, 13, and 14 were removed (Table 5). For the third run of the model the sum of precipitation over 1 day was also removed, and Table 6 below highlights the improved summary of the multiple linear regression model. The number of variables analysed dropped from 15 to 7.

Table 6. Summary of the Improved Multiple Linear Regression Model for Dwell Delay (4)

Coefficients:		**Estimate**		**St. E.**	**Pr(>|t|)**	
	(Intercept)	β_0	0.268	0.003	<2e-16	***
x_1	prec_sum_7	β_1	0.006	0.001	<2e-16	***
x_2	temp_min_1	β_2	-0.004	0.000	<2e-16	***
x_3	temp_min_7	β_3	-0.001	0.000	<2e-16	***
x_4	temp_min_14	β_4	-0.002	0.000	<2e-16	***
x_5	temp_min_28	β_5	-0.001	0.000	<2e-16	***
x_6	temp_max_1	β_6	0.005	0.000	<2e-16	***
x_7	temp_max_28	β_7	0.004	0.000	<2e-16	***

R^2= 0.001

Notes: Variables statistically significant at the 0 level are marked by a triple asterisk

Based on Table 6 above all variables are deemed as statistically significant based on the p-value. Since the p-value is less than 0.05 (significance level) the null hypothesis can be rejected, and it can be said that these 7 variables have an impact on dwell delay size. β_0 in this model is 0.268, which represents the expected dwell delay size when all predictor variables are 0.

When looking more closely at the coefficients, the sum of precipitation over 7 days, and the maximum temperature over 1 and 28 days is positively related to dwell delay size. In contrast the, the minimum temperature over 1 and 28 days is negatively related to dwell delay size. The coefficients represent the rate of change. In order words, the amount that dwell delay size changes when a predictor variable increases by 1 unit amount. In regard to precipitation this means that as the precipitation increases as does the dwell delay size. This relationship is also similar to what was plotted in Figure 14. Additionally, only the precipitation over 7 days was indicated as statistically significant implying that the sum of precipitation over 1, 14, 21, and 28 days have little impact on dwell delays.

As mentioned, temperature is more complex to handle as they are both positive and negative integers. The minimum temperature is statistically significant over all time periods expect for 21 days. This means that low temperatures have a big impact on dwell delays, which is similarly depicted in Figures 19-23. The relationship is negative indicating that dwell delay size decreases with temperature. The maximum temperature looks at the highest temperature experienced and therefore, it can be expected that it is rare for this number to be significantly below freezing, with the expectation of maximum temperature over 1 day. Table 6 indicates that maximum temperature is statistically significant over 1 and 28 days. This means that the biggest delays most likely occur when it is extremely cold or hot over 1 day or constantly over one month. The positive relationship between maximum temperature and dwell delay size depicted in Table 6 indicates that dwell delay size increases as the maximum temperature increases. This is different than what is depicted in Figures 29 and 33 but this can most likely be explained by the complexity of including temperature values that are both below 0˚C and above. Additionally, temperature is typically viewed as an exponential relationship as opposed to a linear one.

The R^2 explains the variance for the independent variable which is explained by one or more independent variables. In other terms how well, the data fits the model. However, it cannot determine any biases and it does not determine if a regression model is sufficient or not. Since the R^2 of this model is almost 0 the model does not explain any variability of the response variable data around the mean. However, as mentioned R^2 is not always a good indicator of how data fits the model. In Table 6 the data has small standard error values and small p-values that do suggest the model is statistically significant. The data dealt with in this model most likely is not truly linear especially since temperature is typically an exponential relationship, which may be able to explain the low value of R^2 and low coefficient numbers.

4.2.2 Run Delay Multiple Linear Regression Model

The regression coefficients were estimated using the Linear Model (lm) in R. In this case the response variable is the size of the run delay. A linear regression model assumes there is a linear relationship between the predictor and response variables. The results for the linear regression model for a run delay are indicated in Table 7 below. The regression is based off equation 4 from Chapter 2.5.1.

Table 7. Summary of the Multiple Linear Regression Model for Run Delay (4)

Coefficients			Estimate	St. E.	Pr(>|t|)	
	(Intercept)	β_0	0.236	0.003	< 2e-16	***
x_1	prec_sum_1	β_1	0.017	0.002	< 2e-16	***
x_2	prec_sum_7	β_2	0.004	0.001	3.65e-06	***
x_3	prec_sum_14	β_3	0.003	0.001	3.16e-05	***
x_4	prec_sum_21	β_4	0.003	0.001	0.65	
x_5	prec_sum_28	β_5	0.002	0.001	0.000	***
x_6	temp_min_1	β_6	-0.005	0.000	< 2e-16	***
x_7	temp_min_7	β_7	-0.002	0.000	4.07e-07	***
x_8	temp_min_14	β_8	-0.002	0.000	0.002	**
x_9	temp_min_21	β_9	0.001	0.001	0.11	
x_{10}	temp_min_28	β_{10}	-0.002	0.000	9.01e-05	***
x_{11}	temp_max_1	β_{11}	0.005	0.000	< 2e-16	***
x_{12}	temp_max_7	β_{12}	0.000	0.000	0.300	
x_{13}	temp_max_14	β_{13}	0.002	0.001	8.04e-05	***
x_{14}	temp_max_21	β_{14}	-0.002	0.001	0.009	**
x_{15}	temp_max_28	β_{15}	0.002	0.001	9.23e-06	***

R^2= 0.000

Note: Variables statistically significant at the 0 level are marked by a triple asterisk (***), variables statistically significant at the 0.001 level are marked by a double asterisk (**), variables statistically significant at the 0.01 level are marked by a single asterisk (*), and variables statistically significant at the 0.05 level are marked by a dot (.)

Based on Table 7 above 10 out of the original 15 were deemed as statistically significant. Therefore, in order to get a better indication of these variables' effects on run delays, the regression model was run once again, removing all statistically insignificant variables. The new model results are highlighted in Table 8 below.

Table 8. Summary of the Improved Multiple Linear Regression Model for Run Delay (4)

Coefficients:			**Estimate**	**St. E.**	**Pr(>|t|)**	
	(Intercept)	β_0	0.236	0.003	< 2e-16	***
x_1	prec_sum_1	β_1	0.017	0.002	< 2e-16	***
x_2	prec_sum_7	β_2	0.004	0.001	2.61e-06	***
x_3	prec_sum_14	β_3	0.003	0.001	8.88e-07	***
x_4	prec_sum_28	β_4	0.002	0.000	3.63e-10	***
x_5	temp_min_1	β_5	-0.005	0.000	< 2e-16	***
x_6	temp_min_7	β_6	-0.002	0.000	1.20e-15	***
x_7	temp_min_28	β_7	-0.002	0.000	1.34e-13	***
x_8	temp_max_1	β_8	0.005	0.000	< 2e-16	***
x_9	temp_max_14	β_9	0.002	0.000	8.54e-06	***
x_{10}	temp_max_28	β_{10}	0.001	0.000	4.96e-05	***
R^2= 0.000						

Variables statistically significant at the 0 level are marked by a triple asterisk

Based on Table 8 above all variables are deemed as statistically significant based on the p-value. Since the p-value is less than 0.05 (significance level) the null hypothesis can be rejected, and it can be said that these 10 variables have an impact on run delay size. β_0 here is 0.236, which represents the expected dwell delay size when all predictor variables are 0.

Regarding the coefficients, the sum of precipitation over 1, 7, 14, and 28 days; and the maximum temperature over 1, 14, and 28 days are positively related to run delay size. The minimum temperature over 1, 7 and 28 days is negatively related to run delay size. Here the coefficients indicate that run delay size increases with the sum of precipitation. The sum of precipitation over 1 day has the largest coefficient and smallest p-value indicating that 1 day of precipitation most likely has the most effect on run delays. The sum of precipitation over 21 days was not deemed as statistically significant indicating that precipitation over shorter and longer periods of time have more influence over run delay size compared to the middle time period of 21 days.

Temperature again is more complicated as when the coefficient is positive and temperature is positive run delay size increases, however when the temperature is negative with a positive coefficient the delay size decreases. In contrast when a coefficient is negative, and temperature is positive the size of delay decreases but when temperate and coefficient is negative the delay size increases. This relationship is clear as Minimum temperature over 1, 7, and 28 days has a negative coefficient indicating that temperatures below zero cause more delays. Additionally, the time periods indicate that there are more problems with delays when temperatures are below freezing for a short period of time, 1 day and a week; and also, when it is below freezing over a month long. Maximum temperature over 1, 14, and 28 days have positive coefficients which indicate delays increase with temperature. Since it is the maximum temperature over a certain time period it can be expected that these values are rarely below zero.

Once again, the R^2 value is 0.000 when rounded to the nearest thousandth, however all the p-values are under the significance level of 0.05 and the standard error values are low; this indicates that the model is still add value and the null hypothesis can be rejected. This means that precipitation and temperature have an impact on run delays.

4.2.3. Comparison of the Two Multiple Linear Regression Models

When comparing the multiple linear regression models between dwell and run delays the first major difference is the number of statistically significant predictor variables. Table 9 below visualises this difference.

Table 9. Statistically Significant Predictor Variables for Dwell Delay vs. Run Delay of the Linear Regression Model (4)

Dwell Delay	Run Delay
• **Sum of precipitation over 7 days**	• Sum of precipitation over 1 day
• **Minimum temperature over 1 day**	• **Sum of precipitation over 7 days**
• **Minimum temperature over 7 days**	• Sum of precipitation over 14 days
• Minimum temperature over 14 days	• Sum of precipitation over 28 days
• **Minimum temperature over 28 days**	• **Minimum temperature over 1 day**
• **Maximum temperature over 1 day**	• **Minimum temperature over 7 days**
• **Maximum temperature over 28 days**	• **Minimum temperature over 28 days**
	• **Maximum temperature over 1 day**
	• Maximum temperature over 14 days
	• **Maximum temperature over 28 days**

Note: Variables bolded are statistically significant in both dwell and run delays

According to Table 9 above, 6 variables are statistically significant for both dwell and run delays. This highlights how interconnected run and dwell delays are. There are more run times scheduled compared to dwell times. However, if there is a run delay on a single line track the effect can be transferred to the dwell time, resulting in a dwell delay if a train has to wait longer at the station before the run delay ahead is resolved. For dwell delays the sum of precipitation is only statistically significant over 7 days compared to run delays where it is statistically significant over 1, 7, 14, and 28 days. This may be due to that accumulative precipitation can be high over short and long periods of time. In the short term it can cause flash flooding, however in the long-term flooding can also occur. If a track is flooded trains may have to run at slower speeds which may not necessarily affect the dwell delays. In extreme cases where the train cannot drive anymore, or infrastructure is damaged due to a flood event this can also affect dwell delays. In the case of minimum and maximum temperatures they are statistically significant for up to 2 weeks and again over 1 month. The results indicate that minimum temperatures below freezing and high maximum temperatures result in the biggest dwell and run delay sizes.

4.2.4. Dwell Delay Multiple Logistic Regression Model

As the relationship between weather and delays is complex perhaps a linear regression model may not fully represent the true relationship. A multiple logistic regression was computed next in order understand the maximum likelihood of a delay. As explained in Chapter 2, the response variable is binomial, 1 representing the presence of a delay, and 0 representing the absence of a delay. The regression coefficients were estimated using the maximum likelihood method provided by the Generalised Linear Model (glm) command in R, and the Wald chi-square statistic test determined the statistical significance of each individual regression coefficient (β_n)

Table 10. Summary of the Multiple Logistic Regression Model for Dwell Delay (7)

Variable:		Estimate		St. E	z value	Pr(>|t|)	
	(Intercept)	β_0	-1.739	0.006	-291.007	< 2e-16	***
x_1	prec_sum_1	β_1	0.030	0.003	8.833	< 2e-16	***
x_2	prec_sum_7	β_2	0.014	0.001	9.528	< 2e-16	***
x_3	prec_sum_14	β_3	0.005	0.001	3.479	0.001	***
x_4	prec_sum_21	β_4	-0.006	0.001	-4.396	1.1e-05	***
x_5	prec_sum_28	β_5	0.012	0.001	13.225	< 2e-16	***
x_6	temp_min_1	β_6	-0.024	0.001	-47.134	< 2e-16	***
x_7	temp_min_7	β_7	-0.007	0.001	-10.870	< 2e-16	***
x_8	temp_min_14	β_8	-0.007	0.001	-8.664	< 2e-16	***
x_9	temp_min_21	β_9	-0.005	0.001	-5.266	1.4e-07	***
x_{10}	temp_min_28	β_{10}	-0.007	0.001	-8.920	< 2e-16	***
x_{11}	temp_max_1	β_{11}	0.025	0.001	45.508	< 2e-16	***
x_{12}	temp_max_7	β_{12}	0.001	0.001	0.981	0.327	
x_{13}	temp_max_14	β_{13}	-0.001	0.001	-0.597	0.550	
x_{14}	temp_max_21	β_{14}	-0.012	0.001	-9.498	< 2e-16	***
x_{15}	temp_max_28	β_{15}	0.038	0.001	39.991	< 2e-16	***

Note: Variables statistically significant at the 0 level are marked by a triple asterisk (***), variables statistically significant at the 0.001 level are marked by a double asterisk (**), variables statistically significant at the 0.01 level are marked by a single asterisk (*), and variables statistically significant at the 0.05 level are marked by a dot (.)

Based on Table 10 above it is evident that precipitation over 1, 14, 21, and 28 days are statistically significant and have an impact on dwell delays. In regard to temperature, the minimum temperature over 1, 7, 14, 21, and 28 days are statistically significant; and maximum temperature is statistically significant over 1, 21, and 28 days. In order to better understand the statistical significance, the regression model was run once more until all variables were deemed statistically significant. The number of variables analysed dropped from 15 to 13. The table also includes the calculated odds ratio which was also computed in R and the lower and upper limits of the 95% confidence intervals, which is shown in Table 11 below.

Table 11. Summary of the Improved Multiple Logistic Regression Model for Dwell Delay (7)

Coefficients:			Estimate	OR	C.I. 95% Lower	C.I. 95% Upper	St. E.	z value	
	(Intercept)	β_0	-1.784	0.176	0.174	0.178	0.006	-292.656	***
x_1	prec_sum_1	β_1	0.030	1.030	1.023	1.037	0.003	8.877	***
x_2	prec_sum_7	β_2	0.014	1.014	1.010	1.017	0.001	9.519	***
x_3	prec_sum_14	β_3	0.005	1.005	1.002	1.008	0.001	3.424	***
x_4	prec_sum_21	β_4	-0.006	0.994	0.991	0.997	0.001	-4.343	***
x_5	prec_sum_28	β_5	0.012	1.012	1.011	1.014	0.001	13.204	***
x_6	temp_min_1	β_6	-0.024	0.976	0.975	0.977	0.001	-47.501	***
x_7	temp_min_7	β_7	-0.007	0.993	0.992	0.994	0.001	-10.952	***
x_8	temp_min_14	β_8	-0.007	0.993	0.991	0.994	0.001	-8.782	***
x_9	temp_min_21	β_9	-0.005	0.995	0.993	0.997	0.001	-5.325	***
x_{10}	temp_min_28	β_{10}	-0.007	0.993	0.991	0.974	0.001	-8.934	***
x_{11}	temp_max_1	β_{11}	0.025	1.025	1.024	1.026	0.000	52.438	***
x_{12}	temp_max_21	β_{12}	-0.017	0.988	0.987	0.990	0.001	-12.143	***
x_{13}	temp_max_28	β_{13}	0.038	1.038	1.036	1.040	0.001	40.108	***

Note: Variables statistically significant at the 0 level are marked by a tripled asterisk (***).
Pr(>|z|) <2e-16

In a logistic model, increasing x_n by one unit will increase the log odds of an event happening by β_n. According to the improved model summary in Table 11 above some variables have a positive relationship and some a negative. Dwell delay increase is positively related to the sum of precipitation over 1, 7, 14 and 28 days, and to the maximum temperature over 1, and 28 days. In contrast, dwell delay increase is negatively related to the sum of precipitation over 21 days; the minimum temperature over 1, 7, 14, 21, and 28 days; and the maximum temperature of 21 days. The Wald chi-square test indicates that a large absolute z value and small p-value can reject the null hypothesis, demonstrating that these 13 variables play a role in dwell delays.

In R-studio the log odds are computed into odds and then odds ratio in order to understand the changes in delays from a probability point of view. Odds is known as the probability of a train being delayed divided by the probability that it is not delayed. In simple the odds ratio is the ratio of one odds against the other. For instance, the odds ratio is the odds of a train being delayed beyond a certain predictor variable threshold divided by the odds of the train being delay if it is not beyond that threshold. An odds ratio greater than 1 indicates a positive relation of a predictor variable towards the likelihood of a dwell delay increase. If it is lower the odds of a dwell delay is lower. The sum of precipitation over 21 days; the minimum temperature over 1, 7, 14, 21, and 28 days; and the maximum temperature over 21 days all have odds ratios below 1, which corresponds to the negative coefficients they display. However, they are quite close to 1.00 which may indicate that the odds of a delay are similar across all conditions. The odd ratios with values above 1 are the sum of precipitation over 1, 7, 14, and 28 days, and the maximum temperature over 1 and 28 days. Here it is seen that delays are most impacted by precipitation and minimum temperature over all time periods analysed, and the maximum temperature over short periods, one day and over extremely long periods, 21 days and one month.

4.2.5. Run Delay Multiple Logistic Regression Model

The regression coefficients for run delay were estimated using the maximum likelihood method provided by the Generalised Linear Model (glm) command in R, and the Wald chi-square statistic test determined the statistical significance of each individual regression coefficient (β_n). The results are highlighted in Table 12 below.

Table 12. Summary of the Multiple Logistic Regression Model for Run Delay (7)

Coefficients:		**Estimate**		**St. E.**	**Z-Value**	**Pr(>|t|)**	
	(Intercept)	β_0	-1.384	0.004	-366.537	< 2e-16	***
x_1	prec_sum_1	β_1	0.013	0.002	5.974	2.31e-09	***
x_2	prec_sum_7	β_2	-0.005	0.001	-0.585	0.558	
x_3	prec_sum_14	β_3	0.008	0.001	8.182	2.79e-09	***
x_4	prec_sum_21	β_4	0.000	0.001	0.272	0.786	
x_5	prec_sum_28	β_5	0.005	0.001	8.731	< 2e-16	***
x_6	temp_min_1	β_6	0.006	0.003	17.008	< 2e-16	***
x_7	temp_min_7	β_7	-0.001	0.000	-0.231	0.817	
x_8	temp_min_14	β_8	-0.002	0.001	-3.136	0.002	**
x_9	temp_min_21	β_9	0.003	0.001	4.091	4.30e-05	***
x_{10}	temp_min_28	β_{10}	0.007	0.001	12.354	< 2e-16	***
x_{11}	temp_max_1	β_{11}	-0.000	0.000	-0.172	0.864	
x_{12}	temp_max_7	β_{12}	0.002	0.001	3.819	0.000	***
x_{13}	temp_max_14	β_{13}	-0.001	0.001	-2.040	0.041	*
x_{14}	temp_max_21	β_{14}	-0.007	0.001	-0.878	0.380	
x_{15}	temp_max_28	β_{15}	-0.010	0.001	-15.065	< 2e-16	***

Note: Variables statistically significant at the 0 level are marked by a triple asterisk (***), variables statistically significant at the 0.001 level are marked by a double asterisk (**), variables statistically significant at the 0.01 level are marked by a single asterisk (*), and variables statistically significant at the 0.05 level are marked by a dot (.)

Based on Table 12 above it is evident that precipitation over 1 day, 14 days, and 28 days are statistically significant and have an impact on run delays. In regard to temperature, minimum temperature statistically significant over 1, 21, and 28 days; and maximum temperature is statistically significant over 7 and 28 days. In order to better understand the statistical significance, the regression model was run twice more until all variables were deemed statistically significant. In the first rerun variables 2, 4, 7, 8, 11, 13, and 14 were removed (Table 12). For the third run of the model the minimum temperature over 21 days and maximum temperature over 7 days was removed, and the Table 13 below highlights the improved summary of the multiple regression model. The number of variables analysed dropped from 15 to 6. The table also includes the calculated odds ratio which was also computed in R and the lower and upper limits of the 95% confidence intervals.

Table 13. Improved Multiple Logistic Regression Model for Run Delay (7)

Coefficients:		Estimate		OR	C.I. 95% Lower	C.I. 95% Upper	St. E.	z value	
	(Intercept)	β_0	-1.384	0.251	0.249	0.252	0.004	-376.997	***
x_1	prec_sum_1	β_1	0.012	1.013	1.008	1.017	0.002	6.086	***
x_2	prec_sum_14	β_2	0.007	1.007	1.006	1.008	0.001	11.940	***
x_3	prec_sum_28	β_3	0.005	1.005	1.005	1.006	0.000	13.911	***
x_4	temp_min_1	β_4	0.006	1.006	1.006	1.007	0.000	24.819	***
x_5	temp_min_28	β_5	0.008	1.008	1.007	1.008	0.000	33.429	***
x_6	temp_max_28	β_6	-0.010	0.990	0.989	0.990	0.000	-49.839	***

Note: Variables statistically significant at the 0 level are marked by a tripled asterisk (***).
Pr(>|z|) <2e-16

According to Table 13 above the sum of precipitation over 1, 14, and 28 days; and the minimum temperature over 1 and 28 days is positively related to dwell delay increase. This also corresponds with odds ratios that are over 1 which indicate the odds of a delay increasing. On the other hand, the maximum temperature over 28 days is negatively related to delay increase. This corresponds with odds ratios that are below one, and therefore the odds of delay increase are not as great as those with odds ratios above 1. This opposite to what is seen in the linear regression model, where maximum temperature is positively related to delays and minimum temperature is negatively related to delays. This discrepancy is most likely due to complications of temperature having both negative and positive values. Additionally, if all of the odds ratios are rounded to the nearest whole number, they are all 1; indicating the odds of a run delay are equal across all conditions.

4.2.6. Comparison of the Two Multiple Logistic Regression Models

When comparing the multiple logistic regression models between dwell and run delays the first major difference is the number of statistically significant predictor variables. Table 14 below visualises this difference.

Table 14. Statistically Significant Predictor Variables for Dwell Delay vs. Run Delay of the Logistic Regression Model (7)

Dwell Delay	Run Delay
• **Sum of precipitation over 1 day**	• **Sum of precipitation over 1 day**
• Sum of precipitation over 7 days	• **Sum of precipitation over 14 days**
• **Sum of precipitation over 14 days**	• **Sum of precipitation over 28 days**
• Sum of precipitation over 21 days	• **Minimum temperature over 1 day**
• **Sum of precipitation over 28 days**	• **Minimum temperature over 28 days**
• **Minimum temperature over 1 day**	• **Maximum temperature over 28 days**
• Minimum temperature over 7 days	
• Minimum temperature over 14 days	
• Minimum temperature over 21 days	
• **Minimum temperature over 28 days**	
• Maximum temperature over 1 day	
• Maximum temperature over 21 days	
• **Maximum temperature over 28 days**	

Note: Variables bolded are statistically significant in both dwell and run delays

The table above highlights that more predictor variables for dwell delay are statistically significant compared to run delays. There are 6 variables which are statistically significant for both dwell and run delays. They are precipitation over 1, 14, and 28 days; the minimum temperature over 1, and 28 days; and maximum temperature over 28 days. The dwell delays are impacted across all time periods, while run days only over 1, 14, and 28 days. Precipitation may be more statistically significant over more time periods because the longer time periods accumulate more precipitation. Large amounts of precipitation can be problematic whether it is over 1 day or 28 days. The minimum temperature is statistically significant over all time periods analysed for dwell delays but only over 1 and 28 days for run delays. This may highlight that minimum temperatures have a big impact on both dwell delays or run delays, signifying that low temperatures have a large impact on delays and can cause many issues. In the case of maximum temperature, it is only statistically significant over 28 days for run delays and over 1, 21, and 28 days for dwell delays. This indicates that the maximum temperature has the greatest impact over short periods of time or very long.

4.2.7. Comparison of Multiple Linear Regression Model vs. Multiple Logistic Regression Modelling

When comparing the linear and logistic models the first main difference is the number of variables which are deemed as statistically significant for dwell delays compared to run delays. In the linear model 7 variables are statistically significant compared to 10 variables for run delays. The logistic model is opposite in that 13 variables are significantly significant for dwell delays and 6 for run delays. Additionally, in the logistic model there are some discrepancies with the coefficients and therefore odds ratios. In the linear model all precipitation and maximum temperatures have positive coefficients while the minimum temperature coefficients are all negative. This can be justified with the fact that the high dwell and run delay sizes should occur when the sum of precipitation and/or maximum temperature is high, and when minimum temperatures are below freezing. However, in the logistic regression this does not hold to be true for all cases. This is most likely caused by a number of things. First, the dataset contains a lot of weather data that is deemed as “normal.” Meaning conditions that do not cause delays, therefore, the extreme conditions that effect delays the most are rare. Secondly, with logistic regressions it is typical to calculate the probability of a delay increase. Since temperature is both positive and negative this is harder to calculate because some sort of threshold needs to be established in order to calculate the probability. For instance, knowing the odds of a train being delayed if the temperature is beyond a certain threshold, divided by the odds it is not gives the odds ratio. Using the odds, the train is not delayed beyond a certain temperature threshold can help calculate the probability. Therefore, the odds ratios are a bit challenging to interpret since, the temperature and precipitation thresholds are not defined as this is outside the scope of this thesis.

Temperature most likely has more of an exponential relationship with run and dwell delays and therefore, a logistic model may be more accurate. However, since the thresholds are missing it makes it more challenging to interpret the results. The results from the linear regression are more in line with the expected results, and the figures are plotted linearly in Chapter 4.1, so therefore the linear model is chosen as the model of choice for this thesis.

4.3 Analysis Results Sub-conclusion

The results show that the most statistically significant variables for dwell delays are the sum of precipitation over 7 days, the minimum temperature over 1, 7, 14, and 28 days, and the maximum temperature over 1 and 28 days. The sum of precipitation over 1, 7, 14, and 28 days, the minimum temperature over 1, 7, and 28 days, and the maximum temperature over 1, 14, and 28 days are the most statistically significant variables for run delays. The regression models show that the relationship between weather and delays is that the size of a dwell or run delay increases with the sum of precipitation, and as maximum temperature increases; and that the size of a run dwell delay decreases as minimum temperature increases. These relationships indicate that with increased amounts of precipitation, extremely high temperatures, and temperatures below freezing the likelihood of a delay is higher. Additionally, it can be assumed that weather variables that are most statistically significant above have strong relationships and can more accurately predict delays. With these results in mind, it is possible to consider resilience and what these results mean for the further in regard to climate change, which will be discussed in the next chapter.

DISCUSSION

5.0 Discussion

In this chapter, some reflections on the findings are presented along with future implications, the importance of resilience, the limitations of this thesis, and some suggestions for future research.

5.1 Discussion of the Results

The analysis highlighted that the linear regression model was best suited to the scope of this thesis because it best matches the expected results, the graphical evaluations, and the results from similar studies. The results indicate that over the past 11 years, the most significant variables for dwell delays are the sum of precipitation over 7 days, the minimum temperature over 1, 7, 14, and 28 days, and the maximum temperature over 1 and 28 days. While the most statistically significant variables for run delays are the sum of precipitation over 1, 7, 14, and 28 days, the minimum temperature over 1, 7, and 28 days, and the maximum temperature over 1, 14, and 28 days. Moreover, the minimum temperature for both run and dwell delays are negatively related to delay size, meaning as temperature decreases delay size increases. The sum of precipitation and maximum temperature are positively related to delay size, meaning as precipitation and maximum temperature increase so does delay size.

More variables are statistically significant to run delays than dwell delays. This may be because there are more scheduled run times than dwell times. Meaning, as a train is moving between stations it is more likely to encounter harsh weather conditions compared to a train stationary at the station. In addition, when a run delay occurs it does not always mean the effect moves to the dwelled train. Dwell and run delays can occur together or independently. 21 days was not considered to be a statistically significant time period for any of the variables. This indicates that both precipitation, minimum and maximum temperature have a greater impact on rail delays over 1 day, 1 week, 2 weeks, and 1 month; highlighting that weather also has an impact on rail delays over longer periods of times, not only instantaneously.

5.1.1 Precipitation

The results suggest that dwell and run delay size increases with the sum of precipitation. Precipitation does not only include rain but also includes snow, sleet, hail, etc. One of the most severe implications of the sum of precipitation is flooding. Railway infrastructure and services are very vulnerable to flooding and can result in trains having to drive lower speeds, and/or cause damage to infrastructure (Bubeck, et al., 2019). Snow can also play a vital role in delays by causing visibility issues, and slower speeds (Figure 8). Skåne today is quite subject to cloudbursts, which occur when there are vast amounts of rainfall recorded in very short periods of only a few hours (Belusic, et al., 2019). Table 7 in Chapter 4.2.2 indicates that the sum of precipitation over 1 day has the largest coefficient, therefore coinciding with the largest rate of change. This means that cloudbursts likely currently have an impact on delays. However, longer time periods are also statistically significant suggesting that longer time periods of precipitation also can cause issues, as flooding can also occur long term. Michaelides, et al. (2014) reviewed three major EU funded projects that look at weather extreme impacts on transportation systems. The Extreme Weather impacts on European Networks (EWENT) established that when rainfall is equal to or exceeds 30mm a day harmful impacts are possible, when rail exceeds 100mm per day harmful impacts are likely, and when rainfall exceeds 150mm a day harmful impacts are certain. The dataset used for this thesis has recorded cumulative temperatures between 00mm-230mm, indicating that Skåne has periods of time with a lot of precipitation. Since the relationship suggests that both dwell and run delay size increase with sum of precipitation and is statistically significant the null hypothesis can be rejected, and it can be concluded that precipitation does affect rail delays. Several studies have come to similar conclusions (Xia et al.; 2013, Brazil et al., 2017). In Sweden, Palmqvist, Olsson, & Hiselius (2017b) have similar results indicating that punctuality drops 1.8% points when a quarter of all trains analysed in the study accumulated at least 30mm of precipitation.

5.1.2. Minimum Temperature

The results suggest dwell and run delay size decreases as the minimum temperature increases, implying that colder temperatures are correlated to higher delay sizes. The minimum temperature was found to be significantly statistical to dwell delays over 1, 7, 14, and 28 days, and over 1, 7, and 28 days for run delays. This suggests that minimum temperature makes the biggest impact over short time periods and long, highlighting that cold temperatures have a large impact on rail delays. The biggest increase in delay size occurs with temperatures below freezing. Looking at 7, 14, and 28 days gives a better indication that cold snaps have an impact on train delays. In simple terms a cold snap is defined as a period of consecutive days with a temperature below a certain threshold, based typically on minimum temperatures over at least 3 days (Belusic, et al., 2019). Cold snaps can lead to track breakages, freezing of signals and other equipment, and icy conditions on the tracks. The longer the duration of these conditions the longer the impacts on delays will maintain. EWENT established that when temperatures are below 0˚C harmful impacts are possible, when temperatures are below -7˚C harmful impacts are likely, and when temperatures are below <-20˚C harmful impacts are certain (Michaelides, et al. 2014). Although defining the thresholds was outside the scope of this thesis when looking at Figures 11-20 in Chapter 4.1.2 the trends seem to follow a similar trend and the regression indicates that the null hypothesis can be rejected, and that minimum temperature plays a role in rail delays. Furthermore, the lowest minimum temperature recorded in the dataset over the 11 years studied is -19˚C, indicating that temperatures in Skåne are well below freezing, leading to delays. These findings are similar to what is found in other studies highlighting that freezing temperatures are correlated to rail delays (Zakeri & Olsson, 2017; Zakeri & Olsson 2018). In Sweden, Palmqivst, Olsson, & Hiselius (2017) discovered that punctuality falls exponentially when temperatures fall below 0˚C, and the punctuality continues to fall as temperature does; they suggest that at -5˚C the punctuality in Skåne already drops by 7.5%.

5.1.3. Maximum Temperature

In regard to maximum temperature the results indicate that both dwell and run delays increase as maximum temperature increases, implying that higher temperatures are correlated with more rail delays. The maximum temperature was found to be statistically significant over 1 and 28 days for dwell delays and 1, 14, and 28 days for run delays. This suggests that the maximum temperature has the greatest impact on rail delays instantaneously over 1 day, and longer term over 14 and 28 days. This means that heatwaves most likely negatively impact punctuality of trains in Southern Sweden. Similar to a cold snap, a heatwave can be defined as a period of consecutive days, typically at least 3 days where the temperature is above a certain threshold. Extremely high temperatures can lead to many infrastructure issues such as track buckling, and the overheating of equipment which can also lead to fire. Like cold spells, the longer a heatwave lasts the more negative implications on rail delays the event can have. The EWENT project concluded that harmful impacts are possible when temperatures exceed 25˚C, harmful impacts are likely when temperatures are above 32˚C, and harmful impacts are certain when temperatures exceed 43˚C (Michaelides, et al. 2014). The maximum temperature in the dataset recorded over 11 years is 33˚C suggesting that Skåne today may have more issues with temperatures below freezing than extreme high temperatures due to its northern geographic location. The results here are similar to other studies which quantify the effects of heat on railways (Dobney, et al., 2009; Forzieri, et al., 2018). In Sweden in particular, Palmqvist, Olsson, & Hiselius, 2017b found that punctuality drops by 5% at 23˚C and by 26% at 27˚C.

5.2 Future Implications of the Findings and Resilience Thinking

The results from the graphical evaluations and linear regression analysis demonstrate the railway industry's current vulnerability to the current climate. The climate change projections for Skåne indicate that this sensitivity is expected to increase. Understanding the effects of weather over the past 11 years gives a good basis for infering how the effects will differ as climate change progresses. Climate change is expected to raise temperatures in Skåne in both the summer and winter months. This is expected to lead to more frequent heatwaves, which can lead to track buckling, overheated signals, and fires. The amount of precipitation is expected to increase more in the summer months compared to winter months (SMHI, 2015), and for more precipitation to shift from snow to rain, which increases the risk of flooding. As historic weather extremes become more common, it is important to increase the resilience of railways. Better understanding these historic patterns can thus help identify, prioritise, and motivate adaptation measures.

5.2.1 The Future Implications of Precipitation

In Northern Europe the IPCC has confidence that in the future there will be an increase in daily precipitation extremes in Northern Europe for all seasons, due to a robust poleward shift of circulation patterns (Belusic, et al., 2019). It is expected that in Sweden, and Skåne in particular, there will be a greater increase in precipitation in the winter months compared to summer (Belusic, et al., 2019; SMHI, 2015; Lindgren, Jonsson, & Carlsson-Kanyama, 2009). With the added expected increase in temperature the amount of precipitation that falls as snow is expected to decrease (SMHI, 2015).

As mentioned, the greatest implication of increased precipitation is related to flooding. Currently in some parts of Sweden spring snowmelt contributes to the highest chance of flooding; however, snowmelt induced floods are expected to decrease, and floods occurring in the autumn and winter as a consequence of more rainfall are expected to increase (Belusic, et al., 2019). The probability of extreme 100 and 200-year floods is also expected to increase (Belusic, et al., 2019), indicating that not only are flood events expected to increase but their severity will as well. In urban areas flash and urban floods which are triggered by local intense precipitation events are also more likely to increase across Europe in general (Kellermann, et al., 2016). Increased precipitation is also associated with erosion and landslides which can put railways at risk (Belusic, et al., 2019). However, rain is not the only form of precipitation. The increasing temperatures and moisture content of the atmosphere allows favorable conditions for the increase of hailstorms and thunderstorms and their intensity across Sweden (Belisuic, et al., 2019).

Currently, the linear regression model shows that both dwell and run delays increase as precipitation increases. Due to the projected increase in precipitation, it can be expected that these delays will only get more frequent and severe. It is known that increases in heavy precipitation events have a negative impact on transportation and infrastructure (Michaelides, et al., 2014). Flooding and intense storms can cause damage to infrastructure and also result in visibility issues. Flooding can also affect telecommunications as power failure can reduce railway capacity (Belusic, et al., 2019). If these events are expected to increase it can be expected that the delays will also increase if nothing is done to increase the resilience of railways to climate change.

5.2.2 The Future Implications of Increased Temperatures

Overall, climate change projections predict an increase in temperatures in both summer and winter months (SMHI, 2015). In Skåne, projections show an increase in up to 6˚C by the end of this century under a business-as-usual scenario (SMHI, 2015). The frequency, duration, and severity of heatwaves is expected to increase across all of Europe (Belsuic, et al., 2019); and in Skåne the number of consecutive days above 20˚C is also expected to increase (SMHI, 2015). This implies that Skåne will experience more frequent and extreme heatwaves in the area, occurring for longer periods of time. In contrast, the extremes associated with cold temperatures are projected to decrease (Belusic, et al., 2019; SMHI, 2015). Although there is a general decreasing trend in winter extremes in Sweden, the winter extremes are still expected to impact railway transportation and would still need to be considered in investment in preparedness and maintenance strategies in the future (Michaelides, et al., 2014).

The results from the linear regression and graphical evaluations show that freezing temperatures have a greater impact on rail delays compared to extreme high temperatures. However, it can be argued that in the future freezing temperatures will still play a role in Skåne. The lowest temperature recorded in the dataset is -19˚C, and if the projections of a 6˚C increase in temperature becomes prevalent, then this minimum temperature will still be below freezing. Studies have suggested that harmful impacts can still occur with temperatures around -5˚C or -7˚C, which is not significantly below the freezing point (Michaelides, et al., 2014; Palmqvist, Olsson, & Hiselius, 2017b). Especially, if these freezing temperatures occur in interaction with heavy precipitation such as snow or freezing rain, these events can have serious impacts on railway infrastructure and therefore delays.

Although the graphical evaluations highlighted a greater probability of delays during freezing temperatures compared to high temperatures it can be argued that Skåne has yet to fully experience the effects of extreme high temperatures. The highest maximum temperature value in the dataset is 33˚C, and close to this point the graphical evaluations start to indicate an increase in the probability of a delay. Some studies suggest that the likelihood for failures associated with high temperatures do not occur until 32˚C and even up to over 40˚C (Michaelides, et al. 2014). Countries in Northern Europe are typically not yet set up with the coping range to deal with extreme heat (Lindgren, Jonsson, & Carlsson-Kanyama, 2009). The increased risk of fires, rail buckling, the installation of cooling systems, as well as passenger comfort are some concerns related to heat that will have to be dealt with more often in the future (Lindgren, Jonsson, Carlsson-Kanyama, 2009).

The future implications go in line with the results of the regression analysis under current conditions. The maximum temperature is positively related to both dwell and run delays, and statistically significant over 1, 14, and 28 days. This means that heatwaves are already impacting delays and it can be expected that the relationship between maximum temperature and rail delays will get stronger as heatwaves become more frequent, last longer, and are more extreme. Similarly, the minimum temperature is negatively related to dwell and run delays, and is statistically significant over 1, 7, 14, and 28 days. Although winter temperatures are expected to increase, they can still be below freezing in the future which is shown to still cause delays. Furthermore, when minimum temperatures are above freezing, they have a positive relationship with delays. This means that in the future, days where the minimum temperature is extremely hot, delays still can occur.

5.2.3 Resilience of Skåne Railways in the Future

Based on the future implications of climate change it is evident that Skåne will experience more precipitation and increasing temperatures, which is likely to reduce resilience if nothing is done. Therefore, despite the projected increase in interruptions, increasing the resilience of railways to weather phenomena is essential to ensure that railways remain punctual. Railways need to be reliable in order for them to be the first-choice mode of transportation of customers (Lindgren, Jonsson, & Carlsson-Kanayama, 2009). From the findings of the analysis high and low temperatures, and increased amounts of precipitation have large impacts on the punctuality, which is similar to other studies which have also indicated that more attention should be given to ensuring that railways are resilient to the future impacts of climate change (Palmqivst, Olsson, & Hiselius, 2017; Diab & Shalaby, 2020; Michaelides, et al., 2014; Bešinović, 2020).

In this thesis the resilience of a railway transportation system is defined as "*the ability of a railway system to provide effective services in normal conditions, as well as to resist, absorb, accommodate, and recover quickly from disruptions or disasters"* (Bešinović, 2020, p. 461). This thesis has focused on the disruptions from temperature and precipitation. The results indicate that most delays occur during temperatures below freezing, and with increased amounts of precipitation. Therefore, it can be argued that the system needs to increase its ability to withstand extreme weather impacts, to operate in the face of extreme weather, and promptly recover from the effects (Diab & Shalaby, 2020). This holds especially true as climate change predictions project an increase in precipitation and temperature in Skåne. This means the railway system needs to be prepared for even more frequent and severe delays to occur as the extreme conditions that cause delays today are expected to increase.

Trafikverket owns and manages around 80% of all railways in Sweden (Lindgren, Jonsson, & Carlsson-Kanyama, 2009). In 2009, Lindgren, Jonsson, & Carlsson-Kanyama released an interview study about climate adaptation of railways in Sweden. It can be argued that when railways are adapted to climate change this in turn also increases the resilience as railways are able to operate efficiently under normal conditions as well as recover quickly from disruptions. The study critiqued Trafikverket's climate adaptation strategies, stating it is unclear if the measures taken today are with intentions of adapting to future climate change or are a result of coping with current climate variations. For instance, the establishment of tree-free zones to protect tracks from falling trees is a direct result of a hurricane in January 2005. The study concluded that proactive, anticipatory, and planned adaptation strategies for climate change projections is lacking. They argue that climate change should already be considered in the early planning stages and that vulnerability assessments aimed at analysing future climate change impacts cannot be based on past events alone.

In 2018, Sweden released a new Climate Adaptation Strategy (Liljegren, 2019). Climate adaptation strategies up until 2018 in Sweden has been decentralised with responsibility lying predominately with individual actors and municipalities (Liljegren, 2019). However, with the implementation of this new strategy responsibility is given to government organisations and agencies such as Trafikverket. This led Trafikverket to release their own report on climate adaptation and resilience strategies. They highlight the urgency for their transportation to be adaptable to climate change as the extreme weather events are predicted to be more common in the future (Liljegren, 2019). The document highlights that climate adaptation is a cross-sectoral issue with many actors involved (Liljegren, 2019). The overall goal of transport policies within Trafikverket ensures socio-economic, efficient, long-term sustainable transportation for all businesses and citizens across Sweden, and to reduce vulnerabilities (Liljegren, 2019). Trafikverket has 21 objectives in three parts to achieve more robust systems and a strategy map to "*plan, maintain, enable, and build*" (Liljegren, 2019, pg. 19). Part 1 includes creating conditions for effective work with climate adaptation. Here, Trafikverket plans to collaborate nationally and internationally, establish how climate change will impact transportation, and establish methods for determining what measures are cost-effective for climate adaptation. Part 2 involves preventing the negative effects of climate change by creating resilient and robust facilities. Here, Trafikverket has planned to adapt maintenance methods to the impact of climate change and evaluate the risk points in existing railway facilities. Lastly, part 3 involves managing the effects of climate change. Here they plan to have a high level of preparedness and knowledge on how to handle the extreme effects of climate change (Liljegren, 2019). In their report it is highlighted that increasing existing facilities' resilience to climate change has no activity yet (Liljegren, 2019).

Based on this it is evident that Trafikverket has started thinking more about climate change adaptation strategies. The future implications of climate change projections highlight that extreme weather events will become more common and therefore the resilience of railways in Skåne must be increased. Compared to the 2009 study, now that Trafikverket has more responsibility in climate change adaptation and resilience strategies more planned, proactive, and anticipatory strategies are in place. However, now it is time for Trafikverket to put their plans into action and start working on their goals. Both Lindgren, Jonsson, & Carlsson-Kanyama (2009) and Trafikverket state the importance of understanding the different vulnerabilities, climate threats and their consequences of the past in order to guide the implementation of strategies and measures in the future. The results of this thesis highlight the vulnerability of Skåne railways over the past 11 years to high amount of precipitation, high temperatures, and temperatures below freezing;

indicating that more resilience strategies should focus on mitigating the effects of these weather conditions.

The Austrian Railway operator ÖBB is an example of a railway system that is brought up in the literature that is striving towards resilience in the future. Their risk management strategy puts a great emphasis on precautionary, non-structural, and preparatory risk mitigation strategies (Kellermann, et al., 2016). In particular they have a weather monitoring and warning system which has been in effect since 2005 (Kellermann et al., 2016). The system is programmed with thresholds for very high and low temperatures as well as intensive precipitation to warn ÖBB of any extreme weather events (Kellermann, et al., 2016). This allows the ÖBB to predict extreme weather events before they happen and take action to ensure the safety of the passengers and staff and ensure that delays are kept to a minimum. If weather events can be forecasted beforehand it may be easier to reroute trains or give passengers other options in order to arrive to their destination on time; especially since railways are more challenging to reroute after a disruption compared to road networks due to single tracks. Although Austria has a different physical geography compared to Skåne, the country is subject to similar flooding events triggered by heavy precipitation, temperatures below freezing, and rising temperatures in the future. A warning system such as this one is an example of a non-structural and preparatory risk mitigation strategy Skåne railway systems could use. Currently, Trafikveket's strategies are in the early stages of first understanding the effects of climate change on Sweden's transportation systems. Additionally, it was highlighted in Trafikverket's strategies the opportunity for international collaboration. Looking towards other countries with active resilience strategies already in place could be beneficial for Trafikverket in order to increase the resilience of their own railways in Skåne.

Sweden is a part of the EU and therefore has signed several international agreements and frameworks that deal with climate adaptation of transportation infrastructure and resilience (Liljegren, 2019). For instance, the EU "Sustainable and Smart Mobility Strategy" is a strategy within the EU and therefore, Sweden is subject to reach these goals. Moreover, 2021 is the European Year of Rail which also aims at cutting carbon emissions in the transportation sector and to encourage more people to use railways. Although weather is not the only factor that impacts delays in Sweden or in other countries within the EU, the results of this thesis indicate it still has a significant impact and is expected to worsen in the future as climate changes towards warmer temperatures and increased amounts of precipitation. Railways are already highlighted as a mode of transportation which reduces carbon emissions within the transportation system. Therefore, railway resilience to climate change is essential in ensuring trains remain punctual in order for the EU to achieve their sustainable transportation goals and to increase rail ridership.

In 2015 Sweden also adopted the UN Agenda 2030 which consists of 17 Sustainable Development Goals (Liljegren, 2019; UN, n.d). In the agreement, the importance is highlighted that all countries work with climate adaptation and other goals towards a more sustainable future (Liliegren, 2019). SDG 7 *Affordable and Clean Energy*, SDG 11 *Sustainable Cities and Communities*, and SDG 13 *Climate Action* are three SDGs that pertain to this thesis and to the resilience of railways in Skåne. It can be argued that SDG 7 *Affordable and Clean Energy* is already well on its way due to the high number of electrified railways which lower carbon emissions within the transportation sector. However, there is still work to be done in some other goals such as SDG 11 *Sustainable Cities and Communities* and SDG 13 *Climate Action.* SDG 11 promotes the shift towards sustainable transportation in urbanised areas, and SDG 13 promotes the need for climate change adaptation and resilience. SDG 13 also consists of two sub-goals that aim to increase resilience to climate related hazards and to integrate climate change measures into national strategies, polices, and planning (Liljegren, 2019; United Nations, n.d.). The results of this thesis indicate that the impacts of temperature and precipitation on railway delays is already prevalent, and this is expected to become more prevalent in the future. As mentioned, the electrification of railways in Sweden already promotes clean energy and works towards climate action. However, if railways are unable to adapt to climate change and become unreliable, then people will seek other modes of transportation which may not exhibit clean energy and counteract climate action. Initiatives such as the SDGs and EU sustainable transportation goals act as an important step towards Skåne railways becoming more resilient to the effects of climate change in the future as these hold Trafikverket accountable to increasing the resilience of railway infrastructure and systems in order to achieve the goals they have agreed to achieving.

5.3 Assumptions and Limitations

The ability to carry out an analysis such as this one has greatly improved compared to studies in the past due to the amount of data accessible. However, there are still some assumptions and limitations of the study that should be considered when dealing with large amounts of data.

When downloading weather data from SMHI, there were many decommissioned temperature stations, especially along major railways lines in the Eastern part of Skåne. This led to a significant gap between the number of temperature and precipitation stations. This means that there was data from more precipitation stations than temperature stations which may slightly skew results. Furthermore, sometimes there are issues with weather stations and therefore measurements are occasionally missed. Since there is 11 years of data, we can assume that the gaps affect the results very minimally, however it is still important to mention.

Another limitation of the study is that the regression models were run with data that mostly coincides with "normal" conditions, which most likely do not cause delays. The more extreme cases which can lead to flooding, heatwaves, or cold snaps are less likely to occur but have the greater impact on delays. Therefore, the R^2 values for the linear regression models were close to 0 and the coefficients were also smaller than originally expected. The regression models also only model the relationship between delays and the chosen weather variables. In reality, there are many other factors which can influence delays such as other infrastructure faults, passenger behaviour, or issues with operations.

Many studies have focused on only 1 year or relatively short time spans compared to this thesis which looks at 11 years. Therefore, the results may have differed if more years were added to the analysis, for instance 30 or 50 years. Furthermore, results may have different if only one year was chosen with particularly "abnormal" weather; for instance, a year with an extreme cold winter.

5.4 Suggestions for Future Research

This thesis aimed to act as a basis for understanding the current effects of weather on rail delays. Future studies may include modelling the future effects of climate change in order to quantify how climate change will impact future rail delays by forecasting future climate. For future research, the interactions between temperature and precipitation and other weather variables such as wind and snow depth may be considered. For instance, a blizzard is an extreme event which occurs when temperatures are below freezing, it is snowing and windy. Droughts are associated with low amounts of precipitation and high temperatures. Understanding these interactions can be beneficial in further understanding the effects of weather on railways. Conducting similar studies in all of Sweden can allow for a comparison to Skåne to indicate which weather variables affect other regions of Sweden and therefore which adaptation strategies should be implemented. It may also be worth performing similar studies on freight trains. Finally, this study could be enhanced by identifying the extreme thresholds to understand what temperatures and amounts of precipitation exactly lead to the most severe delays. This thesis discusses about the implications of “cold” and “hot” therefore, being able to define these definitions with thresholds can improve the research in the future.

CONCLUSION

6.0 Conclusion

Railways have a large potential in the race to mitigate climate change due to their low emissions. However, this potential can only be reached if they are resilient and able to adapt to the increasingly extreme weather phenomena that are associated with climate change projections. This thesis' aim was to examine the Skåne railway industry's current vulnerability to temperature and precipitation and to discuss how this vulnerability is expected to increase with climate change, due to the sensitivity railways exhibit towards disruptions.

The results from the multiple linear regression model revealed that the most statistically significant variables for dwell delays are the sum of precipitation over 7 days, the minimum temperature over 1, 7, 14, and 28 days, and the maximum temperature over 1 and 28 days. Additionally, the most statistically significant variables for run delays are the sum of precipitation over 1, 7, 14, and 28 days, the minimum temperature over 1,7, and 28 days, and the maximum temperature over 1, 14, and 28 days. The models highlight that the size of a dwell or run delay increases with the sum of precipitation and maximum temperature. In contrast the size of a dwell or run delay decreases as minimum temperature increases. This indicates that delay size is biggest with increased amounts of precipitation, extreme high temperatures, and when temperatures are below freezing. Graphical evaluations revealed similar trends with the highest probability of dwell and run delays occurring when temperatures are below freezing and with increased precipitation.

Understanding this vulnerability over the past 11 years gives a better indication of how this will change with future climate change projections. In Skåne, it is projected that temperatures and precipitation are expected to increase. The results indicate that railways already have issues with providing reliable services under current conditions. Therefore, it is expected that as climate changes and temperatures and precipitation are more extreme the delay likelihood will increase with it. Looking towards the future, it is important that railways become more resilient to climate change. The EU has goals to promote railways as a sustainable mode of transportation and to increase ridership. Therefore, punctuality is essential in order to achieve these goals. Increasing the resilience of railways in Skåne to the effects of extreme weather both today and in the future is vital in ensuring that trains arrive on time, operate efficiently, and act as the first-choice mode of transportation for passengers.

The ability to carry out an analysis such as this one has improved greatly compared to studies in the past due to the amount of data accessible. This thesis contributes to this field of research by investigating the sum of precipitation, minimum temperature, and maximum temperature from the past 11 years over a range of time-periods, specifically, 1, 7, 14, 21, and 28 days; not only investigating the daily effects which have mainly been studied so far. Analysing the accumulative effects, allows for studying the impacts of more extreme weather events such as heatwaves, cold snaps, and periods of intensive rainfall which may lead to flooding. Furthermore, this thesis focused on Skåne, an area where fewer studies have been conducted. It is an important area to study as it is one of the most populated areas in Sweden and many passengers are commuting for work and leisure, and therefore are reliant on punctual services.

The theoretical framework used in the analysis was Resilience Thinking. This framework provided a basis for understanding how resilience is defined and used in other fields. Then it was narrowed down specifically to uses within the transportation sector to define resilience of railways as a system that is able to resist and bounce back quickly from interruptions. The main methods used throughout this thesis were graphical evaluations, regression modelling, and a literature review in order to determine the current relationship between precipitation, temperature, and delays; and to discuss the future implications of this relationship in light of climate change.

REFERENCES

References

Adjetey-Bahun, K., Birregah, B., Châtelet, E., & Planchet, J-L. (2016). A Model to Quantify the Resilience of Mass Railway Transportation Systems. Reliability Engineering and System Safety 153(2016): 1-14. Retrieved from: http://dx.doi.org/10.1016/j.ress.2016.03.015

Aparicio, Á., Leitner, M., Mylne, K., Palin, E., & Sobrino, N. (2013). Support to transport and environment assessments. Adaptation to Climate Change in the Transport Sector. ETC/CCA Technical Paper 03/2013, Bologna, Italy.

Armstrong, J., Preston, J., & Hood, I. (2017). Adapting Railways to Provide Resilience and Sustainability. Engineering Sustainability 170(4): 225-234. Retrieved from: https://doi.org/10.1680/jensu.15.00017

Armstrong, J. & Preston, J. (2011). Alternative Railway Futures: Growth and/or Specialisation? Journal of Transport Geography 19(2011): 1570-1579. Doi: 10.1016/j.jtrangeo.2011.03.012

Belusic D., Berg P., Bozhinova, D., Bärring, L., Döscher, R., Eronn, A., Kjellström, E., Klehmet K., Martins, H., Nilsson, C., Olsson, J., Photiadou, C., Segersson, D., Strandberg, G. (2019). Climate Extremes for Sweden. *Svergies meterologiska och hydrologiska institut*. Retrieved from: http://smhi.diva-portal.org/smash/get/diva2:1368107/FULLTEXT01.pdf [20.05.21]

Bešinović, N. (2020). Resilience in Railway Transport Systems: A Literature Review and Research Agenda. *Transport Reviews 40*(4): 457-478. Doi: 10.1080/01441647.2020.1728419

Brazil, W., White, A., Nogal, M., Caulfield, O'Connor, A., & Morton, C. (2017). Weather and Rail Delays: Analysis of Metropolitan Rail in Dublin. *Journal of Transport Geography 59* (2017): 69-76. Retrieved from: https://doi.org/10.1016/j.jtrangeo.2017.01.008

Brons, M.R.E., & Rietveld, P. (2008) Rail Mode, Access Mode and Station Code: The Impact of Travel Time Unreliability. (Research Report of the project TRANSUMO BTK). VU University. Retrieved from: https://research.vu.nl/en/publications/rail-mode-access-mode-and-station-choice-the-impact-of-travel-tim

Bryman, A. (2012). Social Research Methods. 4th Edition. Oxford University. New York, New York.

Bubeck, P., Dillenardt, L., Alfieri, L., Feyen, L., Thieken, A.H., & Kellermann, P. (2019). Global Warming to Increase Flood Risk on European Railways. *Climatic Change 155*: 19-36. Retrieved from: https://doi.org/10.1007/s10584-019-02434-5

Chan, R. & Schofer, J.L. (2016). Measuring Transportation System Resilience: Response of Rail Transit to Weather Disruptions. *Natural Hazards 17*(1): 05015004-1-8. Doi: 10.1061/(ASCE)NH.1527-6996.0000200

Chen, Z., Wang, Y., & Zhou, L. (2021). Predicting Weather-Induced Delays of High-Speed Rail and Aviation in China. *Transport Policy 101* (2021) 1-13. Retrieved from: https://doi.org/10.1016/j.tranpol.2020.11.008

Diab, E., & Shalaby, A. (2020). Metro Transit System Resilience: Understanding the Impacts of Outdoor Tracks and Weather Conditions on Metro System Interruptions. *International Journal of Sustainable Transportation 14*: 657-670. Doi: 10.1080/15568318.2019.1600174

Dobney, K., Baker, C.J., Quinn, A.D., & Chapman, L. (2009). Quantifying the Effects of High Summer Temperatures due to Climate Change on Buckling and Rail Related Delays in South-East United Kingdom. *Meteorological Applications 16*: 245-251. Doi: 10.1002/met.114

European Commission (2013). Adapting infrastructure to climate change. SWD (2013) 137 final of 16/4/2013, Brussels. Retrieved from: https://ec.europa.eu/clima/sites/clima/files/adaptation/what/docs/swd_2013_137_en.pdf [12.03.21]

European Commission (n.d.). *A Little Bit of History!* European Year of Rail. Retrieved from: https://europa.eu/year-of-rail/why-rail_en [22.03.21]

European Commission (2020). *The Journey Begins - 2021 Is the European Year of the Rail!* European Commission Retrieved from: https://ec.europa.eu/commission/presscorner/detail/en/IP_20_2528 [22.03.21]

European Union, European Parliament and Council (2020) Directive COM (2020) 789 final. Communication from the Commission to the European Parliament, The Council, The European Economic and Social Committee and the Committee of the Regions Sustainable and Smart Mobility Strategy - Putting European Transport on Track for the Future. Retrieved from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0789 [22.03.21]

European Union (n.d.) *Skåne.* URBACT. Retrieved from: https://urbact.eu/skane [23.02.21]

EU Rail (n.d.). *SJ High-speed Rail*. EU Rail. Retrieved from: https://www.eurail.com/en/get-inspired/trains-europe/high-speed-trains/sj [02.02.21]

Fleming, J. & Ledogar, R.J. (2008). Resilience, an Evolving Concept: A Review of Literature Relevant to Aboriginal Research. *Pimatisiwin 6*(2): 7-23. Retrieved from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2956753/pdf/nihms387.pdf

Flyvbjerg, B. (2006). Five Misunderstandings About Case-Study Research. Qualitative Inquiry 12(2): 219-245. Retrieved from: http://dx.doi.org/10.1177/1077800405284363

Forzieri, G., Bianchi, A., Batista e Silva, F., Herrera, M.A.M., Leblois, A., Lavalle, C., Aerts, C.J.H., Feyen, L. (2018). Escalating Impacts of Climate Extremes on Critical Infrastructures in Europe. *Global Environmental Change 48*(2018): 97-107. Retrieved from: doi.org/10.1016/j.gloenvcha.2017.11.007

Greater Copenhagen (n.d.). *About Greater Copenhagen*. Greater Copenhagen. Retrieved from: https://www.greatercph.com/about [12.02.21]

Intergovernmental Panel on Climate Change. (2014a). Barros, V.R., Field, C.B., Dokken, D.J., Mastrandrea, M.D., Mach, K.J., Bilir, T.R., Chatterjee, M., Ebi, K.L., Estrada, Y.O., Genova, R.C., Betelhem, G., Kissel, E.S., Levy, A.N., MacCracken, S., Mastrandrea, P.R., & White, L.L. (eds.) Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects.

Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. Retrieved from: https://www.ipcc.ch/site/assets/uploads/2018/02/WGIIAR5-PartB_FINAL.pdf

Intergovernmental Panel on Climate Change. (2014b). Core Writing Team, Pachauri, R.K., Meyer, L. (eds). Climate Change 2014: Synthesis Report. Geneva, Switzerland. Retrieved from: https://www.ipcc.ch/site/assets/uploads/2018/05/SYR_AR5_FINAL_full_wcover.pdf

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning: with Applications in R. Springer.

Kellermann, P., Bubeck, P., Kundela, G., Dosio, A., & Thieken, A.H. (2016). Frequency Analysis of Critical Meteorological Conditions in a Changing Climate - Assessing Future Implications for Railway Transportation in Austria. *Climate 4*: 25. Doi: 10.3390/cli4020025

Liljegren, E. (2019). *Regeringsuppdrag om Trafikverkets klimatanpassningsarbete.* Trafikverket. Retrieved from: https://trafikverket.ineko.se/Files/sv-SE/60329/Ineko.Product.RelatedFiles/2018_195_regeringsuppdrag_om_trafikverkets_klimatanpassningsarbete.pdf [22.05.21]

Lindgren, J., Jonsson, D.K., & Carlsson-Kanyama, A. (2009). Climate Adaptation of Railways: Lessons from Sweden. EJTIR 9(2): 164-181. Retrieved from: https://journals.open.tudelft.nl/ejtir/article/view/3295/3462

Ling, X., Peng, Y., Sun, S., Li, P., & Wang, P. (2018). Uncovering Correlation Between Train Delay and Train Exposure to Bad Weather. *Physica A 512* (2018): 1152-1159. Retrieved from: https://doi.org/10.1016/j.physa.2018.07.057

Leviäkangas, P., Tuominen, A., Molarius, R., Kojo, H., Schabel, J., Toivonen, S., Keränen, J., Ludvigsen, J., Vajda, A., Tuomenvirta, H., Jurga, I., Nurmin, P., Rauhala, J., Rehm, F., Gerz, T., Muehlhausen, T., Schweighofer, J., Michaelides, S., Papadakis, M., Dotzek, N., & Groenemeijer, P. (2011). D1 Review on Extreme Weather Impacts on Transport Systems. VTT Working Papers 168. ISBN: 978-951-38-7509-1. Retrieved from: https://www.vttresearch.com/sites/default/files/pdf/workingpapers/2011/W168.pdf

Mattsson, L-G., Jenelius, E. (2015). Vulnerability and resilience of transport systems – A discussion of recent research. Transportation Research Part A, 81: 16-34. Retrieved from: http://dx.doi.org/10.1016/j.tra.2015.06.002

Michaelides, S., Leviäkangas, P., Doll, C., Heyndrickx, C. (2014). Foreward: EU-Funded Projects on Extreme and High-Impact Weather Challenging European Transport Systems. *Natural Hazards 72*: 5-22. Doi: 10.1007/s11069-013-1007-1

Misnevs, B., Melikyan, A., & Bazaras, D. (2015). Hazard Assessment of Weather Factors for the Occurrence of an Emergency on the Railway. Procedia Computer Science, 77: 40–47. Retrieved from: https://doi.org/10.1016/j.procs.2015.12.357

Ochsner, M. (2021). *Climate and Urban Development Factors Affecting the Vulnerability of Malmbanan.* Aalborg University. Retrieved from: https://projekter.aau.dk/projekter/files/402545895/Michelle_Ochsner_P3_Project.pdf [05.03.2021]

Olsson, N.O.E. & Haugland, H. (2004). Influencing Factors on Train Punctuality - Results from Some Norwegian Studies. *Transport Policy 11*(2004): 387-397. Doi: 10.1016/j.tranpol.2004.07.001

Oslakovic, I.S., ter Maat, H., Hartmann, A., & Dewulf, G. (2013). Risk Assessment of Climate Change Impacts on Railway Infrastructure. Engineering Project Organization Conference, Devil's Thumb Ranch, Colorado, USA, 2013-07-09/2013-07-11. Retrieved from: https://library.wur.nl/WebQuery/wurpubs/451524

Palmqvist, C.W. (2019). Delays and Timetabling for Passenger Trains (PhD thesis). Retrieved from: https://www.trafikverket.se/contentassets/27b69216a6b845efb48ec7dd2bf049ce/carl-william_palmqvist_komplett.pdf

Palmqvist, C.W., Olsson, N.O.E., & Hiselius, L. (2017a). Delays for Passenger Trains on a Regional Railway Line in Southern Sweden. Journal on Transport Development and Integration 1 (3): 421-431. DOI: 10.2495/TDI-VI-N3-421-431

Palmqvist, C.W., Olsson, N.O.E., & Hiselius, L. (2017b). Some Influencing Factors for Passenger Train Punctuality in Sweden. International Journal of Prognostics and Health Management 8 (Special Issue on Railways & Mass Transportation. Retrieved from: https://lup.lub.lu.se/record/0f5b6553-212d-47b6-bedb-b8e11add8246

Regionfakta (2018*). Skåne County: Facts and Perspectives*. Region Fatka. Retrieved from: https://www.regionfakta.com/skane-lan/in-english/geography-/towns/ [11.02.21]

Region Skåne och Helsingborg Stad (2017). *Skåne: Facts and Key Trends*. Retrieved from: https://utveckling.skane.se/siteassets/publikationer_dokument/skane_facts-and-key-trends.pdf [11.02.21]

Skånetrafiken (2020). *Fler resmöjligheter med tåg. Skånetrafiken*. Retrieved from: https://www.skanetrafiken.se/globalassets/kartor/ovriga-kartor-2020/december-2020/skanekarta_tag_dec2020.pdf [11.02.21]

Stéphan, M., & Blayac, T. (2021). Are Retrospective Rail Punctuality Indicators Useful? Evidence From Users Perceptions. Transportation Research Part A 146(2021): 193-213. Retrieved from: https://doi.org/10.1016/j.tra.2021.01.013

Sveriges Meteorologiska och Hydrologiska Institut (2021). Lufttemperatur timvärde [Data File]. Retrieved from: https://www.smhi.se/data/meteorologi/ladda-ner-meteorologiska-observationer/#param=airtemperatureInstant,stations=all [Downloaded 05.02.2021]

Sveriges Meteorologiska och Hydrologiska Institut (2021). Nederbörd, summa 1 dygn [Data File]. Retrieved from: https://www.smhi.se/data/meteorologi/ladda-ner-meteorologiska-observationer/#param=precipitation24HourSum,stations=all [Downloaded 05.02.2021]

SMHI (2020). *Vad är normalperioder?* SMHI. Retrieved from: https://www.smhi.se/kunskapsbanken/klimat/normaler/vad-ar-normalperioder-1.4087 [03.02.21]

Sveriges meterologiska och hyrologiska institut (2015). *Framtidsklimat i Skånes län. Klimatologi Nr 29.* Retrieved from: http://smhi.diva-portal.org/smash/get/diva2:948120/FULLTEXT01.pdfzX [11.03.21]

Trafikverket (2017). *Network Statement 2017.* Trafikverket. Retrieved from: https://www.trafikverket.se/contentassets/b6f27615be234f1fababa0b1f25196dd/network_statement_2017_edition_20151210.pdf

United Nations (n.d.) Sustainable Development Goals [Online]. Retrieved from: https://sustainabledevelopment.un.org/sdgs [08.04.21]

United Nations (2018). World Urbanization Prospects: The 2018 Revision. Retrieved from: https://www.un.org/en/events/citiesday/assets/pdf/the_worlds_cities_in_2018_data_booklet.pdf [15.04.21]

Xia, Y., Van Ommeren, J.N., Rietveld, P., & Verhagen, W. (2013). Railway Infrastructure and Train Operator Performance: The Role of Weather. *Transportation Research Part D 18*(2013): 97-102. Retrieved from: http://dx.doi.org/10.1016/j.trd.2012.09.008

Yin, R. K. (2009). Case study research: Design and methods (4th Ed.). Thousand Oaks, CA: Sage.

Zakeri, G., & Olsson, N.O.E. (2018). Investigating the Effect of Weather on Punctuality of Norwegian Railways: A Case Study of the Nordland Line. *Journal of Modern Transportation 26* (4): 225-267. Retrieved from: https://doi.org/10.1007/s40534-018-0169-7.

Zakeri, G., & Olsson, N.O.E. (2017). Investigation of Punctuality of Local Trains - The Case of Oslo. *Transportation of Research Proceedia 27* (2017): 373-379. Doi: 10.1016/j.trpro 2017.12.080

Økland, A., & Olsson, N.O.E. (2020). Punctuality Development and Delay Explanation Factors on Norwegian Railways in the Period 2005-2014. *Public Transport 13*: 127-161. Retrieved from: https://doi.org/10.1007/s12469-020-00236-y