UiT The Arctic University of Norway

Faculty of Engineering Science and Technology

Department of Computer Science and Computational Engineering

Engineering methods for enhancing railway geometry and winter road assessment: A safety and maintenance perspective

Tanita Fossli Brustad

A dissertation for the degree of Philosophiae Doctor - May 2020

Acknowledgement

There are a number of people I would like to thank, who have helped and supported me throughout this Ph.D work.

Firstly, I would like to express my gratitude to my main supervisor Rune Dalmo. His guidance and sharing of knowledge, within a number of topics, has given me a deeper insight into the theory than I ever would have obtained alone. He has spent numerous hours assessing my work, giving me valuable feedback and advice, and is always available for a discussion.

Then I would like to thank my two co-supervisors, Børre Bang and Arne Lakså. Børre is the one who lit the fire on the topics resulting in this thesis. He always shares his ideas, has an answer for every question, and has been involved in all parts of my work. Arne has been there to continue the discussions and guidance when all else fails. His deep understanding and knowledge can solve any problem.

My research partner Aleksander Pedersen was an essential asset for me to be able to finish my dissertation. Our joint work, with creative brainstorming sessions and interesting conversations, has been invaluable, to say the least. Thanks are due to my colleague Jostein Bratlie for helping to dig up necessary information and for discussing useful theory with me, even though your thoughts have been elsewhere. My other friends and colleagues at UiT The Arctic University of Norway provide a knowledgeable and entertaining work environment, where all ideas are welcome with enthusiasm.

I would like to thank my family for believing in me and always encouraging me to do my best. My parents, Tove and Kenneth, are the number one reason I am where I am, and who I am, today. Their unconditional love and support has given me confidence that I can do anything I put my mind to. My sister, Annette, my best friend and better half, has shared this journey with me more closely than anyone else. She keeps me calm and makes me laugh, even on the most stressful days, and loves me no matter what. And last, but not least, a big kiss to my pet Rufus who reminds me every day how important I am in someone else's life.

Abstract

In many areas around the world there are limited transportation possibilities when travelling between key cities. If these areas also experience demanding weather conditions or geography, getting from A to B, during difficult conditions, is usually not optimal in regards to accessibility, safety, and comfort. Under challenging conditions, two essential elements in strengthening accessibility, safety, and comfort are: information and maintenance. Northern Norway, especially above the Arctic Circle, can be categorised within the previously described type of area, with its challenging geography and winter conditions. Transportation of people and goods to, from, and within Northern Norway is often done by road or railway. This puts pressure on the limited road and railway capacity in the region, making it important to keep them maintained and open for traffic.

In this study, challenges are identified within the areas of railway geometry and winter road assessment, and these two topics form the motivation behind the work. This thesis examines the two topics, railway geometry and winter road assessment. In railway geometry, a focus is on transition curves and how they can be improved in existing railways, to increase safety and reduce maintenance. In winter road assessment, a focus is on sensor data and the possibility to create a hybrid sensor, that can help maintenance personnel do their job more efficiently, and be used as a tool to warn other road users of dangerous conditions. The aim of the research is to evaluate possible changes or additions that can be implemented in road and railway, in order to improve accessibility, safety, and comfort, while reducing maintenance needs. Methods in the thesis include evaluation and analyses of curve properties in regards to given criteria, as well as experimentation and analyses of data from various sensors with winter road assessment in mind. The main results that contribute to the fields can be summarised as: increased experimental knowledge, increased technology knowledge, new approaches, extensive testing, extensive analyses, and identification of challenges.

Contents

List of Figures

List of Tables

List of Abbreviations

3D	Three-dimensional
ABS	Anti-lock Braking System
B-function	Blending function
B-spline	Basis spline
ERB-function	Expo-Rational Blending function
GMlib	Geometric Modeling Library
GMLib	Google Maps Library
GPS	Global Positioning System
IEEE	Institute of Electrical and Electronics Engineers
LCA	Lateral Change of Acceleration
MARWIS	Mobile Advanced Road Weather Information Sensor
OBD	On-board Diagnostics
RCM	Road Condition Monitor
SPTC	Symmetrically Projected Transition Curves
USB	Universal Serial Bus
VLSI	Very Large-scale Integration
VR	Virtual Reality

Part I

Introduction

1 Background

In this chapter the history of geometry and modern sensors are derived, together with specific research history on railway geometry and sensors in winter road assessment.

1.1 Geometry

This section presents briefly the long history of geometry, and describes how railway track geometry has changed over the years.

1.1.1 A brief history of geometry

Origin and early history

The geometry branch of mathematics has an early history that has been discussed and fought over since the early 1900's [1]. The greatest disagreement among historians is centered around the origins of geometry. For a long time ancient Greece was thought to be the cradle of geometry because of their written sources on definitions, theorems, axioms, and proofs [2]. However, with other discoveries dating back further than Greek sources it is now clear that the assumptions of a Greek origin is wrong [3, 4]. Today there are multiple views on the beginning of geometry with the most studied being an Egyptian origin [5], a Mesopotamian origin, an Indian origin [3], and a common origin for several regions [4]. The reason why historians find it hard to pinpoint the exact origin is mainly because it is complicated to extract geometry from early history of mathematics, and because it is difficult to give an exact dating of many written sources (including how long it was in practical use before being documented). Some of the earliest recorded sources known to this day are the two Egyptian papyri, the Moscow Mathematical Papyrus and the Rhind Mathematical Papyrus, and the Mesopotamian clay tablet Plimpton 322 [6], all dated roughly around 2000 B.C. Other well known texts are the Indian *Śulbasūtras* (Approx. 800 B.C.) [7] and the Chinese collection of mathematical problems "Nine Chapters on The Mathematical Art" (Approx. 200 B.C. - 200 A.D.) [8]. The texts show that geometry was used in a variety of areas: the Indians used the theorem of Pythagoras for altar construction [3], the Egyptians and Babylonians had area and volume calculations [2, 9], and the Chinese used the theorem of Pythagoras to measure sun height and earthly distances [2]. Of course the previous examples are just narrow set of the many uses of geometry in these ancient civilizations. For a more complete history see [2, 4, 9]. In addition to the former cultures, with their well documented geometry history, other cultures show a versatile use of geometrical elements even though they have not named or recorded it as such. For instance, there is evidence of intentionally geometrically shaped ornaments existing as far back as 40 000 B.C. [2], and the construction of Stonehenge (Approx. 3000 B.C. - 2000 B.C.) and Woodhenge (Approx. 1800 B.C.), by the English Beaker people, are said to have been built using the Pythagorean Theorem [2].

Although geometry did not originate in Greece there is not doubt that the Greek have made the biggest impact in this field. Where other cultures included geometry in a general mathematical theory, the Greek made the study of geometry the crown jewel of their sciences. From around

600 B.C. Thales of Miletus [10] (referred to as the father of geometry) changed geometry from a trial-and-error approach to using logic deduction. Thales believed that logic and reasoning were more important than intuition and experimentation, thus he was among the first to created theorems and axioms, and write deductive proofs. After Thales initial discoveries a number of Greek geometers added more theorems, axioms, and proofs to this geometrical knowledge during the next 1000 years. Some of the most known names are; Pythagoras of Samos, Hippocrates of Chios, Euclid of Alexandria, Archimedes of Syracuse, and Apollonius of Perga. For an in-depth overview of Greek geometry, [9, 11, 12] may be of interest.

Geometry during the Middle Ages and Renaissance (500 A.D. - 1700 A.D.)

During the Middle Ages and the Renaissance geometry developed very individually around the world. In Europe the geometric development [2, 9] went into hibernation in the early and high Middle Ages. Apart from applied geometry used in construction, with little to no written sources, the main focus in geometry was to understand and think through independently already existing work from ancient authors. Many ancient geometric texts (or parts of them) were available translated from Arabic to Latin, and later translated directly from Greek to Latin. In the late Middle Ages and the beginning of the Renaissance the development of geometry began to flourish once again. It was during this period that the first printed books became available, among them the books of Euclid ("Elements") in 1482. His books on geometry became a "bestseller" and was used in mathematics education in many European universities. Renaissance geometry was mainly divided in two: theoretically oriented and practically oriented. Both classifications experienced progress in this era, with new theorems and knowledge in theory, and with a transfer of practical geometry to other areas (e.g. astronomy, geodesy, cartography, and art). The ending of the renaissance brought with it two important advancements in the field, analytic geometry and projective geometry [13]. Analytic geometry (geometry with coordinates and equations) was created by Descrates and Fermat and became important in other mathematics fields, and in physics. Projective geometry (geometry without measurements) was studied by Desargues, however, it did not flourish until the work of Poncelet in the 19th century. It is also worth mentioning that in the late 17th century Newton and Leibniz developed the beginning of, what we today know as, analysis, which had an impact on geometry, solving problems that previously were troublesome. China experienced much of the same as Europe with development in geometry (and mathematics in general) being slow [2]. Since geometry was not its own discipline but an integrated part of mathematics, geometry grew in line with mathematics. In the 13th and 14th centuries (at the end of the Song dynasty and during the Yuan dynasty) Chinese mathematics experienced its first growth, and in the 16th century (during the Ming dynasty) its second growth. In America little is know of the history of geometry and what we do know is mostly constricted to the Aztecs, the Mayans, and the Incas [2]. Since written sources are scarce the knowledge of geometry in these Mesoamerican cultures comes from archaeological research of town ruins and other relics. This research shows that their expertise in geometry was within architecture, where they used geometrical shapes in construction of buildings and temples, layouts of towns, and as ornaments on buildings, ceramics, and clothes. As a contrast to the slow development of geometry in many countries in the Middle Ages, the geometry in Islamic countries prospered [2, 9]. Islamic geometry (also referred to Arabic geometry) was largely based on Greek geometry, which is understandable seeing that all Greek classics were accessible in good Arabic translations as far back as 900 A.D. The key aspects of their investigations lay within the branches of theoretical geometry, applied geometry, and trigonometry. In theoretical geometry the work of Euclid was the core of the research, which they further developed and wrote their own books on. In applied geometry the written work is assumed to have been intended for craftsmen, providing them with instructions on how to design surfaces of curvature. Works in this area included complicated calculations and constructions of Arabic architecture, including the characteristic stalactites. Lastly, in trigonometry the works of the Indians and Greeks were further developed angling it

more towards applied methods.

Modern day geometry

Starting from the 18th century geometry developed rapidly. Many scientists contributed to various areas in the field, and ground breaking discoveries were made. During this period "known" geometry was split into several different sub-fields and new branches, of previously "unknown" geometry, emerged. Some examples are: the descriptive geometry of Monge (1798), the continuation of projective geometry with Poncelet (1822), the modern analytic geometry of Plücker (1828), and the two important developments of non-Euclidean geometry (by Gauss, Bolyai, and Lobachevsky (1829)) and Riemann geometry (by Riemann (1854)) [13]. The two last fields are noteworthy discoveries in modern geometry. Non-Euclidean geometry came from trying to develop a self-consistent geometry where one of Euclid's postulates (the parallel postulate) was false. Ever since Euclid's works became available, the parallel postulate had drawn the eyes of many people trying to deduce it from other postulates. Not until the 18 century serious attempts at proving it were made. This of course turned out to be impossible, as shown by Gauss, Bolyai, and Lobachevsky, independently, when they proposed examples of non-Euclidean geometry [2, 14]. Riemann geometry came from the study of non-Euclidean geometry, however, this was a different non-Euclidean geometry than that of Gauss, Bolyai, and Lobachevsky, where the intrinsic geometry of surfaces was generalised. With this, Riemann set forth three kinds of geometries, which Klein called the elliptic (Riemann), parabolic (non-Euclidean), and hyperbolic (Euclidean) geometry [13]. The new realisation of different kinds of geometry presented the need for new axioms in Euclidean geometry, and in 1899 Hilbert presented 21 statements, known as Hilbert's axioms, that replaced Euclide's axioms and postulates [9]. In the 20th century the field of geometry had become a complex area, with all its sub-areas, and research and activity in other parts of mathematics. During this era geometrical notions and problems were also used to develop theories in other fields (e.g. Riemann geometry was a key element in Einstein's general theory of relativity in physics). At the end of the 20th century and to this day a new development in geometry has come with the advent of the computer, where new disciplines in need of investigation have unfolded. Among these are computational geometry and digital geometry [2].

1.1.2 Railway track geometry

Railway track geometry referrers to the horizontal and vertical layout of the railway track. Although track geometry is considered three dimensional, the horizontal and vertical separation is common in research. In this thesis the focus will be on the horizontal layout, so further descriptions on railway tracks are assumed to mean the horizontal alignment unless otherwise specified. The history of track geometry can be separated into different areas. Here the history of railway segments will be presented, because this is an important part of railway track history and the most relevant for the rest of the thesis. Early railway tracks usually consisted of a connection of straight sections and circular arcs. With the low speed of trains and the wide radius of the circular sections the transition between straight and circular sections was considered easy and safe with regards to the motion of the train [15], and was well within the scope of current surveying technologies at that time. When the evolution of railway tracks made the radius of the circular tracks smaller and the speed of trains increased, transitions became jerky and unpleasant, and sometimes even dangerous [15, 16]. The previous realisation opened up the door to easement curves, or transition curves as they are more known by, that made the transition from straight track to curved track safer and more comfortable even when higher speed was involved [15, 16, 17]. Because of this there are today three main track types in the horizontal layout: tangent tracks (straight lines), circular tracks with a constant radius, and transition-curve tracks [18]. Different types of transition curves have been used in railway over the years. Two such curves are addressed by professor W. J. Macquorn Rankine in [19]: The "harmonic curve"

or "curve of sines" by William Gravatt from 1829, and the "curve of adjustment" by William Froude from 1842. In the late 19th century a railway transition spiral with linear curvature, in relation to arclength, was derived by multiple engineers independently, and without knowing that Leonhard Euler had already derived the curve many years prior. Because of this the curve has many names i.e. Euler spiral, clothoid, Cornu spiral, and Glover's spiral, but clothoid is the most common. The first person to draw an equivalence between the railway transition spiral and Euler's definition was A. L. Higgins in 1922 [20]. With focus on the curvature, transition curves can be divided into two groups: transition curves with linear curvature and transition curves with non-linear curvature [21]. From its invention and to this day the clothoid, from the first group, is the most used transition curve in railway. This is because its linear curvature was considered an optimal transition for a long time. In addition to the clothoid, many non-linear curvature curves were suggested as transition curves early on. Various examples are given in [22]; the cosine curve (1868), the Helmert curve (1872), the Ruch curve (1903), the Watorek curve (1907), the Bloss curve (1936), and the sinusoidal curve (1937). Some of these curves were implemented in certain railway tracks, but they never obtained the popularity of the clothoid. The introduction of lateral change of acceleration (LCA) in 1996 [23] showed that with the increasing speeds in rail transport, the clothoid was not as optimal any more (more specifically for speeds above $120\frac{km}{h}$ [24]). This realisation opened the door to research on new types of transition curves including parabolas [25], sinusoids [26], polynomial curves [27], and spline curves [28], that had a smoother LCA than the clothoid. Today's research on railway transition curves can be divided into two areas (regarding curve analysis): simple evaluation methods and advanced evaluation methods. Research using simple evaluation methods involves analysing curves with linear, non-linear, and mixed curvature. Examples of linear curvature research include Symmetrically Projected Transition Curves (SPTC) [29]. Non-linear research include log-aesthetic curves [24], a remodeled cubic parabola [30], sinusoids [26], a new design of the Bloss curve [21], and the Wiener Bogen (Viennese curve) [31]. Mixed curvature research include curves with linear curvature in the middle region and smoothed start/end [32], and curves with linear start/middle region and smoothed end [33, 34]. Research using advanced evaluation methods focus on forming polynomial curves with a combined use of optimization methods and advanced vehicle-track models [27, 35, 36], and investigating the non-linear phenomena that occur in the motion of railway vehicles in transition curves with speeds close to the vehicle's critical velocity [37]. As a last comment it is worth mentioning recent work in relation to highway transition curves, since railway and highway transition curves are closely connected and often referred to in the same setting. The latest research on highway transition curves focuses on curves that replace multiple segments of the original highway. Examples here are two new families of the general transition curves [38] (their generalisation [39], applications [40], and possible other representations [41]), flexible polynomial transition curves [42], and a general formulation of an optimization problem to create the horizontal alignment over a longer stretch [43].

1.2 Sensor technology

This section presents the history of the modern sensor, and describes how it has been an interest in winter road assessment since the 1930s.

1.2.1 History of the modern sensor

When we talk about sensors or transducers throughout this thesis we use the following definition "a device which provides a usable electrical quantity (output) in response to a specific physical quantity, property, or condition which is measured (measurand)" [44]. This may also be referred to as a modern sensor. One of the first modern sensors was built by Wilhelm Von Siemens in 1860. He created a thermometer based on a copper resistor, to further understand how the

temperature sensitivity of the resistor affected its electrical resistance. In 1871 Siemens described the resistance thermometer before the Royal Society in London with the proposition of using it in metallurgical furnaces [45]. The next development was done by Warren S. Johnson over the course of three years (1880-1883). He invented a device that controlled and regulated room temperature, the first thermostat, and got it patented in 1883 as the "electric tele-thermoscope" [46]. Since then sensors have been developed in multiple areas. The 20th century gave rise to an extensive variety of sensors with different use cases; e.g. infrared sensors came about in the 1940s, the first motion sensor for an alarm system was developed in the 1950s by Samuel Bogno [47], the biosensor was first described in 1962 [48], and Naoyoshi Taguchi put in a tremendous amount of work to create the first gas sensor (1968) [49]. The term smart sensor became known in the 1980s [50] and is an important advancement in sensor technology. Credit of the term is frequently given to Joseph Giachino from Ford Motor Company based on his paper, "Smart sensors", from 1986 [51]. However, Wen Ko of Case Western Reserve University was the first to introduce intelligent sensors, some years earlier, in 1982 with the paper "VLSI and intelligent transducers" [52]. The definition of a smart sensor, as given by the IEEE Standard 1451.0-2007 [53], is "A transducer that provides functions beyond those necessary for generating a correct representation of a sensed or controlled quantity. This functionality typically simplifies the integration of the transducer into applications in a networked environment". Early sensors were simple devices and with time complexity grew. In the last two decades (the beginning of the 21th century) the capabilities of sensors have expanded due to the development in computing, connectivity to the web, mobile smart devices, cloud integration, and so on, which have given new interesting opportunities to the sensor technology field, and increased the quantity of sensors exponentially [54] (see Figure 1.1).

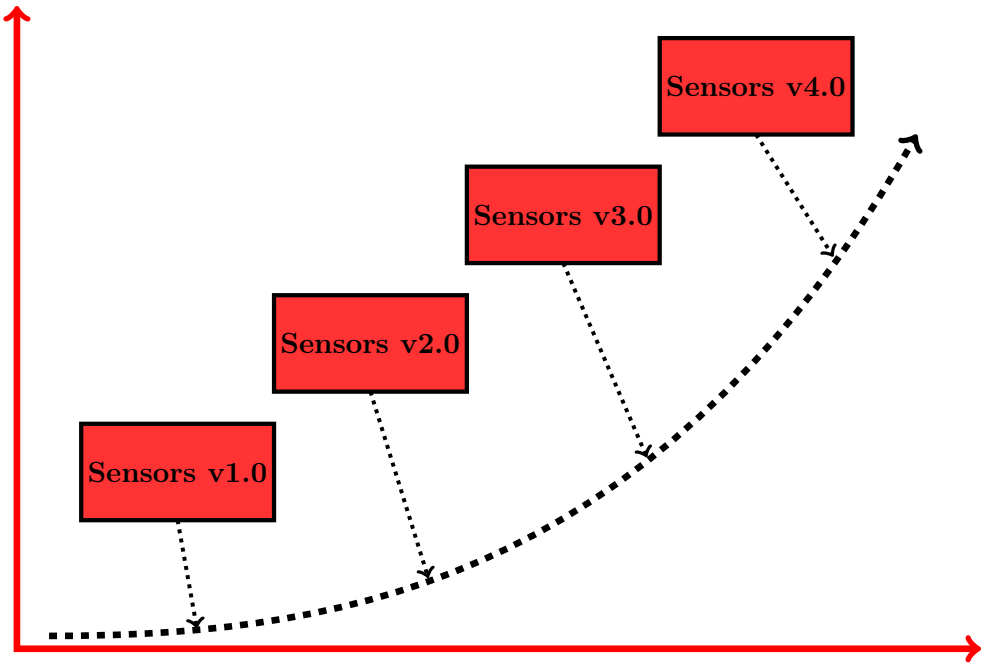


Figure 1.1: Evolution of sensors.

1.2.2 Sensors in winter road assessment

The increase in number and speed of vehicles on the road have given rise to research with focus on safety and maintenance. An important topic her is assessment of winter roads. In many countries the winter months can be unpredictable, in regards to weather conditions, which again changes the road conditions (sometimes very quickly). The interest for sensors in winter road assessment goes back to the 1930s when Torben Bronée patented an apparatus for indicating ice formation

on roads at a distance [55]. Other early inventions include Louis Clark's Highway condition indicating system (1954) [56], and Marcel Boschung's early ice-warning device (1977) [57]. A common factor for the previous inventions is that they were all stationary devices measuring from the side of the road. One of the first on-vehicle winter road sensors was proposed by Peter Decker in 1979 [58]. This sensor measured the possibility of glazed ice formation based on road temperature and relative humidity, and notified the driver with a warning device. In the 1990s the interest in winter road assessment grew and a number of papers were written. Ice detection measurements continued to be a topic of research [59, 60], and a new direction of friction measurement arose [61, 62, 63], together with the possibility for state predictions (e.g. icy, dry, wet, snowy) [64, 65, 66]. In the 20 years that have gone by of the 21st century the advancement in technology has produced new and interesting developments in winter road assessment. A variety of sensor and methods have been used for measuring specific variables [67, 68, 69, 70, 71], and some sensors have been developed specifically for winter road sensing (two well known examples are RCM411 [72] and MARWIS [73]).

2 Intention

In this chapter the motivation of the thesis is presented, challenges in railway geometry and winter road assessment are discussed, and objectives and research questions are described.

2.1 Motivation

Northern Norway is an area with high mountains, long fjords, and shifting weather conditions. In addition, there is a limited number of main roads and railroads when traveling north of the Arctic Circle. Every day a considerable amount of transported goods are moved using these roads and railways, with Narvik as a key city. The geographical location, as well as the access to rail, road, sea, and air, makes Narvik a natural logistics intersection for moving goods from the region and within the region [74]. It is estimated that 90 % of the groceries north of the Arctic Circle arrive with the Ofoten railway line (Ofotbanen) before being transported by road north/south (E6), and east/west (E10) in the region, with an increasing number of fresh fish returning (the other way) [75]. This puts pressure on the limited road and railway capacity making it important to keep them well maintained and open for traffic to ensure accessibility and safety.

During the winter months snowfall and ice formation make road conditions very unstable, where the conditions can vary from day to day and deteriorate considerably on short notice. When the friction on the road is reduced the risk of accidents increases [76, 77, 78, 79]. With normal road users and transport trucks traveling side by side on narrow roads the consequences of a potential accident are high, and every year accidents lead to fatalities and blocked roads [80]. There is even research arguing that the common type of heavy goods vehicles traveling in the Nordic countries often are a safety risk on icy conditions [81]. Travelling in Northern Norway is mostly done by car or bus, since other transportation possibilities are few. Apart from Ofotbanen, where you can only travel to Narvik, there are no railway lines north of Fauske, which means that almost $\frac{1}{3}$ of the country is without train connections (see Figure 2.1). When it comes to travelling by airplane it is usually expensive and requires multiple stops to reach the final destination (sometimes it is even necessary going south to get north). With these restrictions in transportation it is important that travelling by road is possible and safe for all road users even under harsh conditions.

Ofotbanen is a railway line running from the port of Narvik to the Norway-Sweden border. The line is single tracked and was built from 1898 to 1902. Today Ofotbanen is an essential railway line in Norway, transporting ore, goods, and passengers, with the main traffic being iron ore haul with about 10-12 trains a day [83]. The mountains and fjords surrounding Narvik give the railway a rough terrain to venture through, alongside a risk of avalanches and rockslides. From Narvik port the railway has a constant large climb up the mountain sides all the way to the Swedish border, with many sharp turns along the way. Ofotbanen is known for its challenging railway geometry; mainly connected to the small radius curves and the rail-section setup, which causes unwanted wear in certain vulnerable areas of the rail. The railway is an important part

Figure 2.1: A map of the Norwegian railway [82].

of the Narvik industry, and since it is single tracked minimizing the need for maintenance, while increasing safety, will be an advantage.

The motivation behind this thesis is to investigate the possibilities of improving railway geometry and winter road assessment to increase safety and reduce the need for maintenance, and show the effects of these improvements with analysis.

2.2 Challenges in railway geometry and winter road assessment

The challenges in railway geometry are today connected to increased train speeds, while also trying to increasing safety and comfort, and reducing maintenance needs. A popular topic in railway geometry has evolved around transition curves. A transition curve is a curve that connects two railway segments of different curvature together to create a smoother and safer ride. An example of a straight segment connected to a circular segment with a transition curve is given in Figure 2.2. As mentioned in section 1.1.2 the introduction of LCA [23] as an important criterion in

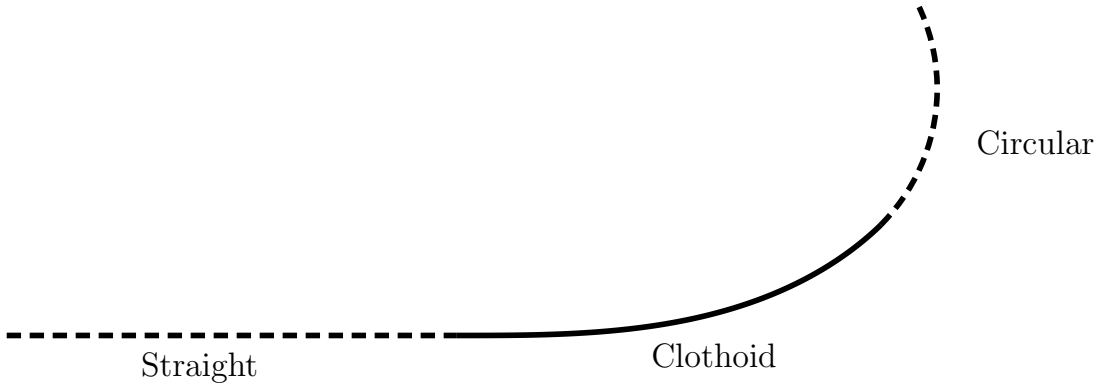


Figure 2.2: A straight curve and a circular curve connected by a transition curve.

vehicle-road dynamics showed that the transition curves used when constructing railways today, clothoid curves, are not optimal. The reason for this is the jump in LCA that occurs in the beginning and at the end of a transition curve, causing a sudden change in lateral force. The LCA function is closely connected to the curvature (or rather the change in curvature) along the railway, meaning that if the change in curvature is smooth then the LCA is smooth. Today's clothoid transition curves have a linear change in curvature between two constant curvature segments, which gives a geometric continuity [84] of G^0 for the curvature in the joints, causing the jump in LCA. To obtain a smoother LCA (remove the jumps) the degree of the geometric continuity in the joints has to be higher than 0. An important observation here is that by raising the continuity degree, the extreme values of the LCA will also increase. This means that there has to be a balance between the smoothness and the extreme values, so that the transition is as smooth as possible while keeping the LCA within a certain threshold.

The overall challenges in winter road assessment is mainly connected to three questions:

1. What do we need to measure?
2. How do we measure it?
3. How do we use the obtained data in the best way?

In order to increase safety and reduce maintenance costs it is important to be able to accurately measure areas with dangerous road conditions, and alert the driver, other road users, and maintenance personnel of the danger. Today, most vehicles have embedded sensors that tells the driver

of possible difficult conditions, either by sound and display, or by activating a countermeasure (for instance, ABS, ice warning, or anti skid). The possibility to alert others and to pass valuable information to maintenance personnel is not yet implemented in full. In order to maintain roads in vulnerable areas, instead of the whole network of roads, information has to be gathered, possibly stored, and passed on. The first issue that comes up is what information that should be gathered. From research, an assortment of variables (e.g. road temperature, water thickness, and surface status) and methods (e.g. image recognition, and video analysis) have been suggested for road assessment, as well as indirect variables (e.g. acceleration, and GPS). Choosing the variable/method should be a trade off between the trustworthiness of the assessment, the cost of the implementation, and the ethical aspect of collecting the data from road users. The second issue is deciding what to do with the collected data, how/if it should be stored, and to what extent it should be accessible to others. One example, that is already implemented, is storing data on a server and making the results (not raw data) available to everyone online, whereas raw data is only available to the person owning the sensor. If the decision to store data is made, questions regarding storage time and storage space will arise. The amount of data collected from sensors can be large, especially if multiple vehicles are measuring. This means that a possible compromise between the number of measuring vehicles and how long data is stored, may have to be considered. When it comes to sharing data, ethical and legal aspects are important to reflect on, as well as the questions on who the data should be available for and what should be available.

A more specific challenge connected to winter road assessment is the possibility to perform self help in situations where conditions are critical. A problem for heavy haulers driving on winter conditions is that sometimes they struggle in areas that are not categorized as difficult to traverse. Mainly when going uphill or downhill, because they are pulling heavy trailers. In Northern Norway this usually means blocked roads over longer periods of time with detours adding, in the worst case, hours to the travel time. When Heavy haulers have already gotten in a position where they struggle the possibility of getting out of the situation on their own is slim, and help can take a long time to arrive, because of the distances between key cities in the region. A way of preventing many of these mishaps is to give the drivers means to help themselves before the situation gets out of hands. One example that is available today is sand spreaders from Autoline AS [85], custom made for the truck industry, mounted in front of the drive wheel. The sand spreaders are operated from inside the vehicle, and increase the safety for all road users by spreading sand in the wheel tracks if a critical situation occurs.

2.3 Objectives and research questions

The main objective of the thesis is to investigate if geometry, sensors, and data analysis can be used as tools to increase safety and reduce maintenance on roads and railways under unideal conditions. In this context unideal conditions can be described as slippery winter roads, railways with restrictions on placement of the track, and railways with high speed trains. The research has been divided into three parts. Part one is a preliminary study connected to geometry and sensors with focus on increasing the understanding of the fields. This provides the building blocks for the other two parts; railway geometry and winter road assessment.

The three specific objectives for railway geometry are to perform a state-of-the-art study on railway transition curves, to investigate the suitability of blending splines as transition curves, and compare curvature transition between a segmented railway track and a B-spline curve. The state-of-the-art study is conducted in order to increase the author's knowledge on the subject, and to identify current research challenges and future research opportunities in an important field. The investigation of blending spline transition curves is performed to reveal if the flexibility of the blending spline can be used to obtain a smoother transition curve compared to the

clothoid, in the end points. In the comparison between an already existing railway track and a B-spline curve, the objective is to examine if the B-spline curve gives a better transition of curvature between segments than the original railway track.

The two specific objectives in winter road assessment are to investigate various sensors in a field study setting, and perform experiments to test the effects the Autoline sand spreaders have on the vehicle and the road. The field study experiments are performed in order to increase the knowledge of commercial and experimental sensors in relation to reliability, usability, benefits in winter road assessment, and how they can complement each other, as a step towards possible hybrid sensor technology. The Autoline sand spreader experiments are performed in order to map the performance of the sander as a tool to help heavy haulers that experience problems under slippery road conditions.

The thesis aims at answering the following research questions linked to each part:

1. Railway geometry
    (a) What challenges can be observed in railway transition curve research?
    (b) What are the impacts of a transition curve on a vehicle, in connection to lateral forces and path traced?
    (c) Can blending splines be used as segments in existing railways, replacing clothoids, to increase the smoothness in the connection points to the adjacent segments?
    (d) By replacing multiple segments of an existing railway curve with a B-spline curve, can curvature transitions be improved without making extensive changes to the alignment?
2. Winter road assessment
    (a) Can various sensors benefit each other in a way that defends creating a hybrid sensor solution to improve winter road assessment?
    (b) What information can we get from available winter road sensors and how do some of them compare to each other and to laboratory experiments?
    (c) Can targeted sanding increase surface friction enough to impact heavy hauler pull force?

Part II

Methods

3 Methodology

The research methodology in the thesis is based on a known approach in the development of science. When modern science began its rise in the 17th century, the standard for generating new knowledge and updating existing knowledge became what is today known as the scientific method [86, 87, 88]. The scientific method is an empirical cyclic process that involves making hypotheses based on observations, conduct experiments to test the hypotheses, and analyse the results to decide further actions. A common comparison is to compare the scientific method with common sense in every day life, with the difference that the steps in the scientific method are well expressed and controlled. The cyclic method can be divided into six steps [89], showed in Figure 3.1, (in literature the number and names of steps can differ, but here we use the following):

1. State problem
2. Formulate hypothesis
3. Design experiment
4. Collect data
5. Analyse data
6. Draw conclusion

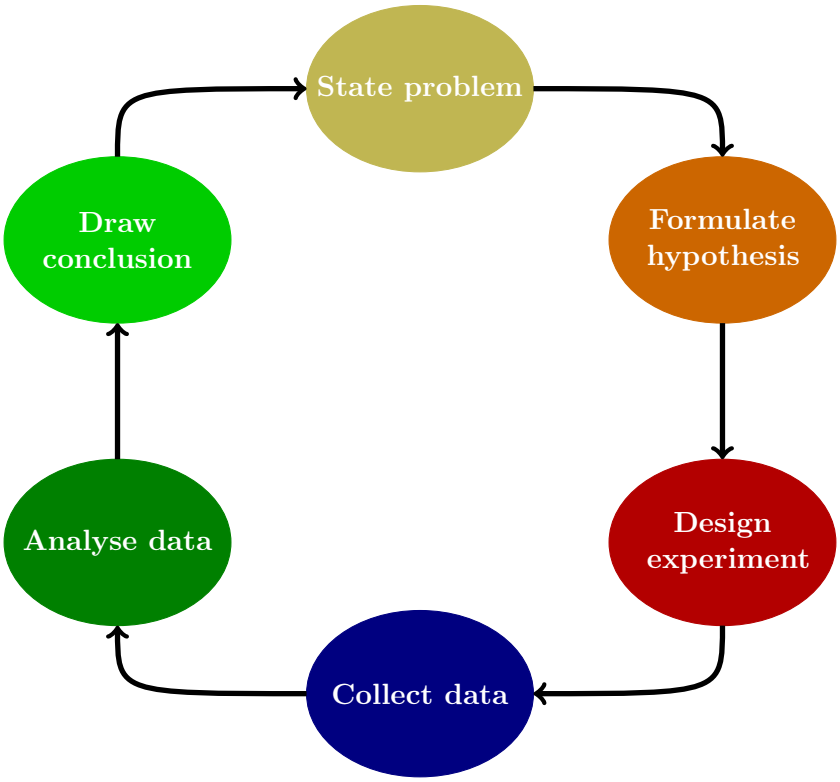
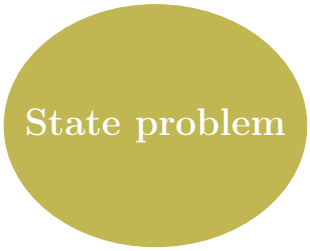


Figure 3.1: The cyclic steps in the scientific method.

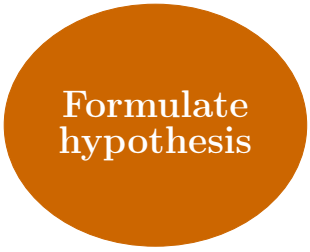
1. State problem

The first step includes questioning an occurrence that we want to find an explanation to. Often, this step involves investigating and evaluating personal observation and research, or research of other scientists. The process of forming a research question is an important part that may affect the further process and the obtained results.



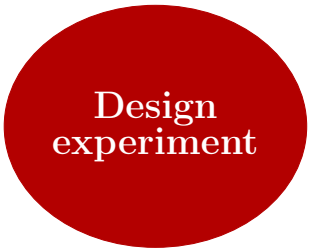
2. Formulate hypothesis

Based on the previous research question a hypothesis is developed in order to try and explain the observation. A hypothesis can be specific or broad, but should always be testable. This means that the results from an experiment testing the hypothesis should have the capability of differing with the predicted positive outcome (account for the possibility that the hypothesis can be wrong).



3. Design experiment

The design of experiments depends on the aim of the study. The experiments should be possible to carry out under controlled conditions, and should minimize possible errors. When choosing the design an important property is to obtain results that are reliable, valid, and replicable. Experiments are usually divided into two major areas, quantitative methods and qualitative methods, depending on the data that is to be collected. Quantitative methods are methods that gather numerical data which can be used directly to construct graphs or tables. Qualitative methods are methods that gather non-numerical data (i.e. features, definitions, characteristics, and symbols) which usually need processing to make sense of and sort into patterns.



4. Collect data

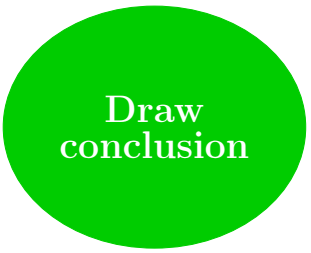
The collect data step involves executing the experiments, decided on in the previous step, to collect enough and valid data for further analysis. The step usually involves altering independent variables (can only be changed by the researcher) to generate data for dependent variables (can be changed by other variables). There may also be variables that have to be kept constant to prevent external factors from influencing the result.

5. Analyse data

After collecting data, the data has to be analysed. In this step the results are investigated and evaluated using, for instance, graphs, plots, or tables, and a conclusion is made in relation to the reliability and outcome of the results. When trying to reach a conclusion regarding the results it is necessary to have in mind that a true prediction cannot prove a hypothesis. Just like a false prediction can not disprove a hypothesis.

6. Draw conclusion

In the last step of the method a conclusion is made regarding the results and the next action to be taken. If the hypothesis is contradicted it can be necessary to start the cycle again from step 2 and formulate a new hypothesis. If the evidences of a positive result are weak the cycle can be repeated from step 3. And if the evidences strongly support the hypothesis the research can be presented and published. Usually the cycle is repeated several times for a problem statement depending on its complexity.



Limitations

The scientific method is evidence and logic based, which by themselves can not conclude if a hypothesis is absolutely true or absolutely false. The method only gives a probability of something being true or false, where the more evidence pointing in favour of a prediction makes it more probable of being correct. This is an important limitation to be aware of, but does not mean we should discard the approach. The objective of the scientific method is not to establish a theory with absolute certainty, but rather take into account that science can be inconsistent and these inconsistencies have to be studied because they can lead to new discoveries.

4 Methods in railway geometry

The methods used in steps 3, 4, and 5, of the scientific method, for the railway geometry part consists of computer programming, and analysing graphs of curve properties. Next follows a description of each individual part that constitutes the process of going from a chosen hypothesis to a conclusion (designing the experiment, collecting data, and analysing data). The process is divided into

- Experimental setup
- Technology selections
- Variable selections, considerations, and simplifications
- Analyses and comparisons

4.1 Experimental setup

The railway geometry part of the thesis, including the preliminary study paper on geometry and virtual reality, consisted of three different experimental setups. Two of them were strictly computer experiments, and the last was a mix between physical laboratory experiments and computer experiments. In the first computer experiment, the blending spline experiment, blending splines were implemented and analysed as transition curves in an already existing railway. In the second computer experiment, the B-spline experiment, curves were created and their curvature properties analysed according to a wanted threshold. In the mixed experiment the traced curve of a physical vehicle model was compared to the same traced curve of a virtual vehicle model, both to analyse the similarities between them but also to investigate the differences between various transitions when going from a straight section to a curved section.

The experimental setup of the blending spline experiment used three segments from Ofotbanen as a starting point, two circular arc segments connected with a clothoid. The clothoid was then replaced with various blending splines [90, 91, 92], to analyse the suitability of the blending spline as a transition curve compared to the clothoid. Multiple parameters were adjusted in the blending spline, including local curves and blending functions, in order to find a transition curve that was smoother than the clothoid in the end points. A set of smoothness criteria was present and used as a tool in the comparisons.

The experimental setup of the B-spline experiment included implementation of a composite curve with data gathered from Ofotbanen. The data consisted of three types of sections: straight sections, transition sections (clothoids), and arc/circular sections, making up parts of the railway track along Ofotbanen. From this track data, selected points, along the straight and circular sections, were evaluated and used as interpolation points in a cubic B-spline approximation of the original composite curve.

The mixed experiment was setup with a physical model, a Lego Mindstorms set (Figure 4.1),

and a virtual model, in Blender (Figure 4.2). The two models were run through three different tracks: a track with no transition between straight and curved sections, a track with a clothoid transition, and a track with an S-shaped curvature. In order to create the tracks with the different transitions, the front wheels in both models were turned a given angle every half second according to Table 4.1.

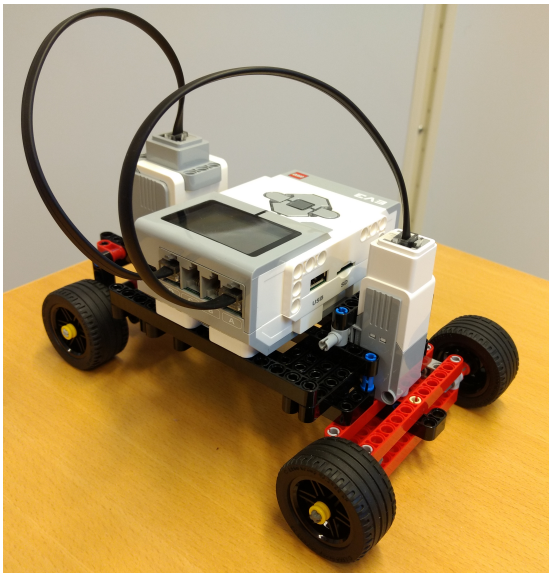


Figure 4.1: The physical model built with a Lego Mindstorms set.

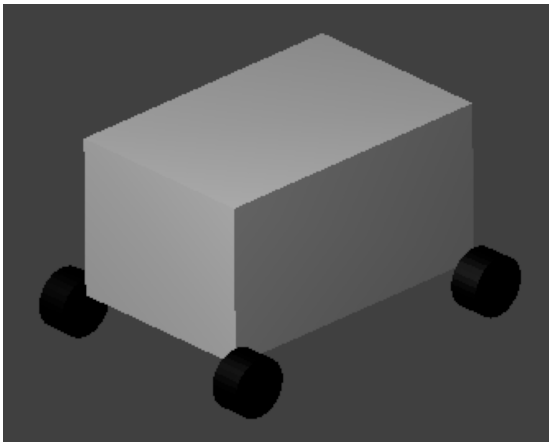


Figure 4.2: The virtual model created in Blender with Bullet physics.

Table 4.1: Change in turn-angle to create different transitions. Discretized into 11 steps over 5.5 seconds.

Turn-angle in degrees for the transition tracks		
No transition	Clothoid	S-shaped
0	0	0
0	3	0.6
0	6	1.8
0	9	4.2
0	12	8.4
0	15	15
30	18	21.6
30	21	25.8
30	24	28.2
30	27	29.4
30	30	30

4.2 Technology selections

The selection of technology for the computer programming was based on, availability, previous experience, knowledge of the problem at hand, reliability, and the ability of the technology. In both the blending spline experiment, and the B-spline experiment, Qt was used together with an in-house geometric modelling library (GMlib) (not to be confused with Google Maps Library (GMLib)).

In the mixed experiment, the technology included a Lego Mindstorms EV3 set with accompanying sensors, and the 3D computer graphics toolset, Blender. The specific sensor used with the Lego model was a gyroscope, and programming was performed in Lego Mindstorms' own block-based programming software. The virtual model was created in Blender Game Engine with Bullet as the physics engine, where the setup of the vehicle and the various transition experiments were programmed in Python scripts, run in the game engine.

4.3 Variable selections, considerations, and simplifications

The blending spline experiment had three main variable selection topics, entry points, local curves, and ERB-function parameters. Since existing railway segments were used in the experiments, it was decided that the position of the two circular segments would remain fixed, yet the entry points of the blending spline transition curve onto the circular segments could vary within given boundaries. Two curves were explored as local curves in the blending spline, arcs and Bézier curves. Arcs were chosen because of their natural implementation possibilities by extending the already existing circular segments, and Bézier curves were chosen for the flexibility in altering their shapes, while at the same time being intuitive to constrain in the ends. The shape of the ERB-function (blending function) was varied through the tightening parameters, α and γ, the slope parameter, β, and the balance parameter, λ. In the tests, only one parameter was modified at a time, with the other three set to default values. Seven unique ERB-functions were tested as blending functions in the experiments. Considerations and simplifications that were made in the testing consisted of manually choosing entry points and ERB-function parameters, and manually modifying the control points in the local Bézier curves opposed to creating an algorithm that could find a best approximation.

The B-spline experiment used curvature in the comparisons and assumed a constant railway vehicle speed. A focus was given to cubic B-spline curves because of their minimization property of the second derivative. Tests were performed with: varying number of data samples (between $500-2500$), uniform and non-uniform samples and knot vectors, and different numbers of control points (between $30-274$); in order to find a flexible enough composition. The samples were placed on the straight and arc segments of the existing railway, leaving the clothoid segments free to be approximated by the system in hopes of achieving smoother transitions.

The selected variables in the mixed experiment consisted of rotation rate and traced path. Rotation rate was chosen because of its close connection to curvature and LCA of a curve, while the traced path was chosen mainly for comparison purposes between the physical and virtual model. Both models were run at constant speed during measurements, where the Lego model's max speed was used as a template. Considerations had to be taken when creating the virtual model because of Blender's weakness in working with small collision shapes [93]. To compensate for this, the virtual model was scaled up by a factor of 10 compared to the physical model. One noteworthy simplification was made in the experiments, the use of a Lego model. However, for the purpose of the study a Lego vehicle was considered adequate since the main property needed in the vehicle was an ability to uphold a constant speed, which the Lego vehicle was capable of.

4.4 Analyses and comparisons

In the blending spline experiment a selection of blending splines (with varying entry points, local curves, and ERB-functions) were implemented and analysed as transition curves between two circular arc segments of different radii. The analysis of the new transition curves, and the comparisons against the clothoid, was based on the five smoothness criteria given in [24], in addition to the overall shape of the curvature functions. Comparison of the blending splines with varied

parameters were also conducted where the focus was on the shape of the curvature functions, and how they were altered.

In the B-spline experiment a collection of B-spline curves were created, with a varying number of data points, uniform and non-uniform sampling, different numbers of spline coefficients, and uniform and non-uniform knot vectors. Then, for all curves the curvature was plotted and compared to the curvature of the segmented curve. In addition, an error measurement, mean squared error and maximum distance, was conducted on the segmented curve against each of the B-spline curves.

The results for the mixed experiment were divided into two parts: geometry analyses and VR comparisons. The geometry analyses included testing the transition geometry for the physical model by using a gyroscopic sensor that measures the rotation rate. The rotation rate was then plotted for the three test cases and compared against each other to observe the differences in properties they possessed. In the VR comparisons a comparison of the physical model vs the virtual model to look at the accuracy between them was conducted. This was performed by considering the rotation rate and by tracing the path while driving, for both models, and then visually comparing the plotted graphs between the models for each test case.

5 Methods in winter road assessment

The methods used in steps 3, 4, and 5, of the scientific method, for the winter road assessment part consists of field experiments, and analysing graphs of sensor variables. Next follows a description of each individual part that constitutes the process of going from a chosen hypothesis to a conclusion (designing the experiment, collecting data, and analysing data). The methods described in this section does not include the preliminary study experiments since they are similar to the main experiments in the winter road assessment research (performed on a smaller scale). The process is divided into the following four sections:

- Field experiment setup
- Sensor selections
- Data collection and storage
- Variable selections and comparisons

5.1 Field experiment setup

For the winter road assessment part of the thesis two different field experiments were performed. In the first field study various sensors were tested in the field, compared to laboratory experiments, and analysed against each other. This study included testing and analysis of two known winter road sensors (RCM411 and MARWIS), an experimental radar sensor (Walabot), and also a discussion on the use of vehicle data (collected with an OBD-II adapter) as a supplementary tool in winter road assessment. Data was collected, from all sensors, over 5 days of driving from Narvik to Vadsø and back, via Sweden and Finland (approximately 1729 km). Weather conditions ranged from sunny to heavy snow, with temperatures between -9 and +5 degrees Celsius, and road conditions switching between icy, snowy, wet, moist, slushy and dry asphalt. The sensor setup on the vehicle, seen in Figure 5.1, was done in the following way: the RCM411 sensor was mounted on the tow ball pointing at the right wheel track, the MARWIS sensor was secured with a rack to the truck bed pointing towards the center of the road behind the vehicle, and the Walabot sensor was attached below the back left passenger door pointing straight down at the wheel track.

In the second field study, experiments were performed in collaboration with Autoline AS to test the effects of their sand spreaders on heavy hauler pull force. This study included experiments and analysis on pull force between a truck and a trailer, measured with a weight, and a discussion on road friction before and after sanding. The experiments were carried out at Nord-Norsk Trafikksenter on Finnsnes in Troms, Norway, in -14 degrees Celsius on a clear winter day. To measure the pull force a weight was attached between a truck and a trailer, see Figure 5.2. Then experiments were performed by locking the wheels on the trailer and moving the truck forward with increasing torque and accelerator deployment until the truck wheels lost grip, while collecting pull force data from the weight.

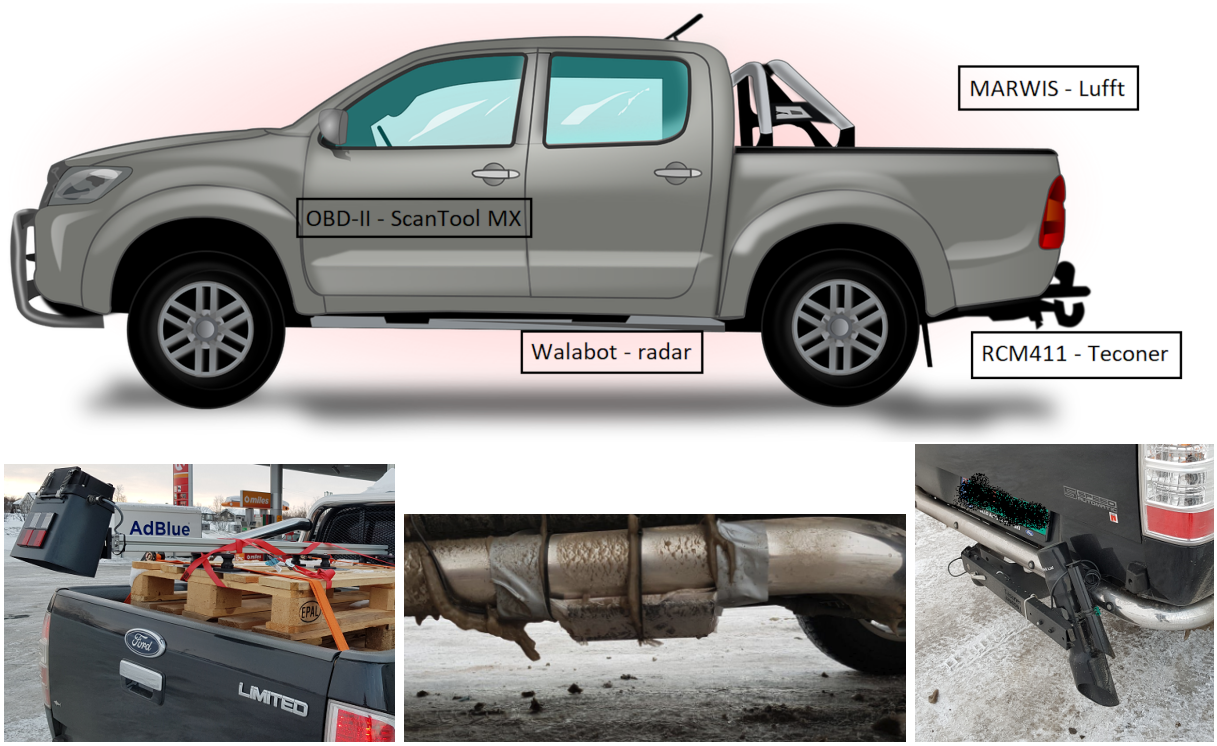


Figure 5.1: Placement of the sensors on the vehicle. Top image by OpenClipart-Vectors from Pixabay.

5.2 Sensor selections

The sensors were selected based on the objectives of the research that was to be performed. In the field study of the various sensors RCM411 and MARWIS were chosen because they are well known in winter road research, and a number of laboratory tests have been performed on them. The reason for choosing the Walabot sensor was its small size and low cost, as well as it being a radar. This was interesting since radar technology has not been tested extensively for winter road assessment, which in theory it should be suited for given that radar is not dependent on sight, being less affected by snowy or foggy conditions where, for instance, optical sensors have a problem. The main reason for selecting more than one sensor in this research, and the reason for also including collection of data from the vehicle (with an OBD-II), is the possibility of using hybrid measurements in the future with various data that compliment each other in a beneficial way.

In the sand spreader experiments the weight used to measure the pull force was a Dynafor LLX2 digital load indicator 10 tons + 13.5 tons shackles [94]. It was chosen for the following four reasons:

1. Its high working load limit, 10 tons, which we calculated was necessary for the planned experiments.
2. Its operating temperature, ranging between: -20 to +40 degrees Celsius.
3. Its high accuracy of 0.1% of the maximum measuring range.
4. The possibility to combine it with monitoring software for simultaneously reading, managing and storing data.

We also used an RCM411 sensor to measure the road friction before and after sanding. This was

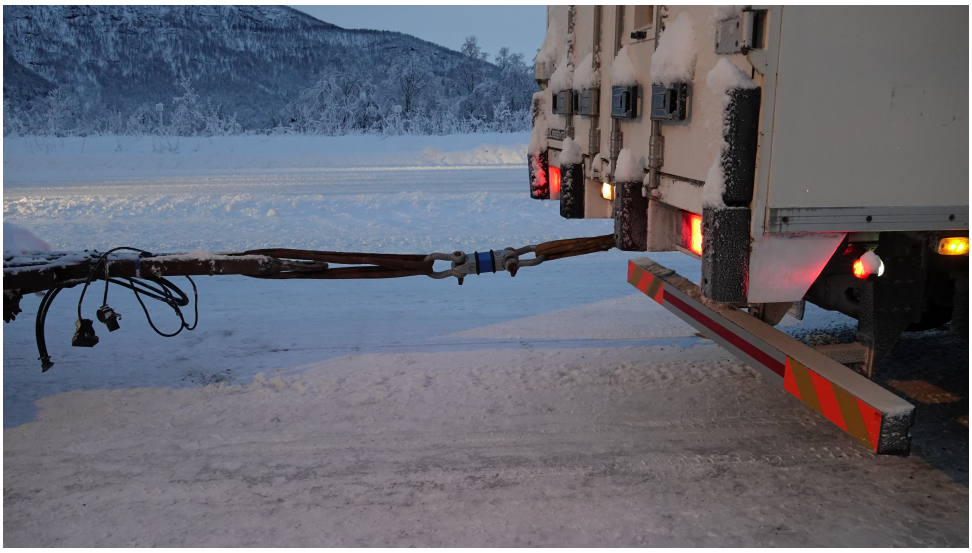


Figure 5.2: Connection of the weight between the truck and the trailer.

chosen because of our previous experience with the sensor, connected to accuracy and reliability when measuring.

5.3 Data collection and storage

In the field study of various sensors the setup for collecting and saving data was done as follows for each individual sensor:

- The RCM411 sensor sends data via Bluetooth to an Android phone and to servers showing the data online on `https://roadweather.online`.
- The MARWIS sensor sends data via Bluetooth to an Asus Zen Pad with an in-house developed app that shows real-time updates and saves the data.
- OBD-II data is collected via Bluetooth from an OBD-II adapter to a mobile app where real-time data is displayed and logged.
- The Walabot sends data via USB to a Raspberry Pi that saves the data.

In addition to the individual sensor saving data to its own space, all measurements were cleaned, merged to one file for each sensor, and stored together in Box [95] for easy access and sharing.

In the sand spreader experiments collecting and saving data from the weight was done by using the remote display delivered with the weight. The remote display has a two-way communication with the weight by a permanent radio connection (2.4 G.Hertz), and can be plugged into a computer by USB where the data is transferred and saved.

5.4 Variable selections and comparisons

In the field study of various sensors, a large amount of data was collected over a long test route. A number of variables were measured from the sensors; RCM411 and MARWIS returned data from 14 variables each, Walabot returned data from 1 variable, and OBD-II collected data from over 20 variables (OBD-II can return data from over 100 parameters, but we selected around 20 for our purpose). For the comparisons four road segments were selected to analyse RCM411 and MARWIS, three segments were selected to compare Walabot on different road conditions, and

two segments were specifically extracted to show certain phenomena of the Walabot. The four segments in the RCM411/MARWIS comparisons, can be seen in Table 5.1. They were selected based on the road condition connected to the segment, and uptime of the sensors. The segments had a timescale of 30 minutes, corresponding to about 40 km in length. The three segments

Table 5.1: Description of the analysed road segments in the comparison between RCM411 and MARWIS.

Segment description			
Date	Location (on the stretch between)	Length [km]	Road conditions
March 28, 2018 10:10 - 10:40	Karasjok-Utsjoki	39.4	Mostly snow
April 2, 2018 16:45 - 17:15	Karasjok-Kautokeino	39.0	Dry and wet
April 3, 2018 09:45 - 10:15	Kautokeino-Enontekiö	40.3	Mostly ice
April 3, 2018 15:30 - 16:00	Kiruna-Narvik	38.0	Mixed

in the Walabot comparisons were also chosen based on road conditions. These segments had a timescale of 10 minutes, and were extracted from the three first segments in Table 5.1. The reason for choosing a smaller timescale was because of the Walabot's high sampling rate (4-5 measurements per second). The two last segments were chosen based on two phenomena that occurred in specific places along the route. The timescales of the segments were 3.5 minutes and 9.5 minutes. For the comparison between RCM411 and MARWIS three variables, were analysed; friction, water thickness, and road temperature. The variables were chosen because both sensors measured them, and because they were the most relevant information in regards to winter road assessment. A comparison was performed between the two sensors by analysing the values of the variables against each other, and against two laboratory experiments done on the same variables from the same types of sensors.

In the sand spreader experiments data was collected as one whole segment for all trials. Three trials were run without sand, and two with sand. The reason for doing only two trials with sand was connected to safety issues, and three trials before sanding was considered enough, given that the values did not change much between tests. A comparison was made between the measurements before and after sanding by analysing the maximum values of the pull force, and the steepness/slope of the graphs in regards to an exponential function. In addition, the friction of the road beneath the truck was measured before and after sanding, and compared in view of the pull force.

Part III

Results and Contributions

6 Results

In this chapter a list of the research papers is presented, together with paper summaries for each individual paper, listing the main findings.

6.1 List of research results

The thesis consists of preprints of 7 papers. A list of the papers follows:

1. Brustad, T.F. and Dalmo, R. Railway transition curves: A review of the state-of-the-art and future research. *Accepted by Infrastructures after minor revisions.*

2. Brustad, T.F. Preliminary Studies on Transition Curve Geometry: Reality and Virtual Reality. *Emerging Science Journal.* 2020; 4. 1-10. 10.28991/esj-2020-01204.

3. Pedersen, A. and Brustad, T.F. A study on Hybrid Sensor Technology in Winter Road Assessment. *Safety.* 2020; 6(1):17.

4. Brustad, T.F and Dalmo, R. Exploring benefits of using blending splines as transition curves. *Submitted for review* (May 7, 2020).

5. Pedersen, A., Brustad, T.F., Bang, B. and Dalmo, R. Alternative representation of railway track geometry. In *Proceedings of IHHA* 2019, Narvik, Norway, 2019.

6. Brustad, T.F., Pedersen, A. and Bang, B. Targeted sanding and its impact on heavy hauler pull force and surface friction. *Submitted for review* (March 19, 2019).

7. Brustad, T.F., Pedersen, A. and Bang, B. A field study of sensors for winter road assessment. *Submitted for review* (March 18, 2020).

In addition to the papers 1 to 7, which can be found in Chapters 9 to 15 in the same order as outlined above, the following scientific works are related to the thesis, but not included:

8. Brustad, T.F. GPU-based rendering of blending spline surface lattices. Communicated at: *9th International Conference on Mathematical Methods for Curves and Surfaces*, Tønsberg, Norway, June 23 - 28 2016.

9. Brustad, T.F. Comparison of intrinsic properties between B-splines and blending splines. Communicated at: *9th International Conference on Curves and Surfaces*, Arcachon, France, June 28 - July 4 2018.

10. Brustad, T.F., Pedersen, A., Bang, B. and Dalmo, R. Representation of flexible transition curves in railway. Poster presented at: *IHHA* 2019, Narvik, Norway, June 10 - 14 2019.

6.2 Paper summary

This section contains a summary of the previously listed papers 1 to 7 with focus on the main findings. The papers are divided as mentioned in Section 2.3; preliminary study, railway geometry and winter road assessment.

6.2.1 preliminary study

Chapter 9 - Railway transition curves: A review of the state-of-the-art and future research

The paper presents the state-of-the-art of railway transition curves. The study highlights relevant definitions, historical events, present research, research challenges, and future opportunities. And aims at collecting and discussing the latest work, in order to provide a better understanding of research challenges and identify future research directions. The main findings are as follows:

- Challenges in railway transition curve research include:
  - Finding evaluation criteria that are "good enough" to evaluate the suitability of the curve as a transition curve.
  - Deciding whether to create curves that can replace one or more original railway segment. A curve replacing several segments may be more flexible in finding an optimal shape, but less flexible in avoiding obstacles. While curves replacing one segment will have opposite flexibility.
  - Linking research and industry closer to find curve properties that are beneficial in vehicle dynamics as well as implementation and maintenance, in particular with respect to wear of tracks and vehicles.
- Future research opportunities include:
  - Finding new evaluation criteria, from an industrial point of view, that can complement the vehicle dynamics criteria and add guidelines connected to implementation and maintenance.
  - Documenting the effect of replacing one segment compared to multiple segments in an existing railway, with a focus on cost weighed against intervention to nature and existing rail elements, as well as the obtained improvements to the alignment.
  - Analysing new curves in a transition perspective that has not been studied in that setting. Examples here include blending splines [90, 91, 92] and circle splines [96, 97].

Chapter 10 - Preliminary studies on transition curve geometry: Reality and virtual reality

The paper describes the experiments and results of laboratory tests conducted on a physical down-scaled vehicle to increase the understanding of the effect of transition curves on lateral force and traced vehicle path, when driving from a straight path to a circular path. Three paths are studied: a path with no transition curve, a path with a clothoid transition curve, and a path with an S-shaped curvature transition curve. In addition a comparison of the physical vehicle model against a virtual model is given in order to increase the validity of the results and to test if simple VR technology resembles the physical model in behaviour. The plots of lateral forces and vehicle paths, for the physical and virtual model, are analysed and compared, with the main focus on two topics: vehicle behaviour when driving through various transition curves and similarity between physical and virtual model. The main findings are as follows:

- The lateral forces working on a vehicle driving along a curve is equal in shape to the curvature of that curve, as long as the vehicle speed is constant.
- The geometric continuity of the curvature in the joints of the transition curve affects the size of the jump in lateral force. Higher continuity means smaller jumps.
- The physical model and the Blender model have similar behaviour in lateral force and traced path, however, the Blender model is a little too perfect lacking, at least, wheel-ground friction to be more realistic.

Chapter 11 - A Study on Hybrid Sensor Technology in Winter Road Assessment

The paper presents preliminary studies on various sensors in relation to creating possible hybrid sensor technology in winter road assessment. The tested sensors include two commercial sensors, RCM411 and MARWIS, one experimental sensor, Walabot, and four supplementary sensors/devices, OBD-II, video, sound, and smartphone sensors. Three main analyses are performed based on data collect from three test routes: a consistency test of the RCM411, a comparison between MARWIS and RCM411 friction, and a comparison between Walabot image energy and RCM411 friction. In addition, the supplementary sensors/devices are tested and discussed. The main findings are as follows:

- The RCM411 sensor variables: friction, surface temperature, and road state; behaves consistently when driving the same test route twice.
- MARWIS and RCM411 friction gave dissimilar results on the same test route.
- Walabot image energy and RCM411 friction appeared to correlate.
- Using OBD-II, video, sound, and smartphone data in winter road assessment can bring a new dimension to the research for a low cost.
- The preliminary experiments showed promising results towards a hybrid sensor because the tested sensors/devices had differences that could complement each other in a beneficial way.

6.2.2 Railway geometry

Chapter 12 - Exploring benefits of using blending splines as transition curves

The paper examines blending splines as transition curves in railway. Three segments are extracted from Ofotbanen, two circular segments connected by a clothoid segment, with the intention of replacing the clothoid with a blending spline in the existing railway. Blending splines with various local curves and ERB-functions are analysed and compared against the clothoid, based on given smoothness criteria. The main findings are as follows:

- The blending spline can give a higher smoothness in the end points, compared to the clothoid.
- The choice of local curves, in the blending spline, has a big impact on the behaviour of the curvature.
- The choice of ERB-function parameters, in the blending spline, shifts the amplitude of the bumps in the curvature function and changes extreme values.
- Using Bézier curves as local curves makes it easy to control the smoothness in the end points, and provides the opportunity to manipulate the curve through the control points.
- A blending spline with 4th degree local Bézier curves gives a curvature function close to a monotonically increasing function, and may be possible to use with a monotonically increasing superelevation ramp, just like the Wiener Bogen transition curve.

Chapter 13 - Alternative representation of railway track geometry

The paper compares a B-spline curve to a segmented railway curve (consisting of straight, clothoid, and circular segments) to examine if the first curve gives a better transition, in the curvature domain, than the last curve. A part of the Ofotbanen railway is analysed and a B-Spline curve is created with focus on achieving a smoother curvature transition than what is

present in the original segmented curve. The curvature plots for the segmented curve and various B-Spline curves are compared and analysed, with special consideration to the joints of the segmented curve. The main findings are as follows:

- The curvature of the B-Splines, both with a uniform and non-uniform parameterization, matched the curvature of the segmented curve, in full for the clothoid segments and closely for the other two segments.
- In the joints of the segmented curve, the B-Spline curves had a tendency to create wiggles (arising from the polynomial interpolation) in the curvature domain, moving along the constant curvature segments (straight and circular segments).
- The curvature of the segmented curve had a G^0 continuity in the joints, while the B-Spline curves had a geometric continuity of higher degree (this is also the reason for the wiggles) in the same areas, which in theory gives a better transition.
- Changing the degree and parameterization of the B-Spline showed that some combinations created less wiggles and a smoother curve than others.

6.2.3 Winter road assessment

Chapter 14 - Targeted sanding and its impact on heavy hauler pull force and surface friction

The paper describes the experiment and results of field work conducted to show the impact targeted sanding has on heavy hauler pull force. The tests are performed by placing a weight between a heavy hauler and a trailer, and data is collected before and after sanding when the truck is trying to pull the trailer (brakes are locked on the trailer). The weight and friction data is then examined to look at the increase in both variables after sanding, compared to before sanding. The main findings are as follows:

- The friction was estimated to have increased from 0.24 before sanding to 0.48 after sanding, for this particular test case.
- The weight measurements showed that there was an increase in maximum pull force of about 2000 kg after sanding, and that the derivative/slope of the force had a time dependent increase in pull force that was much higher after sanding compared to before.

Chapter 15 - A field study of sensors for winter road assessment

The paper gives a comparison and analysis of sensors for winter road assessment, in a real world setting. Four sensors are investigated through field experiments: RCM411, MARWIS, OBD-II, and Walabot. The sensor data is then examined in three groups: RCM411 and MARWIS are compared to each other and to laboratory experiments, OBD-II is discussed as a supplementary tool in the assessments, and initial tests are performed with the experimental Walabot radar. The main findings are as follows:

- Results from the RCM411/MARWIS comparison revealed that the laboratory experiments and the field experiments behaved in a similar manner with enough of the same characteristics to increase the reliability of both investigations. Compared against each other, the sensors showed much of the same assessments, with the biggest differences being that MARWIS behaved more sensitive to changes in conditions, and had unexplainable spikes for some variables.
- The analysis of OBD-II indicated that data collected from the vehicle had good potential as a supplementary tool in road assessment, and could be used to calculate output variables not directly connected to the input variables.

- Testing of the experimental Walabot radar showed that the differences in the data between various road conditions were so small that a constructive feedback was impossible without further testing. In addition, two phenomena of outside forces affected the measurements; vibration and water/ice covering the casing.

7 Contributions

In this chapter the contributions to railway geometry and winter road assessment are listed.

7.1 Railway geometry

The main contributions to the study of transition curves and railway geometry are the extensive state-of-the-art study, with discussions, on railway transition curves, and the experimentation into smoother transitions in an already existing railway. The contributions are divided into two categories: transition curves and railway segments. A list of the categories with relevant contributions follows:

- Transition curves:
  - Conducting experiments on a down-scaled vehicle model to observe the effects of various transition curves and the lack of a transition curve, in regards to lateral force working on the vehicle and the path traced while driving.
  - Conducting comparisons of a down-scaled vehicle model against a Blender vehicle model to consider the resemblance, in regards to behaviour, between the two models.
  - Identifying current research challenges, surfacing from the state-of-the-art research, connected to transition curves in railway.
  - Discussing future research opportunities based on research challenges and experimental knowledge.
  - Suggesting two new types of transition curves based on known challenges and properties, blending splines and circle splines.
  - Implementing and analysing blending splines as transition curves, replacing a clothoid between two circular arc segments, in an already existing railway.
- Railway segments:
  - Testing a new approach of creating a smoother curve, with minor changes, on top of an already existing railway track.
  - Implementing a B-spline curve on top of a segmented railway curve with enough freedom to chose its own path over the transition segments.
  - Conducting comparisons between the B-spline curve and the segmented curve with focus on the transition in the curvature domain.

7.2 Winter road assessment

The main contributions to the study of winter road assessment are the diverse field experiments that were conducted in genuine winter environments, as opposed to laboratory experiments or test site experiments. The contributions are divided into three categories: testing known sensors, testing experimental sensors, and validating tools connected to increasing safety on winter roads. A list of the categories with relevant contributions follows:

- Testing known sensors:
  - Collecting a large amount of data from field experiments performed under a variety of road and weather conditions, over a long test route.
  - Conducting extensive comparisons of field experiments against laboratory experiments, for the RCM411 and MARWIS sensors.
  - Conducting comparisons between RCM411 and MARWIS with field test data.
- Testing experimental/supplementary sensors:
  - Suggesting and analysing multiple sensors that can be used in hybrid measurements, based on field experiments and experience.
  - Testing an experimental radar sensor in a field study setting, that has never before been tested for this purpose.
  - Analysing vehicle data as a supplementary tool in winter road assessment.
- Validating tools:
  - Conducting field experiments on a heavy hauler, equipped with an Autoline sand spreader, on slippery winter conditions.
  - Analysing measurement data to validate if the Autoline sand spreader can increase pull force of the heavy hauler by sanding beneath the drive wheels.
  - Analysing friction data to examine how much the Autoline sand spreader increases the friction on the road.

Part IV

Conclusion

8 Conclusions and future work

In this chapter concluding remarks are given, and some ideas and suggestions for future work based on the results in the thesis are presented.

8.1 Concluding remarks

In this thesis two main topics have been considered: railway geometry and winter road assessment. In railway geometry the focus has been on transition curves, their impact on railway safety and maintenance, and ways to improve transitions. In winter road assessment the focus has been on sensor data, their measurements and capabilities, and the possibility of creating a hybrid sensor. Both topics are important to the northern regions of Norway, because of its challenging geography and winter conditions, putting pressure on the limited road and railway capacity. Improving transitions on the railway line can reduce the need for maintenance, and improve safety and comfort of passengers, while accurately assessing winter roads means road users and maintenance personnel can be notified of specific dangerous conditions.

The research in the study was conducted by evaluating and analysing curve properties in regards to curvature and lateral forces, and performing extensive field experiments with data analyses of various sensors. Based on the research questions for railway geometry the results show that: the challenges that can be observed in railway transition curve research are connected to evaluation criteria, replacing one or multiple segments, and linking research and industry; the impact of a transition curve can be seen in the size of the jumps in lateral forces, and the shape of the traced path; and replacing multiple segments in an existing railway curve can improve the smoothness of the transitions. Based on the research questions for the winter road assessment the results show that: various sensors seems to be able to benefit each other enough to defend a hybrid solution; available sensors can give information on road state, friction, road and air temperature, water thickness, and ice percentage, however, the sensitivity to certain conditions seems to differ between sensors; and targeted sanding seems to increase friction enough to increase pull force for heavy haulers.

8.2 Suggestions for future work

The research within railway geometry, and transition curves, has opened up multiple possibilities for future work. The thesis focused on the horizontal alignment of the railway track, so the vertical track alignment or considering the railway as a 3D system, are possible future topics. In addition, there are some indications, from Bane NOR, that applying chain theory to simulate railway wagons through a turn will show wear on the rail in unexpected areas. When it comes to using LCA to evaluate the suitability of a transition curve, the variable of speed in the equation was kept constant. By varying the speed along the curve new results may emerge in regards to choosing a proper transition curve. A first attempt at using blending splines as transition curves was tested. In this area there are some interesting future work opportunities. One example is using higher degree ($d > 4$) Bézier curves as local curves, in order to make an algorithm that

can find an optimal placement of the control points, without decreasing the smoothness in the end points. Another example is to test if a blending spline with non-monotonically increasing curvature can be paired with a monotonically increasing superelevation ramp to create comfortable transitions, just like the Wiener Bogen. Other future research opportunities can be found in Paper 1. They include: finding new evaluation criteria for transition curves that also benefit the industry in regards to implementation and maintenance, documenting the effect of replacing one segment compared to multiple segments in an already existing railway, and analysing the suitability of circle splines as transition curves.

In the winter road assessment research some ideas for future work include more testing of the sensors, ways of processing data, and how to best display and share information. Testing of the sensors should be done under more controlled conditions with a focus on creating a hybrid technology with the various sensor's data that complement each other in a beneficial way. In the future, processing of data should be given more though, preferably with hybrid technology in mind. Algorithms dealing with multiple inputs from different sensors that return one collected assessment would be of interest. The information that is obtained from the sensors has to be displayed to the driver and, in addition, shared with other vehicles and maintenance personnel. In this thesis, displaying and sharing information have not been addressed, however, both are important topics that should be considered for future research.

Bibliography

[1] A. Seidenberg. The origin of mathematics. *Archive for History of Exact Sciences*, 18(4):301–342, 1978.

[2] C.J. Scriba, P. Schreiber, and J. Schreiber. *5000 Years of Geometry: Mathematics in History and Culture*. Springer Basel, 2015.

[3] A. Seidenberg. The ritual origin of geometry. *Archive for History of Exact Sciences*, 1(5):488–527, Oct 1961.

[4] B. L. van der Waerden. *Geometry and Algebra in Ancient Civilizations*. Springer-Verlag Berlin Heidelberg, 1983.

[5] John R. Silvester. *Geometry: Ancient and Modern*. Oxford University Press, 2001.

[6] John P. Britton, Christine Proust, and Steve Shnider. Plimpton 322: a review and a different perspective. *Archive for History of Exact Sciences*, 65(5):519, Aug 2011.

[7] Bibhutibhusan Datta. *The Science Of The Sulba: A Study In Early Hindu Geometry*. University of Calcutta, 1932.

[8] Kangshen Shen, John N. Crossley, Anthony Wah-Cheung Lun, and Hui Liu. *The Nine Chapters on the Mathematical Art: Companion and Commentary*. Oxford University Press, 1999.

[9] Audun Holme. *Geometry : Our Cultural Heritage*. Springer Berlin Heidelberg, 2nd edition, 2010.

[10] P.F. O'Grady. *Thales of Miletus: The Beginnings of Western Science and Philosophy*. Western philosophy series. Ashgate, 2002.

[11] George Johnston Allman. *Greek geometry from Thales to Euclid*. Dublin Printed at the University Press, 1877.

[12] Viggo Brun. *Alt er tall : matematikkens historie i oldtid og middelalder*. U-bøkene. Universitetsforl, Oslo, 1964.

[13] David Eugene Smith. *History of mathematics : 2 : Special topics of elementary mathematics*, volume 2. Dover, New York, 1958.

[14] J. Gallier. *Geometric Methods and Applications: For Computer Science and Engineering*. Texts in Applied Mathematics. Springer New York, 2011.

[15] J. Glover. Transition curves for railways. *Proc. Inst. C. E.*, 140, 1900.

[16] E. M. Horsburgh. XXIV.—The railway transition curve. *Proceedings of the Royal Society of Edinburgh*, 32:333–347, 1913.

[17] A. W. Miller. The transition spiral. *Australian Surveyor*, 7(8):518–526, 1939.

[18] J. Nam. *How Railway Systems Work*. AuthorHouse UK, 2014.

[19] Institution of Engineers and Shipbuilders in Scotland. *Transactions of the Institution of Engineers and Shipbuilders in Scotland*. Number v. 4. 1861.

[20] Arthur Lovat Higgins. *The transition spiral and its introduction to railway curves with field exercises in construction and alignment*. Van Nostrand company, New York, 1922.

[21] Constantin Ciobanu. Bloss transition - a short design guide. *PWI Journal*, 133:14–18, 01 2015.

[22] Björn Kufver. *Optimisation of horizontal alignments for railways : Procedures involving evaluation of dynamic vehicle response*. PhD thesis, Department of Vehicle Engineering, Royal Institute of Technology, Stockholm, Sweden, 2000.

[23] Orhan Baykal. Concept of lateral change of acceleration. *Journal of Surveying Engineering*, 122:132–141, 08 1996.

[24] Abdullah Arslan, Ergin Tari, Rushan Ziatdinov, and Rifkat Nabiyev. Transition curve modeling with kinematical properties: Research on log-aesthetic curves. *Computer-Aided Design and Applications*, 11:509–517, 10 2014.

[25] Ergin Tari. The new generation transition curves. *ARI The Bulletin of the Instanbul Technical University*, 54:35–41, 01 2004.

[26] A. Pirti, M.A. Yücel, and T. Ocalan. Transrapid and the transition curve as sinusoid. In *Tehnički vjesnik*, volume 23, pages 315–320, 2016.

[27] K. Zboinski and P. Woznica. Optimization of polynomial transition curves from the viewpoint of jerk value. *Archives of Civil Engineering*, 63, 01 2017.

[28] Azhar Ahmad and Jamaludin Ali. G3 transition curve between two straight lines. pages 154–159, 08 2008.

[29] N. Eliou and G Kaliabetsos. A new, simple and accurate transition curve type, for use in road and railway alignment design. *European Transport Research Review*, 6:171–179, June 2014.

[30] SA Shebl. Geometrical analysis of non-linear curvature transition curves of high speed railways. *Asian Journal of Current Engineering and Maths*, 5(4):52–58, Aug 2016.

[31] Robert Wojtczak. *Charakterystyka krzywej przejściowej Wiener Bogen®*. Publishing House of Poznan University of Technology, 2017.

[32] Wladyslaw Koc. Transition curve with smoothed curvature at its ends for railway roads. *Current Journal of Applied Science and Technology*, 22:1–10, 07 2017.

[33] W Koc. Smoothed transition curve for railways. *Przeglad Komunikacyjny*, (7):19–32, 2019.

[34] W Koc. New transition curve adapted to railway operational requirements. *Journal of Surveying Engineering*, 145, 08 2019.

[35] Krzysztof Zboinski and Piotr Woznica. Optimisation of polynomial railway transition curves of even degrees. *Archives of Transport*, 35:71–86, 01 2015.

[36] Krzysztof Zboinski and Piotr Woznica. Combined use of dynamical simulation and optimisation to form railway transition curves. *Vehicle System Dynamics*, 56(9):1394–1450, 2018.

[37] Krzysztof Zboinski and Milena Golofit-Stawinska. Investigation into nonlinear phenomena for various railway vehicles in transition curves at velocities close to critical one. *Nonlinear Dynamics*, pages 1–47, 10 2019.

[38] Andrzej Kobryń. New solutions for general transition curves. *Journal of Surveying Engineering*, 140:12–21, 02 2014.

[39] Andrzej Kobryń. Universal solutions of transition curves. *Journal of Surveying Engineering*, 142:04016010, 02 2016.

[40] Andrzej Kobryń and Piotr Stachera. S-Shaped transition curves as an element of reverse curves in road design. *The Baltic Journal of Road and Bridge Engineering*, 14:484–503, 11 2019.

[41] J. Sanchez-Reyes and J.M. Chacon. Nonparametric Bézier representation of polynomial transition curves. *Journal of Surveying Engineering*, 144, 05 2018.

[42] Andrzej Kobryń. Use of polynomial transition curves in the design of horizontal arcs. *Roads and Bridges - Drogi i Mosty*, 16(1):5–14, 2017.

[43] Gerardo Casal Urcera, Duarte Santamarina, and Miguel Vázquez-Méndez. Optimization of horizontal alignment geometry in road design and reconstruction. *Transportation Research Part C: Emerging Technologies*, 74, 01 2017.

[44] Instrument Society of America. ISA-37.1-1975 (R1982), Electrical Transducer Nomenclature and Terminology. Standard.

[45] J. Scholz, T. Ricolfi, W. Göpel, J. Hesse, and J.N. Zemel. *Sensors: A Comprehensive Survey, Thermal Sensors*. Sensors. Wiley, 2008.

[46] Warren S. Johnson. Electric tele-thermoscope. `https://patentimages.storage.googleapis.com/83/96/21/7f61e638f6b543/US281884.pdf`. Patent, 1883.

[47] Samuel M. Bagno. Approach alarm system with unwanted signal elimination. `https://patentimages.storage.googleapis.com/13/5a/f1/959d7d03bb540b/US2767393.pdf`. Patent, 1953.

[48] Jon S. Wilson. *Sensor Technology Handbook*. Newnes, 2005.

[49] J. Vetelino and A. Reghu. *Introduction to Sensors*. CRC Press, 2017.

[50] R. Frank. *Understanding Smart Sensors*. Artech House sensors library. Artech House, 2000.

[51] Joseph M. Giachino. Smart sensors. *Sensors and Actuators*, 10(3):239 – 248, 1986.

[52] Wen H. Ko and C.D. Fung. VLSI and intelligent transducers. *Sensors and Actuators*, 2:239 – 250, 1981.

[53] IEEE standard for a smart transducer interface for sensors and actuators - common functions, communication protocols, and transducer electronic data sheet (TEDS) formats. *IEEE Std 1451.0-2007*, pages 1–335, Sep. 2007.

[54] Michael J. McGrath, Cliodhna Ní Scanaill, and Dawn Nafus. *Sensor Technologies: Healthcare, Wellness, and Environmental Applications*. Apress, Berkeley, CA, 2013.

[55] Torben Bronée. Apparatus for indicating at a distance the danger of slippery ice being formed on roads. `https://patentimages.storage.googleapis.com/4c/76/d0/760ea22dafbf82/US2301247.pdf`. Patent, 1939 (Granted 1942).

[56] Louis P. Clark. Highway condition indicating system. https://patentimages.storage.googleapis.com/df/7f/49/4eb1b5a5fc6274/US2849701.pdf. Patent, 1954 (Granted 1958).

[57] Marcel Boschung. Early ice-warning device. https://patentimages.storage.googleapis.com/79/8b/03/be85b0a783afb6/US4222044.pdf. Patent, 1977 (Granted 1980).

[58] Peter W. Decker. Road surface ice detector and method for vehicles. https://patentimages.storage.googleapis.com/37/7a/3a/d3a8d87b16e3d6/US4274091.pdf. Patent, 1979 (Granted 1981).

[59] R. Finkele. Detection of ice layers on road surfaces using a polarimetric millimetre wave sensor at 76 ghz. *Electronics Letters*, 33(13):1153–1154, June 1997.

[60] Bent H. Sass. A numerical model for prediction of road temperature and ice. *Journal of Applied Meteorology*, 31(12):1499–1506, 1992.

[61] Wilfrid A. Nixon. The potential of friction as a tool for winter maintenance. Technical Report 392, Iowa Department of Transportation and Iowa Highway Research Board, The University of Iowa, Iowa City IA 52242-1585, February 1998.

[62] H. Rudolf, G. Wanielik, and A. J. Sieber. Road condition recognition using microwaves. In *Proceedings of Conference on Intelligent Transportation Systems*, pages 996–999, Nov 1997.

[63] Izumi Sakai and Claire Stenson. Optical sensors for inferring road adhesion. In P. McGeehin, editor, *Fibre Optics '90*, volume 1314, pages 236 – 243. International Society for Optics and Photonics, SPIE, 1990.

[64] N. Kees and J. Detlefsen. Road surface classification by using a polarimetric coherent radar module at millimeter waves. In *1994 IEEE MTT-S International Microwave Symposium Digest (Cat. No.94CH3389-4)*, pages 1675–1678 vol.3, May 1994.

[65] Andreas Kuehnle and Wilco Burghout. Winter road condition recognition using video image classification. *Transportation Research Record*, 1627(1):29–33, 1998.

[66] G. Magerl, W. Pritzl, and P. W. Frohling. Remote sensing of road condition. In *[Proceedings] IGARSS'91 Remote Sensing: Global Monitoring for Earth Management*, volume 4, pages 2137–2140, June 1991.

[67] Y. Hou, A. Gupta, T. Guan, S. Hu, L. Su, and C. Qiao. Vehsense: Slippery road detection using smartphones. In *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, pages 1–5, June 2017.

[68] P. Jonsson, J. Casselgren, and B. Thörnberg. Road surface status classification using spectral analysis of NIR camera images. *IEEE Sensors Journal*, 15(3):1641–1656, March 2015.

[69] S. Kawai, K. Takeuchi, K. Shibata, and Y. Horita. A method to distinguish road surface conditions for car-mounted camera images at night-time. In *2012 12th International Conference on ITS Telecommunications*, pages 668–672, Nov 2012.

[70] R. Omer and L. Fu. An automatic image recognition system for winter road surface condition classification. In *13th International IEEE Conference on Intelligent Transportation Systems*, pages 1375–1379, Sep. 2010.

[71] V. V. Viikari, T. Varpula, and M. Kantanen. Road-condition recognition using 24-GHz automotive radar. *IEEE Transactions on Intelligent Transportation Systems*, 10(4):639–648, Dec 2009.

[72] Taisto Haavasoja, Juhani Nylander, and Pauli Nylander. Experiences of mobile road condition monitoring. In *Proceedings of SIRWEC 2012, Helsinki, Finland*, May 2012.

[73] Lufft. *USER Manual MARWIS/StaRWIS*, 2018 (Accessed 2019-10-01). `https://www.lufft.com/download/manual-lufft-marwis-starwis-en/`.

[74] Narvik Havn KF. Port of narvik: Arctic-hub. (Online) `https://www.narvikhavn.no/media/26759/narvikhavnkf.pdf`. Accessed: 2019-07-03.

[75] Narvik Havn KF. Ofotbanealliansen. (Online) `https://www.narvikhavn.no/knutepunkt-narvik/satsningsomraader/ofotbanealliansen.aspx`. Accessed: 2019-07-03.

[76] Jean Andrey, Brian Mills, and Jessica Adams. Weather information and road safety, 08 2001.

[77] Daniel Eisenberg and Kenneth E. Warner. Effects of snowfalls on motor vehicle collisions, injuries, and fatalities. *American Journal of Public Health*, 95(1):120–124, 2005.

[78] Keith K. Knapp, Dennis Kroeger, and Karen Giese. Mobility and safety impacts of winter storm events in a freeway environment. Technical report, Iowa State University. Center for Transportation Research and Education, 2000.

[79] Carl-Gustaf Wallman and Henrik Åström. Friction measurement methods and the correlation between road friction and traffic safety : A literature review. Technical Report 911A, Swedish National Road and Transport Research Institute, 2001.

[80] Per Andreas Langeland and Ross Owen Phillips. Heavy vehicles and traffic accidents – Norway versus other European countries. Technical report, Institute of Transport Economics, 2016.

[81] J Granlund and P Thomson. Traffic safety risks with EU tractor-semitrailer rigs on slippery roads. HVTT14: International Symposium on Heavy Vehicle Transport Technology, 14th, 2016, Rotorua, New Zealand, 2016.

[82] Jernbanedirektoratet. Jernbanenettet i norge. Online. `https://www.jernbanedirektoratet.no/no/jernbanesektoren/jernbanenettet-i-norge/`.

[83] Bane Nor. Ofotbanen. (Online) `https://www.banenor.no/jernbanen/banene/ofotbanen/`. Accessed: 2019-07-03.

[84] Thomas Garrity and Joe Warren. Geometric continuity. *Computer Aided Geometric Design*, 8(1):51 – 65, 1991.

[85] Autoline AS. Market leader in scandinavia, 2017 (Accessed 2019-03-05). (Online) `http://www.autoline.no/About-us`.

[86] P. Kosso. *A Summary of Scientific Method*. SpringerBriefs in Philosophy. Springer Netherlands, 2011.

[87] Mario Perez-Zepeda, Carmen Garcia-Peña, and Luis Gutiérrez-Robledo. *Aging Research-Methodological Issues*. 07 2015.

[88] Martyn Shuttleworth and Lyndsay T. Wilson. What is the scientific method? (Online) `https://explorable.com/what-is-the-scientific-method`. Accessed: 2019-11-05.

[89] John Misachi. What is the scientific method? (Online) `https://www.worldatlas.com/articles/what-is-the-scientific-method.html`. Accessed: 2019-11-05.

[90] L. T. Dechevsky. Expo-rational B-splines. Communicated at the Sixth International Conference on Mathematical Methods for Curves and Surfaces, Tromsø, Norway, 2004.

[91] Lubomir T. Dechevsky, Børre Bang, and Arne Lakså. Generalized expo-rational B-splines. *International Journal of Pure and Applied Mathematics*, 57(6):833–872, 2009.

[92] Lubomir T. Dechevsky and Peter Zanaty. Smooth GERBS, orthogonal systems and energy minimization. *AIP Conference Proceedings*, 1570(1):135–162, 2013.

[93] Erwin Coumans. Bullet 2.83 physics sdk manual, 2015.

[94] Tractel. Dynafor llx2, 2019 (Accessed 2019-03-05). (online) `https://www.tractel.com/us/series.php?id_serie=7`.

[95] Box. Simplify how you work, Accessed 2019-11-25. (Online) `https://www.box.com/home`.

[96] Carlo H. Séquin, Kiha Lee, and Jane Yen. Fair, G2- and C2-continuous circle splines for the interpolation of sparse data points. *Computer-Aided Design*, 37(2):201 – 211, 2005.

[97] Carlo Sequin and Kiha Lee. Fair and robust circle splines. In *SIGGRAPH 2003 Conference on Sketches and Applications: in conjunction with the 30th annual conference on Computer graphics and interactive techniques*, January 2003.

Part V

Papers

9 Railway transition curves: A review of the state-of-the-art and future research

Tanita Fossli Brustad and Rune Dalmo

Abstract

Transition curves are a useful tool for lateral alignment of railway segments. Their design is important to ensure safe and comfortable travel for passengers and cargo. Well designed transition curves can lead to reduced wear of tracks and vehicles, which is beneficial from a maintenance point of view. Extensive studies have been performed through decades to find transition curves that can replace existing railway segments for the purpose of enhancing certain properties. Those studies seek to form curves that satisfy desired evaluation criteria, which are often connected to geometric continuity between the curve segments, and vehicle dynamics, to secure a smooth ride. This research topic is still ongoing and active at present. Recent results and findings are in line with the developments on the topic of vehicle dynamics and within the railway industry. For this reason it is appropriate to collect and discuss the latest work, since there are no up-to-date detailed literature reviews available. This paper explores the present state-of-the-art of railway transition curves, and identifies some of the research challenges and future research opportunities in the field.

9.1 Introduction

Most railways are described in terms of piece-wise curves. The two main types of segments, namely, straight lines and circular arcs, are connected together via transition curves. The utilization of suitable transition curves is crucial in the processes of constructing new and refurbishing existing railways. This is, in particular, important to facilitate safe and comfortable train travel, and to reduce the need for maintenance [1, 2].

The main purpose of a transition curve is to enable a smooth transition between straight and curved railway segments by preventing sudden jumps in lateral forces [3], which could be the case if two main segments were coupled directly together. The advent of high speed trains and the development in heavy haul railways have triggered new requirements for transition curves. One such requirement, identified in [4], is lateral change of acceleration (LCA). The LCA function combines curve properties and vehicle parameters, to give an evaluation criterion that includes both geometry and vehicle dynamics constraints, since both are important when analysing transition curves.

Exploring new transition curves for horizontal railway alignment with more favourable properties than the classical ones is an active area of research which is highly relevant and progressing. Although the main goal of replacing one or more railway segments in both new and existing tracks is coherent in most of the research, the methods and evaluation criteria differ. Some examples of contrasts include the cover of the new curve (replacing one [5] or several [6] original segments), the applied model (simple [5] or advanced [7] vehicle model), and the principles for evaluation (lateral forces [8], center of gravity of the vehicle [9], design simplicity [10], or cost [11]). Considerable opportunities are available to the railway industry, however, they do not come without challenges.

An important realization when it comes to optimization of railway travel and transport is that the geometric properties of the transition curve is only one part in a complex system. Transition curves alone are important, but they need to be included in a bigger picture. Research regarding safe and comfortable railways is conducted in several areas. For instance, the topics of active steering of railway vehicles [12], bogie suspension [13], and tilting trains [14] are investigated. Additionally, we note the focus on increasing the knowledge on railway dynamics [15, 16], improving health monitoring systems [17, 18], and further developing simulation models [19, 20, 21].

In this paper a survey of the state-of-the-art of railway transition curves is presented. The study highlights important definitions, historical events, present research, research challenges, and topics for future work. The scope of the paper is to collect and discuss the latest work on railway transition curves, in order to provide an overview of research challenges and to identify relevant potential future research directions in this field. The paper focuses on the state-of-the-art of transition curves within railway applications. Nevertheless, some aspects of transition curves in highway research are also covered briefly since the two are closely connected and often referred to in the same setting.

The paper is organized as follows. Section 9.2 gives an overview of relevant definitions; transition curves, the classical clothoid, and lateral change of acceleration, and introduces briefly the history of transition curves in railway. Section 9.3 presents the state-of-the-art. Challenges related to the state-of-the-art are then discussed in Section 9.4, and future research opportunities are considered in Section 9.5. Finally, some concluding remarks are given in Section 9.6.

9.2 Overview

This section provides some relevant definitions, to make this paper self-containing and readable for a broad audience, and covers briefly the history of the transition curve in railway.

9.2.1 Definitions

Below follows short explanations of transition curves, the classical clothoid curve, and the lateral change of acceleration concept, since these definitions and concepts are used in the sequel.

Transition curves

The transition curve fulfils the role of connecting straight sections to curved sections, as well as coupling curved sections together, in railway- and highway design [22, 23]. In the first case, the task of the transition curve is to gradually decrease the radius of curvature from infinity at the straight section to that of the circular curve at the curved section, and at the same time provide a change in superelevation from zero to maximum elevation. The curvature- and superelevation functions usually share the same behaviour (both are often linear), however, some research addresses the case of difference [24].

The described properties are important in order to counteract sudden jerks in the centrifugal forces, and gently apply them over the course of the transition curve. The most common transition, called simple transition curve, takes place between a straight and a curved section (see Figure 9.1). Transition curves are also used to enhance the change between two subsequent circular curves where the difference in radius is large. This kind is called a segmental transition curve [22], as illustrated in Figure 9.2. The main advantages of using transition curves are, according to [1, 2], the following:

- Providing a comfortable ride for passengers.
- Providing a safer ride for passengers.
- Enabling the vehicle to drive at a higher speed.
- Reducing wear and tear on wheels and rails, decreasing maintenance and repair costs.

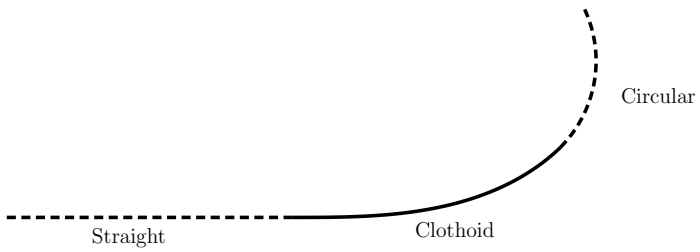


Figure 9.1: A straight curve and a circular curve connected by a transition curve.

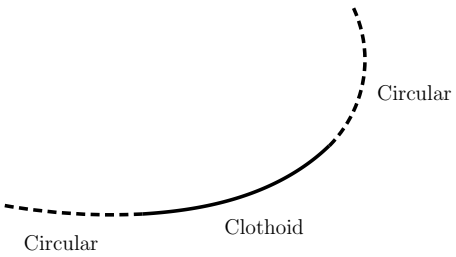


Figure 9.2: Two circular curves with different radius connected by a transition curve.

The clothoid is the most commonly used transition curve in road and railway design. Its curvature properties were for a long time considered to be optimal in railway alignment. However, it was regonized later, by examining how the LCA works on the vehicle, that the clothoid is not optimal for high speed trains with regards to passenger comfort.

The clothoid

The clothoid of length l and end radius r is a spiral defined parametrically as

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} C(t) \\ S(t) \end{pmatrix},$$

where $C(t)$ and $S(t)$ are the Fresnel integrals

$$\begin{aligned} C(t) &= \frac{1}{a}\int_0^{\hat{t}} \cos\left(\frac{\pi}{2}u^2\right) du \\ S(t) &= \frac{1}{a}\int_0^{\hat{t}} \sin\left(\frac{\pi}{2}u^2\right) du, \end{aligned} \tag{9.1}$$

where $a = \sqrt{\frac{1}{\pi r l}}$ is a scaling factor, and $\hat{t} = at$, with the parameter $-\infty < t < \infty$. The following power series expansion of the integrals is given in [25]

$$\begin{aligned} C(t) &= \frac{1}{a}\sum_{i=0}^{\infty} \frac{(-1)^i (\frac{\pi}{2})^{2i} \hat{t}^{4i+1}}{(2i)!(4i+1)} \\ S(t) &= \frac{1}{a}\sum_{i=0}^{\infty} \frac{(-1)^i (\frac{\pi}{2})^{2i+1} \hat{t}^{4i+3}}{(2i+1)!(4i+3)}. \end{aligned}$$

The clothoid has been proposed, solved, and reinvented by a number of people throughout history. For this reason it is known by many names, such as Clothoid, Euler spiral, and Cornu spiral. The history of the spiral is covered in [26]. One of the first instances of the clothoid as a transition curve in railway was by Arthur Talbot [27], who was among the first to approach the transition curve as a mathematical problem. Since then the clothoid has become very popular and is the most widely used curve in railway geometry, mainly because of its curvature properties, which changes linearly with the curve length.

Lateral change of acceleration

Lateral change of acceleration (LCA) is a criterion that combines the geometry of the transition curve with variables from the vehicle to generate a vehicle dynamics measurement of the analysed curve. Baykal [4] expressed LCA as

$$z = \frac{pv}{\sqrt{u^2+p^2}}\left(3ka_t + v^2\frac{dk}{dl} - \frac{kv^2u + gp}{u^2+p^2}\frac{du}{dl}\right), \tag{9.2}$$

where p is the horizontal width of the platform (m), v is the velocity of the train $(\frac{m}{s})$, u is the superelevation, which is the elevation of the outer rail according to the inner rail (m), k is the curvature along the curve $(\frac{1}{m})$, a_t is the tangential acceleration $(\frac{m}{s^2})$, and g is the gravity constant $(9.81\frac{m}{s^2})$. The LCA function describes the change in resultant acceleration occurring along the curve normal with respect to time. It can be thought of as the lateral jerk of the railway wagon. From expression (9.2) it can be observed that if the velocity of the vehicle is constant, then the LCA is a scaled change in curvature. This shows that the curve geometry and the vehicle dynamics are closely connected when evaluating the suitability of a transition curve by means of LCA.

There are mainly three criteria [28] which are used in the comparison of curves with a basis in the LCA fuction. These are, in prioritized order,

1. discontinuity (jumps),

2. magnitude (extreme values), and

3. discontinuity (breaks).

The absence of jumps in the LCA function is the most important criterion of the three. This is because discontinuities (in the form of jumps) affect travel comfort and cause wear on wheels and rails. Any curve without jumps in the LCA is considered to be superior to curves with jumps.

If neither of the curves have jumps, then they are considered to be equivalent in regards to criterion 1, and criterion 2 can be used. It compares the extreme value (the largest absolute value) of the LCA function, for each curve, against an upper boundary value. Some references for the maximum allowed value are listed in [29]:

- $z_{max} = 0.3\frac{m}{s^3}$ for highways [30];
- $z_{max} = 0.3\frac{m}{s^3}$ for railways [3];
- $z_{max} = 0.4\frac{m}{s^3}$ for railways [31];
- $z_{max} = 0.6\frac{m}{s^3}$ for highways [32];
- $z_{max} = 0.6\frac{m}{s^3}$ [33].

Any curve satisfying these conditions is superior to those that do not. If the curves are equivalent after criterion 2, as well, then the third criterion can be tested. Criterion 3 considers discontinuities in the form of breaks at the start and end of the transition curve. Any curve without breaks at the end-points is superior to those which has such breaks.

As noted above, the clothoid is not optimal when its LCA is examined. This can be seen from the LCA plot of a clothoid connected to a straight and circular segment depicted in Figure 9.3. By visual comparison of the LCA in the figure against the three criteria, it can be observed that the clothoid does not satisfy criterion 1 (since it has jumps in the graph). As consequences, passengers' comfort can be affected, in particular for high speed railways, and wheels and rails can be exposed to more wear.

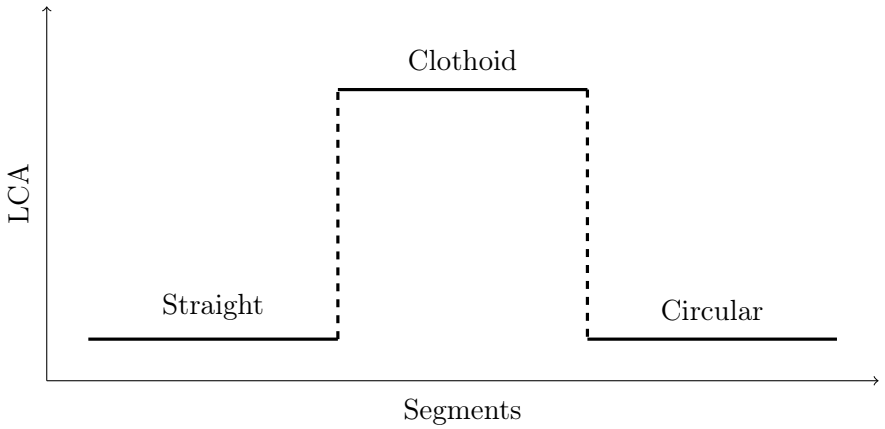


Figure 9.3: LCA plot of a clothoid connected to a straight and circular segment.

9.2.2 History

Early railway tracks used to consist of straight sections and circular arcs only. With the low train speed and the wide radius of the circular sections the direct transition between straight and

circular sections was considered to be easy and safe with regards to the motion of the train [34], as well as within the scope of current surveying technologies at that time. With the evolution of railway the radius of the circular tracks decreased and the travel speed of the trains increased, and the transitions became jerky and unpleasant, and sometimes even dangerous [34, 35]. This recognition paved the way for easement curves, or transition curves, which they are more known as, that enhanced the transition from a straight track to a curved track. As a result, the ride became safer and more comfortable even when higher speed was involved [34, 35, 36]. Because of this there are today three main track types in the horizontal layout, namely, tangent tracks (straight lines), circular tracks (with a constant radius), and transition-curve tracks [37].

Different types of transition curves have been used in railway through the years. Two early examples of such curves are addressed by prof. W. J. Macquorn Rankine in [38]: The "harmonic curve" or "curve of sines" by William Gravatt from 1829 and the "curve of adjustment" by William Froude from 1842. In the late 19th century a railway transition spiral with linear curvature, in relation to arc length, was derived independently by multiple engineers, who seemed to be unaware that Leonhard Euler had already derived that curve more than a century in prior. Consequently the curve has many names, i.e., Euler spiral, clothoid, Cornu spiral, and Glover's spiral, of which clothoid is most commonly used. A. L. Higgins was the first to draw an equivalence between the railway transition spiral and Euler's definition in 1922 [39]. With emphasis on curvature, transition curves can be divided into two groups: transition curves with linear curvature and transition curves with non-linear curvature [40]. From its invention and to this day the clothoid, which belongs to the first group, is the most widely used transition curve in railway because its linear curvature property made it considered to be an optimal transition for a long time.

In addition to the clothoid, many curves with non-linear curvature were suggested as transition curves early on. Various examples are provided in [24]; the cosine curve (1868), the Helmert curve (1872), the Ruch curve (1903), the Watorek curve (1907), the Bloss curve (1936), and the sinusoidal curve (1937). Some of these curves were implemented in certain railway tracks, but they never gained the popularity of the clothoid.

The introduction of lateral change of acceleration (LCA) in 1996 [4] showed that with the increasing speeds in rail transport, the clothoid was not as optimal any more (more specifically for speeds above $120\frac{km}{h}$ [29]). This realisation opened the door to research on new types of transition curves, including parabolas [41], sinusoids [42], polynomial curves [43], and spline curves [44], which enjoy smoother LCA than the clothoid.

9.3 State-of-the-art

This section presents the state-of-the-art research on transition curves. The presentation is assessed in two parts. The first one addresses transition curves emerging from railway research, whereas the second part deals with transition curves that origin from research within a broader range of application areas, including the combination of highway- and railway design. Some characteristics of the curves are tabulated below in Table 9.1.

9.3.1 Railway transition research

In the recent research on transition curves in railway, the main topic revolves around finding and analyzing curves that may be suitable as transition curves. Within this topic, one subject which characterizes the work can be identified: evaluation method of curve properties. On one hand, there are simple methods involving analysis of the lateral forces that are utilized to evaluate curve properties, and on the other hand, advanced vehicle-track dynamics are exploited to evaluate

Table 9.1: A summary of the characteristics connected to the presented transition curves.

Curve	Characteristics			
	Type	**Curvature**	**Cover**	**Evaluation**
Cubic parabola [45]	Polynomial	Non-linear	Single	Curvature and jerk smoothness
Sinusoid transition curves [42]	Sinusoid	Non-linear	Single	Curvature and jerk smoothness
Bloss transition curves [40]	Bloss	Non-linear	Single	Software implementation
Wiener Bogen [46]	Curvature based	Non-linear	Single	Curvature and cant smoothness
Parametric transition curves [47]	Polynomial	Linear/non-linear hybrid	Single	Curvature smoothness
Smoothed transition curves [48, 49]	Polynomial	Linear/non-linear hybrid	Single	Horizontal ordinate
Optimized polynomial transition curves [50, 43, 51]	Polynomial	Linear or non-linear	Single	Quality function, boundary condition
General transition curves [52, 53]	Polynomial	Linear or non-linear	Multiple	Suitable LCA
Universal transition curves [54, 55]	Polynomial	Linear or non-linear	Multiple	Suitable LCA
Non-parametric transition curves [56]	Bézier	Linear or non-linear	Multiple	CAD implementation
Optimal layout curves [11]	Line, arc, clothoid	Linear and constant	Multiple	Cost function
Log-aesthetic curves [29]	Log-aesthetic	Linear	Single	Suitable LCA
Symmetrically Projected Transition Curves [10]	Polynomial	Linear	Single	Simplicity, accuracy

the properties of the transition curve.

Simple evaluation methods

In the group that utilizes simple evaluation methods we find work by a number of authors. The research can be divided in two main areas: research on transition curves with non-linear curvature [40, 42, 45, 46], and research on curves possessing a linear curvature property in the middle region combined with smoothed curvatures at the ends [47, 48, 49].

In the first case, with non-linear curvature, the motivation is to improve the smoothness of lateral acceleration and lateral jerk (closely connected to LCA) in the start and end of the transition curve, when compared to linear curves. This is achieved in various ways. For example, in [45] a re-modelling of a nonlinear curvature cubic parabola is explained, where the formulation of the curve relies on a pre-evaluation of the lateral jerk. The nonlinear cubic parabola is given as

$$y(x) = \frac{x^4}{4RL^2} - \frac{x^5}{10RL^2}, \tag{9.3}$$

where R is the radius of the circular segment, and L is the length of the curve. By re-modelling the nonlinear cubic parabola (9.3) and considering a desired lateral jerk diagram, the proposed curve becomes the following expression

$$y(x) = \begin{cases} \frac{4x^5}{15RL^2} & \text{if } 0 \leq x \leq \frac{L}{4}, \\ -\frac{4x^5}{15RL^2} + \frac{2x^4}{3RL^2} - \frac{x^3}{3RL} + \frac{x^2}{12R} - \frac{Lx}{96R} + \frac{L^2}{1920R} & \text{if } \frac{L}{4} < x \leq \frac{3L}{4}, \\ \frac{4x^5}{15RL^2} - \frac{2x^4}{3RL^2} + \frac{8x^3}{3RL} - \frac{13x^2}{6R} + \frac{5Lx}{6R} - \frac{L^2}{8R} & \text{if } \frac{3L}{4} < x \leq L. \end{cases} \tag{9.4}$$

The new transition curve (9.4) removes the jump in lateral jerk, and smooths the transition, while keeping the maximum jerk value within a reasonable limit.

In [42] an investigation of the sinusoid as a transition curve in high speed railways is investigated. The curve formulated as

$$\Theta(l) = \frac{l^2}{2RL} + \left(\frac{L}{4\pi^2 R}\right)\left[\cos\left(\frac{2\pi l}{L}\right) - 1\right] \tag{9.5}$$

yields an S-shaped curvature diagram that removes the jump in lateral jerk. One reason why curves such as (9.4) and (9.5) are not widely used is because they are difficult to tabulate and stake out. Consequently, in the research, formulas are given to obtain coordinates along the sinusoid, based on the Fresnel integrals in (9.1) from the clothoid.

Another interesting nonlinear curvature curve is the Bloss transition curve combined with a new design approach providing additional information based on Network Rail standards and international practice [40]. The Bloss curve (9.6) is already available in railway design software, however, important knowledge about the curve has not been implemented.

$$y(l) = \frac{1}{R}\left(\frac{l^4}{4L^2} - \frac{l^5}{10L^3}\right). \tag{9.6}$$

In this work conditions regarding minimum curve length of the Bloss transition curve is derived with a basis in design rules from the Track Design Handbook [57]. The new formulas are essential towards providing accurate and good design guidance when using Bloss transitions.

Lastly, a new transition curve considering the center of gravity of the vehicle as a criterion is presented [46]. This curve is called the Wiener Bogen and it is defined with a curvature function linked to a cant function and the center of gravity,

$$\kappa(l) = \kappa_1 + (\kappa_2 - \kappa_1)f(l) - h(\psi_2 - \psi_1)\frac{d^2 f}{dl^2},$$

where κ_1 and κ_2 are the curvature at each end, ψ_1 and ψ_2 are the cant values at each end, h is the height to the center of gravity, and $f(l)$ is the cant shaping function (chosen among 6 types, of which the most common is a seventh degree polynomial). The Wiener Bogen is implemented in Austrian railway and has proven to offer great comfort and reduced maintenance costs. Compared to the other nonlinear curvature transition curves the Wiener Bogen is unique in the way that its curvature is S-shaped (like the others), but with additional bends at the ends (which means that the curvature is not monotonically increasing). A visual explanation is given in Figure 9.4. A similar characteristic is not found in any of the other commonly used transition curves.

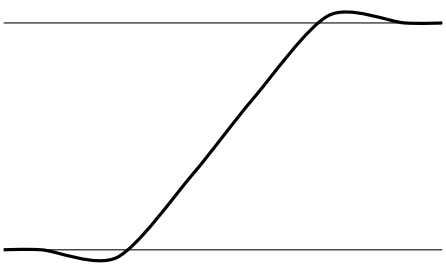


Figure 9.4: A visual explanation of the additional bends in the Wiener Bogen curvature.

All nonlinear curvature transition curves share the property that the curves are always longer than their clothoid counterparts, in order to fulfil the criteria imposed on them. This can sometimes be a disadvantage, especially in renewals of existing railway lines. However, the disadvantage of additional length is compensated by the fact that the LCA of nonlinear curvature curves tends to be of better quality, according to the three criteria give in Section 9.2.1, than the LCA of the clothoid, most notably by guaranteeing a smoother ride.

For the research on linear curvature curves with smoothed ends the motivation is similar to the case of non-linear curvature. However, the linear middle section (from the clothoid) is kept, since this part yields no problem with vehicle dynamics, and only the ends are changed. The research so far is conducted by one author, Koc, starting with a transition curve with linear curvature middle part and smoothed start/end [47], and later moving to a transition curve with linear start/middle part and smoothed end [48, 49]. Both examples are transitions between a straight section and a circular section. In the research of a transition curve with smoothed start and end [47] a new curve is proposed where each region (initial region, middle region, and final region) is treated separately. The purpose of the new curve is to remove the discontinuities (smooth the bends) in the curvature diagram in the initial and final regions, which is present in clothoids, to improve the dynamic properties of the rail. The parametric equations for each region are given as follows (where parameters C and D control the length and shape of the curves, and should be given according to Table 1 in [47]).

Initial region: $l \in [0, Cl_k]$

$$x(l) = l, \tag{9.7}$$

$$y(l) = \frac{A_{11}}{4} l^4 + \frac{A_{12}}{5} l^5, \tag{9.8}$$

where

$$A_{11} = \frac{3 - 3D(1-C) + CD}{6C^2 R l_k^2}, \text{ and } A_{12} = -\frac{1 - D(1-C)}{4C^3 R l_k^3}.$$

Middle region: $l \in [Cl_k, l_k - Cl_k]$

$$\begin{aligned} x(l) =& x(Cl_k) + \cos A_{20}(l - Cl_k) - \frac{1}{2} F \sin A_{20}(l - Cl_k)^2 \\ & - \frac{1}{6}(G \cos A_{20} + 2A_{23} \sin A_{20})(l - Cl_k)^3, \end{aligned} \tag{9.9}$$

$$\begin{aligned} y(l) =& y(Cl_k) + \sin A_{20}(l - Cl_k) + \frac{1}{2} F \cos A_{20}(l - Cl_k)^2 \\ & - \frac{1}{6}(G \sin A_{20} - 2A_{23} \cos A_{20})(l - Cl_k)^3, \end{aligned} \tag{9.10}$$

where $x(Cl_k)$ and $y(Cl_k)$ are from the initial region,

$$\begin{aligned} A_{20} =& A_{21} + A_{22} Cl_k + A_{23}(Cl_k)^2, \\ A_{21} =& \frac{3D - CD - 3}{12R} Cl_k,\ A_{22} = \frac{1 - D}{2R},\ A_{23} = \frac{D}{2Rl_k}, \\ F =& A_{22} + 2A_{23} Cl_k, \text{ and } G = A_{22}^2 + 4A_{22}A_{23} Cl_k + 4A_{23}^2 (Cl_k)^2. \end{aligned}$$

Final region: $l \in [l_k - Cl_k, l_k]$

$$\begin{aligned} x(l) =& x[(1-C)l_k] + \cos A_{30}(l - l_{30}) - \frac{1}{2}L \sin A_{30}(l - l_{30})^2 \\ & - \frac{1}{6}(M \cos A_{30} + N \sin A_{30})(l - l_{30})^3, \end{aligned} \tag{9.11}$$

$$\begin{aligned} y(l) =& y[(1-C)l_k] + \sin A_{30}(l - l_{30}) + \frac{1}{2}L \cos A_{30}(l - l_{30})^2 \\ & - \frac{1}{6}(M \sin A_{30} + N \cos A_{30})(l - l_{30})^3, \end{aligned} \tag{9.12}$$

where $x[(1-C)l_k]$ and $y[(1-C)l_k]$ are from the middle region,

$$\begin{aligned} l_{30} =& (1-C)l_k,\ A_{30} = A_{31} + A_{32}l_{30} + A_{33}l_{30}^2 + A_{34}l_{30}^3 + A_{35}l_{30}^4, \\ A_{31} =& -\frac{3 - 6C + 6C^3 + 9CD - 8C^2D - 3D}{12C^3R}, \\ A_{32} =& \frac{6 - 9C + 6C^3 + 15CD - 12C^2D - 6D}{6C^3R}, \\ A_{33} =& \frac{1}{2}\frac{12 - 12C + 24CD - 16C^2D - 12D}{4C^3Rl_k}, \\ A_{34} =& \frac{1}{3}\frac{6 - 3C + CD - 4C^2D - 6D}{2C^3Rl_k^2}, \\ A_{35} =& \frac{1}{4}\frac{1 + CD - D}{C^3Rl_k^3}, \\ L =& A_{32} + 2A_{33}l_{30} + 3A_{34}l_{30}^2 + 4A_{35}l_{30}^3, \\ N =& 2A_{33} + 6A_{34}l_{30} + 12A_{35}l_{30}^2, \text{ and} \\ M =& 4A_{33}^2l_{30}^2 + 12A_{33}A_{34}l_{30}^3 + 16A_{33}A_{35}l_{30}^4 + 9A_{34}^2l_{30}^4 \\ & + 24A_{34}A_{35}l_{30}^5 + 16A_{35}^2l_{30}^6 + 8A_{32}A_{35}l_{30}^3 \\ & + 6A_{32}A_{34}l_{30}^2 + 4A_{32}A_{33}l_{30} + A_{32}^2. \end{aligned}$$

Koc remarks in his work [47] that there should be given considerations to the occurrence of very small horizontal ordinates in the initial region (in the connection to a straight section) because of practical implementation of the transition curve. His suggestion is to reduce the length of the initial region, and choosing C and D accordingly. The work conducted in [48, 49] further considers horizontal ordinates. This new transition curve is based on the same parametric formulas as previously described ((9.7) - (9.12)), however, the curve is smooth in only one end, the final region, while the remainder of the curve has a linear curvature. The smoothness in the initial region is removed with the argument that it is completely ineffective to implement and maintain in the field. Figure 9.5 shows approximations of the curvature behaviour for a transition curve with smoothed initial and final regions, and a transition curve with only a smoothed final region, compared to the curvatures of a clothoid and a Bloss curve.

Advanced evaluation methods

In the group that uses advanced vehicle-track dynamics the research is performed by Zboiński together with Woźnica or Golofit-Stawinska. The two topics they focus on are forming of polynomial transition curves with a combined use of optimization methods and advanced vehicle models [50, 43, 51], and investigating the non-linear phenomena that occur in the motion of railway vehicles in transition curves with speeds close to the vehicle's critical velocity [58]. The authors identify a change in content of papers concerning transition curves where an attempt to get away from the standard approach of evaluating curve properties and finding new methods is

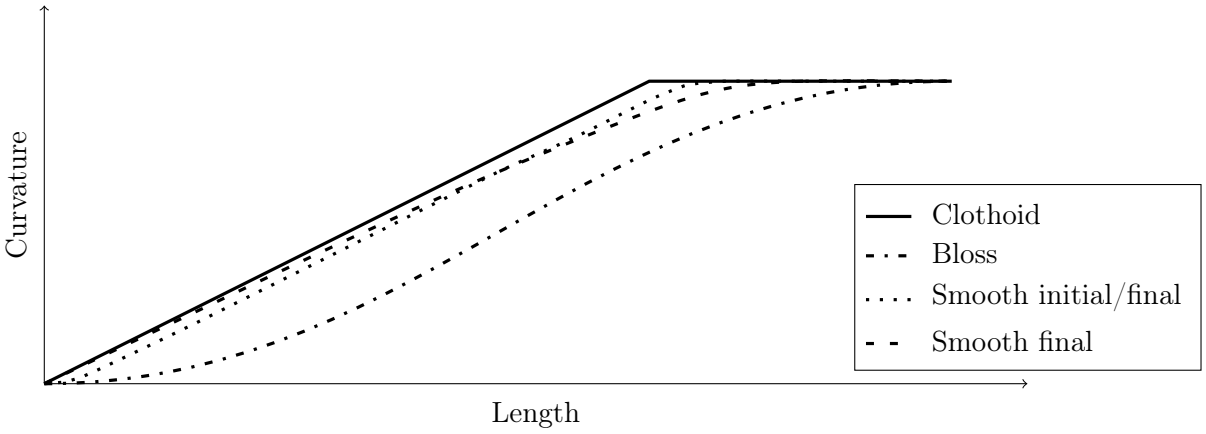


Figure 9.5: Curvature diagram for the clothoid, Bloss curve, Koc's curve with smoothed initial and final regions, and Koc's curve with smoothed final region.

made [51]. Based on this they observe a lack of research where advanced dynamics and optimization methods are exploited to investigate interesting phenomena of railway vehicles in transition curves and form optimal transition shapes. The curves chosen for optimization in [50, 43, 51] are polynomial curves of degree $n \geq 4$, and the objective is to find the polynomial coefficients A_i, that define the curve's shape, in the following expressions (curve, curvature, superelevation, inclination of superelevation ramp, and velocity of wheel vertical rise along the curve)

$$
\begin{aligned}
y(l) &= \frac{1}{R}\left(\frac{A_n l^n}{L^{n-2}} + \frac{A_{n-1} l^{n-1}}{L^{n-3}} + \cdots + \frac{A_4 l^4}{L^2} + \frac{A_3 l^3}{L}\right), \\
\kappa(l) &= \frac{1}{R}\left(n(n-1)\frac{A_n l^{n-2}}{L^{n-2}} + (n-1)(n-2)\frac{A_{n-1} l^{n-3}}{L^{n-3}} + \cdots + 3\cdot 2\frac{A_3 l}{L}\right), \\
h(l) &= H\left(n(n-1)\frac{A_n l^{n-2}}{L^{n-2}} + (n-1)(n-2)\frac{A_{n-1} l^{n-3}}{L^{n-3}} + \cdots + 3\cdot 2\frac{A_3 l}{L}\right), \\
i(l) &= H\left(n(n-1)(n-2)\frac{A_n l^{n-3}}{L^{n-2}} + (n-1)(n-2)(n-3)\frac{A_{n-1} l^{n-4}}{L^{n-3}}\right. \\
&\qquad \left. + \cdots + 4\cdot 3\cdot 2\frac{A_4 l}{L^2} + 3\cdot 2\cdot 1\frac{A_3}{L}\right), \\
f(l) &= Hv\left(n(n-1)(n-2)\frac{A_n l^{n-3}}{L^{n-2}} + (n-1)(n-2)(n-3)\frac{A_{n-1} l^{n-4}}{L^{n-3}}\right. \\
&\qquad \left. + \cdots + 4\cdot 3\cdot 2\frac{A_4 l}{L^2} + 3\cdot 2\cdot 1\frac{A_3}{L}\right),
\end{aligned}
$$

where H is the maximum superelevation, and v is the velocity. The optimization process takes into account a model of vehicle dynamics, track-vehicle and vehicle-passenger dynamical interactions, and a wide set of physical quantities to evaluate the curve properties, which are untypical in transition curve research. The method is still in development and many factors are not yet taken into account, including track irregularities, type of vehicle, and vehicle load. However, the authors state that these factors, and more, are possible to account for in the simulation model. An advanced evaluation method offers novel opportunities in finding new transition curves, and can also reveal unknown phenomena affecting vehicles travelling through a transition curve as presented in [58]. Some results observed in the research are:

- disappearance of vibrations in the transition curve for 25TN bogie of freight cars, despite vibrations in the connected straight and circular segments;

- vibration amplitudes in the transition curve for a bogie of average parameters that exceeds amplitudes in the connected straight and circular segments;
- maximum vibration amplitude in the transition curve for a 4-axle passenger car MKIII, much higher than in the connected segments at some configurations of the suspension parameters.

9.3.2 General transition research

The recent research on transition curves in a general perspective (connected to both railway and highway) focus on two specific topics: replacing multiple segments (and not only the original transition curves), and finding new transition curves with better properties than the clothoid or the cubic parabola.

Replacing multiple segments

The research starts with the introduction of two new families of the general transition curves (curves with smooth and nonsmooth curvature diagrams) [52], that can be used, between two straight sections, as an alternative to the segments: first transition curve - circular arc - second transition curve, and first transition curve - second transition curve. The advantages of the new curves are stated to be the capability of describing the original segmented curve with only one equation, that the curve ensures continuous change of curvature along the entire length, and that the LCA is of better quality in the new curves compared to the clothoid. Both families are polynomial transition curves based on the function

$$y(x) = \sum_{i=0}^{n} a_i x^i, \tag{9.13}$$

between the start point P and the end point K. For the family with nonsmooth curvature diagrams the polynomial in (9.13) is of degree $n = 4$ and gives the following expression

$$y(x) = x_K(N_1 \tan u_P + N_2 \tan u_K),$$

where

$$N_1 = t - t^3 + \frac{1}{2}t^4,$$

$$N_2 = t^3 - \frac{1}{2}t^4,$$

and $t = \frac{x}{x_K}$, $t \in [0, 1]$. For the family with smooth curvature diagrams the polynomial in (9.13) is of degree $n = 6$ and gives the following expression

$$y(x) = x_K(F_1 \tan u_P + F_2 \tan u_K),$$

where

$$F_1 = t - \frac{5}{2}t^4 + 3t^5 - t^6,$$

$$F_2 = \frac{5}{2}t^4 - 3t^5 + t^6,$$

and $t = \frac{x}{x_K}$, $t \in [0, 1]$.

With a basis in the previous work a number of papers have emerged. In [54] a generalization of the general transition curves was presented as universal transition curves, UTC-1 and UTC-2, arising from expression (9.13) for $n = 4$ and $n = 5$, respectively. These curves enjoy the same advantages as the general transition curves, and also offers the possibility to create curves

beyond the U-shape with straight start and end, e.g., S-shaped transition curves which are useful for routing of reverse curves [55], resulting from expression (9.13) for $n = 5$ and $n = 7$. Another example is the work done in [56] where Kobryń's work [6, 52, 54] is represented in nonparametric Bézier and nonparametric B-spline forms, making them suitable to implement into commercial software. The main goal of the work is to show that the polynomial transition curves, presented on power form, admit a straight forward representation on Bézier form as a linear combination of degree-n Bernstein polynomials ($B_k^n(t)$) and $n+1$ control points ($\{b_k\}_{k=0}^n$), see formula (9.14).

$$b(t) = \sum_{k=0}^{n} b_k B_k^n(t), \quad B_k^n(t) = \binom{n}{k}(1-t)^{n-k}t^k, \quad t \in [0,1]. \tag{9.14}$$

With geometric constraints on the control points $b_k = \{x_k, y_k\}$ the Bézier curve coincides with the graph of a degree-n polynomial function (9.13) in Bernstein basis

$$y(x) = \sum_{k=0}^{n} y_k B_k^n(t), \quad t = \frac{x}{x_n} \in [0,1]. \tag{9.15}$$

In this representation few formulas describe all possible cases, and end conditions are constrained by the arrangement of the control points.

Another example of recent work on the same topic is [53] where flexible polynomial transition curves are discussed together with formulas for setting them out in the field. The two polynomial curves are the curve defined in [59] with a smooth curvature diagram

$$y(x) = \frac{x_K \tan u_P}{C}\left(Ct + \frac{2-5C}{2}t^4 - \frac{7-15C}{5}t^5 + \frac{1-2C}{2}t^6\right), \tag{9.16}$$

and the curve defined in [60] with a nonsmooth curvature diagram

$$y(x) = \frac{x_K \tan u_P}{C}\left(Ct + \frac{1-3C}{3}t^3 - \frac{1-2C}{4}t^4\right), \tag{9.17}$$

where $t = \frac{x}{x_K}$ for $x \in [0, x_K]$. The flexibility of (9.16) and (9.17) comes from the variable

$$C = \frac{R_K \tan u_P}{x_K},$$

and its permitted value options that shape the curves. In addition to the polynomial equations, formulas connected to setting out the curves in the field are defined. Among these are formulas for intermediate points $\{x', y'\}$ along the curve based on the tangent line drawn in the start point P,

$$x' = x\cos u_P + y\sin u_P, \tag{9.18}$$

and

$$y' = x\sin u_P - y\cos u_P, \tag{9.19}$$

with $x = tx_K$ for $t \in [0,1]$ and y is determined from curve (9.16) or (9.17).

Lastly, the work in [11] aims to give a simple and general formulation of an optimization problem to create the horizontal alignment over a longer stretch, where every geographical point has a price. The new curve is composed of straight sections and circular arcs connected with clothoids, given by vertices (x_i, y_i), radii (R_i), and angles (ω_i), together with start and end points

$$\boldsymbol{x}^N = (x_1, y_1, R_1, \omega_1, \ldots, x_N, y_N, R_N, \omega_N), \quad N \in \mathbb{N}.$$

In addition, a cost function is given, which brings together all existing costs of passing through a point (x, y) in the domain. The goal is to minimise the cost function for each $N = 1, 2, \ldots$, and choose the curve with the lowest value. The cost function is described as a general function that can represent various parameters, e.g. building costs, environmental restrictions, political restrictions, or terrain constraints. To test the optimization model, a section of the road joining Larraga-Lerín (NA601) in Navarra, Spain, has been considered. This particular road is of interest because its road alignment was improved through a reconstruction project. The results of NA601 compared to the curves from the optimization problem show that the curve with $N = 3$ is similar to NA601, which makes the authors believe that their proposed method can be a good way to improve alignments in road construction projects.

New transition curves

In [29] the family of log-aesthetic curves are examined for implementation as transition curves with a focus on revealing if they are suitable according to given criteria, and to analyse if they have a better LCA when compared to clothoids. This is the first time these curves have been studied in terms of road and railway dynamics (apart from the clothoid which is part of the log-aesthetic family), and this in itself makes them interesting. The basic formula of aesthetic curves is obtained by assuming that the logarithmic curvature histogram (LCH) is a linear function [61]

$$\log\left(\rho\frac{ds}{d\rho}\right) = \alpha\log\rho + C, \tag{9.20}$$

where ρ is the curvature, C is a constant, and α is a slope parameter in the LCH. In the study, three log-aesthetic curves with α in (9.20) set to 1, 2, and 3 are analysed in regards to smoothness criteria and LCA, and then compared to the clothoid. The results show that the log-aesthetic curves have similar dynamic properties as the clothoid, making them suitable as transition curves on the same level as clothoids.

A different approach is taken in [10] where a curve with the simplicity of the cubic parabola and the accuracy of the clothoid is introduced, the Symmetrically Projected Transition Curve (SPTC). The formula of a SPTC is as follows

$$y(x) = \frac{x^3}{2A^2}\left(\frac{1}{3} + \frac{1}{14}(\frac{x^2}{2A^2})^2 + \frac{3}{88}(\frac{x^2}{2A^2})^4 + \frac{1}{48}(\frac{x^2}{2A^2})^6 + \ldots\right), \tag{9.21}$$

with $A^2 = RX$, where R is the radius at the end of the curve, and X is the curve's projection onto the x-axis. Formula (9.21) gives the intermediate points of the transition curve along the projection x. Via only considering the first term in the expression, a cubic parabola is obtained. Thus, the cubic parabola can be thought of as a first approximation of the SPTC. In the research the SPTC is compared to the clothoid and the cubic parabola, not based on vehicle-track dynamics, but rather design simplicity and accuracy, and it is thought to be a better option in cases where the cubic parabola is preferred (over the clothoid) as a transition curve.

In addition to the previous work a state-of-the-art paper has been written with focus on line to circle spiral transition curves in railway, highway, robotics, and computer imaging [62]. Challenges, in railway, highway, and robotics, are noted there in the design of curves that fulfil smoothness criteria while also account for limitations in curvature, speed, and obstacles, and consider energy and time efficiency. In computer imaging challenges are connected to unwanted singularities in image interpolation when using C^1 or C^2 continuous splines. Other research challenges discussed in the same study are Hermite end conditions and shape control, and arc-length parameterization.

9.4 Challenges in railway transition curve research

Based on the state-of-the-art in Section 9.3 a number of research challenges can be identified. They are connected to evaluation criteria, flexibility of the new curve, and linking research and industry.

9.4.1 Evaluation criteria

From the state-of-the-art it can be observed that the evaluation criteria for evaluating new transition curves consist of both simple and more advanced methods. Analysis of curvature, lateral acceleration, lateral jerk, and LCA are common approaches, and in recent research more advanced methods using vehicle-track dynamics have emerged. Although LCA is categorized as a simple method in this paper, it is more advanced than methods comparing only geometric properties of the curve, e.g. curvature. The LCA takes into account parameters from the track and the train, in addition to curve properties. However, since the method considers a railway wagon as a particle, and not a system, on the curve, it is labelled as simple.

The possibilities in evaluating curve properties in relation to its suitability as a transition curve makes it difficult to decide on an approach that is "good enough". Is an appropriate and smooth LCA enough to characterize a curve, or is a more advanced method that takes into consideration the whole system of the vehicle chain and track preferable? The question is an open one. However, it can be argued that as a first and important assessment of a new transition curve, using a simple method, e.g. LCA, is adequate, while a more advanced method may strengthen the argumentation and bring new information regarding the curve.

A strong reason for using an advanced vehicle-track model is presented in [51]. The authors state there that some dynamical effects observed in complete vehicle models are not present in models represented by a particle. Unexpected behaviour of various railway vehicles is also observed in [58] with advanced vehicle model simulations. In that study differences can even be seen between the leading and trailing bogies when they are compared against each other for one railway car. Another reason is the knowledge that one single railway wagon most likely acts differently through a curve than a chain of wagons.

9.4.2 Flexibility of the new curve

There are two relevant areas to consider in regards to implementing transition curves: implementing transition curves in new railway lines, and implementing transition curves in the renewal of existing railway lines. Building new railway lines implies that constrictions on the rail alignment are limited to geography and terrain challenges, which gives a larger freedom to improve the alignment when compared to renewal of old lines, where additional obstacles present in the old line must be taken into account. This means that a new transition curve has to be flexible enough to handle both of these scenarios. From the state-of-the-art research it was observed that some of the work included new transition curves that could replace either a single segment (such as the old transition curve), e.g. [10, 29, 46, 47], or multiple segments (such as an entire turn or even more), e.g. [52, 54, 55].

A transition curve which replaces multiple segments may be more flexible when it comes to finding an optimal shape, but at the same time less flexible when it comes to avoiding obstacles, since a larger area of the track is involved. The opposite will be the case for transition curves replacing one segment; more flexible in avoiding obstacles, since a smaller area is involved, and less flexible in finding an optimal shape.

9.4.3 Linking research and industry

The link between railway transition curve research and the industry that has to implement and maintain them is sparsely covered in recent literature. Properties of a curve that are beneficial in vehicle dynamics may not be as beneficial to implement or maintain. This can be observed in the work of Koc [48, 49] where the author argues that what is considered to be an optimal curve in relation to LCA is in practice often impossible to implement and maintain because of the very low values of horizontal ordinates (and ordinates of the superelevation) in the region connected to a straight line. Koc states that implementing smooth curves in practice leads to a shortening of the transition curve, with an extension of the straight line. The conclusion is thus a recommendation to give up the condition of a smooth curvature in the end connected to a straight line, and rather keep conditions similar to those of the clothoid.

A common trend in recent papers on transition curves is to evaluate their suitability based on curve properties through analysis and simulations. Few full-scale tests, where the curve is implemented and analysed in a real world setting, have been performed. One exception is the Wiener Bogen transition curve which is implemented in Austrian railway with successful outcomes [9]. Exploring the link between simulations and full-scale experiments is relevant and important in order to fully understand the characteristics of a possible new transition curve.

Another property that should be considered is which form the curve is given on. A form that can be directly implemented into existing commercial railway software is more preferable than forms that are hard to implement. This is recognised as important in [56], where a non-parametric Bézier form (9.15) of the polynomial transition curves given on power form in [6, 52, 54] is presented. The Bézier form is commonly used in computer-aided geometric design (CAGD) and computer graphics because of its superior geometric and numerical properties [63, 64] when compared to the power form. It is also simpler to enforce continuity conditions between segments for curves on Bézier form. A different approach is taken in [53] where appropriate formulas are derived, based on the curves in (9.16) and (9.17), enabling calculation of data required for setting out the transition curves in the field. The formulas include tangents, normals, and coordinates, in addition to expressions ((9.18) and (9.19)) for calculating intermediate points $\{x', y'\}$ along the curves. Formulas for finding points along the transition curve is also presented in [42] for sinusoids. The method uses Fresnel integrals to obtain coordinates because the original expression (9.5) is difficult to tabulate.

9.5 Future research opportunities

There are a number of interesting research opportunities connected to railway transition curves. Most of them are related to the state-of-the-art in Section 9.3 and the research challenges in Section 9.4. Here some of them are highlighted and discussed based on the authors' interests.

The problem addressed in [48, 49] shows that there is a need for linking research and industry more closely. Finding curves that have beneficial properties in regards to vehicle dynamics while at the same time are possible to implement and maintain in a sustainable way can change the railway industry for the better. Future work in this area will have to include finding new evaluation criteria for the curves, from an industrial point of view, that can complement the vehicle dynamics criteria and add guidelines connected to implementation and maintenance.

The research seems to be fairly well balanced between addressing transition curves which replaces one rail segment and those dealing with multiple rail segments. However, the advantages of using one approach over the other, in existing railways, is not well documented. In theory, replacing one segment would require less work and provide more flexibility in avoiding of obstacles

than what would be the case of replacing multiple segments. In practice the argumentation is not as straight forward. Replacing one segment in an already existing railway means that parts of the adjacent curves have to be shortened or extended, since the new transition curve is either longer or shorter. If the new curve, with enhanced properties, has the same length as the old segment, at least one of the adjacent segments have to be moved in order to maintain a smooth transition. This means that multiple segments will be affected either way. Research regarding the two methods with a focus on cost weighed against intervention to nature and existing railway, as well as the obtained improvements to the alignment, is a possible future work opportunity.

Although transition curve research is overflowing with families of curves suited for transition purposes, it is always interesting to analyse new curves in a transition perspective. Blending splines [65, 66, 67] are constructions where local functions at the knots are blended together by C^k-smooth basis functions, see Figure 9.6. They are defined in [68] as

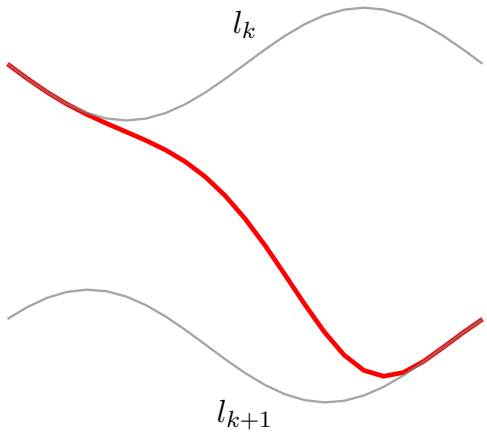


Figure 9.6: Two local curves (l_k and l_{k+1}) blended together to create a blending spline (red curve).

$$f(t) = \sum_{k=1}^{n} l_k(t) B_k(t), \;\; t \in (t_1, t_n],$$

where $l_k(t)$ are scalar-, vector-, or point-valued local functions defined on (t_{k-1}, t_{k+1}), $t = \{t_k\}_{k=0}^{n+1}$ is an increasing knot vector, and $B_k(t)$ are the blending functions (B-functions). Blending splines may be suitable as transition curves because of their flexibility in the blending process, connected to the choice of local functions, B-function, and number of knot intervals. The curves can provide a decided degree of smoothness (which is linked to the B-function and local curves) in the knots and can be used as a replacement for one or several segments. Some initial experiments have been conducted, by the authors of this paper, which shows promising results regarding curvature smoothness and the flexibility of the blending spline. In Figure 9.7 the curvature of a blending spline constructed by blending two Bézier curves with an S-shaped blending function can be seen. The curve is placed between two circular segments, of different radii, in an existing railway, as a substitute for a clothoid. The analyses of the curvature in this blending spline transition curve show that the curvature smoothness is G^1 at the points of connection to the adjacent circular segments, where the clothoid is G^0. From the initial experiments it was observed that the curvature smoothness can be increased to G^2 and higher by choosing appropriate local curves, for instance by using arcs, or Bézier curves of degrees higher than 3. The research is still at an early stage, and although the end points are smooth, there remains to be put more efforts into attempts for smoothing out the overall shape of the blending spline curvature.

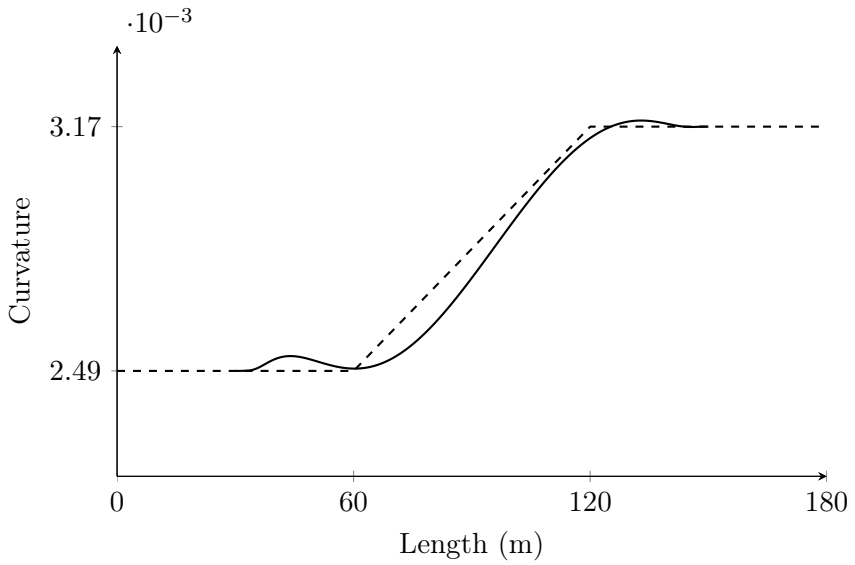


Figure 9.7: Curvature plots of a clothoid with parts of connected circular segments (dashed) and a blending spline, with local Bézier curves (solid).

Another example of interesting curves are circle splines [69, 70]. They are created from a blending of two circular arcs based on point positions and angles [70] as follows. Given a sequence of interpolatory constraint points $P_0, P_i, \ldots, P_i, \ldots, P_n$, a circle spline is created by blending two circular arcs (arc_i and arc_{i+1}) for every segment (P_i, P_{i+1}), see Figure 9.8. Each of the two arcs

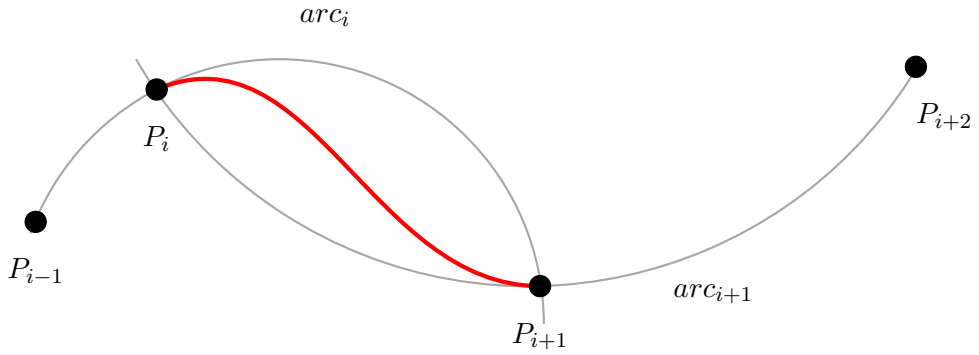


Figure 9.8: Two arcs blended together to obtain a circle spline (shown as a red curve).

are defined to go through three points, $arc_i \to P_{i-1}, P_i, P_{i+1}$, and $arc_{i+1} \to P_i, P_{i+1}, P_{i+2}$. From the arcs the tangent vectors t_i and t_{i+1} and the curvature of the composite curve in the points P_i and P_{i+1} are obtained. In the blending process, as the point $P(u)$ travels across an arc, $arc(u)$ ($u \in [0, 1]$) from P_i to P_{i+1}, the arc changes smoothly from arc_i to arc_{i+1}. The intermediate $arc(u)$ is defined by P_i and P_{i+1} as well as the tangent $t(u)$ at P_i. To parameterize the changing process of the arcs, the directional angle

$$\tau(u) = \tau_i \cos^2\left(u\frac{\pi}{2}\right) + \tau_{i+1} \sin^2\left(u\frac{\pi}{2}\right),$$

created by blending τ_i and τ_{i+1} given by t_i and t_{i+1}, is used.

A trigonometric blending function is utilized to obtain G^2 continuity. The result is a robustly produced fair-looking G^2-continuous curve without cusps or kinks. Connections between straight lines and circular arcs are also possible, where the line is considered as an arc with infinite radius.

The circle spline is interesting in this setting mainly because of the way it is constructed, with arcs, and because of its smoothness.

9.6 Conclusion

In this paper, the state-of-the-art of transition curve in railway has been explored and presented, including a discussion on research challenges and future research opportunities. The paper covers essential information concerning where the research on railway transition curves is standing today. It is meant to be a source which can bring better understanding of present research challenges and inspiration for further work within the field. Transition curves in horizontal railway alignment is an important subject in relation to increase passengers' safety and reduce the need for maintenance, in particular when high speeds are involved. Geometrical properties of the curve together with vehicle dynamics can form a strong basis for the suitability of a curve as a transition curve in railway, both for the case of designing new railways, as well as for replacing one or several segments in an existing railway.

The research on the topic is active and it is far from over. New ideas and approaches emerge in line with the developments in vehicle dynamics and the railway industry. Various branches have formed where the scientists employ different methods and criteria to solve problems in the field. Challenges in ongoing research can be observed in areas connected to evaluation criteria for a new curve, flexibility of the curve, and linking the research to the railway industry. Based on these challenges and the state-of-the-art research, future research opportunities are identified to include finding new evaluation criteria connected to the industry, research regarding advantages of replacing one or more rail segments, and investigating new families of curves, exemplified by blending splines and circle splines, as possible new transition curves.

References

[1] Thomas F. Hickerson. *Route Location And Design.* McGraw-Hill, 5 edition, 1964.

[2] Martin Lipicnik. New form of road/railway transition curve. *Journal of Transportation Engineering*, 124(6):546–556, 1998.

[3] Coenraad Esveld. *Modern Railway Track.* MRT-Productions, Delft University of Technology, Netherlands, 2 edition, 2001.

[4] Orhan Baykal. Concept of lateral change of acceleration. *Journal of Surveying Engineering*, 122:132–141, 08 1996.

[5] Ergin Tari and Orhan Baykal. A new transition curve with enhanced properties. *Canadian Journal of Civil Engineering*, 32(5):913–923, 2005.

[6] Andrzej Kobryn. Polynomial solutions of transition curves. *Journal of Surveying Engineering*, 137(3):71–80, 2011.

[7] Krzysztof Zboinski and Piotr Woznica. Optimisation of the railway transition curves' shape with use of vehicle-track dynamical model. *Archives of Transport*, 22:387–407, 01 2010.

[8] Orhan Baykal, Ergin Tari, Zeki Coskun, and Muhammed Sahin. New transition curve joining two straight lines. *Journal of Transportation Engineering*, 123(5):337–345, 1997.

[9] H. Hasslinger and Stockinger H. Measurement proof of the superiority of a new track alignment design element, the "viennese curve". *ZEVrail*, pages 61–71, February 2005.

[10] N. Eliou and G Kaliabetsos. A new, simple and accurate transition curve type, for use in road and railway alignment design. *European Transport Research Review*, 6:171–179, June 2014.

[11] Gerardo Casal Urcera, Duarte Santamarina, and Miguel Vazquez-Mendez. Optimization of horizontal alignment geometry in road design and reconstruction. *Transportation Research Part C: Emerging Technologies*, 74, 01 2017.

[12] Shiqiao Tian, Xiangping Luo, Lihui Ren, and Chunyu Xiao. Active steering control strategy for rail vehicle based on minimum wear number. *Vehicle System Dynamics*, pages 1–26, 2020.

[13] Seyed Milad Mousavi-Bideleh and Viktor Berbyuk. Multiobjective optimisation of bogie suspension to boost speed on curves. *Vehicle System Dynamics*, 54(1):58–85, 2016.

[14] F. Hassan, A. C. Zolotas, and R. M. Margetts. Optimised pid control for tilting trains. *Systems Science & Control Engineering*, 5(1):25–41, 2017.

[15] Antonio Enrique Blanco-Saura, José Luis Velarte-González, Fran Ribes-Llario, and Julia Irene Real-Herráiz. Study of the dynamic vehicle-track interaction in a railway turnout. *Multibody System Dynamics*, 43(1):21–36, 2018.

[16] S. Bruni, J. P. Meijaard, G. Rill, and A. L. Schwab. State-of-the-art and challenges of railway and road vehicle dynamics with multibody dynamics approaches. *Multibody System Dynamics*, 49(1):1–32, 2020.

[17] Chunsheng Li, Shihui Luo, Colin Cole, and Maksym Spiryagin. An overview: modern techniques for railway vehicle on-board health monitoring systems. *Vehicle System Dynamics*, 55(7):1045–1070, 2017.

[18] David Lebel, Christian Soize, Christine Funfschilling, and Guillaume Perrin. High-speed train suspension health monitoring using computational dynamics and acceleration measurements. *Vehicle System Dynamics*, 58(6):911–932, 2020.

[19] Sönke Kraft, Julien Causse, and Aurélie Martinez. Black-box modelling of nonlinear railway vehicle dynamics for track geometry assessment using neural networks. *Vehicle System Dynamics*, 57(9):1241–1270, 2019.

[20] Han Wu, Xiao-Hui Zeng, Jiang Lai, and Yang Yu. Nonlinear hunting stability of high-speed railway vehicle on a curved track under steady aerodynamic load. *Vehicle System Dynamics*, 58(2):175–197, 2020.

[21] Wenlin Wang, Zirong Zhou, Weihua Zhang, and Simon Iwnicki. A new nonlinear displacement-dependent parametric model of a high-speed rail pantograph hydraulic damper. *Vehicle System Dynamics*, 58(2):272–289, 2020.

[22] Andrzej Kobryn. *Transition Curves for Highway Geometric Design*, volume 14. Springer International Publishing, 1 edition, 02 2017.

[23] J.S. Mundrey. *Railway Track Engineering, Fourth Edition.* Tata McGraw-Hill, 1988.

[24] Björn Kufver. *Optimisation of horizontal alignments for railways : Procedures involving evaluation of dynamic vehicle response.* PhD thesis, Department of Vehicle Engineering, Royal Institute of Technology, Stockholm, Sweden, 2000.

[25] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables.* Dover Publications, Inc., New York, USA, 1972.

[26] Raph Levien. The euler spiral: a mathematical history. Technical report, University of California, Berkeley, 2008.

[27] Arthur Talbot. The railway transition spiral. Reprinted from The Technograph No. 13, 1899. U. Illinois.

[28] Ergin Tari. *New curve trends in alignment design.* PhD thesis, Istanbul Technical University, 1997.

[29] Abdullah Arslan, Ergin Tari, Rushan Ziatdinov, and Rifkat Nabiyev. Transition curve modeling with kinematical properties: Research on log-aesthetic curves. *Computer-Aided Design and Applications*, 11:509–517, 10 2014.

[30] W. (Wilfred) Schofield. *Engineering surveying : theory and examination problems for students.* Oxford, [England] ; Boston : Butterworth-Heinemann, 5th ed edition, 2001.

[31] Johan Förstberg. *Ride comfort and motion sickness in tilting trains*. PhD thesis, Department of Vehicle Engineering, Royal Institute of Technology, Stockholm, Sweden, 2000.

[32] N. Yayla. Highway engineering. Lecture Book, Istanbul, 2009.

[33] John Uren and Bill Price. *Surveying for Engineers*. Macmillan International Higher Education, 5 edition, 2010.

[34] J. Glover. Transition curves for railways. *Proc. Inst. C. E.*, 140, 1900.

[35] E. M. Horsburgh. Xxiv.—the railway transition curve. *Proceedings of the Royal Society of Edinburgh*, 32:333–347, 1913.

[36] A. W. Miller. The transition spiral. *Australian Surveyor*, 7(8):518–526, 1939.

[37] J. Nam. *How Railway Systems Work*. AuthorHouse UK, 2014.

[38] Institution of Engineers and Shipbuilders in Scotland. *Transactions of the Institution of Engineers and Shipbuilders in Scotland*. Number v. 4. 1861.

[39] Arthur Lovat Higgins. *The transition spiral and its introduction to railway curves with field exercises in construction and alignment*. Van Nostrand company, New York, 1922.

[40] Constantin Ciobanu. Bloss transition - a short design guide. *PWI Journal*, 133:14–18, 01 2015.

[41] Ergin Tari. The new generation transition curves. *ARI The Bulletin of the Instanbul Technical University*, 54:35–41, 01 2004.

[42] A. Pirti, M.A. Yücel, and T. Ocalan. Transrapid and the transition curve as sinusoid. In *Tehnicki vjesnik*, volume 23, pages 315–320, 2016.

[43] K. Zboinski and P. Woznica. Optimization of polynomial transition curves from the viewpoint of jerk value. *Archives of Civil Engineering*, 63, 01 2017.

[44] Azhar Ahmad and Jamaludin Ali. G3 transition curve between two straight lines. pages 154–159, 08 2008.

[45] SA Shebl. Geometrical analysis of non-linear curvature transition curves of high speed railways. *Asian Journal of Current Engineering and Maths*, 5(4):52–58, Aug 2016.

[46] Robert Wojtczak. Charakterystyka krzywej przejściowej wiener bogen®. online. Publishing House of Poznan University of Technology, 2017.

[47] Wladyslaw Koc. Transition curve with smoothed curvature at its ends for railway roads. *Current Journal of Applied Science and Technology*, 22:1–10, 07 2017.

[48] W. Koc. Smoothed transition curve for railways. *Przeglad Komunikacyjny*, (7):19–32, 2019.

[49] W Koc. New transition curve adapted to railway operational requirements. *Journal of Surveying Engineering*, 145, 08 2019.

[50] Krzysztof Zboinski and Piotr Woznica. Optimisation of polynomial railway transition curves of even degrees. *Archives of Transport*, 35:71–86, 01 2015.

[51] Krzysztof Zboinski and Piotr Woznica. Combined use of dynamical simulation and optimisation to form railway transition curves. *Vehicle System Dynamics*, 56(9):1394–1450, 2018.

[52] Andrzej Kobryn. New solutions for general transition curves. *Journal of Surveying Engineering*, 140:12–21, 02 2014.

[53] Andrzej Kobryn. Use of polynomial transition curves in the design of horizontal arcs. *Roads and Bridges - Drogi i Mosty*, 16(1):5–14, 2017.

[54] Andrzej Kobryn. Universal solutions of transition curves. *Journal of Surveying Engineering*, 142:04016010, 02 2016.

[55] Andrzej Kobryn and Piotr Stachera. S-shaped transition curves as an element of reverse curves in road design. *The Baltic Journal of Road and Bridge Engineering*, 14:484–503, 11 2019.

[56] J. Sanchez-Reyes and J.M. Chacon. Nonparametric bezier representation of polynomial transition curves. *Journal of Surveying Engineering*, 144, 05 2018.

[57] NR/L2/TRK/2049. *Track Design Handbook*. Network Rail., 2010.

[58] Krzysztof Zboinski and Milena Golofit-Stawinska. Investigation into nonlinear phenomena for various railway vehicles in transition curves at velocities close to critical one. *Nonlinear Dynamics*, pages 1–47, 10 2019.

[59] R. J. Grabowski. Gładkie przejścia krzywoliniowe w drogach kołowych i kolejowych. *Zeszyty Naukowe Akademii Górniczo-Hutniczej*, (82), 1984.

[60] A. Kobryń. Wielomianowe krzywe przejoeciowe w projektowaniu niwelety tras drogowych. *Wydawnictwa Politechniki Białostockiej, Rozprawy Naukowe nr 100*, 2002.

[61] Kenjiro Miura, Junji Sone, Atsushi Yamashita, and Toru Kaneko. Derivation of a general formula of aesthetic curves. *Humans and computers*, pages 166–171, 2005.

[62] Akin Levent, Bayram Sahin, and Zulfiqar Habib. Spiral transitions. *Applied Mathematics-A Journal of Chinese Universities*, 33:468–490, 12 2018.

[63] Gerald Farin. *Curves and surfaces for CAGD: A practical guide*. Morgan Kaufmann Publishers, 5 edition, 2002.

[64] Rida T. Farouki. The bernstein polynomial basis: A centennial retrospective. *Computer Aided Geometric Design*, 29(6):379 – 419, 2012.

[65] L. T. Dechevsky. Expo-rational B-splines. Communicated at the Sixth International Conference on Mathematical Methods for Curves and Surfaces, Tromsø, Norway, 2004.

[66] Lubomir T. Dechevsky, Børre Bang, and Arne Lakså. Generalized expo-rational b-splines. *International Journal of Pure and Applied Mathematics*, 57(6):833–872, 2009.

[67] Lubomir T. Dechevsky and Peter Zanaty. Smooth gerbs, orthogonal systems and energy minimization. *AIP Conference Proceedings*, 1570(1):135–162, 2013.

[68] Arne Lakså. *Basic properties of Expo-Rational B-splines and practical use in Computer Aided Geometric Design*. PhD thesis, University of Oslo, 2007.

[69] Carlo Sequin and Kiha Lee. Fair and robust circle splines. In *SIGGRAPH 2003 Conference on Sketches and Applications: in conjunction with the 30th annual conference on Computer graphics and interactive techniques*, January 2003.

[70] Carlo H. Sequin, Kiha Lee, and Jane Yen. Fair, g2- and c2-continuous circle splines for the interpolation of sparse data points. *Computer-Aided Design*, 37(2):201 – 211, 2005.

10 Preliminary studies on transition curve geometry: Reality and virtual reality

Tanita Fossli Brustad

Abstract

The topic of transition curves is well known in track and road engineering. Over the years the need for easement between straight and curved sections in railway and road design has grown, with an important reason being increase in vehicle speed. Testing of transitions is often done with graph analysis where the property of lateral change of acceleration is compared between varieties of curves. However, graph analysis does not give a clear understanding of the behaviour of a vehicle in a transition phase the way, for instance, field or laboratory experiments might do. In this paper we present an initial study on the behaviour of a down-scaled vehicle model driving through three distinct transitions from straight to curved sections. In addition we show a comparison of the physical model against a virtual model.

10.1 Introduction

Horizontal geometry has been, and is, a popular topic in railway and road design. One focus area is the transition curve and the properties this curve should have. Transition curves are used as easement between straight sections and curved section in track and road engineering [1, 2]. Over the years different criteria have laid the foundation when deciding the type of curve. In the beginning no transitions were needed because of low speed and wide-radius curves, however during the 19th century when the vehicle speed increased the need for easement arose and the clothoid (also referred to as Euler spiral or Cornu spiral) became the most popular [3]. The clothoid is still the most used transition curve today, but with the ever increasing speeds and with the introduction of lateral change of acceleration (LCA) [4] as the most important criterion in vehicle-road dynamics [5], other curves have been recommended in order to increase safety and comfort, and to reduce the maintenance need. The way of testing the properties of a transition curve nowadays is mostly done by considering a graph of the LCA function, and comparing it to the LCA of the clothoid, as in [6, 7]. Field and laboratory testing on transition geometry are not very common in research, with a prominent reason being cost and time, but also because transition research has moved in a direction where a graph of the LCA function is enough information to categorize the quality of the curve. The motivation behind this paper is to increase the understanding of the effect of transition curves by studying the lateral forces working on the vehicle, and by observing the path traced by the vehicle for three different transitions from straight sections to curved sections, in a laboratory study setting. As an initial study and to easily be able to extract path data, a down scaled vehicle model will be used in the laboratory experiments. In addition, a comparison of the physical vehicle model against a virtual model will be given in order to increase the validity of the results and to test if simple VR technology resembles the physical model in behaviour.

In this paper we look at a simple physical vehicle model and examine what observable effects geometry has on a vehicle in a transition phase (going from a straight road into a turn). In addition we create a virtual model of the vehicle to investigate if existing straightforward desktop-based VR technology [8] can give a satisfactory result compared to the physical model when it comes to the behaviour of the vehicle.

10.2 Preliminaries

This section gives a brief description on; which transition geometry we considered for the experiments, the technology to create the physical vehicle model (Lego Mindstorms), and the software used to create the virtual vehicle model (Blender).

10.2.1 Transition geometry

There are many different curves suitable as transition curves. In railways, highways, and robot trajectory design it is desirable to have curves with monotone curvature [9]. As mentioned in the introduction the most commonly used in practice (especially in railway) is the clothoid. A clothoid curve (of length l and end radius r) is defined parametrically, with Fresnel integrals, as

$$x = C(t) = \frac{1}{a}\int_0^{\hat{t}} \cos(\frac{\pi}{2}u^2)du,$$
$$y = S(t) = \frac{1}{a}\int_0^{\hat{t}} \sin(\frac{\pi}{2}u^2)du.$$

Where $a = \sqrt{\frac{1}{\pi r l}}$ is a scaling factor, and $\hat{t} = at,\ -\infty < t < \infty$, see Figure 10.1.

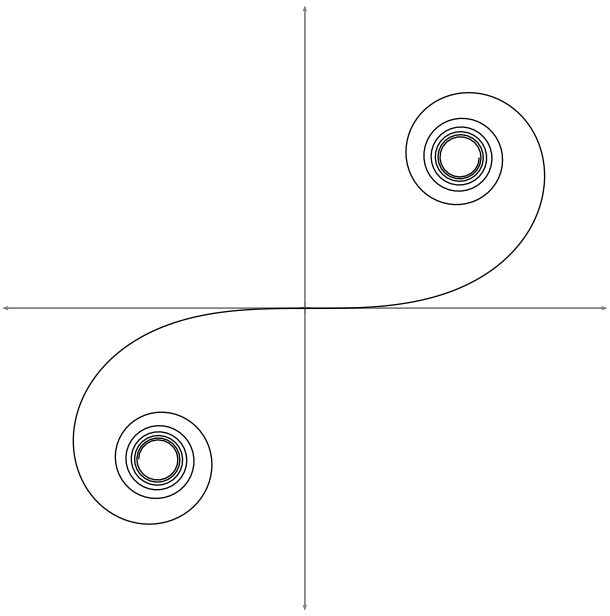


Figure 10.1: The clothoid spiral.

This particular curve is preferred in railway tracks because of some desired properties it possesses: the curvature changes linearly with curve length, and the curve is easy to scale, in relation to curve length and the radii in both ends. An example of the clothoid as a transition between a straight and a circular segment can be seen in Figure 10.2. Taking into account the LCA-function,

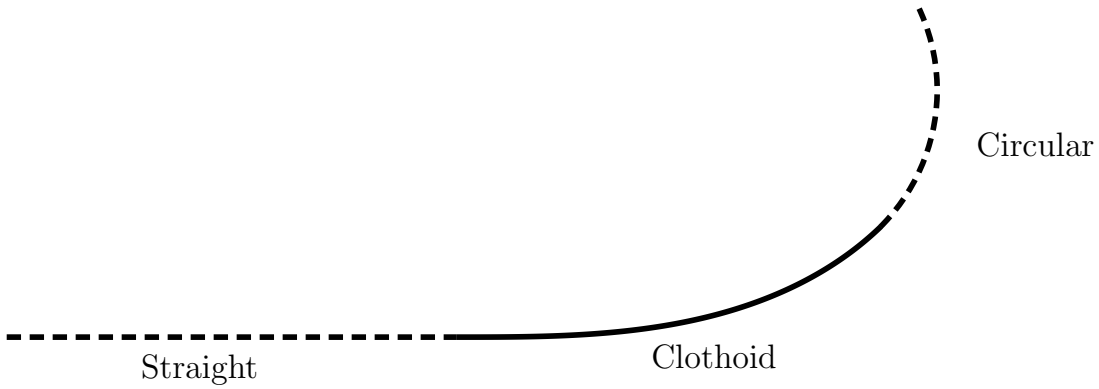


Figure 10.2: A straight curve and a circular curve connected by a clothoid transition curve.

given by [4],

$$z = \frac{pv}{\sqrt{u^2+p^2}}(3ka_t + v^2\frac{dk}{dl} - \frac{kv^2u + gp}{u^2+p^2}\frac{du}{dl}).$$

Where p is the horizontal width between the wheels, v is the velocity of the vehicle, u is the superelevation, k is the curvature along the curve, a_t is the tangential acceleration, and g is the gravity constant, the clothoid is not an optimal transition curve because of its G^0 continuity of the curvature in the joints. However, it is better than no transition curve which will cause a discontinuity in the curvature between segments. A more optimal transition curve would be one where the curvature has a higher geometrical continuity in the joints than G^0, G^n, $n > 0$, while still being monotone and not too steep in the middle section. Ideally, a transition curve should have a curvature closely connected to a sigmoid function (the curvature should be S-shaped), in order to improve the LCA which is closely connected to the curvature. For the comparisons in this paper we take into account what we know about the LCA function and choose to look at

three different scenarios: a track with no transition, a track with a clothoid transition, and a track with a transition that has an S-shaped curvature (we will call it an S-shaped transition). Figure 10.3 shows the curvature of the three transition cases.

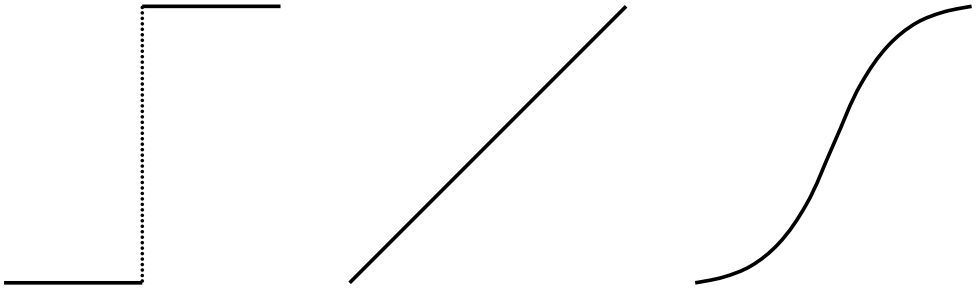


Figure 10.3: The curvature of the transition cases (Left: No transition curve. Middle: Clothoid transition curve. Right: S-shaped transition curve).

10.2.2 Lego Mindstorms

Lego Mindstorms [10] was launched in 1998 and is a software platform for programming Lego inspired robots. Each Mindstorms set consists of Lego building blocks, an intelligent brick that controls the system, and a set of sensors (including motors). Since the start there have been four generations of the Mindstorms platform with the latest being Lego Mindstorms EV3. Lego Mindstorms comes with its own programming software that uses a block-based drag and drop interface where blocks performing a certain task are linked to make the robot run as intended (Figure 10.4 shows a small example program). In addition many other programming languages are supported, both block-based and text-based.

Figure 10.4: An EV3 software program that makes the robot drive forward until the color red or yellow is seen. Then the robot stops and shouts "Fantastic".

The reason for choosing Lego Mindstorms as the test vehicle was mainly because of its small size, making it easy to work with, and because of the various sensors that are available for collecting a number of parameters (the most interesting to us was the gyroscope). We also concluded that a small model would have the same behaviour as a larger model in regards to a transition phase, especially if we consider the LCA function which only includes two parameters from the vehicle; velocity and tangential acceleration. So based on the fact that the Lego motors can maintain a stable constant maximum velocity when running, which means that the tangential acceleration is zero, the lateral forces and the traced path of this small model should, in theory, not differ much from a larger model running with constant speed.

10.2.3 Blender game engine with Bullet physics

Blender is a 3D creation software with tools including modeling, rendering, rigging, simulations, and game creation [11]. The game creation tool, Blender Game Engine [12], can be used for real-time projects. The engine allows the creation of interactive 3D application or simulations, and oversees a game loop which processes logic, sound, physics and rendering simulations in sequential order. A number of powerful libraries are taken advantage of in the Game Engine,

one of them is the Bullet Physics library [13]. This is a physics library that includes 3D collision detection, soft body dynamics, and rigid body dynamics.

The reason for choosing Blender to create a virtual model was because of its physics engine, and because Blender is a well-known tool. Blender offers a physics object called a Vehicle Controller [14] that the engine and Bullet recognize. The Vehicle Controller offers a stable vehicle simulation with simplified physics, that still acts as you would expect a physical vehicle to act, and provides the opportunity to tweak various parameters connected to the vehicle (i.e. maximum velocity). So with the same argumentation as with the Lego vehicle on LCA, the Blender model should, in theory, be suited for experimentation regarding lateral forces and traced paths since the model acts close to a physical vehicle and can hold a constant maximum speed.

10.3 Method

This section describes the setup of the physical and virtual models, how the experiments were created and performed, and how the resulting plots were visually compared.

10.3.1 The physical model

The physical model is constructed from a Lego Mindstorms EV3 set, see Figure 10.5. The model is 21.2 cm long, 14 cm wide, and 11.3 cm high. One EV3 large motor (Figure 10.6) is connected to the two back wheels, making them run in sync, and one EV3 medium motor (Figure 10.7) is connected to the steering of the two front wheels. Both are powered/controlled by an EV3 brick (Figure 10.8). The brick is programmed using the EV3 block-based software where the medium motor's power is manipulated to create the different transitions by turning the front wheels a given angle. Table 10.1 shows the change in motor turn-angle for the three transition cases discretized into 11 steps over 5.5 seconds. A turn-angle of 0 degrees means that the wheels are pointing straight forward making the vehicle drive straight, and a turn-angle above 0 degrees means that the wheels are rotated a given angle, to the left, making the vehicle turn left.

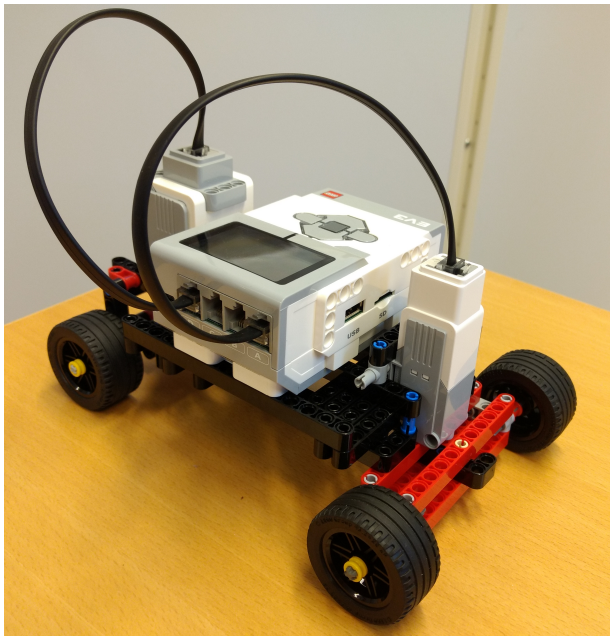


Figure 10.5: The physical model built with a Lego Mindstorms set.

Table 10.1: Change in motor turn-angle to create different transitions. Discretized into 11 steps over 5.5 seconds.

Motor turn-angle in degrees for the transition cases		
No transition	Clothoid	S-shaped
0	0	0
0	3	0.6
0	6	1.8
0	9	4.2
0	12	8.4
0	15	15
30	18	21.6
30	21	25.8
30	24	28.2
30	27	29.4
30	30	30

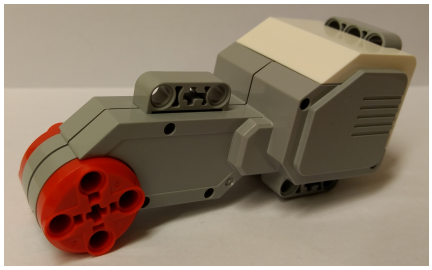


Figure 10.6: The EV3 Large Motor sensor.

Figure 10.7: The EV3 Medium Motor sensor.

Figure 10.8: The EV3 intelligent brick.

10.3.2 The virtual model

The virtual model is created in Blender Game Engine with Bullet as the physics engine, see Figure 10.9. Compared to the physical model the size of the virtual model is scaled up with a factor of 10. This is to compensate for the weakness of the physics engine when working with small collision shapes (object sizes should be above 0.2 units [13]). The setup of the vehicle (with physics, engine power, and steering) is done with Python and a Vehicle Controller. The steering values, which controls the angle of the wheels, are given in the same manner as with the Lego vehicle in Table 10.1, and maximum velocity of the vehicle is set to $3.2\frac{m}{s}$ (about ten times that of the Lego Mindstorms vehicle).

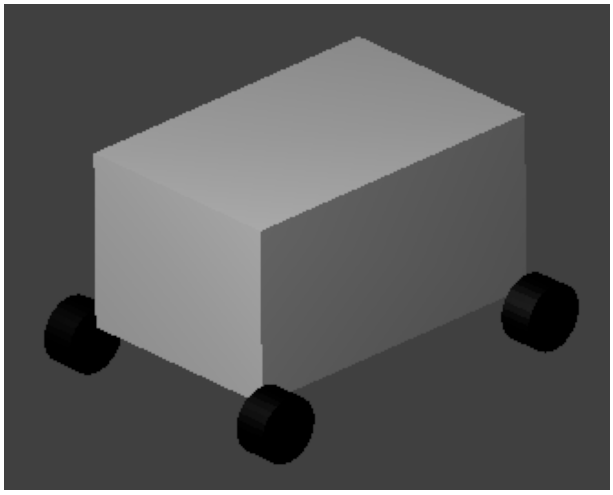


Figure 10.9: The simple virtual model created in Blender with Bullet physics.

10.3.3 Experimental setup

The setup of the vehicle driving experiments is done as follows: For the Lego model a run lasts 8 seconds going from a straight line to a left-hand turn. The first 1.5 seconds is run with a constant wheel turn-angle of 0 degrees (representing the straight section), then the next 5.5 seconds is run with a wheel angle as described in Table 10.1 (representing the transition curve), and finally for the last second the turn-angle is held constant at 30 degrees (representing the curved section). The Blender model is based on the same principle as the Lego model, however since the Blender model uses more time than the Lego model to reach maximum speed, the Lego model reaches maximum speed almost instantly, it is run an extra $7.5 - 8$ seconds in the beginning before starting the 8 second run from straight to curved section.

10.3.4 Analysis

The test cases are divided into two sections; Geometry tests and VR tests. The geometry tests include testing the transition geometry for the physical model by using a gyroscopic sensor. The sensor measures the rotation rate of the Lego robot in degrees per second and should behave differently for the different transitions. For each transition phase the rotation rate is plotted as a graph, and a visual comparison is made between them. The VR tests include a comparison of the physical model vs the virtual model to look at the accuracy between them. This is done by considering the rotation rate and tracing the paths while driving, then visually comparing the graphs for the two models for each transition case.

10.4 Results and discussions

This section shows and discusses the results from the experiments, with a basis in rotation rate and traced path, by using visual comparison of the plots.

10.4.1 Geometry comparison

In Figure 10.10 we see the rotation rate of the three transitions for the Lego model. The figure shows a clear distinction between the transitions where the rotation rate has a jump in the no transition case, is linear in the clothoid case, and is smoother for the S-shaped case. The shape of the graphs is comparable to the shape of the curvature in Figure 10.3, which is to be expected since the lateral forces behave equally to the curvature of the performed curve. This reinforces the credibility of the Lego model as a good test object in spite of its simplicity. The negative graph values just mean that the vehicle is turning left instead of right which would have given equal positive values, and the "jitters" are noise from the sensor. To say something about the observable effects the transitions have on the vehicle we have to look at the spikes in the graphs. The spikes show the moment the wheels are turned and give an indication on how abrupt the change of direction is. The values that are interesting are those moving away from the trend of the graph (the spikes pointing in a positive direction) because they can be interpreted as the force pushing laterally out of the turn. By analysing the spike values in the three graphs it can be seen that the no transition example has the highest value spike with a rotation rate that is approximately 61 $\frac{deg}{s}$ higher than the average values leading up to the spike. For the two other examples, the clothoid has a maximum spike of 18 $\frac{deg}{s}$ higher and the S-shape has a maximum of 8 $\frac{deg}{s}$ higher than the average leading up to the spike. A sudden large jump in lateral forces creates a jerk in the vehicle which can be felt by passengers and may lead to a higher need for maintenance in the case of railway tracks, so a smaller jump is more optimal for both passengers, vehicle, and track. From the previous results this indicates that having a transition curve is superior to not having one, and that an S-shaped curvature curve may be preferable over a linear curvature curve, when the jump in rotation rate is the main criterion.

10.4.2 Physical vs virtual model

For the comparison of the physical model vs the virtual model we start by looking at the rotation rate. Figure 10.11 shows the rotation rate of the Blender model, and as described in the previous sub-section Figure 10.10 shows the rotation rate of the Lego model. When comparing Figure 10.10 and 10.11 the biggest difference between them is the lack of noise and spikes in the Blender measurements. The lack of noise comes from the fact that Blender has "perfect" sensors with only small numerical errors when measuring rotation rate unlike the Lego model, and the absence of spikes can be explained with insufficient friction between the wheels and ground when turning the wheels. A promising result is that the graphs in the figures have the same shape; a jump in the no transition example, linear in the clothoid example, and S-shaped in the S-shaped example. This shows that the "perfect" Blender model and the "natural" Lego model still behave similar even with simplified Blender physics. If we take a closer look on the Blender graphs, we observe that the jumps between plateaus in the rotation rate are -36 $\frac{deg}{s}$ for the no transition case, constantly -4.5 $\frac{deg}{s}$ for the clothoid case, and in the S-shaped case it varies from -0.9 $\frac{deg}{s}$ to -10 $\frac{deg}{s}$, increasing from the smallest jump in the beginning, to the maximum in the middle, then decreasing back to the smallest in the end. Also in this case, as with the spikes in the Lego experiments, a smaller jump is more optimal giving a smoother transition. So a conclusion here is that the clothoid transition is more optimal than no transition, and the S-shaped transition is smoother in the beginning and in the end compared to the clothoid transition, however, for the middle part the clothoid is more optimal. This means that by choosing a transition curve that has a higher continuity curvature in the joints than the linear clothoid, the overall transition will

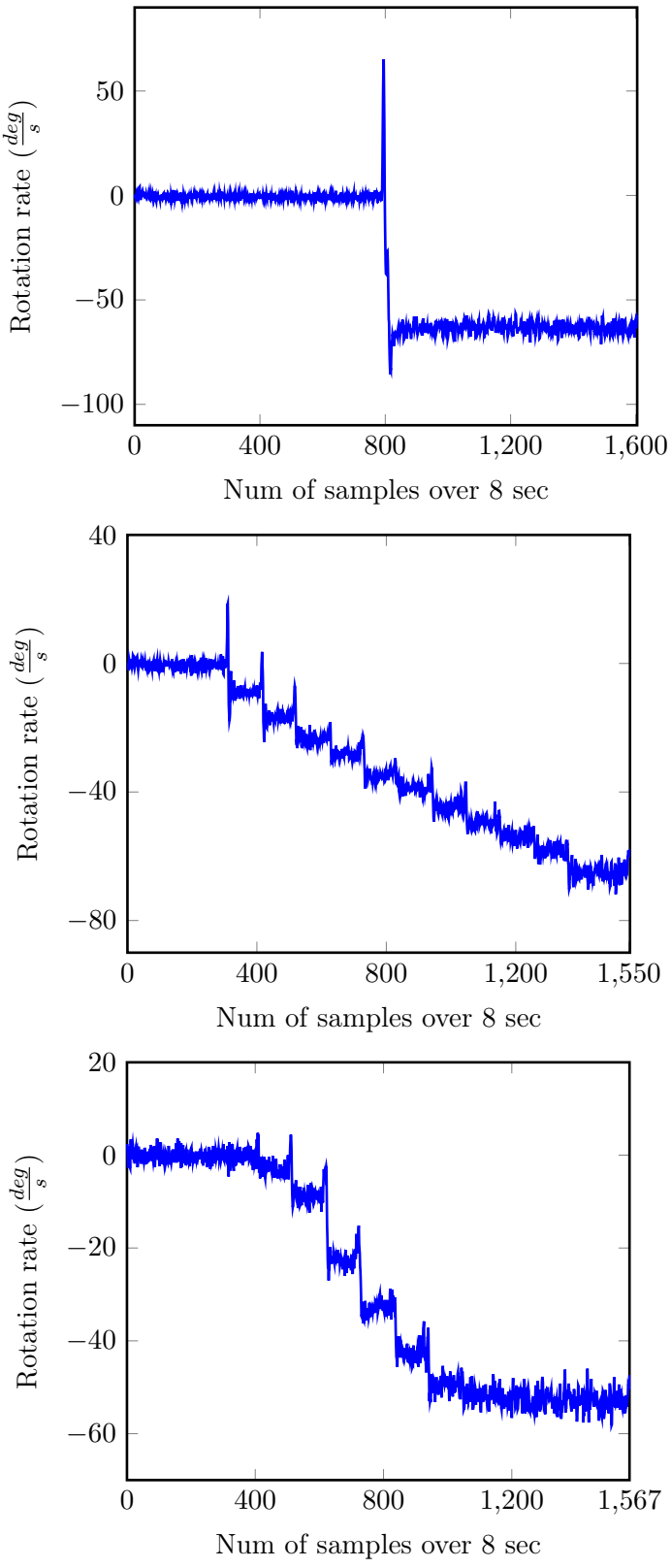


Figure 10.10: The *rotation rate* of the three transition cases for the Lego model. Top: no transition. Middle: clothoid transition. Bottom: S-shaped transition.

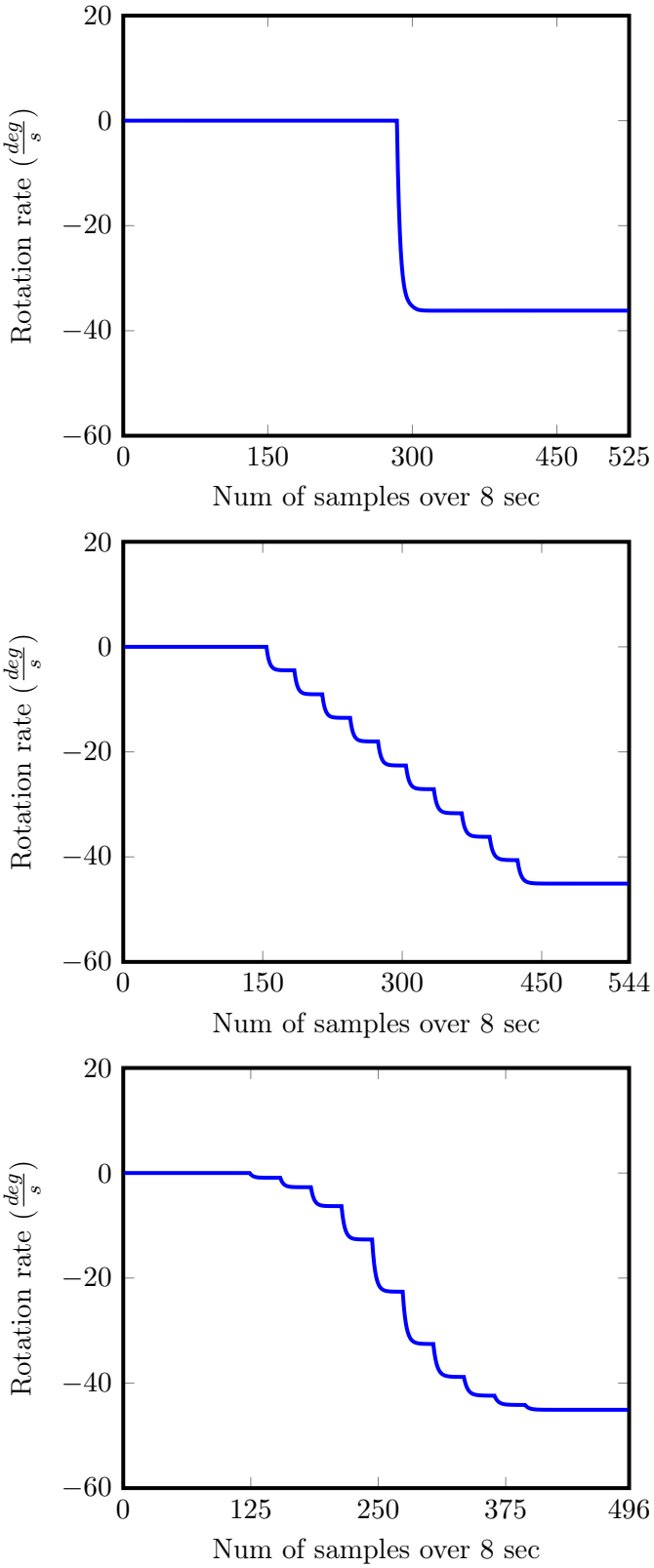


Figure 10.11: The *rotation rate* of the three transition cases for the Blender model. Top: no transition. Middle: clothoid transition. Bottom: S-shaped transition.

be smoother but the maximum jump in lateral force will be higher. So to obtain an optimal transition curve a proper balance between smoothness and maximum jump of lateral force is essential.

The second comparison is done by looking at the traced paths of the models while driving. Figures 10.12, 10.13, and 10.14 show the paths for the three transition cases where the Lego model path is on the top and the Blender model path on the bottom. When comparing them we can clearly see, as expected from the rotation rate comparison, the likeness of the behaviour when driving through the transitions. Each transition example has a characteristic path-shape going into the turn which is clearly observable in both models. The main differences in the graphs are the inability of the Lego model to drive in a perfect straight line, and that the Blender model drives longer and turns wider than the Lego model. The last difference may be due to the wheel-ground friction making it easier for the Blender model to move and turn. Comparing the three transition paths against each other shows that the no transition paths have a distinct spot where the straight segment meets the curved segment, and both the clothoid paths and the S-shaped paths have a more fuzzy transition from the straight segment to the curved segment (which of course makes sense since they have transition curves). The difference between the two last paths is mainly how much they turn when starting with the transition phase, with the S-shaped paths turning more gradual than the clothoid paths.

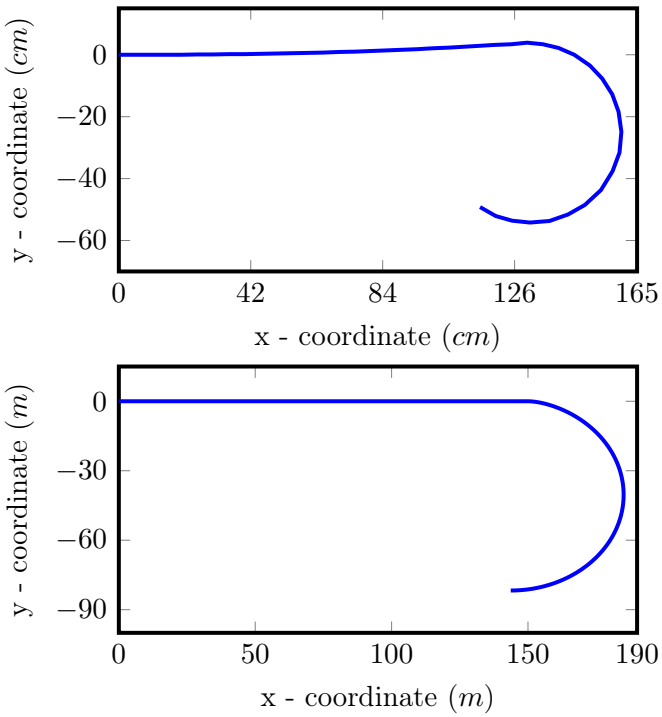


Figure 10.12: The path traced by the vehicle for the no transition example. Top: Lego model. Bottom: Blender model.

10.5 Conclusions

In this paper we have looked at a simple Lego vehicle model and used a gyroscopic sensor to examine the effects of transition geometry on the model. In addition we compared the Lego model with a virtual Blender model by looking at rotation rate and path geometry to investigate if a virtual model can give a satisfactory replication of the physical model behaviour. From the gyroscopic measurements of the Lego model we observed that the rotation rate had the same shape as the curvature for each experiment, and that the turning of the wheels created spikes in the graph. These spikes showed the force pushing laterally out of the turn and told us that

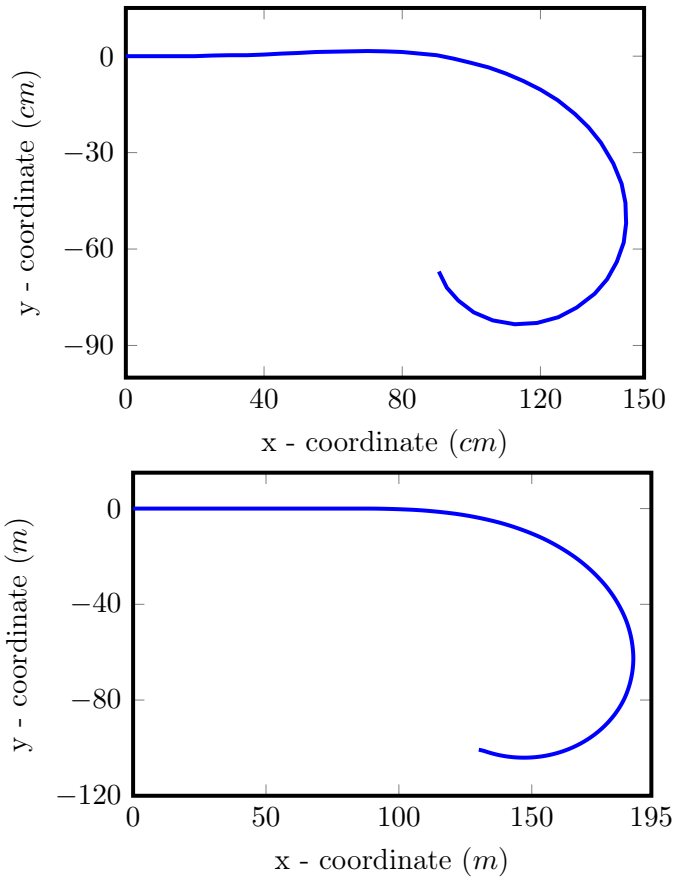


Figure 10.13: The path traced by the vehicle for the clothoid example. Top: Lego model. Bottom: Blender model.

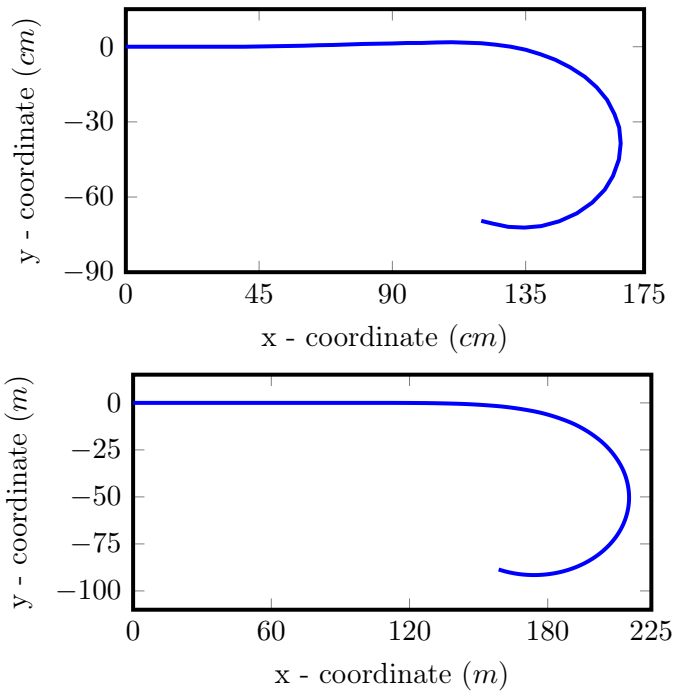


Figure 10.14: The path traced by the vehicle for the S-shaped example. Top: Lego model. Bottom: Blender model.

the no transition turn gave the vehicle a high sudden lateral force compared to the clothoid and S-shaped transitions. When comparing the physical and virtual models we detected that the rotation rates were similar in shape, however the Blender rotation rate did not have any noise or spikes since we were working with a simplified physics model with "perfect" sensors. For the comparison of the traced paths we noticed that the characteristic path-shape of each transition was observable in both models, and that the main difference was connected to distance travelled where the Blender model went a little longer and wider. The results obtained from all experiments tells us that the transition geometry affects the shape of the traced vehicle-path and influence the forces working on the vehicle to a great extent. This means that having a good transition curve is essential in order to minimize the negative effects sudden large changes in force can have on the vehicles, passengers, and tracks (in relation to railway). In addition we conclude that a virtual Blender model gives a good clue on the behaviour of a physical Lego model, with the biggest challenge being wheel-ground friction (especially when turning) and that in some areas (e.g. driving straight) the Blender model is a little too perfect. All in all we think that with small adjustments to the Blender physics a virtual vehicle model can give a satisfactory prediction of the behaviour of a physical vehicle model.

References

[1] Andrzej Kobryń. *Transition Curves for Highway Geometric Design*, volume 14. Springer International Publishing, 1 edition, 02 2017.

[2] J.S. Mundrey. *Railway Track Engineering, Fourth Edition.* Tata McGraw-Hill, 1988.

[3] A. W. Miller M.I.S. (N.S.W.). The transition spiral. *Australian Surveyor*, 7(8):518–526, 1939.

[4] Orhan Baykal. On concept of lateral change of acceleration. *Journal of Surveying Engineering*, 122(3):132–141, 1996.

[5] Orhan Baykal, Ergin Tari, Zeki Coskun, and Muhammed Shain. New transition curve joining two straight lines. *Journal of Transportation Engineering-asce - J TRANSP ENG-ASCE*, 123, 09 1997.

[6] Ergin Tari. The new generation transition curves. *ARI The Bulletin of the Istanbul Technical University*, 54(1):35–41, 2004.

[7] Ergin Tari and Orhan Baykal. A new transition curve with enhanced properties. *Canadian Journal of Civil Engineering*, 32(5):913–923, 2005.

[8] G. G. Robertson, S. K. Card, and J. D. Mackinlay. Three views of virtual reality: nonimmersive virtual reality. *Computer*, 26(2):81–, Feb 1993.

[9] D.J. Walton and D.S. Meek. A planar cubic bèzier spiral. *Journal of Computational and Applied Mathematics*, 72(1):85 – 100, 1996.

[10] Audrey Watters. Lego mindstorms: A history of educational robots. `http://hackeducation.com/2015/04/10/mindstorms`. Accessed: 2019-06-21.

[11] Blender. Blender 2.79 reference manual. `https://docs.blender.org/manual/en/latest/index.html`. 2019-06-26.

[12] Blender. Blender game engine. `https://docs.blender.org/manual/en/latest/game_engine/index.html`. 2019-06-26.

[13] Erwin Coumans. Bullet 2.83 physics sdk manual, 2015.

[14] Blender. Vehicle controller physics. `https://docs.blender.org/manual/en/latest/game_engine/physics/types/vehicle.html`. 2019-08-06.

11 A study on hybrid sensor technology in winter road assessment

Aleksander Pedersen and Tanita Fossli Brustad

Abstract

Road conditions during the winter months in Nordic countries can be highly unstable. Slippery roads combined with heavy haul traffic and ordinary road users can create dangerous, even lethal, situations if road maintenance is unsuccessful. Accidents and critical road conditions may lead to blocked roads, putting strain on a limited number of main roads in many regions, and may in the worst case isolate areas entirely. Using sensors in winter road assessment has been a popular topic for over 20 years. However, with today's developments connected to smaller and cheaper sensors, new opportunities are presenting themselves. In this study, we performed preliminary experiments on a variety of sensors, both commercial and experimental, to evaluate their benefits in possible hybrid sensor technology, which can give a more complete characterization of the road surface than what is possible from just one sensor. From the collected data and visual analysis of the results, the idea of a hybrid sensor seems promising when considering the differences in the tested sensors and how they may complement each other.

11.1 Introduction

Transportation of raw material, goods, and people is a constantly growing factor in the Nordic countries. With this growth, a significant increase in semitrailers for long haul has been observed where research states that many of them are a safety risk on slippery winter roads [1]. To meet the current and future industry with accessible and, more importantly, safe roads, the pressure on successful winter road maintenance is high. Research shows that winter road maintenance is crucial to maintain safe roads [2]. Knowing when and where critical conditions occur, improving risk assessment, can increase the efficiency of maintenance personnel (saving money), and can be used to warn road users of potential risks (increasing safety). The limited number of main roads, combined with long distances between key cities in the Nordic countries, make the road network vulnerable. Critical road conditions can lead to closed roads, which again can lead to areas being isolated over prolonged periods of time. With today's developments connected to faster transfer speed of data, smaller and cheaper sensors, and a focus on smart environments, new opportunities are presenting themselves in winter road maintenance. The new popular area within the winter road assessment research field is that of hybrid measurements with transmission of information to neighbouring vehicles or to road side servers [3, 4]. Hybrid measurements are becoming more and more recognized as a method that can give better overall information on road conditions, often with already available data, compared to lone sensors. This can be seen in recent research, e.g., in [5], where in-car sensor data are incorporated into weather forecast models to improve road weather forecasts; in [6], where a neural network is employed to develop a surface friction prediction model based on historic data from an optical sensor; or, in [7], where acoustic measurements and machine learning algorithms are combined to classify dry, damp, and wet road surfaces.

The research and experiments performed in this study are a part of the Winter Road Maintenance project (WiRMa project). In the project, partners from Norway, Sweden, and Finland conduct research and demonstrate profitable, industrial, and network based systems for decision support in winter road maintenance. The project vision is to demonstrate a viable winter road maintenance decision-support system (DSS) using novel technologies for ambitious ideas by:

- combining measurement techniques for improved icing measurements;
- collecting, communicating, and crunching road weather, vibration, and road surface condition; (snow, ice, wet, slush, etc.) and indirect friction information from vehicles;
- improving forecasts and nowcasts based on vehicle observations; and
- visualizing the information to be used as a DSS in winter road maintenance.

The specific motivation surrounding this work is connected to one of the work packages in the project: Work Package 4 - Winter road condition information. The overall aim of the work package is to investigate the possibility of combining new direct and indirect sensing methods together with known optical sensors to provide a more complete characterization of the road surface (hybrid measurement techniques) by turning vehicles into mobile monitoring stations. Currently, such development is on-going in the automotive industry, but that industry is very closed and the data gathered from modern vehicles are not available for winter road maintenance purposes. Additionally, the data available from vehicles are not based on continuous measurements at the moment, as, for example, slipperiness information is available only when ABS, ESP, or similar are engaged. Mobile systems with continuous measurements will play an important part in monitoring road conditions in Nordic countries, tackling problems related to making rightly timed road infrastructure maintenance decisions. As of now there are simply not enough data and information available to properly handle maintenance on roads that are subjects to harsh conditions and long distances.

Based on the motivations and aims of the WiRMa project, the aim of the study was to select and test a number of sensors to look at the possibility of equipping vehicles with hybrid measurement technology, in order to improve the risk assessment of roads in arctic conditions. The motivation behind this is to increase safety and accessibility for road users during the unpredictable winter months, and help maintenance personnel assess maintenance needs. As a final comment, it is important to remember that sensors will always only be a tool in decision support but ultimately the safety depends on the drivers' behavior of perceiving the sensor alerts, since this is considered to be one of the influential factors in reducing accidents [8]. In this paper, the process and initial results of using certain sensors in winter road assessment are described. The process includes the choices made in selecting sensors (both commercial and experimental), methods for collecting and storing data, and how the measurements were analysed. The presented results are our initial experience with measurement data and the sensor's behaviour, including a discussion where the data are analysed and compared, in relation to reliability, similarity, and usefulness.

11.2 Preliminaries

This section gives a preliminary overview of some of the commercial sensors known in winter road assessment, and gives examples of the possibilities connected to experimental sensors within the field.

11.2.1 Commercial Sensors

When it comes to sensors created for winter road assessment, two categories exist: stationary and mobile. In this study, the focus was on mobile sensors and four optical sensors which were considered: The Road Condition Monitor 411 (RCM411) [9], Mobile Advanced Road Weather Information Sensor (MARWIS) [10], MetRoadMobile [11], and Road Eye [12]. These sensors are measuring the reflected light, from a light source illuminating the road, and analyze the data to classify the surface conditions. The directorate of public roads [13] is also using mobile sensors, but with different measuring types: retardation (deceleration) meter, continuous friction measurement using a mobile wagon attached to the vehicle with either one or two friction wheels, or handheld friction meters. The friction wheels used in Norway are produced by ViaFriction [14], and the directorate of public roads uses two different types, the Road analyser and recorder 5 (RoAR5) [15] and OSCAR [16]. OSCAR is the most reliable of all the friction meters used by the directorate of public roads.

11.2.2 Experimental Sensors

In addition to the sensors that have been specifically created for winter road assessment, experimental sensors, including cameras, have dominated the research field for many years. Four examples showing the variety of experimental sensor types during the last three years (2017-2019) follow. In [17], Pan et al. applied a pre-trained convolutional neural network to automatically detect road surface conditions based on camera images. Hou et al. created a system called VehSense [18], which combined a smartphone and On-Board Diagnostics (OBD-II) data to detect vehicle skidding. Microphones together with neural networks were presented in [19] as a way of detecting road surface wetness. [20] combined piezoelectric and optical sensors to classify road conditions and measure ice/water thickness. These four examples express only a fragment of the creativity and possibilities within the field of road condition assessment (see [21, 22, 23, 24] for additional examples).

11.3 Method

This section describes the choice of sensors used in the experiments, both commercial and experimental. In addition, the placement of the sensors on the vehicle is shown, an overview of how the data for each sensor were collected and stored is given, and lastly how the data were evaluated is explained.

11.3.1 Choice of Commercial Sensors

Based on the project description, one of the key aspects was to assess road condition using measurement techniques. Initially, when we became part of the project, two sensors were already bought and ready to use: RCM411 and MARWIS. The next step was to assemble and test these on a vehicle. Previously, we were discussing four different optical sensors, and since two sensors were already at our disposal and we know that they are commercially available, with extensive support, there was no need to buy any new equipment initially. In addition, the Road Eye sensor (Prototype, Optical Sensors, Gothenburg, Sweden) was to be used in another work package in the project, thus naturally two different sensors were chosen for this work package. Another reason for choosing RCM411 (411, Teconer, Helsinki, Finland) and MARWIS (8900.U04,Lufft, Fellbach, Germany) was because they measure a range of variables, unlike Road Eye and MetRoadMobile (NA, MetSense, Gothenburg, Sweden) that only measure one and two variables, respectively. Both RCM411 and MARWIS are used specifically for measurements on roads, and the directorate of public roads have tested them. As a result, these two commercial sensors are in a much higher price range than the experimental sensors, but then more reliable and suited for their specific purpose.

11.3.2 Choice of Experimental Sensors

The main attributes, in our opinion, which the experimental sensors should possess are: a low price, a reasonable size, should complement the commercial sensor in a beneficial way, and, preferably, be available on the market in some form. Taking these attributes into account, as well as the research done in the literature cited in Section 11.2.2, the main experimental sensor of choice ended up being a Walabot radar [25], with the supplementary selections of vehicle data (OBD data), video, sound, and smartphone sensors.

The Walabot radar is a pocket-sized 3D imaging sensor that uses radio frequency technology to illuminate the area in front of it and sense the returning signals. The sensor supports short range scanning and distance scanning, with the possibility to extract 3D image data, 2D image slices, raw signals, and the sum of raw signals in an image (image pixel energy).

The reason for choosing the Walabot sensor is its small size and low cost, as well as it being a radar. This is of interest since radar technology has not been tested extensively for winter road assessment, which in theory should be suited for given that radar is not dependent on sight, being less affected by snowy or foggy conditions where, for instance, optical sensors have a problem. Thus, since the commercial sensors of choice were both optical sensors, a radar sensor should compliment these sensors in a beneficial way. The other supplementary selections were chosen because of the simplicity to obtain the data, and because they give a variety of information that the other sensors do not give. Vehicle data are already measured by the vehicle and easily acquired by an OBD adapter (we chose a second generation adapter, OBD-II [26, 27]). As for the other data, most people today carry around a smartphone that can both record sound and video, in addition to having a number of other incorporated sensors.

11.3.3 Sensor Placement on the Vehicle

Figure 11.1 shows the placement of the four main sensors (RCM411, MARWIS, Walabot, and OBD-II) on the vehicle. The RCM411 sensor was mounted on the tow ball pointing at the right wheel track, the MARWIS sensor was secured with a rack to the truck bed pointing at the center of the road behind the vehicle, the Walabot sensor was attached below the back left passenger door pointing straight down at the wheel track, and the OBD-II (Model MX 201, OBD Solutions, Phoenix AZ, USA) adapter was plugged into the adapter inside the vehicle. All these sensors had their fixed placement throughout the study, except for the Walabot (Developer Kit, Vayyar Imaging Ltd., Fairfield OH, USA), where various placements were considered before choosing the above described place. The sensors not mentioned here were not part of the main setup, but rather mounted and tested on small intervals under chosen conditions. This included a microphone attached at various places in the dashboard, and a smart phone attached to a phone mount on the dashboard used as a dashcam to capture video of the road in front (which included sound from inside the vehicle). In addition, the dashcam phone also utilized several in house applications for different sensor measurements.

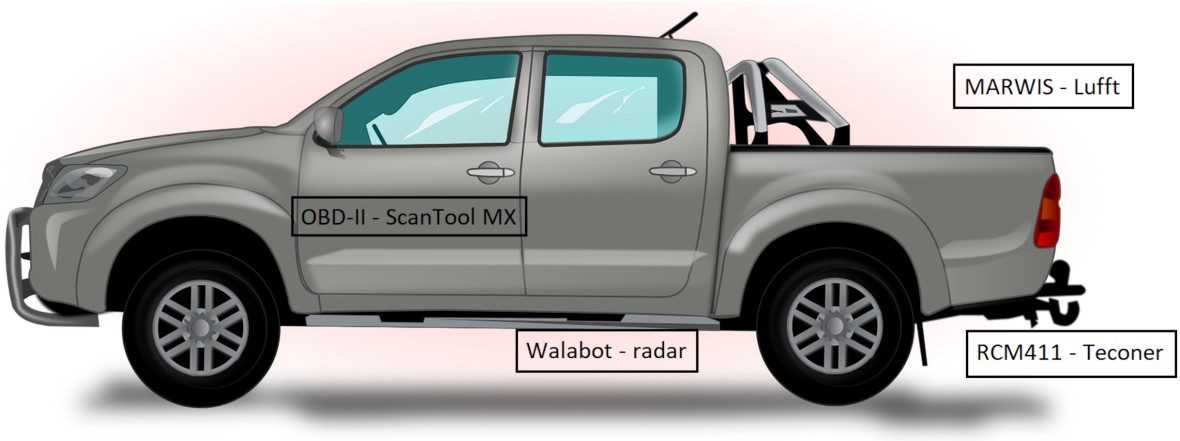


Figure 11.1: Placement of the sensors on the vehicle. Image by OpenClipart-Vectors from Pixabay.

11.3.4 Data Collection and Storage

The data collection and storage were different for each individual sensor. Next follows a description on the methods used, for each sensor, for collecting and storing data. As an overall storage place for all sensors data in all experiments, to ensure easy access and sharing, Box [28] was used.

The RCM411 sensor transferred data to a mobile phone using Bluetooth, which then communicated them to selected servers created specifically for RCM411 sensor data. The servers make it possible to see real-time updates on surface conditions online at `https://roadweather.online`, and download data (in comma-separated values files) from the sensors one owns. Since the online application already existed, was simple to use, and stored the data intuitively, this became the method used for collecting and storing data. In addition, the files from `https://roadweather.online` were downloaded and stored in Box.

The MARWIS sensor was compatible with two existing smartphone applications, were one had equivalent functions as the RCM411 application, the MARWIS application. This app, similar to the RCM411 application, had no internal storage capability and only an option to communicate with a remote server, which required login information which we opted not to use since downloading of raw data seemed impossible. Using the browser application ViewMondo (2.11.0, Informatik

Werkstatt GmbH, Augsburg, Germany), the data from MARWIS are visually good, but there were no available tools for converting the processed data back into a raw format. The other application, ConfigTool (1.1.947, Lufft), was a configuration tool which included a collection of the different sensor variables with an associated visualization. ConfigTool also had an option to save the data to an xml file in the internal storage. Since this was the only suitable option, initially, we decided to store the data in this format. All the measurement data from the sensor were communicated to the phone using Bluetooth, and the ConfigTool files were uploaded to Box.

The Walabot sensor was run from a Raspberry Pi where data were collected and stored, initially. The Raspberry Pi ran a C++ program that collected 4-5 measurement every second from the Walabot, and saved them in comma-separated values files (csv files). To prevent large data loss, in the case of an interrupted connection, a new file was created after a given time interval, storing the data in multiple files instead of just one. Then, after ending a test-run, the data were assembled into one file and stored in Box.

The vehicle data (OBD data) were collected from the vehicle through an OBD adapter (OBD-II) that extracted data from the vehicle and made them available via Bluetooth on a smartphone app. The app lets one select which variables OBD-II should extract from the vehicle, where the possible choices depend on the model and type of vehicle involved. From the app, csv files were available for download, and the downloaded files were then stored in Box.

11.3.5 Measurement Analysis

To evaluate the sensors, in a winter road assessment perspective, the following characteristics were examined and analysed: possible data from the sensors, reliability of the sensors (including existing software), collection and storage experience, and test-run results. An overview of the evaluation criteria is given in Table 11.1. Example data from each sensor are presented and discussed with a focus on relevance in winter road assessment. The reliability of the sensors, including any available software, was considered and evaluated based on our experience when testing them. Both the reliability and experiments presented in this paper are preliminary results recorded when exploring and configuring this type of hybrid setup with multiple sensors and devices being run simultaneously. Collection and storage methods, chosen above, were analysed in regards to fulfilling the task we initially needed them to fulfil. Lastly, the initial test-run results were analysed with visual comparison by observing the differences of two graphs against each other. For the RCM411 sensor, a consistency test was performed to evaluate the stability of the sensor, while, for MARWIS and Walabot, a comparison against RCM411 data was made.

Table 11.1: Description of the evaluation criteria.

Evaluation criteria		Criteria description
Sensor variables:	**S1.**	Number of relevant variables
	S2.	Variable relevance (High, Medium, Low, Exp (Experimental))
Reliability:	**R1.**	Reliability in the initial connection (High, Medium, Low)
	R2.	Reliability during measurements (High, Medium, Low)
	R3.	Reliability of the software (High, Medium, Low)
Data collection and storage:	**D1.**	Format of the collected data
	D2.	Assembly of files required (Yes, No, On stop (If the car power is turned off))
	D3.	Storage (Local, Online, Both)
Test-run results	**T1.**	Initial conclusion of the comparisons

11.4 Results and Discussion

This section includes the results of the initial tests of the three sensors, namely RCM411, MARWIS, and Walabot, with a discussion connected to the evaluation methods described in Section 11.3.5. In addition, the supplementary sensors are considered with a focus on our experience when using them, and a possible hybrid technology is discussed.

11.4.1 RCM411

This section describes the sensor variables available from the RCM411 sensor, gives a discussion on the reliability of the sensor based on our experience, and comments on collection and storage of the data. In addition, a consistency test of the variables friction, surface temperature, and road state (defined here as road surface condition) is shown.

Sensor Variables

The RCM411 sensor offers a range of different variables from the measurements. In the case of winter road assessment, the most interesting are: friction, road state, air temperature, surface temperature, and water layer thickness. For the road state, a number in the range 1-6 is given where 1 = dry, 2 = moist, 3 = wet, 4 = slush, 5 = ice, and 6 = snow. Figure 11.2 shows three examples of extracted data from the RCM411.

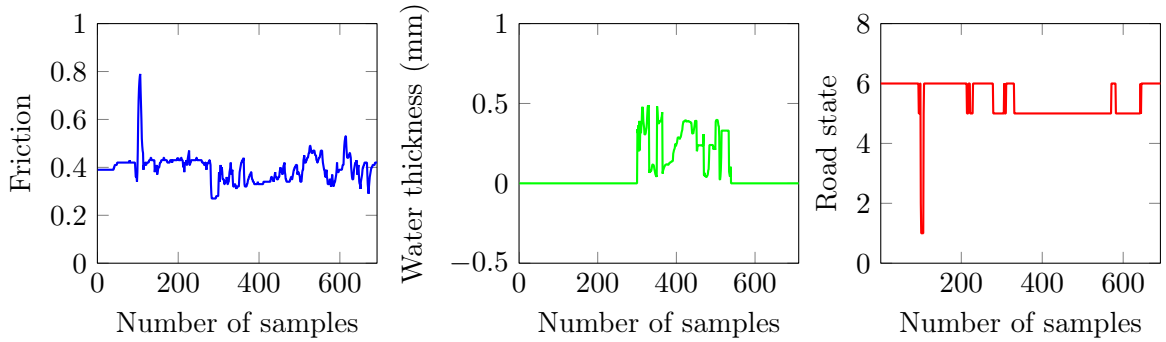


Figure 11.2: Three of the possible data types that can be extracted from the RCM411. Left: Friction. Middle: Water thickness. Right: Road state.

Reliability

The RCM411 sensor was one of the most reliable sensors, in this initial study, in regards to uptime. Through the experiments, the sensor did not once experience disconnection problems or lose data during a run. The sensor was easy to mount on the vehicle and establishing a connection against a smartphone went painlessly. To us,the online resource showing real-time updates is dependable and useful, expressing the variables friction, water thickness, air temperature, and road temperature as a graph, and road state by adding color to the traced path.

Data Collection and Storage

The data collection and storage from the RCM411 sensor, via the online web-page, worked. One csv file was created for each run, with measurements taken once every second, and, by logging into our account, the files, sorted by date and time, were easily downloaded. Since there was only one measurement per second from the sensor during the test-runs, the created files were not overly large and well suited for analysis, even for runs lasting more than an hour.

Consistency Test

The initial test-run of the RCM411 sensor was a consistency test where a decided track was driven twice in order to observe the similarities between the two runs. A comparison was made between three variables, namely friction, surface temperature, and road state, to evaluate the consistency of the measurements. The plots of the frictions can be seen in Figure 11.3, surface temperatures in Figure 11.4, and road states in Figure 11.5. The test-route, located in Narvik, Norway, had a length of about 5 km and took 12 min to drive through, and the weather conditions were sunny with an average air temperature of -4 degrees Celsius. In the comparisons, we do not expect to see two equal graphs for the measured variable, because we know that conditions were not exactly the same for the two runs. The influence of other road users was present, as well as the impact from the sun, and the varying speed of the vehicle. However, we do expect to see similarities between the graphs in some areas.

In the figures, it can be observed that there are similarities between the variables when comparing the two runs. Especially in the friction and road state plots where the runs give equal values for the variables in certain areas, more so for the state than the friction, which is reasonable given that the friction can change within a road state. In the plot of the surface temperature (Figure 11.4), the observation is that they have the same behaviour but the values of the second run are consistently higher than for the first run. This can be explained by the sunny conditions where the air temperature was on average one degree Celsius higher during the second run compared to the first run. An overall conclusion to the consistency test is that the sensor appears to be stable when measuring the variables friction, surface temperature, and road state.

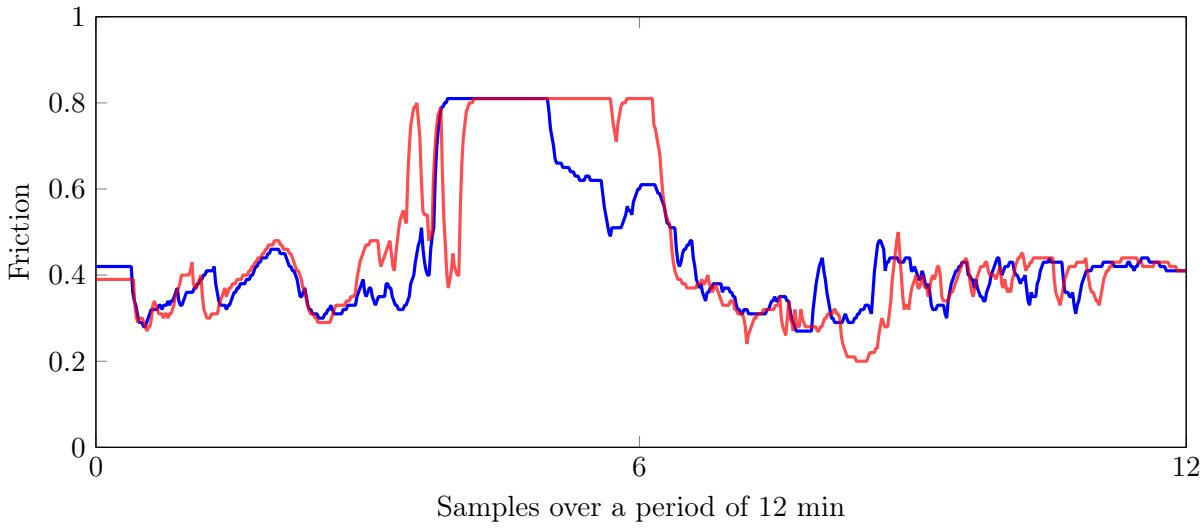


Figure 11.3: RCM411 friction plots from two runs along the same test route: blue, run 1; red, run 2.

11.4.2 MARWIS

This section describes the sensor variables available from the MARWIS sensor, gives a discussion on the reliability of the sensor based on our experience, and comments on collection and storage of the data. In addition, a comparison between MARWIS and RCM411 friction is made.

Sensor Variables

The MARWIS sensor offers many of the same variables as RCM411, and also some additional information, relevant in winter road assessment. The variables equal to RCM411 are: friction,

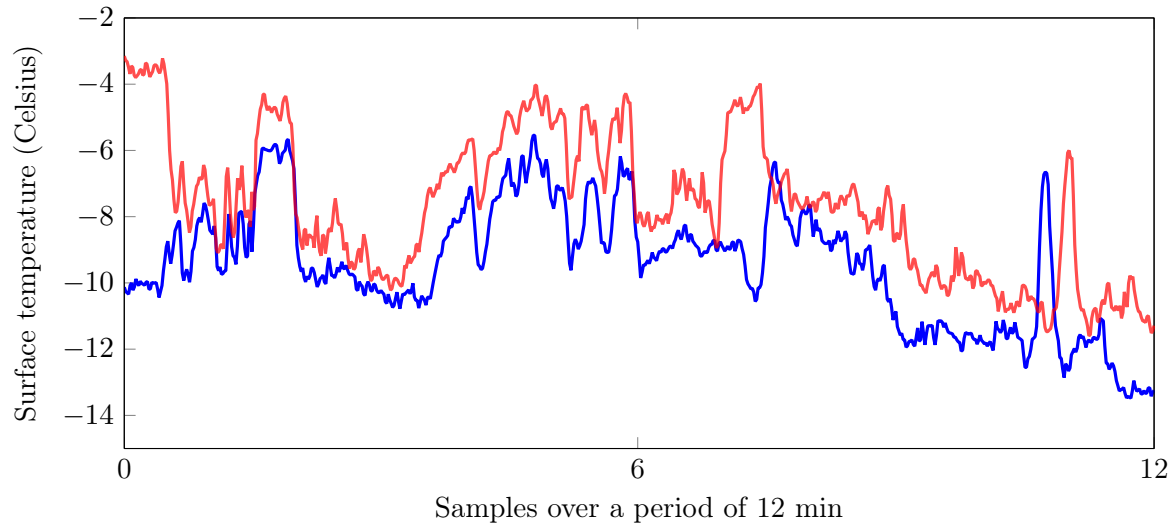


Figure 11.4: RCM411 surface temperature plots from two runs along the same test route: blue, run 1; red, run 2.

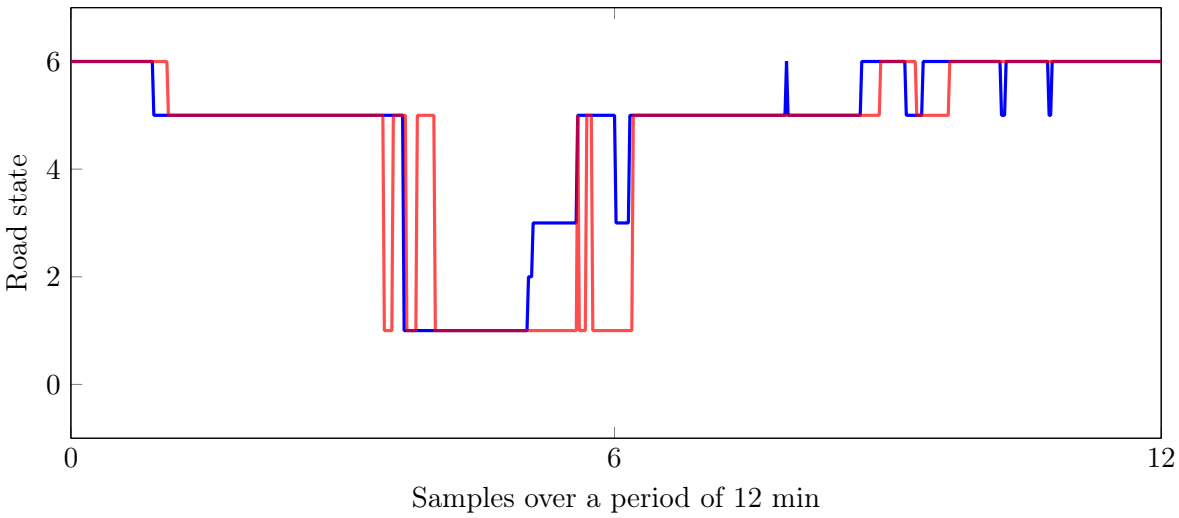


Figure 11.5: RCM411 road state plots from two runs along the same test route: blue, run 1; red, run 2.

road state, air temperature, surface temperature, and water layer thickness. For the road state from MARWIS, a number in the range 1-8 is given where 1 = dry, 2 = damp, 3 = wet, 4 = ice, 5 = snow + ice, 6 = chemically wet, 7 = water + ice, and 8 = snow. The additional variables that MARWIS has are ice percentage, dew point temperature, and relative humidity (air and ground). Three examples of possible data can be seen in Figure 11.6.

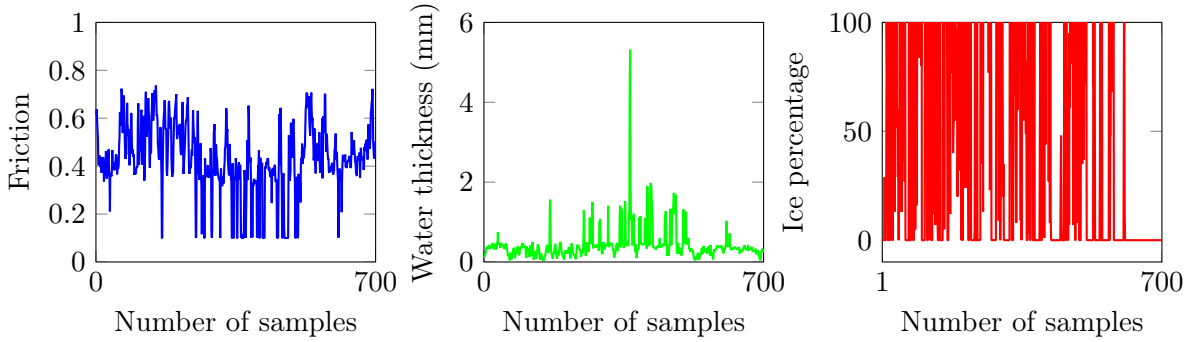


Figure 11.6: Three of the possible data types that can be extracted from the MARWIS: (left) friction; (middle) water thickness; and (right) ice percentage.

Reliability

In the initial study, we recognized problems with MARWIS almost immediately. Initially, we had problems with MARWIS losing power, which was caused by an unstable/broken vehicle power adapter. In addition, some issues where identified with the Bluetooth connection between MARWIS and our smart phone, which caused a disconnect between the two, and the application had to restart before further use. This issue was caused by the application itself, and related to the loss of power. Further, the loss of power disrupted the Bluetooth connection to the sensors, which again led to the phone application crashing without any warnings or reconnects. Another problem included the powering of MARWIS where a converter that came with the sensor converted 12 V to 24 V from the car adapter, which MARWIS required under colder conditions (below a certain temperature, the sensor needs to heat up in order to work properly). The converter did not have any type of safety to limit the voltage, and, when the vehicle was accelerating, the voltage from the adapter increased by a small amount, which led to a higher voltage than MARWIS was made for, frying one of the internal circuits.

Data Collection and Storage

The data collection and storage required a stable connection with no application crashes, and the initial reliability tests led to an evaluation in whether to continue with the commercial application. We had earlier discussed how to save the data for further analysis, and the remote server data upload was discarded since raw data were not available. The other option was to use the configuration tool and save locally, but the format of the data in xml was only byte values, which we agreed was not suitable for our purpose. Since we had no idea if the remote server would let us download data in a useful format, we created our own prototype application for communicating with MARWIS and storing the data. Doing research on the communication protocol, communication settings (Bluetooth) led to specific queries with a bytestream checksum directly to the MARWIS sensor. We got a working application which received data from the sensor based on our query and we identified some of the key problems which we experienced in the initial setup. One reason for the crashes in the commercial application was caused by variables not being initialized after a query, which again was caused by either delays in a response or malfunction in the power/Bluetooth connection. Seeing that we now had control of the data,

we were able to save them in a useful format on the internal storage on the phone for further analysis.

Comparison

The initial test results from the MARWIS sensor was extracted on a test-run together with the RCM411 sensor. Thus, since MARWIS and RCM411 have many equal measurement variables, and they are both optical sensors, it was decided that comparing the two would be interesting in the initial phase. The variable that was compared was surface friction, because this is a key variable in winter road assessment. In Figure 11.7, the graphs of the two frictions plotted, and shows the data collected during a 1 h drive.

In the figure, it can be seen that the two sensors show different friction coefficients for almost the entire test drive. From the setup of the sensors, we know that they do not measure the exact same spot, and that the area measured is larger for the MARWIS sensor than the RCM411 sensor. However, the expectations from our part, before analysing the data, was that areas of equal or similar friction would be observed. Since this is not the case, it would be interesting to do more testing between these two sensors, and compare the results to previous work and comparisons connected to them.

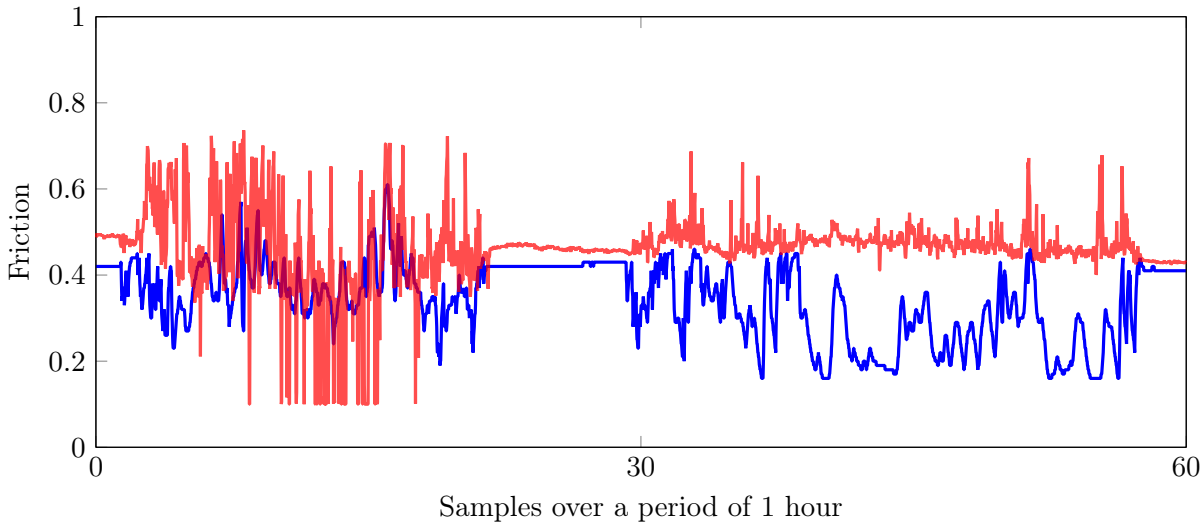


Figure 11.7: The plots of friction data from RCM411 in blue and MARWIS in red, during a 1 h drive.

11.4.3 Walabot

This section describes the sensor variables available from the Walabot sensor, and indicates which variables we found relevant. In addition, the section gives a discussion on the reliability of the sensor based on our experience, comments on collection and storage of the data, and shows a first comparison of Walabot image energy against RCM411 friction.

Sensor Variables

The Walabot sensor is based on imaging data and can offer several types of output variables. The types are: 3D image, 2D image slice, raw signals, image energy (sum of raw signals), and imaging targets (a list of and the number of identified targets). Figure 11.8 shows three examples of possible outputs from the Walabot. In regards to using Walabot data in winter road

assessment, the most suitable data output, having both storage and visual analysis in mind, was predicted to be image energy, since it gives one value every timestep, making it easy to create graphs that can be compared with the other sensor variables.

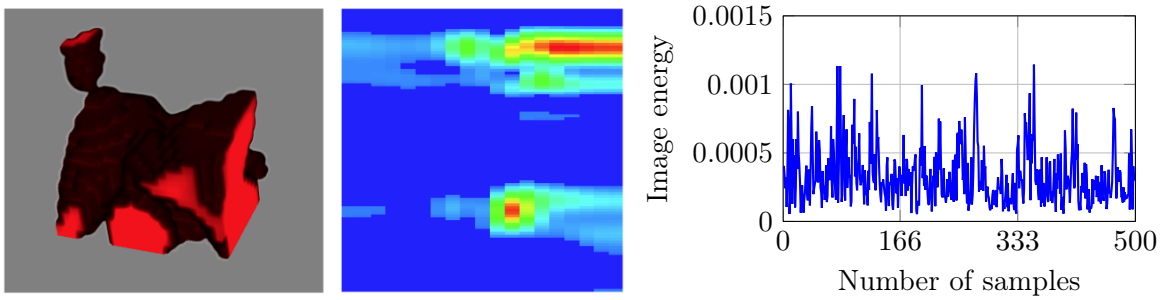


Figure 11.8: Three of the possible data types that can be extracted from the Walabot: (left) 3D image; (middle) 2D image slice; and (right) image energy.

Reliability

Overall, the reliability of the Walabot sensor was adequate. After it was attached to the vehicle, and up and running, no problems occurred when performing the measurements, and data were collected without interruptions. Although collecting data from the initial test-runs went smoothly, the preparations of the Walabot leading up to the test-runs presented some challenges. Firstly, the connection between the Walabot and the Raspberry Pi, via USB, had a tendency to be lost due to a sensitivity in the connection on the Walabot side, where the USB plug needed to be kept stable to secure a connection. This was solved by firmly securing the plug in the Walabot so that it had not room to move. Secondly, the Walabot does not have its own clock, and is therefore dependent on the clock of the device it is connected to. This caused some early headaches seeing that the Raspberry Pi, which the Walabot was connected to, only has a reliable clock as long as it has Internet connection. Thus, when the Raspberry Pi is without Internet access, the clock starts counting from the time it had when the device was last turned off. The strategy used to solve the clock challenge was by either setting up a hotspot on a mobile phone or by manually setting the time on the Raspberry Pi before starting the Walabot.

Data Collection and Storage

The collection and storage method used for the Walabot worked well. Collecting the data in multiple files proved to be necessary given that on some test-runs small interruptions occurred leading to the loss of all data in the file that the Raspberry Pi was saving to in that moment. Since the Walabot gave 4-5 measurements every second a large amount of data was collected when the test-runs lasted an hour or more, thus the method of assembling all the data into one file may not always be desirable. In our case, with the initial tests, putting the data into one file seemed like the best choice, in regards to simplifying analysis and comparisons later on.

A First Comparison

To get a first impression on how the measurements from the Walabot behave when measuring winter roads, an experiment was conducted with the Walabot together with the RCM411. Both sensors were taken for a test drive and a segment over a period of 20 min was extracted from the sensors and compared. The plots of the two measurements can be seen in Figure 11.9, where the blue graph is friction from the RCM411, and the red graph is image pixel energy from the Walabot.

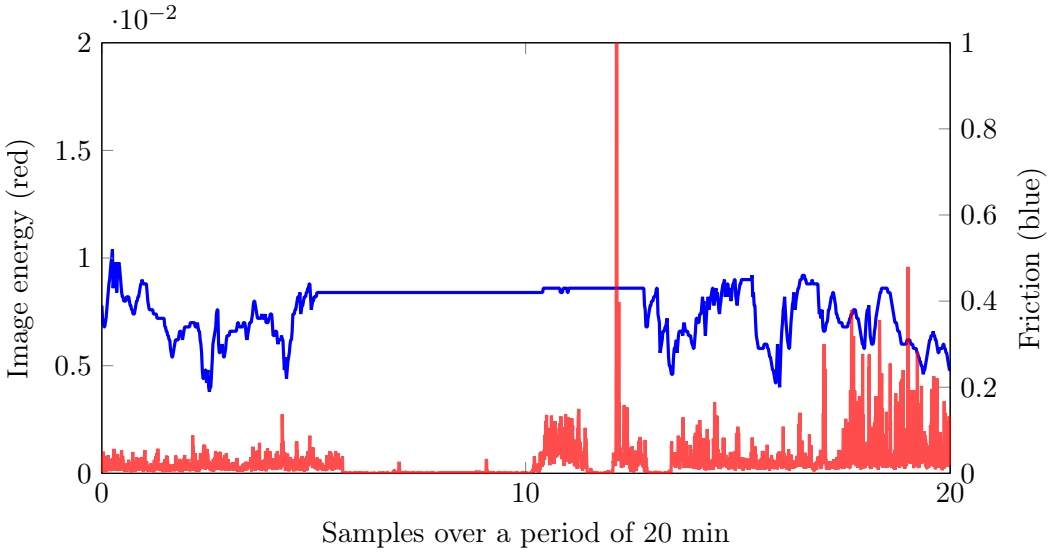


Figure 11.9: The plots of friction data from RCM411 in blue, and Walabot image energy in red, during a 20 min drive.

When comparing the two graphs in Figure 11.9, it can be observed that in the areas where the friction is close to constant, the image energy has the same behaviour giving constant values close to zero, and, for the areas where the friction fluctuates, the image energy also fluctuates. As a first impression, this is a positive result showing that the Walabot radar may have potential in winter road assessment since it seems to have a correlation to, at least, road surface friction. The results from this initial test give strong reasons, and increase the interest, for further investigation of the experimental Walabot sensor for road assessment purposes.

11.4.4 Video, Sound, OBD-II and Smartphone

The measurements and captures collected from the devices/sensors in this section were not the main priority during the startup, and we had no specific goal or ideas of what to expect from the data. Since the measurement setup was mostly based on curiosity and future ideas, no analysis regarding the data captured from the devices/sensors was performed. However, the preliminary experiences are discussed.

Video

One of the ideas was to use video capture from in front of the vehicle as a reference for what to expect from the friction measurements behind the vehicle. Another interesting aspect here for future notice was to combine video with artificial intelligence (AI) and predict the road condition (friction) during the video analysis. An example is shown in Figure 11.10. Capturing this amount of video for an entire trip (several hours) requires both storage capacities on the capturing device as well as storage for further use. Because we were using Box for storage purposes on the other measurement data, we included the video files as well in a folder. With later experiences, this was not a good idea with respect to the capabilities of Box and large file sizes.

Sound

We were also doing some experimental work regarding sound capture using a microphone. The idea was to attach a microphone to the vehicle chassis, preferably something as tough as steel to get more accurate vibrations from the ground travelling through the vehicle frame. There were only some initial sound captures and no post processing or analysis of the data have been undertaken yet. The file sizes were so small that they are negligible when uploading.

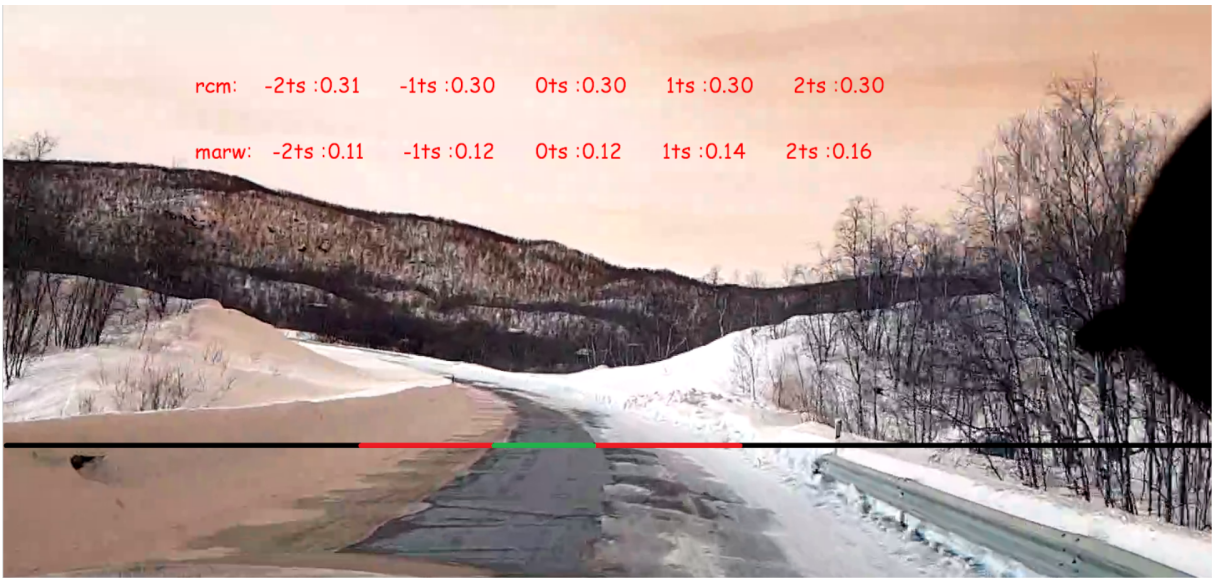


Figure 11.10: Example image on how video can be preprocessed using AI to estimate road conditions. Figure by Kristoffer Tangrand.

OBD-II

Of the experimental sensors and devices, the OBD-II is already being used for car performance analysis and research purposes. The idea here was to extract information about the car, which is already available, and utilize post processing tools to analyse the data with regards to other measurements such as friction. There are many variables available in the application, where the user can select which ones to record the output from. The available variables differ depending on car model and make, and for our use we selected those we found relevant, both for comparison against other sensors and also analysis purposes.

Smartphone

A smartphone is packed with sensors which can record data such as acceleration, orientation, and sound. The commercial application we used for RCM411 had some experimental friction calculations based on measured forces on the phone, that is, using accelerometer to estimate friction and breaking distance. This specific function was optional, and could be uploaded in the same file as the other variables, for comparison against the measured values. Based on these experimental calculations, we wanted to do our own experiments using smartphone sensors only, and no external information. There were some initial experiments, but this requires more work before doing any analysis on the data.

11.4.5 Hybrid Sensor Technology

In Table 11.2, a brief summary of the previous results is displayed, based on the criteria from Table 11.1 together with the price range of each sensor. If we consider the previously discussed sensors in relation to creating a hybrid sensor in the future, there are some interesting observations already. The two optical sensors, RCM411 and MARWIS, have many of the same measured variables so in theory only one of them should be necessary in a hybrid sensor. However, from the friction plots in Figure 11.7 of RCM411 and MARWIS, an observation was made that the measured friction differed for most of the test drive. This opens up a possibility that the two sensors may be sensitive to different types of conditions, and, if so, a combination of both sensors will be beneficial in a hybrid solution.

Table 11.2: Results of the criteria evaluation described in Table 11.1, along with the price range (PR) of each sensor.

Criteria	Sensors						
	RCM411	**MARWIS**	**Walabot**	**Video**	**Sound**	**OBD-II**	**Smartphone**
S1	5	8 (9)	4	1	1	Optional	3
S2	High	High	Exp	High	Exp	Medium	Medium
R1	High	Medium	Low	High	High	High	High
R2	High	Medium	High	High	High	High	High
R3	High	Low (in-house: High)	Medium	High	High	High	High
D1	csv	xml (in-house: csv)	csv	3gp	3gp	csv	txt
D2	On stop	Yes	Yes	No	No	On stop	Yes
D3	Online	Both	Both	Both	Both	Both	Local
T1	Consistent	Dissimilar to RCM411	Interesting results against RCM411	N/A	N/A	N/A	N/A
PR	High	High	Low	Low	Low	Low	Low

The experimental Walabot radar sensor shows, during the first comparison with RCM411 in Figure 11.9, that there may be a correlation between RCM411 friction data and Walabot image energy data. This is of interest since our prediction is that the Walabot is less affected by snowy and foggy conditions compared to the RCM411 sensor, which will make it an asset during tough weather conditions. In addition, the size and price range of Walabot is a major advantage in relation to a hybrid sensor.

For the device/sensors discussed in Section 11.4.4, the overall benefits are their small size and that the technology is available for a reasonable low cost. In addition, winter assessment research has been performed on all of them with positive results. The greatest potential lies, probably, within the use of video in winter road assessment, because it brings a new dimension to the analysis that can involve other research areas. Above we presented an example on combining video with AI to predict friction measurements, but the possibilities are numerous. Thus, in our opinion, video will make a valuable inclusion in hybrid assessment. The OBD-II collects data from the vehicle, in contrast to RCM411, MARWIS, Walabot, and video that collects road data. This in itself makes it an important supplementary in the hybrid measurements. Vehicle data can increase the reliability of the risk assessment, since vehicles may act differently under difficult road conditions. Regarding the two last measurements (sound and smartphone data), more experiments will have to be performed to find the strengths they may possess that the other sensors lack.

Implementation of the technology on the vehicle (placement of the sensors) is a topic that will have to be investigated more in future works. The placements chosen in this study, as described in Figure 11.1, were based on restrictions connected to the sensors, the type of vehicle used, and our opinion based on predictions and ideas. In a hybrid sensor technology, it will be important to consider on which parts of the road measurements should be conducted, and if sensors should be spread out around the vehicle as opposed to gathered in one place. Gathering the sensors in one place can lead to information being lost (e.g., the two wheel tracks can have varying conditions), and spreading them out may require a larger number of sensors (increasing the price). A solution can be to implement sensors that can measure the whole width of the road and then extract condition data for given areas, for instance with a camera, as shown in Figure 11.10; however, this will enlarge the post-processing work. The pros and cons of various sensor placements should be identified and mapped so that a justified decision can be made based on the chosen hybrid

technology.

11.5 Conclusions

Assessment of road conditions in Nordic countries is a necessity due to increasing amounts of transportation and heavy haulers, which enables the need for accurate safety and risk assessment with strong focus on maintenance and estimates on the occurrence of critical conditions. We performed a preliminary study on several sensor technologies with possible connections to hybrid sensor technology. We tested two commercial and one experimental sensors, as well as four supplementary measuring devices or techniques. In addition, a set of classification criteria was developed (see Table 11.1) and used (see Table 11.2) for evaluating the different sensors. The main findings can be summarized as follows:

- The difference in friction for RCM411 and MARWIS, on the test drive, opens up the possibility that the two sensors have different sensitivity to various types of conditions.
- There is a possibility that the RCM411 friction and Walabot image energy are correlated, but the data are currently not sufficient to verify this.
- Video can bring a new dimension to road assessment, by combining it with AI and sensor input.
- The OBD-II data are the only tested sensor data that give information about the vehicle instead of the road, which make them interesting in for hybrid technology.

These early studies established connections between technologies that could potentially decrease risk and, hence, increase safety for all road users. The study also focused on experimental technologies to improve and complement already existing winter road assessment devices. Experimental technologies can present cheaper solutions to replace commercial sensors if the measured variables are correlated. The initial results show promising connections between the tested sensors and possible device combinations which would benefit from a hybrid solution. Implementing a form of hybrid sensor technology could both increase the reliability of risk assessment and have a more complete characterization of the road surface. Having a publicly available joint information source which could identify areas with higher risk that should, if possible, be avoided for safety reasons will increase awareness of difficult conditions. The paper shows several experimental measurements and, as mentioned above, some of them need more work in order to find both strengths and weaknesses. There should also be performed a more thorough field test which incorporates all the above-mentioned equipment to establish a complete view of the measurements compared to each other.

11.6 Funding

This paper was partially funded by the project WiRMa (Winter Road Maintenance), ID 20201092, supported by the Interreg VA NORD-program.

11.7 Acknowledgements

The authors would like to thank Børre Bang for support. The publication charges for this article have been funded by a grant from the publication fund of UiT The Arctic University of Norway.

References

[1] J Granlund and P Thomson. Traffic safety risks with EU tractor-semitrailer rigs on slippery roads. HVTT14: International Symposium on Heavy Vehicle Transport Technology, 14th, 2016, Rotorua, New Zealand, 2016.

[2] Liping Fu and Taimur Usman. *Safety Effects of Winter Weather and Road Maintenance Operations*, chapter 7, pages 101–130. John Wiley & Sons, Ltd, 2018.

[3] Yoshitaka Shibata, Akira Sakuraba, Goshi Sato, and Noriki Uchida. Iot based wide area road surface state sensing and communication system for future safety driving. In Leonard Barolli, Makoto Takizawa, Fatos Xhafa, and Tomoya Enokido, editors, *Advanced Information Networking and Applications*, pages 1123–1132, Cham, 2020. Springer International Publishing.

[4] Yoshitaka Shibata, Akira Sakuraba, Goshi Sato, and Noriki Uchida. Realtime road state decision system based on multiple sensors and ai technologies. In Leonard Barolli, Farookh Khadeer Hussain, and Makoto Ikeda, editors, *Complex, Intelligent, and Software Intensive Systems*, pages 114–122, Cham, 2020. Springer International Publishing.

[5] Yumei Hu, Esben Almkvist, Torbjörn Gustavsson, and Jörgen Bogren. Modeling road surface temperature from air temperature and geographical parameters—implication for the application of floating car data in a road weather forecast model. *Journal of Applied Meteorology and Climatology*, 58(5):1023–1038, 2019.

[6] Ziyuan Pu, Shuo Wang, Chenglong Liu, Zhiyong Cui, and Yinhai Wang. Road surface friction prediction using long short-term memory neural network based on historical data. In *Transportation Research Board 98th Annual Meeting, Washington DC, United States*, January 2019.

[7] M. Kalliris, S. Kanarachos, R. Kotsakis, O. Haas, and M. Blundell. Machine learning algorithms for wet road surface detection using acoustic measurements. In *2019 IEEE International Conference on Mechatronics (ICM)*, pages 265–270, March 2019.

[8] Danish Farooq, Sarbast Moslem, and Szabolcs Duleba. Evaluation of driver behavior criteria for evolution of sustainable traffic safety. *Sustainability*, 11(11):3142, Jun 2019.

[9] Taisto Haavasoja, Juhani Nylander, and Pauli Nylander. Experiences of mobile road condition monitoring. In *Proceedings of SIRWEC 2012, Helsinki, Finland*, May 2012.

[10] Lufft. Marwis specifications. Online, 2019 (Accessed 2019-09-05). `https://www.lufft-marwis.com/en_US/specifications`.

[11] MetSense. Metroadmobile - mobile road friction monitoring. Online, 2017 (Accessed 2019-12-12). `http://www.metsense.com/product?p=MetRoadMobile`.

[12] Sten Löfving. Road eye. Online, 2014 (Accessed 2019-12-12). `http://www.opticalsensors.se/roadeye.html`.

[13] Vegvesenet. The directorate of public roads - winter training for operators. Online, (Accessed 2019-12-13). `https://www.vegvesen.no/_attachment/2365449/binary/1270205?fast_title=SVV+rapport+673+Oppl%C3%A6ring+i+vinterdrift+for+operat%C3%B8rer.pdf`.

[14] ViaTech. Viafiction brukes for å måle friksjonsnivåene på veg og flyplass. Online, (Accessed 2019-01-14). `http://www.viatech.no/products.aspx?lang=no&id=1`.

[15] ViaTech. Roar mk6 oppgradering. Online, (Accessed 2019-01-14). `http://www.viatech.no/products.aspx?lang=no&id=2`.

[16] ViaTech. Oscar. Online, (Accessed 2019-01-14). `http://www.viatech.no/products.aspx?lang=no&id=3`.

[17] G. Pan, L. Fu, R. Yu, and M. Muresan. Evaluation of alternative pre-trained convolutional neural networks for winter road surface condition monitoring. In *2019 5th International Conference on Transportation Information and Safety (ICTIS)*, pages 614–620, July 2019.

[18] Y. Hou, A. Gupta, T. Guan, S. Hu, L. Su, and C. Qiao. Vehsense: Slippery road detection using smartphones. In *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, pages 1–5, June 2017.

[19] Giovanni Pepe, Leonardo Gabrielli, Livio Ambrosini, Stefano Squartini, and Luca Cattani. Detecting road surface wetness using microphones and convolutional neural networks. In *Audio Engineering Society Convention 146*, March 2019.

[20] Kang Gui, Lin Ye, Junfeng Ge, Faouzi Alaya Cheikh, and Lizhen Huang. Road surface condition detection utilizing resonance frequency and optical technologies. *Sensors and Actuators A: Physical*, 297:111540, 2019.

[21] Choongheon Yang, Dukgeun Yun, Jinguk Kim, and Sungpil Shin. Estimation road surface temperature variation using commercial vehicle ambient sensor. *IOP Conference Series: Materials Science and Engineering*, 603:022027, 09 2019.

[22] Jakob Döring, Lakshan Tharmakularajah, Jakob Happel, and Karl-Ludwig Krieger. A novel approach for road surface wetness detection with planar capacitive sensors. *Journal of Sensors and Sensor Systems*, 8:57–66, 01 2019.

[23] Armando Piccardi and Lorenzo Colace. Optical detection of dangerous road conditions. *Sensors*, 19:1360, 03 2019.

[24] Johan Wåhlin, P. Jonsson, Henrik Baad, Johan Edblad, Torgeir Vaa, Jörgen Röien, Johan Casselgren, Taisto Haavasoja, and Kerl Schedler. Laboratory test of five different optical road condition sensors. In *Proceedings of SIRWEC 2016, Ft. Collins, Colorado, USA*, April 2016.

[25] Walabot. Walabot - technical brief. Online, (Accessed 2019-09-05). `https://walabot.com/docs/walabot-tech-brief-416?type=pdf`.

[26] OBD Solutions. *STN1100 Family Reference and Programming Manual*, (Accessed 2019-10-21). `https://www.scantool.net/downloads/98/stn1100-frpm.pdf`.

[27] OBD Solutions. What is obd? Online, (Accessed 2019-10-21). `https://www.obdsol.com/knowledgebase/on-board-diagnostics/what-is-obd/`.

[28] Box. Simplify how you work. Online, (Accessed 2019-11-25). `https://www.box.com/home`.

12 Exploring benefits of using blending splines as transition curves

Tanita Fossli Brustad and Rune Dalmo

Abstract

Track geometry is a fundamental subject in railway construction. With the demand for increased capacity in terms of load and speed the need for suitable transitions between consecutive track sections is highly relevant. Properly constructed transition curves lead to improved travel comfort, increased safety, and reduced wear. The well known clothoid curve is widely used as a transition curve, however, its linear curvature is not sufficiently smooth to meet the requirements for railways carrying high speed trains or heavy haul. Blending spline curves are flexible spline constructions that may be suitable as transition curves because they possess smoothness properties at the end points. This paper demonstrates some selected blending splines applied as transition curves between two existing circular arc segments from the Ofotbanen railway. The results are related to the challenge of improving existing railways without forcing extensive changes to the original track alignment, including comparisons of the new transition curves with the original clothoid curve with respect to smoothness at the end points and the behaviour of the curvature of the curves. This first instance of blending splines as transition curves shows promising results which opens the door to some interesting topics for future work.

12.1 Introduction

Transition curves are an important development in the railway industry. They were introduced to create easement between straight and curved railway sections, and between curved sections of different radii. Properly constructed transition curves increased safety and comfort while travelling, and reduced wear on the rail [1, 2]. Various curves have been recommended as transition curves through the years, among them, the clothoid [3, 4], which has become the most used to this day. The search for new types of transition curves in railway is still relevant, and ongoing, today. This is because of new knowledge related to curve properties [5], where the linear curvature of the clothoid is not smooth enough in the end regions to be optimal. Both in relation to passenger comfort, for speeds above $120\frac{km}{h}$ [6], and in relation to wear on tracks with heavy haul traffic. Some examples of recent research include quartic Bézier spirals [7], log-aesthetic curves [6], a new design of Bloss curves [8], remodelled cubic parabolas [9], sinusoids [10], and polynomial curves [11].

A blending spline curve [12] is a spline construction where local functions are blended together by C^k-smooth blending functions. Blending splines were introduced [13] as an additional tool in computer-aided geometric design (CAGD) with emphasis on user interactivity via its editing capabilities. Because of its flexibility in the blending process, connected to possible local functions and blending functions, the authors suggest that the spline may be suitable as a transition curve.

The motivation behind the work is to find a new railway transition curve that has better properties than the clothoid in regards to smoothness between segments, that can replace the clothoid i an already existing railway without making changes to the original alignment. The need to be able to increase smoothness without making extensive changes to the alignment is relevant for existing railways where large modifications may be invasive or impossible. One example is Ofotbanen, a Norwegian railway line, essential in the transportation of ore from Kiruna in Sweden, and goods between the northern and southern parts of Norway [14]. Ofotbanen is known for its challenging railway geometry, that is vulnerable to wear. Making extensive changes to the alignment of Ofotbanen requires considerable intervention and is impossible in some places because of the mountainous terrain the railway ventures through. The aim of the study is to reveal the suitability of blending splines as a transition curve, replacing one segment (a clothoid) in an already existing railway. A focus is given to making the least amount of intervention to the existing railway by considering the segments attached to the clothoid static, and only allowing the length of the transition curve to increase by moving the entry points along the adjacent segments. In this paper blending splines of various forms are implemented and analysed as transitions between two circular arc segments, of dissimilar radii. The new transition curves are compared to the replaced clothoid by analysing their curvature functions against each other, and drawing conclusions based on smoothness and value of the curvature derivatives.

The remainder of the paper is organised as follows: Section 12.2 gives a short explanation of blending splines and transition curves. In Section 12.3 the methods are described. Section 12.4 presents the obtained results, together with a discussion on the data. Lastly, Section 12.5 draws a conclusion and gives recommendations for future work.

12.2 Preliminaries

The preliminaries section provides an overview of the relevant theory connected to the work. A short presentation of the blending spline is given, together with a brief overview of transition curves in railway.

12.2.1 Blending spline

Blending spline is a collective term of the family of blending-type spline constructions originating from research conducted since 2003. The initial blending spline, called Expo-rational B-spline (ERBS), was presented for the first time in [15] and within two years published in [12, 16]. After that the family grew with new splines derived from the ERBS, including generalized expo-rational B-splines [17] and logistic expo-rational B-splines [18]. The blending spline is a construction where local functions at the knots are blended together by C^k-smooth basis functions, see Figure 12.1. It is defined in [13] as

$$f(t) = \sum_{k=1}^{n} l_k(t) B_k(t), \ t \in (t_1, t_n], \tag{12.1}$$

where $l_k(t)$ are scalar-, vector-, or point-valued local functions defined on (t_{k-1}, t_{k+1}), $t = \{t_k\}_{k=0}^{n+1}$ is an increasing knot vector, and $B_k(t)$ are the blending functions (B-functions). The B-functions can be any function possessing the following set of properties:

1. $B : I \to I (I = [0, 1] \subset \mathbb{R})$,
2. $B(0) = 0$,
3. $B(1) = 1$,
4. $B'(t) \geq 0, \ t \in I$,
5. $B(t) + B(1 - t) = 1, \ t \in I$.

The last property is optional and specifies point symmetry around the point $(0.5, 0.5)$. Figure 12.2 shows a plot of possible B-functions meeting the properties. In this paper we consider a subset of ERBS called the scalable subset, proposed in [13], that uses a specific B-function (the ERB-function) which is C^∞-smooth,

$$B_k(t) = \begin{cases} S_{k-1} \int_0^{\omega_{k-1}(t)} \psi_{k-1}(s) ds & \text{if } t_{k-1} < t \leq t_k, \\ S_k \int_{\omega_k(t)}^{1} \psi_k(s) ds & \text{if } t_k < t < t_{k+1}, \\ 0 & \text{otherwise,} \end{cases} \tag{12.2}$$

where $S_k = (\int_0^1 \psi(s) ds)^{-1}$, $\omega_k(t) = \frac{t - t_k}{t_{k+1} - t_k}$, and $\psi_k(s) = e^{-\beta \frac{|s-\lambda|^{\alpha(1+\gamma)}}{(s(1-s)^\gamma)^\alpha}}$ with the intrinsic parameters restricted to $\alpha > 0$, $\beta > 0$, $\gamma > 0$, and $0 \leq \lambda \leq 1$. We also note that the ERBS shares the minimal support and partition of unity properties of the linear B-spline, which means that only two functions are blended together in every knot interval (i.e. between two knots) and that the two B-functions in the interval sum up to 1 for a given t. By applying the two previous properties the simplified formula of (12.1), over one knot interval, becomes

$$\begin{aligned} f(t) &= B_k(t) l_k(t) + B_{k+1}(t) l_{k+1}(t) \\ &= (1 - B_{k+1}(t)) l_k(t) + B_{k+1}(t) l_{k+1}(t) \\ &= l_k(t) + B_{k+1}(t)(l_{k+1}(t) - l_k(t)). \end{aligned} \tag{12.3}$$

For the experiments in this paper, we use parametric curves as local functions $l_k(t)$. The specific choices are presented and explained in Section 12.3.2.

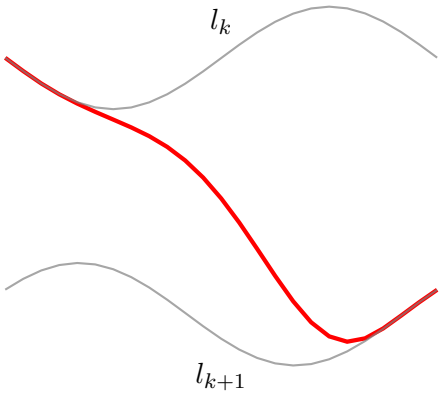


Figure 12.1: Two local curves (l_k and l_{k+1}) blended together to create a blending spline (red curve).

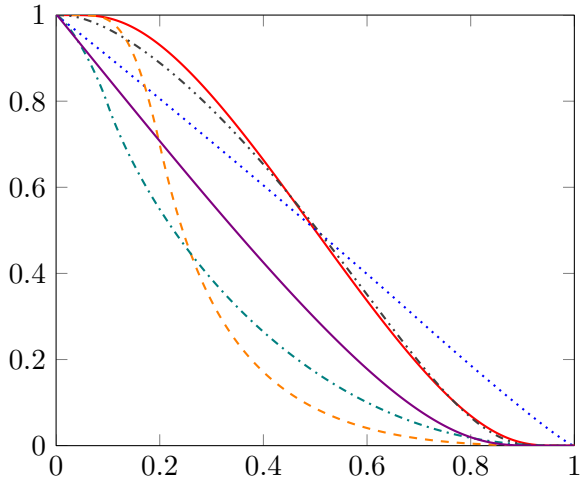


Figure 12.2: Plots of possible blending functions.

12.2.2 Transition curves

In railway design the transition curves fulfil the role of connecting straight sections to curved sections, and curved sections of different radii [19]. The task of the transition curve is to gradually decrease/increase the radius of the curvature from radius, R_1 to R_2, and to provide a change in superelevation following the same behaviour as the curvature. These properties are important in order to counteract a sudden jerk in the centrifugal forces, and instead gently introduce them over the course of the transition curve. The main advantages of using transition curves are [20, 21]:

- Providing a comfortable ride for passengers.
- Providing a safer ride for passengers.
- Enabling the vehicle to drive at a higher speed.
- Reducing wear and tear on wheels and rails, decreasing maintenance and repair costs.

As mentioned in the introduction, the clothoid (Figure 12.3) is the most used transition curve today. One of the first discoveries of the clothoid as a transition curve in railway was by Arthur Talbot in [22], who was among the first to approach the transition problem mathematically. A clothoid (of length l and end radius r) is a spiral defined parametrically as

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} C(t) \\ S(t) \end{pmatrix},$$

where $C(t)$ and $S(t)$ are the Fresnel integrals,

$$\begin{aligned} C(t) &= \frac{1}{a}\int_0^{\hat{t}} \cos\left(\frac{\pi}{2}u^2\right) du, \\ S(t) &= \frac{1}{a}\int_0^{\hat{t}} \sin\left(\frac{\pi}{2}u^2\right) du, \end{aligned} \tag{12.4}$$

with $a = \sqrt{\frac{1}{\pi r l}}$ being a scaling factor, and $\hat{t} = at$, $-\infty < t < \infty$. A main reason for the clothoid's popularity as a transition curve is its curvature, which changes linearly with curve length.

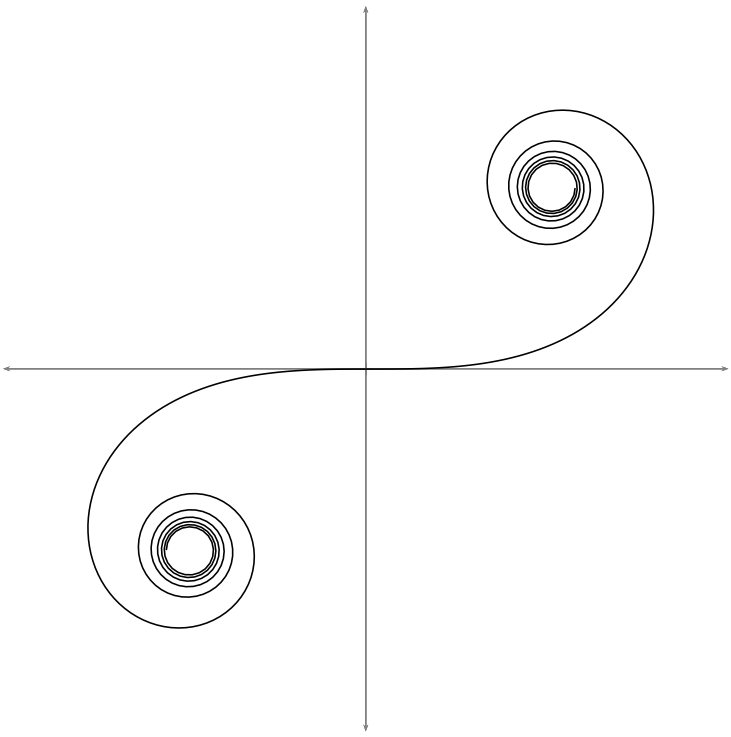


Figure 12.3: The clothoid spiral.

12.3 Method

This section describes the parameters and equations used in setting up the original (existing) railway curve example, relevant parameters for creating the blending spline transition curve and how they are modified, and an explanation of the criteria for comparing the blending spline against the clothoid from the original railway track.

12.3.1 The original railway curve

The original railway curve consists of two arc segments, of different radii, connected by a clothoid transition curve. The example is extracted from the Norwegian railway line, Ofotbanen. Parameters for the segment are presented in Table 12.1, with the layout plotted in Figure 12.4.

Table 12.1: Segment parameters.

	Arc 1	**Clothoid**	**Arc 2**
Start Radius (m)	-	401	315
End Radius (m)	401	315	-
Length (m)	135.335	60	103.481

The clothoid is implemented by exploiting a power series expansion of the integrals in (12.4), given in [23] as,

$$C(t) = \frac{1}{a}\sum_{i=0}^{\infty}\frac{(-1)^i(\frac{\pi}{2})^{2i}\hat{t}^{4i+1}}{(2i)!(4i+1)},$$
$$S(t) = \frac{1}{a}\sum_{i=0}^{\infty}\frac{(-1)^i(\frac{\pi}{2})^{2i+1}\hat{t}^{4i+3}}{(2i+1)!(4i+3)},$$

where the variable a is set up for the clothoid to join two arcs, instead of an arc and a line.

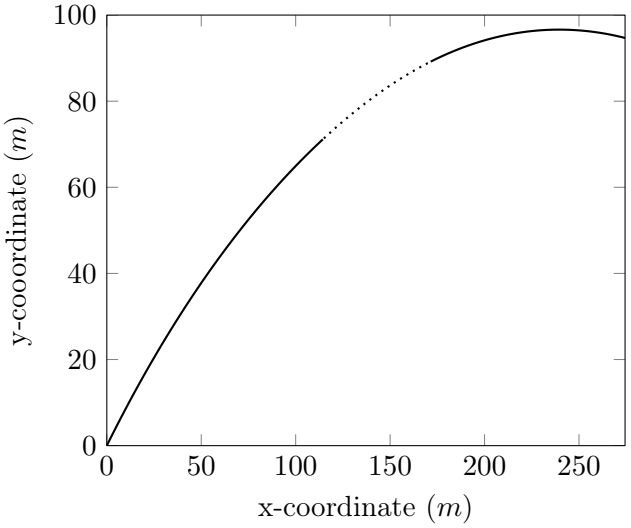


Figure 12.4: Layout of the original railway curve. Two circular arc segments (solid) connected by a clothoid segment (dotted).

12.3.2 Blending spline fitting

The blending spline curve is created for one knot interval, using expression (12.3), in order to preserve maximum flexibility over the entire transition curve. It is relevant to adjust three parameters in this study: entry points, local curves, and B-functions.

Entry points

The entry points for the blending spline on the adjacent circular segments are tested for three positions: $[p_1, q_1]$, $[p_2, q_2]$, and $[p_3, q_3]$, shown in Figure 12.5. $[p_1, q_1]$ is the same entry point as the clothoid, $[p_2, q_2]$ is placed $\frac{1}{4}$ of the circular segment's curve length from $[p_1, q_1]$, and $[p_3, q_3]$ is in the middle of the circular segment. The motivation for choosing entry points only on half of the circular segments is to account for the possibility of also creating blending splines on both sides of the example segments, so that future work can include replacement of multiple clothoid segments.

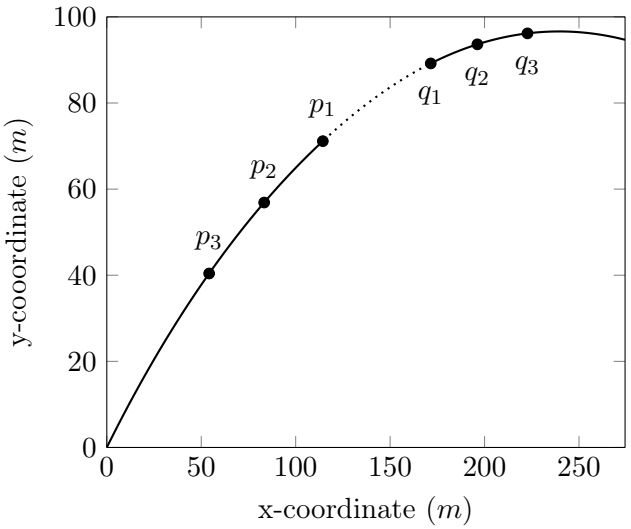


Figure 12.5: The three entry points for the blending spline transition curve on the circular segments.

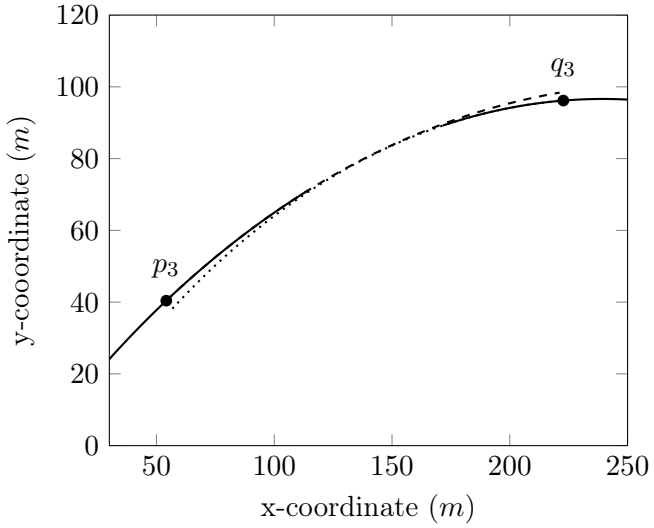


Figure 12.6: An example of two local arc curves (dashed and dotted) starting in $[p_3, q_3]$.

Local curves

The curves that can be used as local curves in a blending spline transition curve have to comply with three criteria. They must have the correct curvature in one end, be possible to connect to adjacent segments with a smoothness equal to or higher than the clothoid, and be scalable in length. The most intuitive curves to use as local curves in the blending spline, for this transition case, are circular arc segments, given by

$$c(t) = (R\sin(t), R(1 - \cos(t))), \quad 0 < t < 2\pi,$$

with the same radius R as the original railway arcs (401 and 315 in this case). They have the correct curvature and smoothness in at least one end, are easy to scale in size, and easy to place correctly against the adjacent segments. An example of two local arc curves starting at entry point $[p_3, q_3]$ is shown in Figure 12.6.

In addition to the circular arcs, Bézier curves are also tested as local curves in the blending spline. Bézier curves are expressed as

$$c(t) = \sum_{i=0}^{n} c_i b_{i,n}(t),$$

where c_i are control points, and $b_{i,n}(t)$ are Bernstein polynomials of degree n. Compared to the arcs, the Bézier curves are more flexible since their shapes can be altered through control points. The Bézier curves in this research are created by approximating the previously described arcs via performing a Taylor expansion of a point and d derivatives in the entry points (which becomes the start of the Bézier curves). To achieve sufficient smoothness at the entry points, while also having the opportunity to manipulate the control point at the opposite end, the degree of the Bézier curves should be at least four ($d > 3$). For the local Bézier curve examples in this paper, the number of derivatives extracted in the start point is 4, which means $d = 4$ and 5 control points are placed. To keep a smooth connection to the adjacent circular arc segments, while at the same time experimenting with the placement of the local curves, the first four control points, from the start of the curves, are fixed and the last control point is moved to the start of the local curve in the opposite end. A visual explanation is given in Figure 12.7. This is expected to give different results than using arcs as local curves, and show if placing the local curves closer to the original clothoid has a significant effect on the curvature.

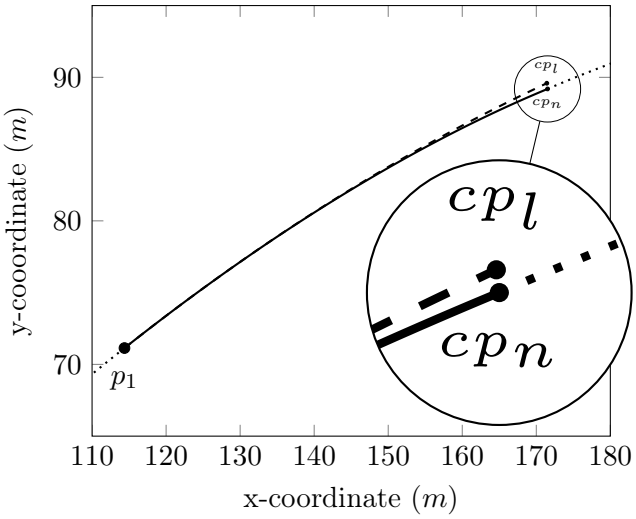


Figure 12.7: An example of how the last control point cp_l is moved to the start of the local curve in the opposite end cp_n, for the left side local Bézier curve starting in p_1. The dashed curve is the first Bézier approximation of the arc, and the solid curve is the one used in the blending.

B-function

The B-funtion used in the study is the one presented in (12.2), an ERB-function. This function can be varied through the following parameters: $\alpha > 0$, $\beta > 0$, $\gamma > 0$, and $0 \leq \lambda \leq 1$, where $\alpha, \gamma \in \mathbb{N}$ and $\beta, \lambda \in \mathbb{R}$. The default ERB-function is given by $\alpha = 1$, $\beta = 1.0$, $\gamma = 1$, and $\lambda = 0.5$, seen in Figure 12.8. α and γ are called the asymmetric tightening parameters. Increasing them tightens the function so that its shape goes towards a linear function on parts of the domain. Figure 12.9 shows plots of ERB-functions with varying tightening values. β is called the slope parameter. It adjusts the steepness of the function. Figure 12.10 shows ERB-functions with varying slope values. λ is called the balance parameter. Varying it moves the function along the x-axis. Figure 12.11 gives plots of ERB-functions with varying balance values. For the comparisons in this paper the parameters are adjusted independently, modifying one while keeping the default values for the rest. The tested values for each parameter are given in Table 12.2.

Table 12.2: ERB-function parameter values used in the comparisons.

	α	γ	β	λ
Default	1	1	1.0	0.5
Tightening 1	4	1	1.0	0.5
Tightening 2	10	1	1.0	0.5
Slope 1	1	1	8.0	0.5
Slope 2	1	1	50.0	0.5
Balance 1	1	1	1.0	0.2
Balance 2	1	1	1.0	0.8

12.3.3 Comparison criteria

To evaluate the blending spline as a transition curve it is compared against the original clothoid with the following criteria, presented in [6].

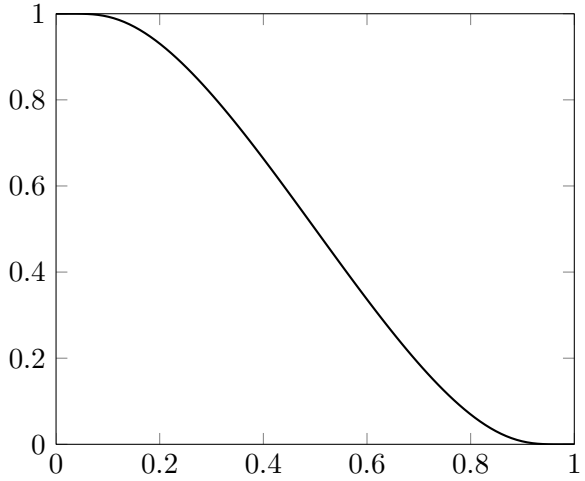


Figure 12.8: Plot of the default ERB-function.

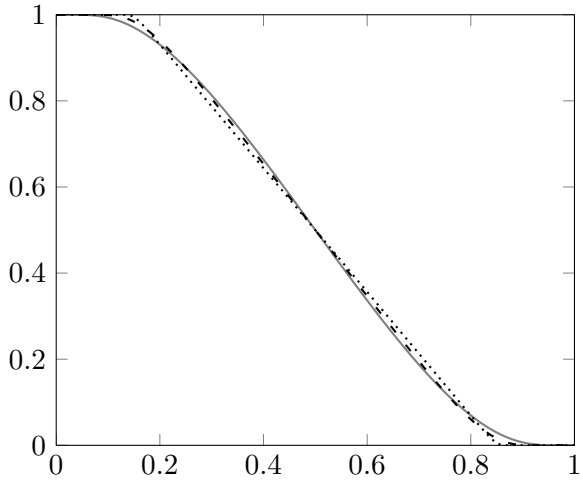


Figure 12.9: Plots of ERB-functions with varying tightening parameters. The solid function is the default ERB ($\alpha = \gamma = 1$), the dashed function has $\alpha = 2$ and $\gamma = 1$, and the dotted function has $\alpha = 10$ and $\gamma = 1$.

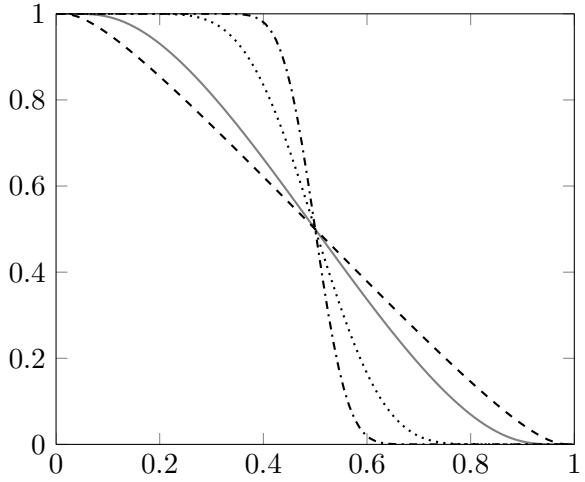


Figure 12.10: Plots of ERB-functions with varying slope parameters. The solid function is the default ERB ($\beta = 1$), the dashed function has $\beta = 0.2$, the dotted function has $\beta = 10$, and the dashdotted function has $\beta = 50$.

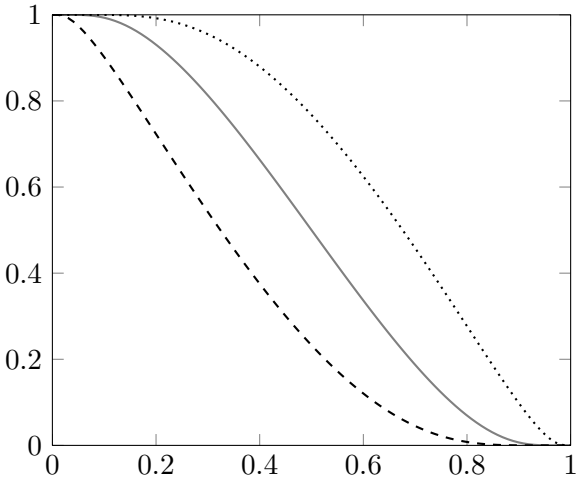


Figure 12.11: Plots of ERB-functions with varying balance parameters. The solid function is the default ERB ($\lambda = 0.5$), the dashed function has $\lambda = 0.2$, and the dotted function has $\lambda = 0.8$.

1. Common connection point: the coordinates of the two curves must be equal in the connection points.
2. Common tangent: the tangents of the two curves must be equal in the connection points.
3. Equal radius of curvature: the radii of the two curves must be equal in the connection points (eliminates discontinuities in the form of jumps in the curvature diagram).
4. Common tangent of curvature functions: the first derivative of the curvature functions of the two curves must be equal in the connection points (eliminates discontinuities in the form of breaks in the curvature diagram).
5. Equal radius of curvature of curvature functions: the second derivative of the curvature functions of the two curves must be equal in the connection points (needed for extremely high-speed railways).

The clothoid fulfils criterion 1 to 3, which means that it has discontinuities in the form of breaks in the curvature diagram, see Figure 12.12.

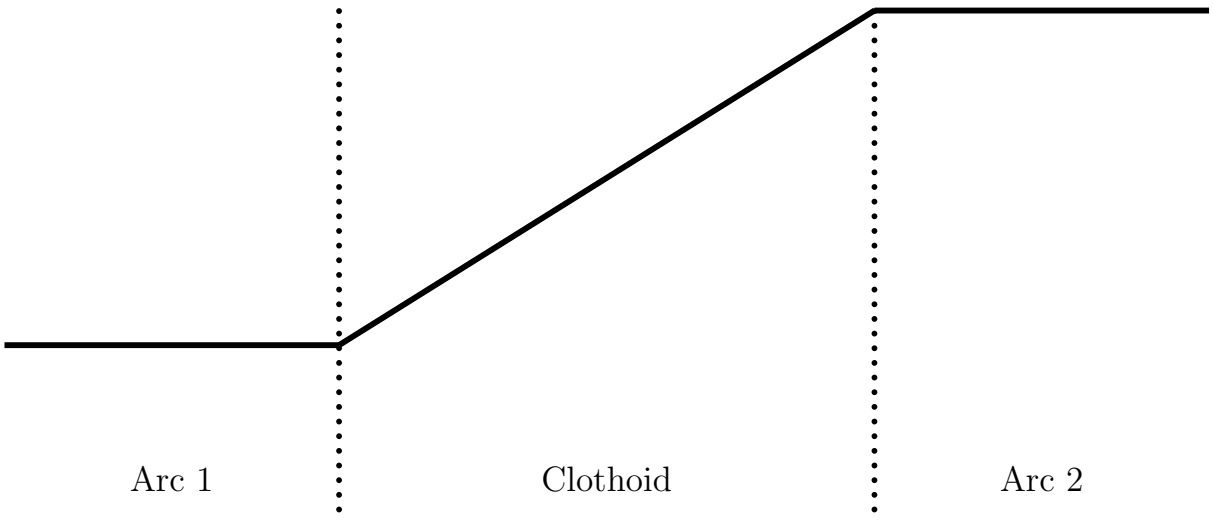


Figure 12.12: Curvature diagram of a clothoid connected to two arc segments of different radii.

12.4 Results and discussions

The results are divided in two, based on the type of local curves used in the blending spline: arcs and Béziers. For each of the local curves the curvature plots of the blending splines are analysed and discussed, and compared with the clothoid, for varying entry points and ERB-function parameters.

12.4.1 Arcs as local curves

Curvature plots for the clothoid and three blending spline transition curves (with entry points p1, p2, and p3), using circular arcs as local curves and a default ERB-function for blending, can be seen in Figure 12.13. The dashed function is the clothoid curvature, while the solid functions are blending spline curvatures ($[p_1, q_1]$ in blue, $[p_2, q_2]$ in red, and $[p_3, q_3]$ in black). From the figure it can be seen that changing the entry points does not have much impact on the behaviour of the blending spline. The main difference is the length of the curves ($L = 60\ m$ for $[p_1, q_1]$, $L = 120m$ for $[p_2, q_2]$, and $L = 180m$ for $[p_3, q_3]$).

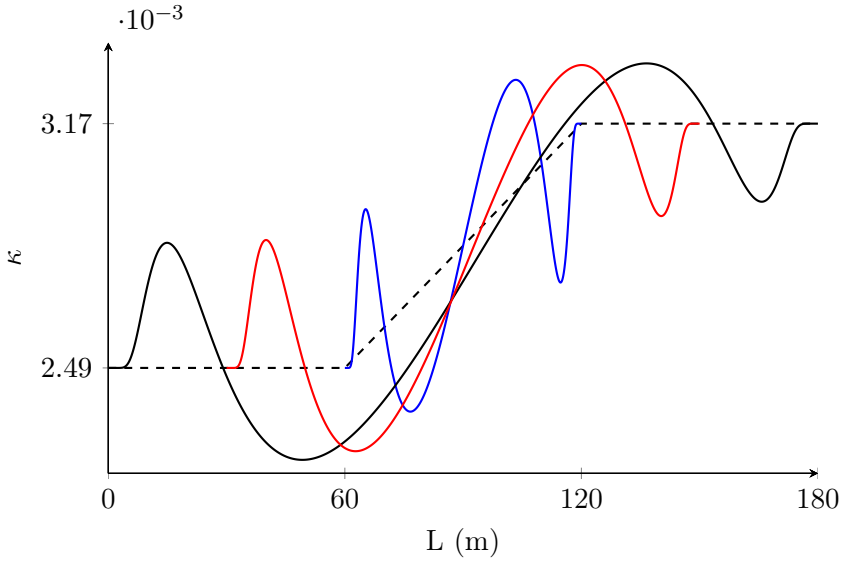


Figure 12.13: Curvature plots of the clothoid with parts of the circular segments (dashed) and default ERB-function blending splines, with local arc curves, for the three entry points $[p_1, q_1]$ (blue), $[p_2, q_2]$ (red), and $[p_3, q_3]$ (black).

Compared to the clothoid's linear curvature, the blending splines have five inflection points in their curvature graphs, which is not optimal and should be attempted smoothed out. A positive characteristics of the blending splines is that they go smoothly towards the end points, where the clothoid has a discontinuity in connection to the constant curvature of the adjacent segments, as described in Figure 12.12. This leads us to believe that the blending spline, at least with a default ERB-function, satisfies additional criteria (4 and 5 presented in Section 12.3.3) when compared to the clothoid. To test our hypothesis, graphs of the first and second derivatives of the curvature for the blending spline in $[p_3, q_3]$ are printed. Since the behaviour is fairly familiar for the three blending splines in the end points only one is checked. The result can be seen in Figure 12.14, where the constant first derivative of the clothoid's curvature (dashed) is plotted against the first (solid) and second (dotted) derivatives of the blending spline's curvature. From the figure it can be observed that the first and second derivatives of the blending spline go to zero in the ends. These conditions are exactly what is needed to fulfil criteria 4 and 5. Hence, the blending spline is better than the clothoid when it comes to smoothness in the connection points to the adjacent segments.

From the previous results an assumption can be made that any blending spline with a smoother curvature diagram towards the end points, than the default ERB-funtion blending spline, will meet all the criteria in Section 12.3.3. Thus, the focus of the discussion when varying ERB parameters will be on smoothing the curvature along the blending spline. In the adjustment of the ERB parameters, the curves that are studied have entry points $[p_3, q_3]$.

By varying the tightening parameters, α and γ, as described in Table 12.2, the resulting curvature diagrams, of the blending splines, become as plotted in Figure 12.15. From the plots it can be observed that by increasing α, the smoothness in the end points increases and the bumps become more prominent, with higher amplitude of the extreme values. By varying the slope parameter, β, as described in Table 12.2, the resulting curvature diagrams, of the blending splines, become as plotted in Figure 12.16. From the plots it can be observed that by increasing β, the bumps are shifted towards the middle of the graph and the smoothness increases in the end points. By varying the balance parameter, λ, as described in Table 12.2, the resulting curvature diagrams, of the blending splines, become as plotted in Figure 12.17. From the plots it can be observed

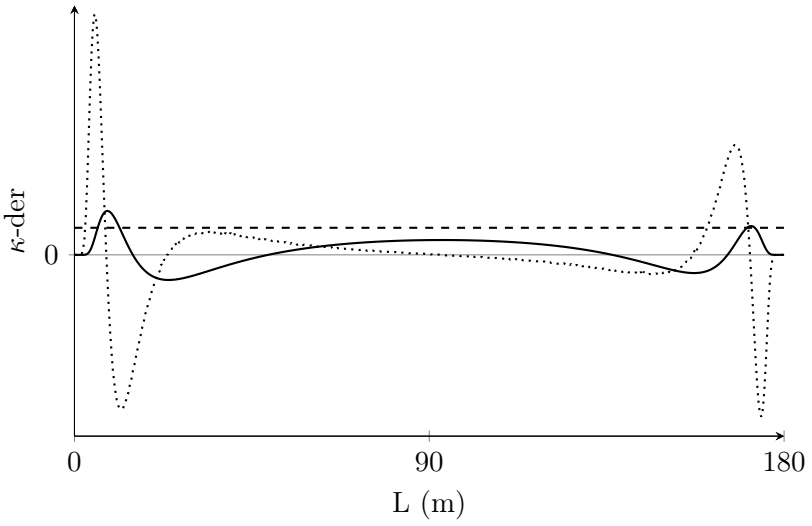


Figure 12.14: Plots of the 1st derivative of the curvature for the clothoid (dashed) and default ERB-function blending splines, with local arc curves, in $[p_3, q_3]$ (solid), and the 2nd derivative of the curvature for the blending spline, scaled by $\frac{1}{10}$ (dotted).

that by moving the balance of the ERB-function, left or right, the smoothness increases in one end and decreases in the other. The same is true for the bumps: in one end they are smoothed, while in the other they become more prominent.

Attempting to smooth the curvature by varying the ERB parameters is not straightforward. It seems that the slope parameter can dampen the fluctuations to some extent, which is also the case in one end when varying the balance parameter. However, all in all, the ERB parameters do not change the characteristics of the curvature, they only affect the extreme values and shift the amplitude of the bumps.

12.4.2 Bézier as local curves

Using Bézier curves as local curves and starting with a default ERB-function for blending, the curvature for the three blending spline transition curves together with the clothoid, and the blending spline in $[p_3, q_3]$ from Figure 12.13, can be seen in Figure 12.18. As before, the dashed function is the clothoid curvature, while the solid functions are blending spline curvatures. In addition, a blending spline (dotted) with local arc curves in $[p_3, q_3]$ is plotted for comparison. From the figure it can be seen that using Bézier curves instead of arcs, as local curves, yields blending splines with curvatures that lie closer to the clothoid, while still having a smooth transition to the end points. With this observation we conclude that the three first criteria from Section 12.3.3 are fulfilled, and that there is a possibility that more criteria are met.

In Figure 12.19 the first (solid) and second (dotted) derivative of the curvature for the blending spline in $[p_1, q_1]$ are plotted, together with the first derivative of the curvature for the clothoid (dashed). From the figure it can be observed that the first derivative of the blending spline goes to zero in the ends, while the second derivative ends a little above and below zero in the two end points. This means that criterion 4 is fulfilled but not criterion 5, since there will be a small jump in the connection points for the second derivative function. However, the blending spline with local Bézier curves is still better than the clothoid when it comes to smoothness in the connection points to the adjacent segments. One way of increasing the smoothness, so that it also fulfils criterion 5 could be to raise the degree of the Bézier curves from $d = 4$ to $d = 5$, and fixing the first four control points.

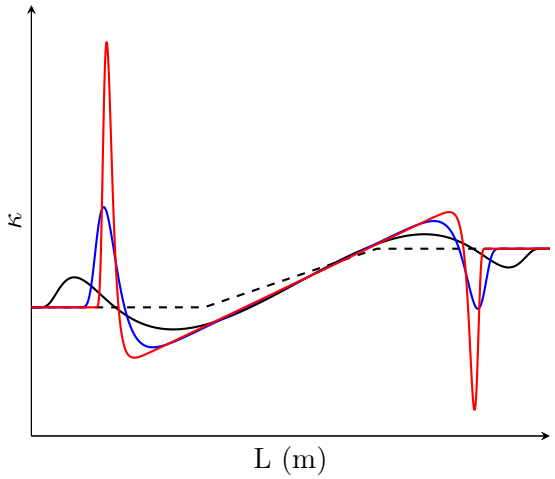


Figure 12.15: Curvature plots of the clothoid with parts of the circular segments (dashed) and blending splines in $[p_3, q_3]$ for tightening parameters $\alpha, \gamma = 1, 1$ (black, default), $\alpha, \gamma = 4, 1$ (blue), and $\alpha, \gamma = 10, 1$ (red).

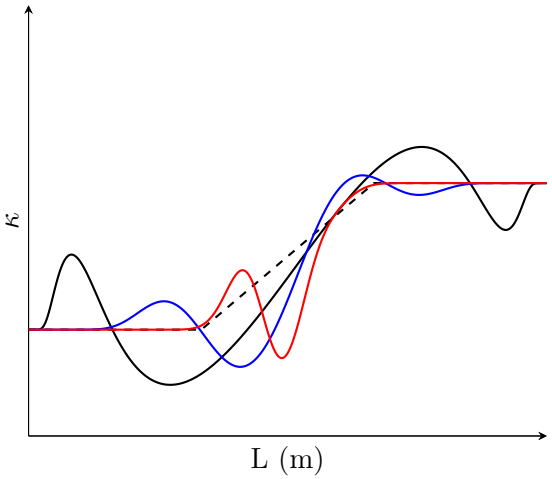


Figure 12.16: Curvature plots of the clothoid with parts of the circular segments (dashed) and blending splines in $[p_3, q_3]$ for slope parameters $\beta = 1.0$ (black, default), $\beta = 8.0$ (blue), and $\beta = 50.0$ (red).

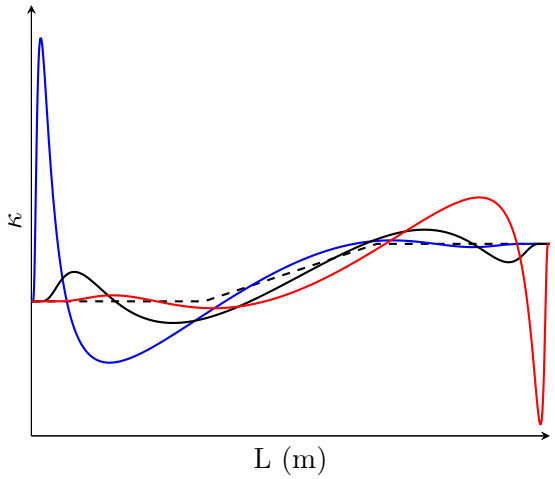


Figure 12.17: Curvature plots of the clothoid with parts of the circular segments (dashed) and blending splines in $[p_3, q_3]$ for balance parameters $\lambda = 0.5$ (black, default), $\lambda = 0.2$ (blue), and $\lambda = 0.8$ (red).

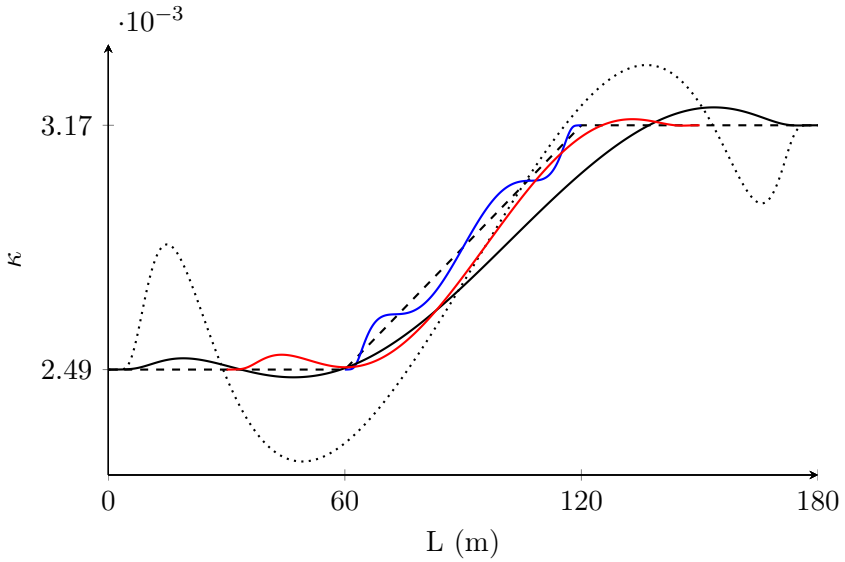


Figure 12.18: Curvature plots of the clothoid with parts of the circular segments (dashed) and default ERB-function blending splines, with local Bézier curves, for the three entry points $[p_1, q_1]$ (blue), $[p_2, q_2]$ (red), and $[p_3, q_3]$ (black). In addition, the curvature of a blending spline with local arc curves, in $[p_3, q_3]$, is plotted (dotted).

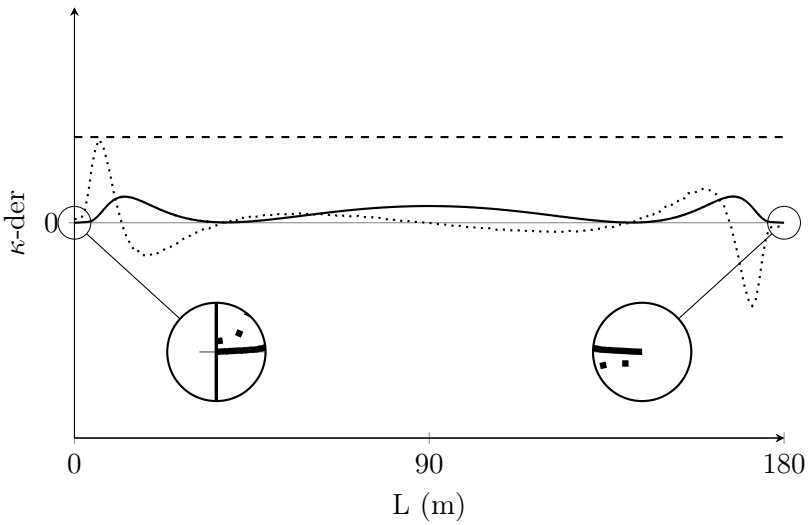


Figure 12.19: Plots of the 1st derivative of the curvature for the clothoid (dashed) and default ERB-function blending splines, with local Bézier curves, in $[p_1, q_1]$ (solid), and the 2nd derivative of the curvature for the blending spline, scaled by $\frac{1}{10}$ (dotted).

The same assumption is made here, as in the discussion with arc curves, that any blending spline with a smoother curvature diagram towards the end points, than the default ERB-function blending spline, will fulfil criteria $1-4$ in Section 12.3.3. Thus, the focus when varying ERB parameters will be on smoothing the curvature along the blending spline. In the variation of the ERB parameters, the curves that are studied have entry points $[p_1, q_1]$, because the curvature lies closest to the clothoid in Figure 12.18. The parameters are varied in the same manner as with the blending spline with local arc curves, described in Table 12.2. In Figure 12.20 the tightening parameters are varied, in Figure 12.21 the slope parameter is varied, and in Figure 12.22 the balance parameter is varied. The results that emerge are very similar to the results obtained for the blending spline with local arc curves: by varying the tightening parameters, the smoothness in the end points increases and the bumps become more prominent, with higher extreme values; by varying the slope parameter, the bumps are shifted towards the middle and the smoothness increases in the ends; and by varying the balance parameter, the smoothness increases in one end and decreases in the other, and the bumps are smoothed in one end, while becoming more prominent in the other.

So, with these results a conclusion can be drawn, again, that the ERB parameters do not change the behaviour of the curvature, they only affect the extreme values and shift the amplitude of the bumps. However, a new result is obtained here showing that by re-shaping the local curves the characteristics of the curvature can be changed, causing smaller fluctuations along the function.

12.5 Concluding remarks

In this paper a family of splines, blending splines, have been applied and analysed as transition curves in an already existing railway. Three segments from the Ofotbanen railway were extracted, two circular arcs connected by a clothoid. The goal was to make the least amount of intervention to the existing railway, by keeping the placement of the arcs unchanged and only allowing the transition curve entry points to move. Under this constraint, the blending splines were compared to the clothoid with regards to smoothness in the end points and the shape of the curvature graph. The results showed that the blending splines, with appropriately selected local curves, were smoother than the clothoid in the ends. However, the curvature plots of the blending splines had fluctuations along its length, which may not be optimal in a transition case if the fluctuations are large. An observation was made that the choice of ERB parameters and local curves had a large impact on these fluctuations. The choice of local curve type and shape had even larger impact. By using Bézier curves as local curves, with a default ERB-function, the fluctuations were relatively small and the curvature function was close to a monotonically increasing function, which seems to be the norm for transition curves, although there exists at least one transition curve that is not, the Wiener Bogen [24]. The Wiener Bogen uses a non-monotonically increasing curvature curve together with a monotonically increasing superelevation ramp, which gives better transition properties than the clothoid. Whether the blending spline curves can be used in the same manner as the Wiener Bogen remains to be tested, and may yield interesting results if it works since the blending splines are smoother in the end points than the clothoid. A different approach, which is very relevant for future work, is to test other local curves that may dampen the fluctuations more (or even remove them). From the results on local Bézier curves, it will be of interest to raise the degree of the Bézier so that more than one control point in the ends can be manipulated. There are strong indications that these extra degrees of freedom can be utilized to smooth the curvature function more than for 4th degree Bézier curves where only one control point is manipulated. In this work, the control point was moved manually in order to test the hypothesis. However, there are indications that it would be possible to develop an algorithmic solution to the placement of the control points in the local Bézier curves, under the constraints of a given optimal curvature. This, together with inspecting ERB-functions where several parameters are modified, with respect to the characteristics of the resulting curvature, is out of scope in the current paper and is considered for future research.

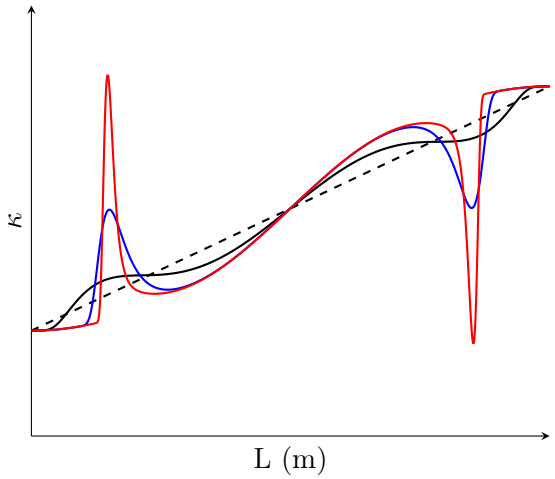


Figure 12.20: Curvature plots of the clothoid (dashed) and blending splines in $[p_1, q_1]$ for tightening parameters $\alpha, \gamma = 1, 1$ (black, default), $\alpha, \gamma = 4, 1$ (blue), and $\alpha, \gamma = 10, 1$ (red).

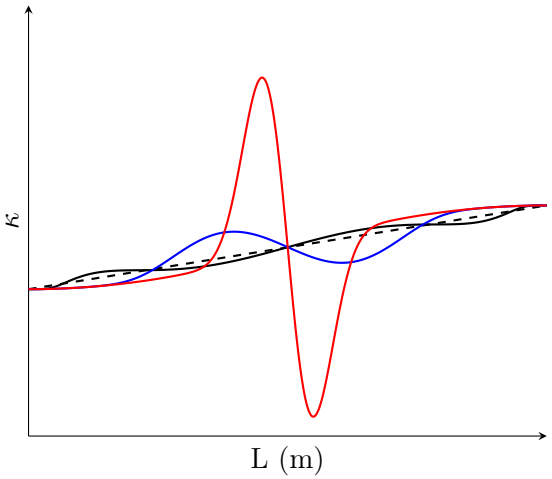


Figure 12.21: Curvature plots of the clothoid (dashed) and blending splines in $[p_1, q_1]$ for slope parameters $\beta = 1.0$ (black, default), $\beta = 8.0$ (blue), and $\beta = 50.0$ (red).

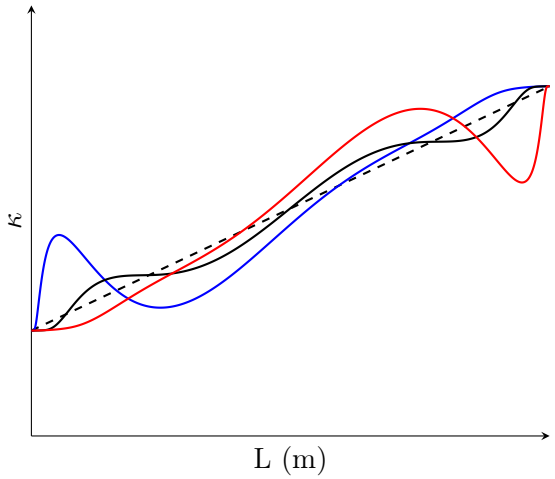


Figure 12.22: Curvature plots of the clothoid (dashed) and blending splines in $[p_1, q_1]$ for balance parameters $\lambda = 0.5$ (black, default), $\lambda = 0.2$ (blue), and $\lambda = 0.8$ (red).

References

[1] J. Glover. Transition curves for railways. *Proc. Inst. C. E.*, 140, 1900.

[2] A. W. Miller. The transition spiral. *Australian Surveyor*, 7(8):518–526, 1939.

[3] R. C. Archibald. Euler integrals and euler's spiral, sometimes called fresnel integrals and the clothoide or cornu's spiral. *American Mathematical Monthly*, 25:276–282, 1918.

[4] Arthur Lovat Higgins. *The transition spiral and its introduction to railway curves with field exercises in construction and alignment.* Van Nostrand company, New York, 1922.

[5] Orhan Baykal. Concept of lateral change of acceleration. *Journal of Surveying Engineering*, 122:132–141, 08 1996.

[6] Abdullah Arslan, Ergin Tari, Rushan Ziatdinov, and Rifkat Nabiyev. Transition curve modeling with kinematical properties: Research on log-aesthetic curves. *Computer-Aided Design and Applications*, 11:509–517, 10 2014.

[7] A. Ahmad and J. M. Ali. G3 transition curve between two straight lines. In *2008 Fifth International Conference on Computer Graphics, Imaging and Visualisation*, pages 154–159, Aug 2008.

[8] Constantin Ciobanu. Bloss transition - a short design guide. *PWI Journal*, 133:14–18, 01 2015.

[9] SA Shebl. Geometrical analysis of non-linear curvature transition curves of high speed railways. *Asian Journal of Current Engineering and Maths*, 5(4):52–58, Aug 2016.

[10] A. Pirti, M.A. Yücel, and T. Ocalan. Transrapid and the transition curve as sinusoid. In *Tehnički vjesnik*, volume 23, pages 315–320, 2016.

[11] K. Zboinski and P. Woznica. Optimization of polynomial transition curves from the viewpoint of jerk value. *Archives of Civil Engineering*, 63, 01 2017.

[12] Lubomir T. Dechevsky, Arne Lakså, and Børre Bang. Expo-rational b-splines. *International Journal of Pure and Applied Mathematics*, 27, 2006.

[13] Arne Lakså. *Basic properties of Expo-Rational B-splines and practical use in Computer Aided Geometric Design.* PhD thesis, University of Oslo, 2007.

[14] Bane Nor. Ofotbanen. (Online) `https://www.banenor.no/jernbanen/banene/ofotbanen/`. Accessed: 2019-07-03.

[15] L. T. Dechevsky. Expo-rational B-splines. Communicated at the Sixth International Conference on Mathematical Methods for Curves and Surfaces, Tromsø, Norway, 2004.

[16] Lubomir T. Dechevsky, Arne Lakså, and Børre Bang. Exploring expo-rational b-splines for curves and surfaces. In M. Dæhlen, K. Mørken, and L. L. Schumaker, editors, *Mathematical Methods for Curves and Surfaces*, pages 253–262. Nashboro Press, 2005.

[17] Lubomir T. Dechevsky, Børre Bang, and Arne Lakså. Generalized expo-rational b-splines. *International Journal of Pure and Applied Mathematics*, 57(6):833–872, 2009.

[18] Lubomir T. Dechevsky and Peter Zanaty. Smooth gerbs, orthogonal systems and energy minimization. *AIP Conference Proceedings*, 1570(1):135–162, 2013.

[19] J.S. Mundrey. *Railway Track Engineering, Fourth Edition.* Tata McGraw-Hill, 1988.

[20] Thomas F. Hickerson. *Route Location And Design.* McGraw-Hill, 5 edition, 1964.

[21] Martin Lipičnik. New form of road/railway transition curve. *Journal of Transportation Engineering*, 124(6):546–556, 1998.

[22] Arthur Talbot. The railway transition spiral. Reprinted from The Technograph No. 13, 1899. U. Illinois.

[23] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*. Dover Publications, Inc., New York, USA, 1972.

[24] Robert Wojtczak. *Charakterystyka krzywej przejściowej Wiener Bogen®*. Publishing House of Poznan University of Technology, 2017.

13 Alternative representation of railway track geometry

Aleksander Pedersen, Tanita Fossli Brustad, Børre Bang and Rune Dalmo

Abstract

Track geometry is important in the design, construction and maintenance of railways. Increasing comfort and faster travel times, as well as decreasing the amount of maintenance needed in railroad tracks, are examples of criteria when rail geometry is constructed or modified. There are several existing methods and commercial tools available which can identify such criteria, such as lateral change of acceleration (LCA). By applying available new interpolation methods to represent the existing chain of curve elements in a railroad construction, we exploit the geometrical properties to study the positional and curvature continuity. In this paper we investigate the composition of curve segments that make up a railway track, and compare the original track geometry and a spline curve representation.

13.1 Introduction

Over the years railway geometry has changed in line with the development of trains and rail transport. Early railway tracks usually consisted of a connection of straight sections and circular arcs. With the low speed of trains and the wide radius of the circular sections the transition between straight and circular sections was considered easy and safe with regards to the motion of the train [1], and was well within the scope of current surveying technologies at that time. When the evolution of railway tracks made the radius of the circular tracks smaller and the speed of trains increased, transitions became jerky and unpleasant, and sometimes even dangerous [1, 2]. This realisation opened up the door to easement curves, or transition curves as they are more known by, that made the transition from straight track to curved track safer and more comfortable even when higher speed was involved [1, 2, 3]. A well-known and much used transition curve is the clothoid curve [4, 5], although other curves, like the cubic parabola and lemniscate [3], were also utilized. For a long time the clothoid curve was a sufficient transition curve in railway, and still is today. However, with the introduction of the lateral change of acceleration (LCA) function [6] in the last quarter of the 20th century, new design criteria were introduced to improve vehicle-road properties in railway. The new criterion of the LCA showed that the high speed requirements in modern rail transport made the popular clothoid curve undesirable because of its discontinuity in LCA at the start and end. This has given rise to a new generation of transition curves [7, 8, 9, 10, 11] with performance superior to the clothoid curve when it comes to LCA. In this paper we consider a composition of curve elements making up a railway track and compare this geometry with a B-spline curve representation with regards to curvature and LCA. We are interested in finding alternative representations of transition curves, which improves the shortcomings of the clothoids, and propose the use of B-splines for this purpose. The railroad geometry that we are using is a list of segments that constitutes the railway known as Ofotbanen, stretching from Narvik to the border of Sweden. The list consists of three different types of curves; lines, arcs and clothoids (euler spiral). Ofotbanen [12] is a railway line with a main traffic of 10-12 daily freight trains hauling iron ore from Sweden to Narvik. The line is also important for transporting goods with container trains and there are two daily passenger trains. Ofotbanen was completed in 1902 and today's track follows that original track to a large extent. A fascinating fact about Ofotbanen, and one of the main reasons for choosing it in this research, is the variety of compositions of the segments, ranging from line-clothoid-arc to arc-clothoid-arc and also arc-clothoid-clothoid-arc. These combinations of curve segments in connection with the heavy load can be related to maintenance problems. We show some results in the comparison part of this paper.

13.2 Theory and method

13.2.1 Lateral change of acceleration

The lateral change of acceleration (LCA) is the change of the resultant acceleration occurring along the curve normal with respect to time. The function is expressed in [6] as

$$z = \frac{pv}{\sqrt{(u^2+p^2)}}\Big(3ka_t + v^2\frac{dk}{dl} - \frac{(kv^2u+gp)}{(u^2+p^2)}\frac{du}{dl}\Big) \tag{13.1}$$

where p is the horizontal width of the platform m, v is the velocity of the train $\frac{m}{s}$, u is the superelevation which is the elevation of the outer rail according to the inner rail m, k is the curvature along the curve $\frac{1}{m}$, a_t is the tangential acceleration $\frac{m}{s^2}$, and g is the gravity constant $9.81\frac{m}{s^2}$. When comparing curves with a basis in the LCA function there are three criteria [9] that are used to compare them by. The continuity of the LCA is the most important of the three criteria. This is because discontinuities (in the form of jumps) affect travel comfort and

cause wear on wheels and rails. Any curve without discontinuity in LCA is considered superior to those that have. If none of the curves have discontinuities they are considered equivalent in regards to criterion 1, and criterion 2 has to be used. Criterion 2 compares the extreme value (largest absolute value) of the LCA function, for each curve, to a boundary value that it should be smaller than. Any curve satisfying this condition is superior to those that do not. If the curves are equivalent after criterion 2 as well the third criterion can be considered. Criterion 3 looks at discontinuities in the form of breaks at the start and end of the transition curve. Any curve without breaks in the end points is superior to those that have. In this work we have restricted the velocity in the LCA function to be constant. A constant velocity means that the LCA becomes a scaled version of the derivative of the curvature. This means that we can look at the curvature and the derivative of the curvature and get a good indication of the behaviour in regards to LCA.

13.2.2 B-Splines

B-splines [13] are commonly used in computer representations of curves and surfaces. A parametric curve can be represented on B-spline form as

$$c(t) = \sum_{i=0}^{n} c_i B_i^d t \tag{13.2}$$

where $B_i^d t$ are the Bernstein basis polynomials of degree d and the coefficients c_i are points defining a control polygon. One well known property of the cubic B-spline is that it minimizes the second derivative. When the speed of the parameterization is constant, then the minimization property applies also to the curvature of the spline curve. As outlined above, the model of our railroad curve consists of a composition of simple geometrical objects. The curvature of the straight line curve segments is zero, whereas the curvature of the circle arcs is constant. The change of curvature in clothoids is constant. This means that the curvature for the global, composite curve is C^0 continuous at the joints between the curve segments. An example of a curvature plot is shown in Figure 13.1.

13.2.3 Method

We evaluate the original composite curve at selected points of interest (POI) to obtain data points. Then we approximate the data points with a cubic B-spline curve via solving a least squares problem, where the idea is to approximate straight lines and (parts of) circle arcs well, while leaving the approximation of transition curves to the "system". Finally, we evaluate the curvatures of the composite curve and the B-spline approximation curve, measure the differences and compare the results. The differences are measured in terms of mean squared error

$$\text{MSE} = \sum_{i=1}^{n} (c_i - s_i)^2 \tag{13.3}$$

where c_i and s_i are samples of the composite curve and the spline curve, and the maximum distance

$$\text{MAX} = \max(c_i - s_i)^2 \tag{13.4}$$

13.3 Results

Here we give our results by showing how the proposed method performs when applied to a collection of representative curve segment. We experiment with varying the number of data points (samples), testing uniform and non-uniform samples, uniform and non-uniform spline knot vectors, and different numbers of spline coefficients. Graphs showing curvature- and difference plots

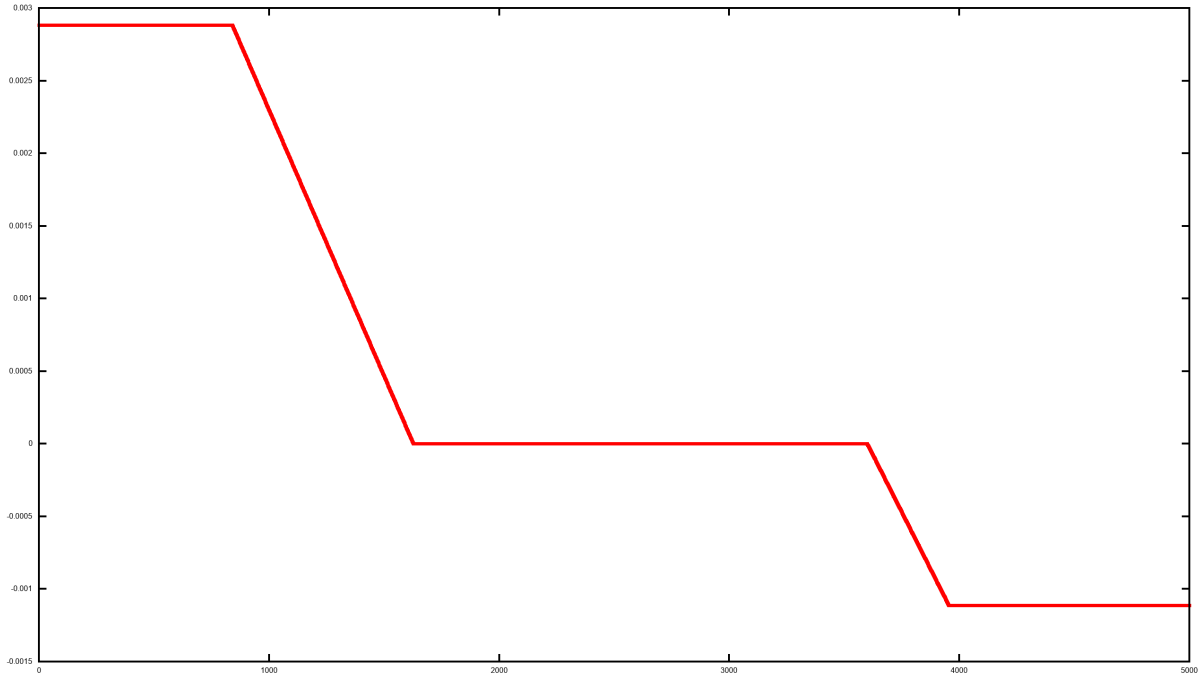


Figure 13.1: Curvature plot of a simple segment curve. Consisting of 5 segments, arc-clothoid-line-clothoid-arc, respectively.

are shown in the figures 13.2, 13.3, 13.4, and 13.5.

Table 13.1 show how the different constructions compare to each other using a mean squared error estimate. MSE (13.3) is the average of the squares of errors and the Max value (13.4) is the largest distance between the two curves.

13.4 Analysis and discussion

The B-spline approximates the composite curve well, as expected, this is shown in Figure 13.1 and 13.5. By considering Figure 13.2, 13.3, and 13.4 we observe that the curvature of the B-spline matches the curvatures of the straight line- and circular arc segments. However, we note that there are some wiggles in the graphs in those areas which are due to noise arising from

Table 13.1: The distance error column is explaining the distribution of control points and sample points of the curve. Measurements are mean squared error and max distance between the two curves

Error measurements		
Distance error	Measurements	
	MSE	MAX
Uniform 30/500	0.180716	2.54563
Uniform 60/500	0.001348	0.01307
Uniform 300/500	0.000001	0.00330
Non-Uniform 31/500	0.263089	4.41375
Non-Uniform 58/500	0.002167	0.03209
Non-Uniform 85/500	0.000264	0.00408
Non-Uniform 274/2500	0.000001	0.00438

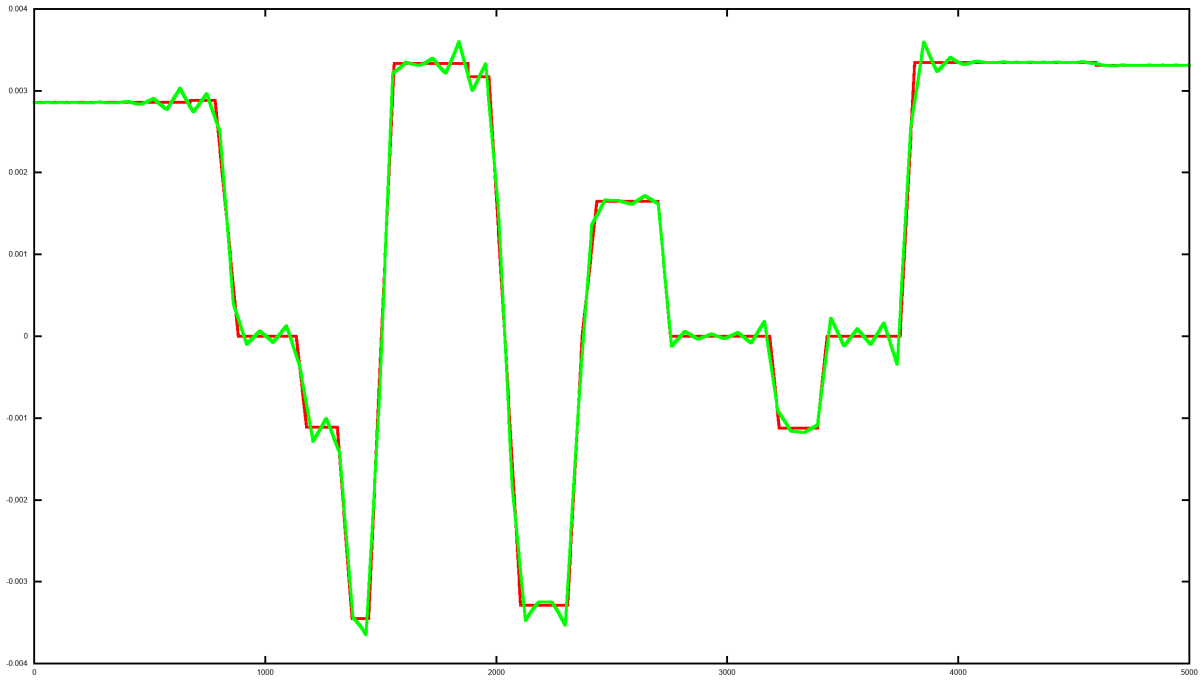


Figure 13.2: Curvature comparison between segment curve (red) and B-spline (green) using uniform knotvector and sampling. Nr. Control points is 90.

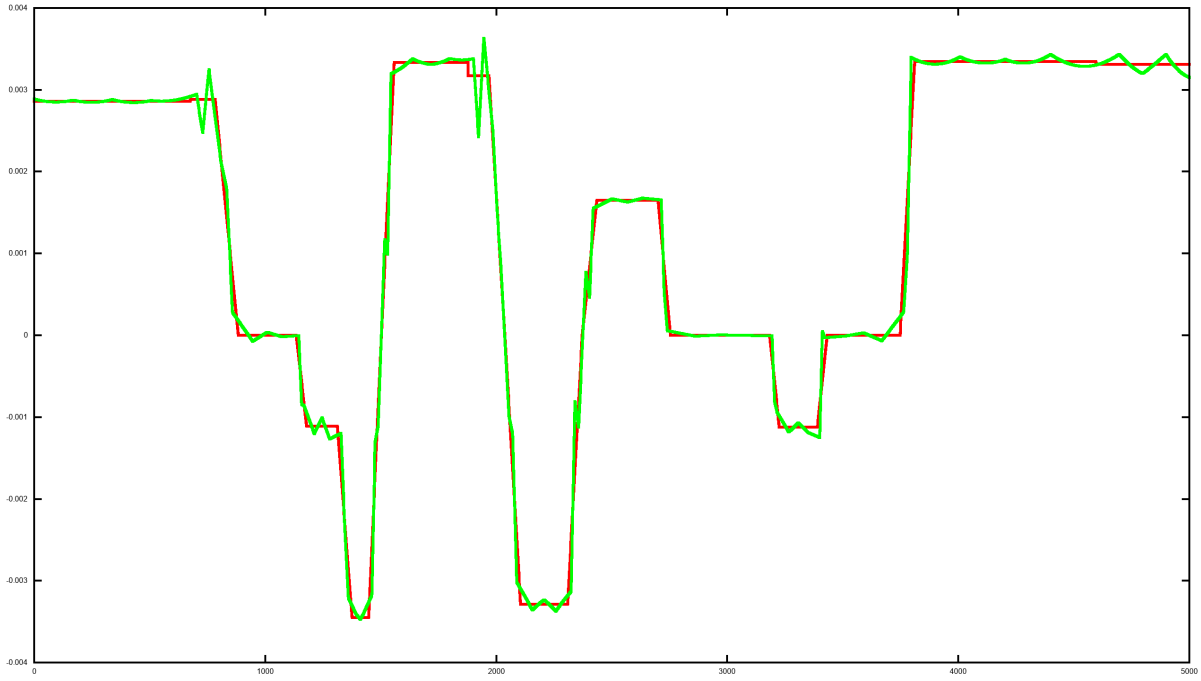


Figure 13.3: Curvature comparison between segment curve (red) and B-spline (green) using non-uniform knotvector with 86 control points and 500 sampling size.

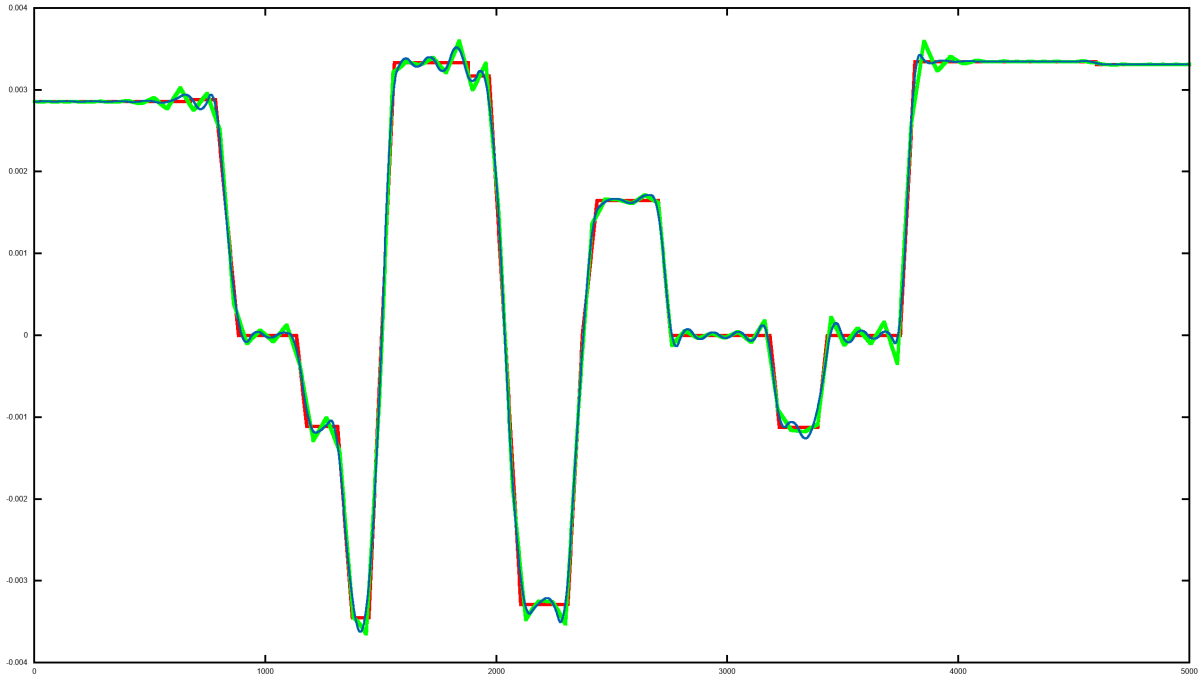


Figure 13.4: Curvature comparison between segment curve (red) and both a 3rd (green) and 4th (blue) degree B-spline with uniform knotvector and sampling.

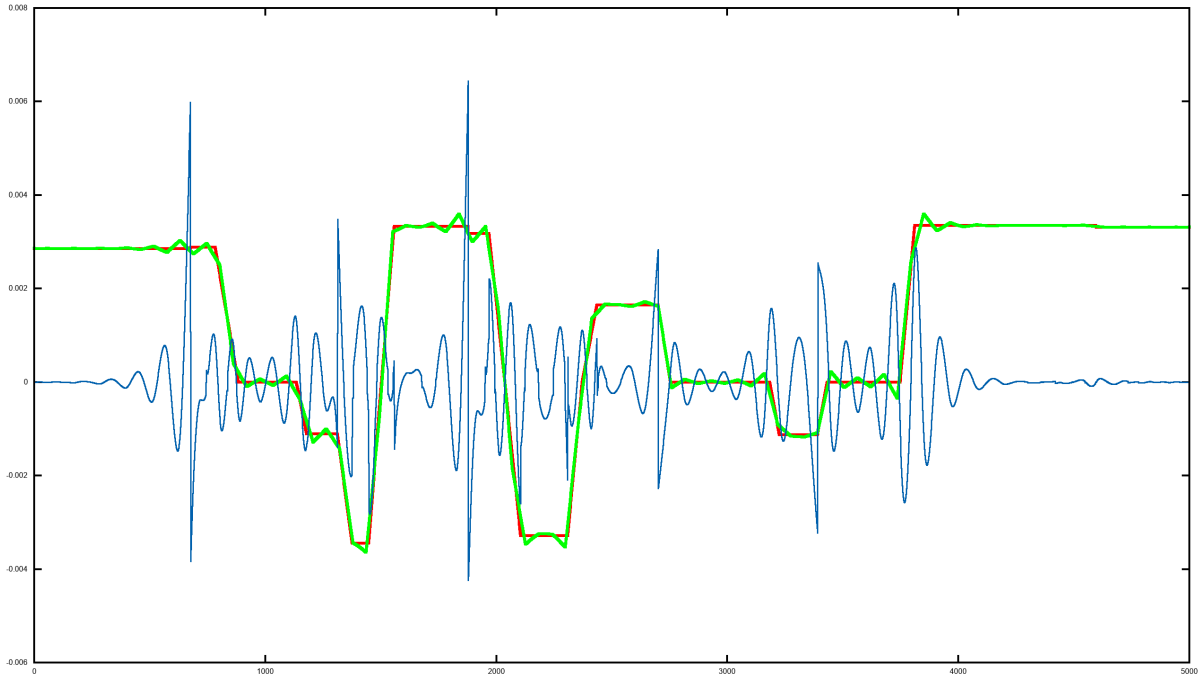


Figure 13.5: Curvature plot against error distance between segment curve and B-spline with uniform knotvector and sampling. Segment curve (red), B-spline (green) and distance difference (blue).

the polynomial interpolation. Here we remark that this is expected, but the curvature numbers for the curve are so small that this noise is shown in the graphs. When comparing the graphs in Figure 13.2 and 13.3, made from a similar number of control points (90 vs 86), we note that a non-uniform parameterization of the B-spline works better with respect to wiggles and smoothness in some areas than a uniform parameterization. The curvature plots for the B-spline curves indicate improved smoothness over the junctions when compared to the composite curve, as shown in Figure 13.2 and 13.3. Figure 13.4 shows how the curvature of a quartic B-spline is smoother than the curvature of a cubic B-spline. With the same number of control points, the curve is also less wiggly.

13.5 Conclusion

The curvature of the proposed B-spline approximation does not vary linearly with arch length, as point-ed out in [14], however, this criterion is challenged with the advent of the LCA [6]. One possible application of this work is within maintenance and trying to reduce wear of the railroad tracks. Alternative transition curves will influence the positioning of the railroad. An improved curvature/LCA can possibly be achieved even with relatively small changes to the layout of the railroad transition curve segments. Another possible application is within software for railroad geometry construction. For future work we note that utilizing non-polynomial splines in the attempt to avoid oscillations near sharp edges would be interesting, and that a possible application to simulation and evaluation of ride comfort parameters based on relevant ISO standards (ISO 2631 etc.) could be considered.

13.6 Acknowledgements

The data used in this work is extracted from the Bane NOR database using Novapoint and helpful background information regarding "Ofotbanen" has been given by employees at Bane NOR. A special thanks to Ingeborg Tulluan, who was kind enough to extract the data needed in this work. Also, thanks to Bjørn Gunnar Larsen with information regarding wear and measurements for some parts of Ofotbanen.

References

[1] J. Glover. Transition curves for railways. *Proc. Inst. C. E.*, 140, 1900.

[2] E. M. Horsburgh. Xxiv.—the railway transition curve. *Proceedings of the Royal Society of Edinburgh*, 32:333–347, 1913.

[3] A. W. Miller. The transition spiral. *Australian Surveyor*, 7(8):518–526, 1939.

[4] R. C. Archibald. Euler integrals and euler's spiral, sometimes called fresnel integrals and the clothoide or cornu's spiral. *American Mathematical Monthly*, 25:276–282, 1918.

[5] Arthur Lovat Higgins. *The transition spiral and its introduction to railway curves with field exercises in construction and alignment.* Van Nostrand company, New York, 1922.

[6] Orhan Baykal. On concept of lateral change of acceleration. *Journal of Surveying Engineering*, 122:132–141, 08 1996.

[7] Ergin Tari and Orhan Baykal. A new transition curve. *Proc Symp 27–29 September, 1st Turkish-German Joint Geodetic Days, Istanbul*, pages 107–119, 1995.

[8] Ergin Tari. Spline as a route element. *Proc 5th Symp on the use of computer in civil engineering, Istanbul*, pages 495–496, 1996.

[9] Ergin Tari. *New curve trends in alignment design.* PhD thesis, Istanbul Technical University, 1997.

[10] Ergin Tari. The new generation transition curves. *ARI The Bulletin of the Instanbul Technical University*, 54:35–41, 01 2004.

[11] Ergin Tari and Orhan Baykal. A new transition curve with enhanced properties. *Canadian Journal of Civil Engineering*, 32(5):913–923, 2005.

[12] Banenord. *Ofotbanen*, (Accessed 2019-02-28). (online)`https://www.banenor.no/Jernbanen/Banene/Ofotbanen/`.

[13] Carl de Boor. *A Practical Guide to Spline*, volume Volume 27. 01 1978.

[14] D.J. Walton and D.S. Meek. Clothoidal splines. *Computers and Graphics*, 14(1):95 – 100, 1990.

14 Targeted sanding and its impact on heavy hauler pull force and surface friction

Tanita Fossli Brustad, Aleksander Pedersen, and Børre Bang

Abstract

Road conditions during wintertime in northern Norway frequently causes problems for heavy haulers transporting goods. The trucks often experience difficulties when travelling uphill or downhill because of snowfall and ice formation, which again constitutes a safety risk and a nuisance for other road users. In this paper we have conducted a series of field tests where we investigate the impact of targeted sanding on a heavy hauler with focus on self-help, pull force and road friction.

14.1 Introduction

Wintertime in the northern parts of Norway offers many challenges within traffic and transport. Access to main roads and well maintained roads decreases the further north you get. Road conditions during the winter can vary from day to day, and may deteriorate considerably on short notice due to ice formation and snowfall. Ice formation on certain parts of the road reduces the friction significantly, which again increases the risk of accidents [1, 2, 3, 4]. A well known problem during the winter months is connected to trucks (heavy haulers) transporting goods. These trucks often experience difficulties when travelling uphill or downhill, because of snow and ice, causing safety risks for other drivers and leading to roads being closed over a period of time [5]. Since northern Norway has a limited amount of roads leading to key cities in the region, trucks blocking roads are a complication for other road users, who will have to wait for the roads to be reopened or use detours that might add several hours to the trip. Part of a solution to this problem is to make it easier for the drivers to help themselves in certain situations. For example, driving uphill can be complicated if the friction on the road surface is low. Low friction can lead to a lack of wheel grip with trucks "spinning" on the spot unable to get up the hill, either because they had to stop unexpectedly, or because the hill is too steep in relation to the friction. A way of increasing the friction is to spread sand [6] in the areas where the wheels are (targeted sanding) so that the truck can get passed the problem zone without having to block the road and wait for help. Autoline AS [7] has created a device that makes it possible for truck drivers to turn on automatic targeted sanding on their own without leaving the cab of the truck. This gives drivers the ability to help themselves in some cases where they otherwise would have to wait for rescue. Using targeted sanding as a self-aid for trucks may lead to cost reductions in winter maintenance and decrease negative environmental effects [8, 9] since only small areas of the road are sanded. As a part of the WiRMa project and in collaboration with Autoline AS and Nord-Norsk Trafikksenter AS we have conducted a number of field tests connected to the Autoline targeted sanding device and the impact this has on pull force of a heavy hauler.

14.2 Method

In this study a field test was performed to look at changes in pull force, before and after targeted sanding, for a heavy hauler pulling a trailer on a snow covered flat surface with a friction of 0.24. The friction on the test surface was measured with a Teconer RCM-411 [10] and the pull force was measured with a weight attached between the truck and the trailer. The experiments were done in order to observe the impact targeted sanding would have on a heavy hauler with focus on the force it was able to pull a trailer with. The field tests were performed at Nord-Norsk Trafikksenter on Finnsnes in Troms, Norway, in -14 degrees Celsius on a clear winter day. Truck driving during the tests was performed by John-Egil Bustamo, and driving with friction measurement was done by Aleksander Pedersen and Børre Bang. The truck used in the experiment was a Scania with a weight of 20 tons towing a trailer weighing 20 tons, see Figure 14.1. This was a standard truck setup from Nord-Norsk Trafikksenter used in education and license purposes. The weight used to measure the pull force was a Dynafor LLX2 digital load indicator 10 tons + 13.5 tons shackles [11] and it was connected as can be seen in Figure 14.2. The experiments were performed by locking the wheels on the trailer and moving the truck forward with increasing torque and accelerator deployment until the truck wheels lost grip, while collecting pull force data from the weight. A more detailed description of the setup is included in Figure 14.3. The device used for sanding was a sand spreader developed by Autoline AS, which is a box filled with gravel, mounted in front of the drive wheels of the truck. The box is operated from the cab and releases a small amount of gravel at a fixed interval while running. The process can be seen in Figure 14.4. Experiments were performed before and after

Figure 14.1: The truck and trailer used in the experiments.

Figure 14.2: Connection of the weight between the truck and the trailer.

sand spreading. Three tests were carried out without sand and two with sand. The two tests after sanding had to be terminated before reaching maximum pull force due to safety issues. Nonetheless interesting results were obtained. To compare the results before and after sanding, graphs of the pull force were created and a visual comparison of the numbers is carried out. The graphs are also compared to an exponential function ($y = ae^x$, $a \neq 0$) that is based on regression of the data set.

14.3 Results and discussion

The result from all five experiments can be seen in Figure 14.5. The first three spikes are measurements before sanding and the two last spikes are after sanding. From the figure it can be observed that the results from the tests before sanding are similar with a small decrease in maximum pull force. The decrease in maximum pull force may come from the fact that the experiments are done in the same spot, leading to a decrease in friction after each run. The two results after sanding was, as mentioned in section 14.2, terminated before reaching a maximum

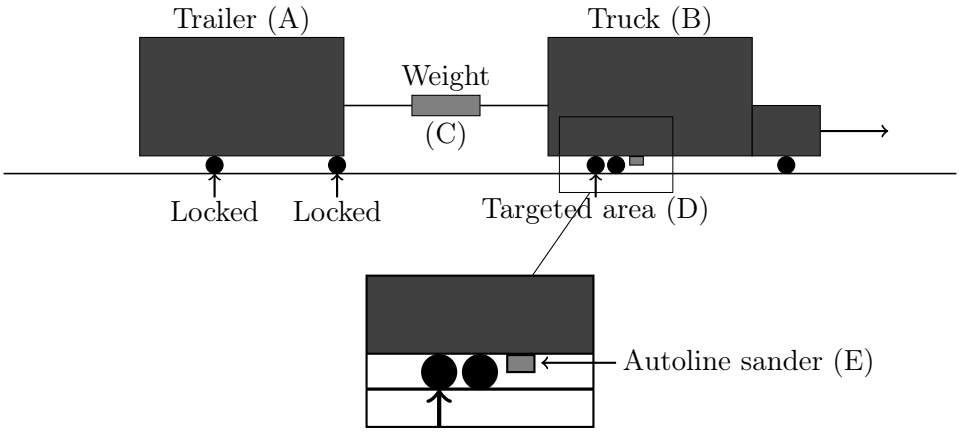


Figure 14.3: The experiment setup: Trailer (A) wheels are locked. A weight (C) is attached between the trailer (A) and the truck (B). Truck (B) is trying to move forward with increasing throttle until the wheels loose grip at targeted area (D). The Autoline sander (E) is only used in two of the experiments, to measure the difference in pull force w/wo sanding.

Figure 14.4: The Autoline sander in action.

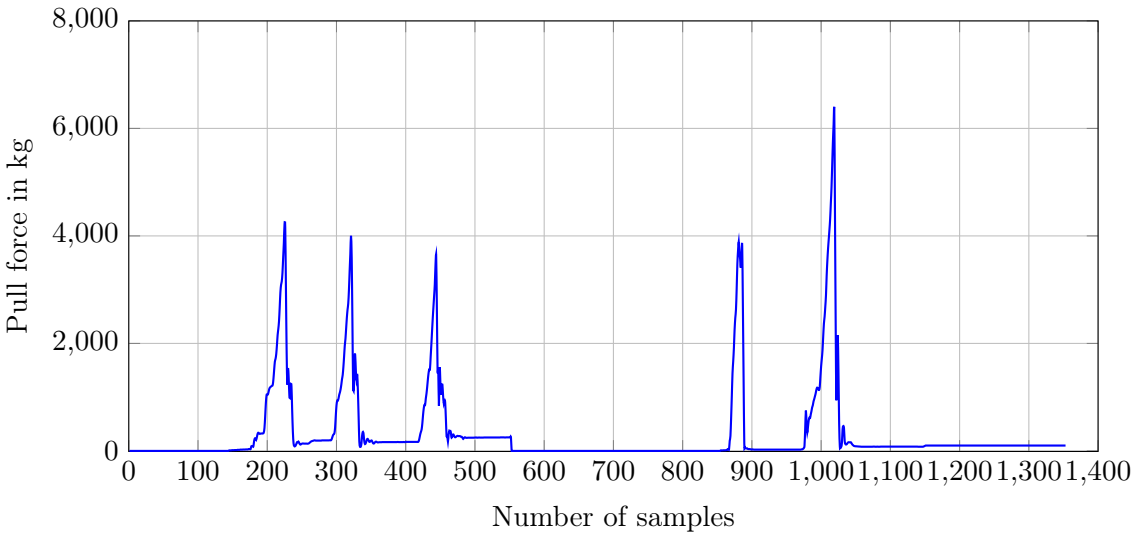


Figure 14.5: The output pull force in kg (y-axis) from the weight at 0.25 s (4 Hz) intervals, from start at sample n=0 and end n=1400 (x-axis). The spikes represents the five experiments, the three first before sanding and the two last after sanding.

pull force. This is easy to see from the graph where one of the tests is almost equal to the results before sanding. However, in the last measurement a significant increase in pull force is obtained where the tests were run over a longer period of time prior to termination. In order to compare the results before and after sanding the authors select one experiment from each and compare the graphs up until maximum pull force. The chosen experiments are the first and the last from

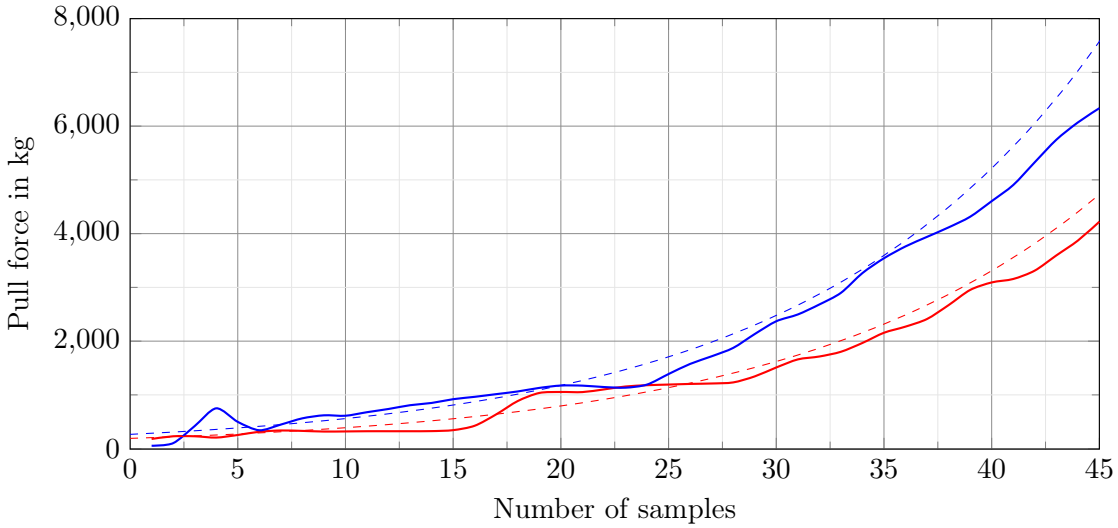


Figure 14.6: The pull force (y-axis) up until maximum force (x-axis) from the first (red) and last (blue) experiments with associated exponential regression curves in stapled lines, see Y_1 and Y_2 in (14.1).

Figure 14.5, and it is shown in Figure 14.6. From the plots in this figure it can be seen that the maximum weight measured from experiment 1 and 5 is 4222 kg and 6334 kg, respectively. This means that there is an increase in maximum pull force of over 2000 kg after sanding compared to before for this specific test case. The interesting part about the previous results is what happens along the curves towards maximum force, i.e. how the curves are developing. By comparing with an exponential regression curve, stapled lines in Figure 14.6, it can be noticed that the blue

curve is a lot steeper than the red curve. Using regression to find the closest match for the data set by an exponential function, the following functions are obtained:

$$\begin{aligned} Y_1 &= 265.68e^{0.2977x}, R^2 = 0.8921 \\ Y_2 &= 190.67e^{0.2854x}, R^2 = 0.9675 \\ Y_1^{'} &= 79.093e^{0.2977x} \\ Y_2^{'} &= 54.417e^{0.2854x} \end{aligned} \tag{14.1}$$

where Y_1 and Y_2 corresponds to blue and red stapled lines in Figure 14.6. The scaled derivative

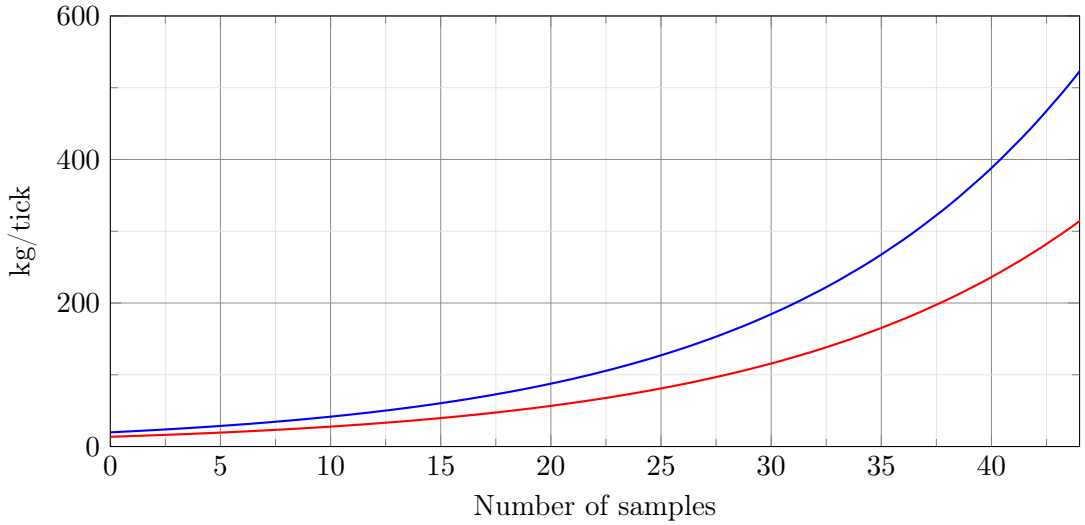


Figure 14.7: Data plot of the derivative of the pull force from the first (red) and last (blue) experiments in Figure 14.6. This is a measure of kilogram per tick, and not per second. Each tick is 0.25 seconds.

plot of $Y_1^{'}$ and $Y_2^{'}$ are shown in Figure 14.7, and as is seen, the increase is around 500-600 kg/s for the last experiment compared to 300-400 kg/s for the first experiment. Since both graphs have a good match with an exponential function, see R^2 in (14.1), we know that they both grow rapidly and if we look at the development 1-2 seconds later, keeping in mind that the test after sanding were terminated prematurely due to safety, the force would be doubled or even tripled. The main result here is that the targeted sanding works well, and will increase the friction of the surface quite good compared to a non sanded surface. Based on the graphs and the maximum forces obtained, an increase in friction can be estimated by computing the ratio of the two results

$$\frac{6334kg}{4222kg} = 1.50,$$

which corresponds to more than 50% increase in friction. Based on the data obtained from other sensors, the average friction coefficient was about 0.24 before sanding, and after sanding it can be estimated to 0.48 in this particular case, which is within the expected range (0.3 - 0.5) [12] for sanded snow/ice road.

14.4 Conclusion

A field test was performed with five experiments to look at the impact targeted sanding would have on a heavy hauler driving on winter roads. The tests were focused around the possible pull force of the truck before sanding compared to after sanding. A limited number of experiments

were made over the course of one day in stable winter conditions of -14 degrees Celsius. This of course limits the collected data to a specific situation and will only give an implication of the sanding effect. However, based on the gathered results a conclusion can be made that targeted sanding is effective with regards to friction and pull force on certain road conditions since it increases the friction on the road in the considered area, giving the truck better grip. The results show an increase of about 2000 kg in maximum pull force, and also revealed some interesting outcomes regarding the development of the derivative/slope of the force, where we see a time dependent increase in pull force that is much higher after sanding compared to before. A little remark on the ratio between maximum values. If we take the average values for each of the two graphs in Figure 14.6 we obtain approximately 3000kg and 2000kg respectively, which computes to $\frac{3000}{2000} = 1.5$. This means that the ratio computed for maximum values is also valid for the average values and the area for each graph. For future work it would be ideal to do more field tests under varying conditions and expand the number of parameters measured during the tests, in order to be more confident in the obtained results. Measured time dependent values for torque on the driving wheels together with accelerator and power values would be of interest. Lastly, road condition monitoring equipment will in near future be able to classify, and possibly give standardized friction estimates. This would open up the possibilities for predictions usable in route planning of heavy haulers, with and without vehicle mounted sand spreaders.

14.5 Acknowledgement

This research was partially financed by the Interreg-Nord project "WiRMa", project ID 20201092. The authors would like to give special thanks to Autoline AS with John-Egil Bustadmo for his ideas, good discussions, and providing the support making this experiment possible, and also thanks to Nord-Norsk Trafikksenter AS for excellent service at their track facilities.

References

[1] Jean Andrey, Brian Mills, and Jessica Adams. Weather information and road safety, 08 2001.

[2] Daniel Eisenberg and Kenneth E. Warner. Effects of snowfalls on motor vehicle collisions, injuries, and fatalities. *American Journal of Public Health*, 95(1):120–124, 2005.

[3] Keith K. Knapp, Dennis Kroeger, and Karen Giese. Mobility and safety impacts of winter storm events in a freeway environment. Technical report, Iowa State University. Center for Transportation Research and Education, 2000.

[4] Carl-Gustaf Wallman and Henrik Åström. Friction measurement methods and the correlation between road friction and traffic safety : A literature review. Technical Report 911A, Swedish National Road and Transport Research Institute, 2001.

[5] J Granlund and P Thomson. Traffic safety risks with EU tractor-semitrailer rigs on slippery roads. HVTT14: International Symposium on Heavy Vehicle Transport Technology, 14th, 2016, Rotorua, New Zealand, 2016.

[6] Torgeir Vaa. Implementation of new sanding method in Norway. Sixth International Symposium on Snow Removal and Ice Control Technology, 2004.

[7] Autoline AS. *Market leader in Scandinavia*, 2017 (Accessed 2019-03-05). `http://www.autoline.no/About-us`.

[8] Christer Johansson, Michael Norman, and Lars Gidhagen. Spatial & temporal variations of pm10 and particle number concentrations in urban air. *Environmental Monitoring and Assessment*, 127(1):477–487, Apr 2007.

[9] Kaarle J. Kupiainen, Heikki Tervahattu, Mika Räisänen, Timo Mäkelä, Minna Aurela, and Risto Hillamo. Size and composition of airborne particles from pavement wear, tires, and traction sanding. *Environmental Science & Technology*, 39(3):699–706, 2005.

[10] Teconer. *Winter maintenance of roads and runways*, 2015 (Accessed 2019-03-05). http://www.teconer.fi/en/winter.html.

[11] Tractel. *Dynafor LLX2*, 2019 (Accessed 2019-03-05). https://www.tractel.com/us/series.php?id_serie=7.

[12] Alex Klein-Paste and Bård Nonstad. Lærebok drift og vedlikehold av veger. Technical report, Vegdirektoratet - Trafikksikkerhet, miljø- og teknologiavdelingen - Vegteknologi, 2015. STATENS VEGVESENS RAPPORTER Nr. 365 Chapter 12.

15 A field study of sensors for winter road assessment

Tanita Fossli Brustad, Aleksander Pedersen, and Børre Bang

Abstract

Winter road assessment is a research field with considerable progress over the last 10 years. Various sensors and methods have been tested and analysed, often in a laboratory setting, in order to come up with robust and valid assessment tools that can be used to warn the driver, road users in general, and maintenance personnel of critical conditions. In this paper we compare the field measurements of an RCM411 and a MARWIS sensor with each other and with previously performed laboratory experiments, we reflect on OBD-II as a tool in winter road assessment, and perform initial field tests with an experimental radar sensor. The results of the RCM411/MARWIS comparison shows significant correlation between our field experiments and the laboratory experiments, OBD-II appears to be fitting as a supplementary tool in the assessments, and the experimental radar tests uncovers a need for more investigation.

15.1 Introduction

Assessing road conditions during the winter season is important for safety and maintenance purposes. Knowing when and where road conditions are critical will help maintenance personnel do their job more efficiently (saving money) and can be used to warn other road users (increasing safety). Monitoring winter roads is essential in countries where the winter months are unpredictable in regards to weather and temperature. Today, research in winter road assessment mostly revolves around the use of sensors and cameras that analyse data collected from (and around) the road or from the car. Various sensors and methods have been proposed over the last 10 years e.g.: [1, 2, 3] used camera images to classify surface conditions, the VehSense system in [4] combined a smartphone and OBD-II to detect vehicle skidding, [5] utilized the latest infrared technology to measure road temperature, and [6, 7] tested 24-GHz radar to recognize road conditions, to mention some. In addition there are sensors that have been developed specifically to collect important data for winter road assessment, two well known examples are RCM411 [8] and MARWIS [9]. A number of laboratory experiments have been performed on these two sensors, see e.g. [10] or [11], and in the last 5 years (2015-2019) the sensors have become more and more common as a tool in road condition research, either for collecting needed data [12, 13, 14], as a reference sensor [15], or to study the reliability against sensors categorized as more reliable [16].

A motivation behind this work, where research material is scarce, is to examine if the RCM411 and MARWIS sensors perform similarly in a field study setting, as they do in a laboratory setting, because this will strengthen the validity and broaden the knowledge of these two sensors, which are specifically developed for winter road assessment. As already mentioned, a variety of sensors have been tested for use in winter road assessment and more are being investigated currently. Just in the past year (2019) a number of papers have been published with focus on new sensors (e.g. an optical sensor in [17], capacitive sensors in [18], microphones in [19], and a piezoelectric sensor togeteher with an optical sensor in [20]) and new developments of known sensors/data, for instance vehicle sensor data [21] and camera images [22]. So the search for "new" sensors is far from over, and experimental technology is still of interest to the field of study. Two desired properties of winter road sensors are that they should be cheap and not too large. Based on these properties, and our interest in experimental technology, our field research includes initial testing of a small and cheap radar sensor with much potential. The main reason for testing more than one sensor in this research, and the reason for also including collection of data from the vehicle (with an OBD-II), is the possibility of using hybrid measurements in the future with various data that compliment each other in a beneficial way.

In this paper we give a brief overview of the popular sensors, RCM411, MARWIS, and OBD-II, for winter road assessment, compare the field measurements of the RCM411 and MARWIS sensors with each other and with previously performed laboratory experiments, reflect on OBD-II as a tool in winter road assessment, and perform initial field tests with an experimental radar sensor.

The reminder of the paper is organised as follows. In section 15.2 a description of the relevant sensors are given; RCM411, MARWIS, OBD-II, and Walabot. In section 15.3 the sensor setup and test route are described and illustrated, together with key information regarding the analysis and comparison of the sensor data. Section 15.4 shows and discusses the results of the comparison between RCM411 and MARWIS for segments of the extensive field test route against laboratory experiments, considers the usefulness of the OBD-II measurements, and reflects on the data from the experimental Walabot. And lastly, conclusions are drawn in section 15.5.

15.2 Preliminaries

In this section a description of the sensors used in our field experiments is presented. The sensors are RCM411, MARWIS, OBD-II, and Walabot (radar). All can be seen in Figure 15.1.

(a) RCM411. (b) MARWIS. (c) OBD-II. (d) Walabot.

Figure 15.1: The four sensors used in the field experiments.

15.2.1 RCM411

Road Condition Monitor RCM411 [8, 23, 24] (Figure 15.1a) is an optical sensor based on spectral analysis that evaluates the surface condition by measuring the optical reflection signal on the road surface. The sensor, combined with a smartphone, analyse the data and gives information on surface conditions, friction, and water layer thickness, together with technical details regarding location and speed. In addition RCM411 uses an IR-thermometer (RTS411), supplied by Teconer, installed to the RCM411-frame. This thermometer gives feedback on air temperature and surface temperature in the same data package as the previous variables. RCM411 is designed to be mounted on the rear of a vehicle on a 50 mm tow ball, with the sensor pointing down towards the wheel track. Data from the sensor is transferred to a mobile phone using Bluetooth, then communicated to selected servers. This provides the opportunity to see real-time updates on surface conditions on a mobile phone app or online at `https://roadweather.online`. Raw data is also available for download on `https://roadweather.online` from the sensor you own.

15.2.2 MARWIS

Mobile Advanced Road Weather Information Sensor (MARWIS) [9, 25] (Figure 15.1b) is an optical sensor that captures the reflecting behaviour of the road surface at varying wavelengths to give feedback regarding the road conditions. In addition MARWIS has three other integrated sensors for measuring temperature and humidity. The data received from the sensor includes surface conditions, friction, water layer thickness, ice percentage, temperature (Surface, ambient, and dew point), and humidity (Relative and relative at ground temperature). MARWIS can be mounted on a truck or a car using a custom rack attached with magnets, and there are three ground-to-sensor distance options depending on the MARWIS model: 0.5 m, 1 m, or 2 m. Data from the sensor is transferred to a mobile phone using Bluetooth and can be viewed in real-time using the MARWIS app [26]. The app has automatic upload of received data to a server (the ViewMondo-Server), which gives the opportunity of downloading, visualising, and analysing the measurements on the ViewMondo software platform [27].

15.2.3 OBD-II

On-Board Diagnostics, Second Generation (OBD-II) [28, 29] (Figure 15.1c) is a computer-based system monitoring the performance of engine components. The system was originally designed to reduce emissions from vehicles and became a requirement in the US in 1996 on all newer vehicles.

Later on other countries adopted similar laws, including Canada, parts of the European Union, Japan, Australia, and Brazil. The OBD-II system in a vehicle can be accessed via a Diagnostic Link Connector using a device called a scan tool or OBD-II adapter. Various information can be obtained from the diagnostics, e.g. status of the "Check Engine" light, diagnostic trouble codes, and hundreds of real-time parameters (i.e. Speed, RPM, coolant temperature, etc.). Most scan tools are compatible with third-party OBD software and mobile apps (e.g. via Bluetooth) that displays and saves information, and allows interaction with the OBD-II in the vehicle.

15.2.4 Walabot

Walabot [30, 31] (Figure 15.1d) is a pocket-sized 3D imaging sensor that uses radio frequency technology to illuminate the area in front of it and sense the returning signals. The sensor supports short range scanning and distance scanning, with the possibility to extract 3D image data, 2D image slices, raw signals, and the sum of raw signals in an image. Data collected from the sensor is processed and sent through a USB cable to a host device (e.g. a Raspberry Pi). Depending on the model, the Walabot can be used in areas such as in-room imaging, in-wall imaging, object detection, motion sensing, and sensing of dielectric properties of materials. In addition many competitions have been held that focus on finding new use cases for the Walabot, some examples can be found at [32].

15.3 Method

In this section the sensor setup and test route are described and illustrated. In addition, a description of key information connected to the analysis and comparison of the sensor data is given.

15.3.1 Sensor setup and field experiments

The sensor setup for collecting and saving data was done as follows: the RCM411 sensor sends data via Bluetooth to an Android phone and to servers showing the data online on `https://roadweather.online`, the MARWIS sensor sends data via Bluetooth to an Asus Zen Pad with an in-house developed app that shows real-time updates and saves the data, OBD-II data is collected via Bluetooth from an OBD-II adapter to a mobile app where real-time data is displayed and logged, and the Walabot sends data via USB to a Raspberry Pi that saves the data.

The sensor setup on the vehicle, seen in Figure 15.2, was done in the following way: the RCM411 sensor was mounted on the tow ball pointing at the right wheel track, the MARWIS sensor was secured with a rack to the truck bed pointing at the center of the road behind the vehicle, and the Walabot sensor was attached below the back left passenger door pointing straight down at the wheel track.

The field experiments were carried out in Northern Norway, Sweden, and Finland over 5 days, 26-28 of March and 2-3 of April 2018. A distance of 1729 km was covered stretching from Narvik to Vadsø, and back (the route is plotted in Figure 15.3). The weather conditions ranged from sunny to heavy snow, with temperatures between -9 and +5 degrees Celsius, and road conditions switching between icy, snowy, wet, moist, slushy and dry asphalt.

15.3.2 Analysis

In the comparison of the RCM411 and MARWIS sensors three variables were compared: friction, water thickness, and road temperature. The results were analysed by studying the graphs of the values obtained from certain stretches of road, chosen based on road conditions and sensor

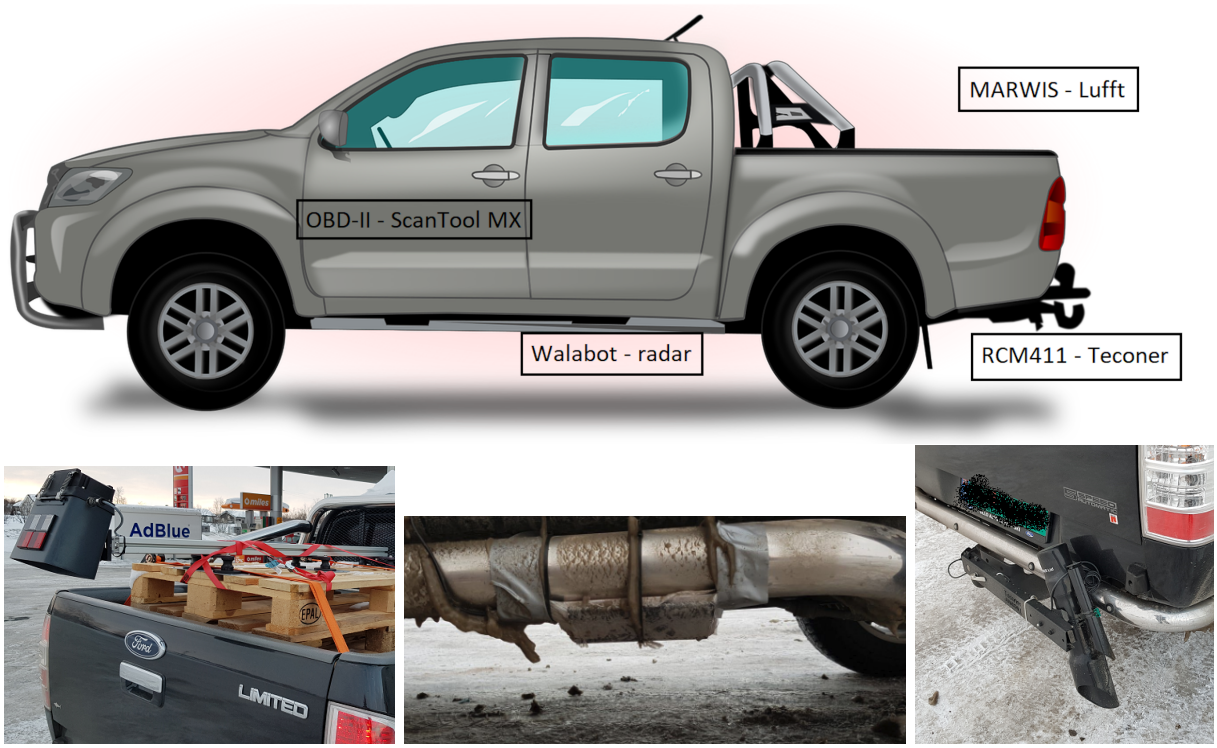


Figure 15.2: Placement of the sensors on the vehicle. Top image by OpenClipart-Vectors from Pixabay.

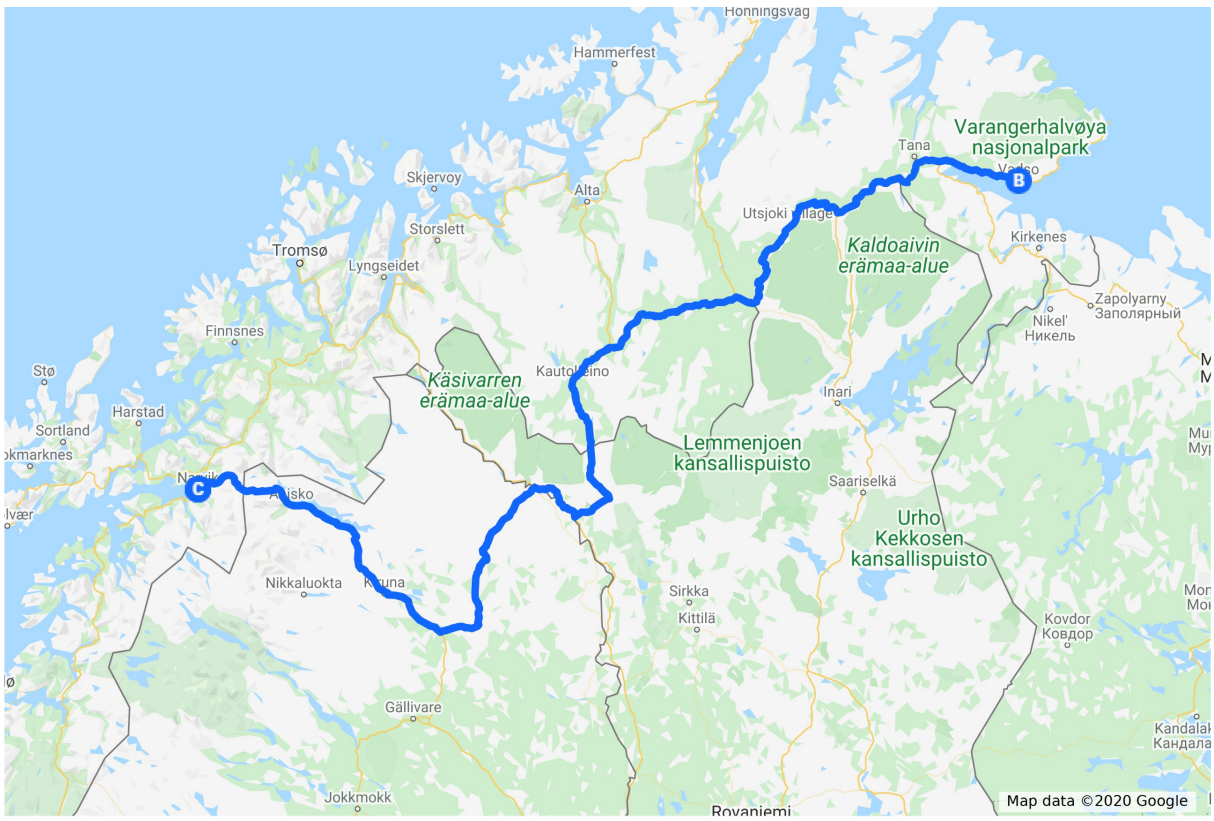


Figure 15.3: The route, in Google maps, where the sensor testing was conducted.

uptimes. Comparisons of the RCM411 and MARWIS sensors have been performed a number of times in laboratory experiments ([10] and [11] are good examples) so the consistencies between the laboratory experiments and field experiments are commented. Calibration of the two sensors was done using a gray asphalt plate, as in [11], before experiments began. Four road segments were used in the comparison and analysis, one on the stretch from Narvik to Vadsø and three the other way, each with a timescale of 30 minutes (about 40 km in length), see Table 15.1. In the comparisons an important thing to keep in mind is that the sensors are not measuring in the exact same spot. The RCM411 measures behind the wheel track while MARWIS is near the middle of the lane. From experience we know that conditions can vary along the breadth of the lane, however, decent comparisons can still be made.

In the Walabot experiments the scan profile of the sensor was set to distance scanning with fast capture rate of low-resolution images, and the collected data was image energy (the sum of all image pixels' raw signal power). As with the RCM411/MARWIS comparison, the results were analysed by studying the graphs of the image energy, comparing them with each other for three segments with different surface conditions (Snow, ice, and dry/wet). In addition two phenomena regarding the measurement data were given a closer look: Walabot vs water/ice and Walabot vs vibration.

For the OBD-II an analysis and discussion is performed based on the available data from the sensor. The main question here being; what are the possibilities in using OBD-II data in winter road assessment?

Table 15.1: Description of the analysed road segments in the comparison between RCM411 and MARWIS.

Segment description			
Date	Location (on the stretch between)	Length [km]	Road conditions
March 28, 2018 10:10 - 10:40	Karasjok-Utsjoki	39.4	Mostly snow
April 2, 2018 16:45 - 17:15	Karasjok-Kautokeino	39.0	Dry and wet
April 3, 2018 09:45 - 10:15	Kautokeino-Enontekiö	40.3	Mostly ice
April 3, 2018 15:30 - 16:00	Kiruna-Narvik	38.0	Mixed

15.4 Results and discussions

In this section we present and discuss our main findings. First we compare the RCM411 and MARWIS because of their similar measurement data and draw similarities to lab experiments, then we consider the usefulness of the OBD-II measurements, and lastly we reflect on the data from the experimental Walabot sensor.

15.4.1 RCM411 vs MARWIS

Friction

The friction measurements, from RCM411 and MARWIS, for the four road segments, in Table 15.1, can be seen in figures 15.4, 15.5, 15.6, and 15.7, respectively. As an overall view of all four

figures the first impression is that the two sensors have detected similar conditions in three of the four segments. However, MARWIS (red line in the graphs) has a tendency to produce spikes reaching min or max values, especially when the friction of the RCM411 (blue line in the graphs) is fluctuating.

In Figure 15.4 with road conditions being mostly snowy, the friction behaviour, and also the values, are similar in both sensors. The biggest difference in this figure is, as commented on above, the spikes that MARWIS has in certain areas, reaching both highest and lowest values. The results from the laboratory tests in [10] and [11] showed that MARWIS on average had lower friction on snow than RCM411, and that the type of snow did not matter for any of the sensors. Comparing the laboratory tests with our results we notice that this behaviour is only present for the first half of the graph, while for the last part the friction of RCM411 is on average lower. When it comes to the spikes in the MARWIS data the lab tests showed no such behaviour.

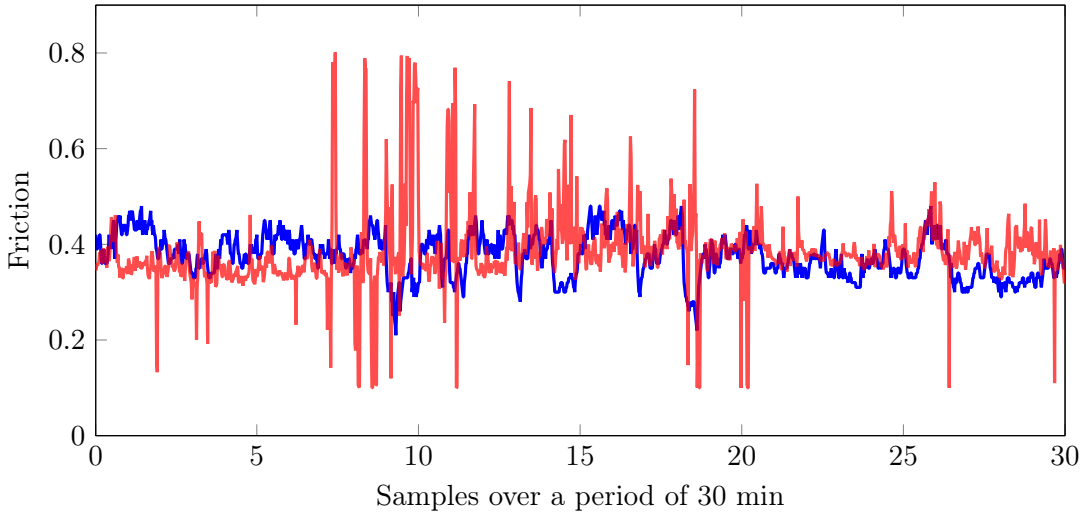


Figure 15.4: A plot of the friction for RCM411 (blue) and MARWIS (red) on snowy road conditions. The x-axis shows the samples along the road segment measured every second, and the y-axis is the friction values.

In Figure 15.5 with road conditions being dry and wet interchangeably the friction behaviour is mostly alike for the dry parts. However, for the wet parts the friction values from MARWIS are considerably lower than RCM411 reaching the lowest value, 0.1, several times (RCM411 has its lowest value around 0.4). In the laboratory tests in [11] the friction for both sensors when measuring on dry asphalt and water was fairly even between, 0.7-0.8, and 0.4-0.7, respectively. Comparing these results with our data shows that the RCM411 follows a similar pattern, while the MARWIS values go far below the lab values for wet conditions.

In Figure 15.6 with road conditions being mostly icy the friction values are not closely related. The values for RCM411 are mainly between 0.3 and 0.6, while the MARWIS values are on average higher, between 0.5 and 0.8. The laboratory tests in [11] showed that MARWIS was giving too high frictions on icy plates compared to RCM411. In these tests the friction measured by RCM411 was between 0.1-0.3, and MARWIS between 0.5-0.65. Although the lab tests were conducted on more slippery ice than our field tests, a conclusion can still be made that MARWIS has a tendency to show too high values for icy conditions when compared to RCM411, which is clearly seen in both experiments.

In Figure 15.7 with mixed road conditions the friction behaviour is more complex. For the dry areas with maximum friction the two sensors agree for the most parts, but when other con-

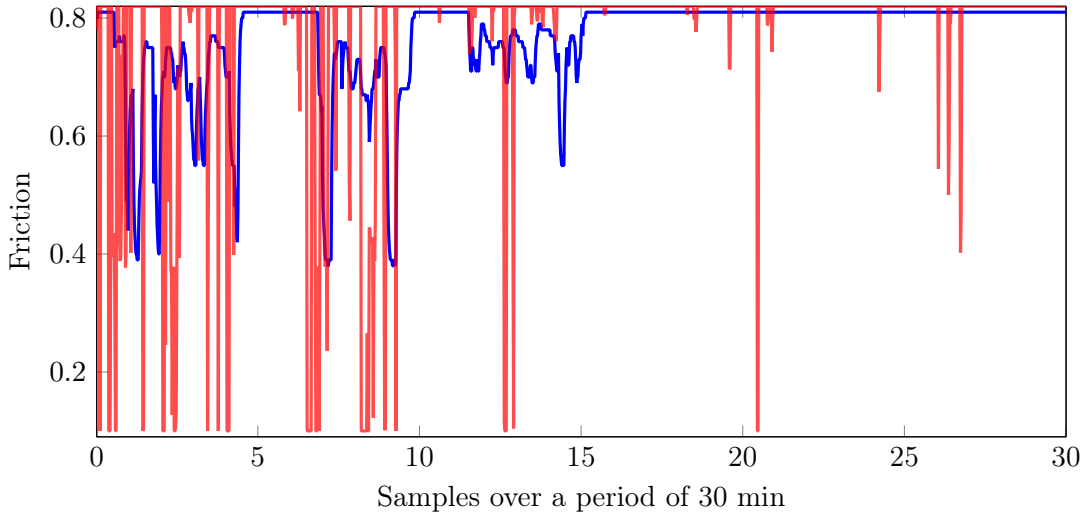


Figure 15.5: A plot of the friction for RCM411 (blue) and MARWIS (red) on dry and wet road conditions interchangeably. The x-axis shows the samples along the road segment measured every second, and the y-axis is the friction values.

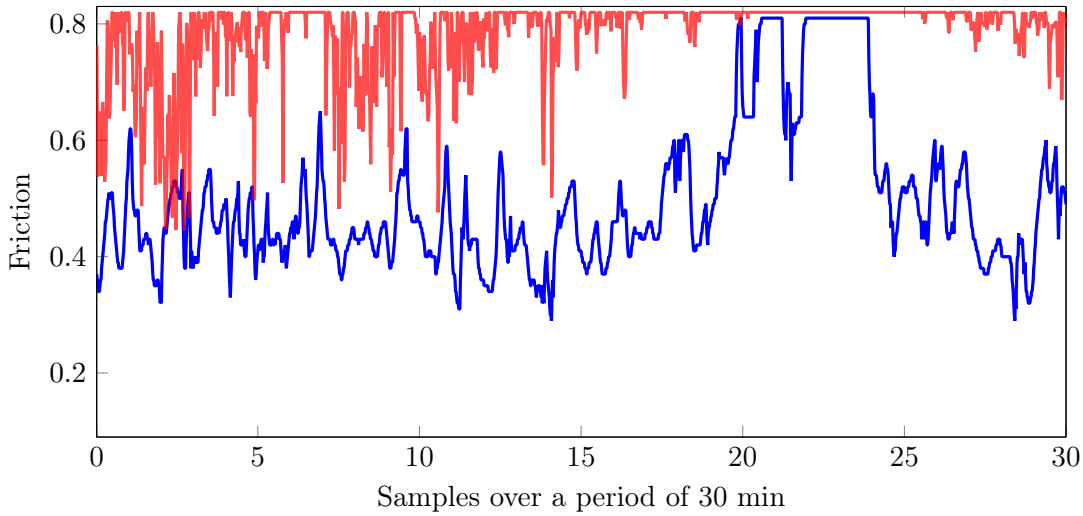


Figure 15.6: A plot of the friction for RCM411 (blue) and MARWIS (red) on icy road conditions. The x-axis shows the samples along the road segment measured every second, and the y-axis is the friction values.

ditions appear MARWIS begins jumping between max and min friction almost everywhere, in contrast to RCM411 which has smaller changes. Taking into account the laboratory test from the three previous friction comparisons it can be argued that the MARWIS behaviour is similar, showing low values for water, high values for ice, and almost the same for snow and dry conditions. If in addition the spike phenomenon of MARWIS is taken into account then the differences between RCM411 and MARWIS on mixed road conditions is what would be expected.

A general conclusion to the friction test results is that the laboratory experiments compared to the field experiments had on average the same trend in the measurement data, which is a good sign that strengthens the reliability of the results in both experiments. As for the spiky behaviour of MARWIS in the field tests it is hard to pinpoint the exact cause without doing specific tests. Some explanations could be related to loss of data (where the value is just set to 0.1 or 0.8), placement and angle of the sensor, or vibrations while driving.

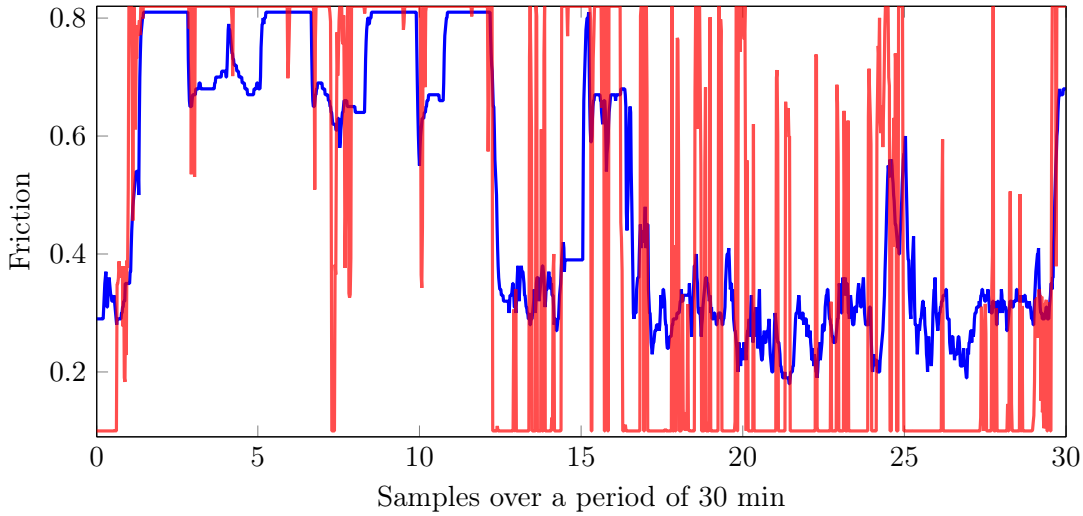


Figure 15.7: A plot of the friction for RCM411 (blue) and MARWIS (red) on mixed road conditions. The x-axis shows the samples along the road segment measured every second, and the y-axis is the friction values.

Water thickness

The water thickness measurements, from RCM411 and MARWIS, for the four road segments, in Table 15.1, can be seen in figures 15.8, 15.9, 15.10, and 15.11, respectively. At first glance of all figures the measurements from the two sensors do not seem to agree when it comes to water thickness. MARWIS (red line in the graphs) seems to show water a lot more frequently than RCM411 (blue line in the graphs), and with higher values.

In Figure 15.8 with road conditions being mostly snowy the water thickness behaviour is somewhat similar in both sensors (staying almost constant), but the values are showing deviating results. The RCM411 has a constant water thickness of zero for the whole segment, while MARWIS shows an almost constant water thickness around 0.3 mm with some scattered values going higher or lower. In the laboratory experiments in [10] the water thickness on snow showed that RCM411 did not detect water for any snow types except loose snow, while MARWIS did detect water for all tested snow types. In addition the lab tests showed that MARWIS was more sensitive to changes in water thickness on a general basis, detecting small changes that RCM411 missed. By comparing the field results with the lab results, we conclude that the deviation in the values between the two sensors in the field tests are exactly reflecting the results from the lab tests, and even describes why MARWIS has scattered outliers.

In Figure 15.9 with road conditions being dry and wet interchangeably the water thickness behaves similar for both sensors, showing water in the same areas. When looking at the values, MARWIS measures higher water thickness than RCM411 almost everywhere on the road segment. From the laboratory tests in [11] for wet asphalt the water thickness measurements are said to be unreliable for both sensors. Two experiments conducted in 2015 and 2016 gave different results regarding over and under estimations, and the color of the asphalt affected the size of the error. By considering this and observing the results from the field tests we conclude that both sensors show water in the same areas, however, using the water thickness values has to be done with care because they can be inaccurate. In the field experiments RCM411 water thickness was on average about 1 mm below MARWIS (which also had some measurements reaching the max of 6 mm), that may have been caused by inaccuracy or because of the sensitive nature of MARWIS.

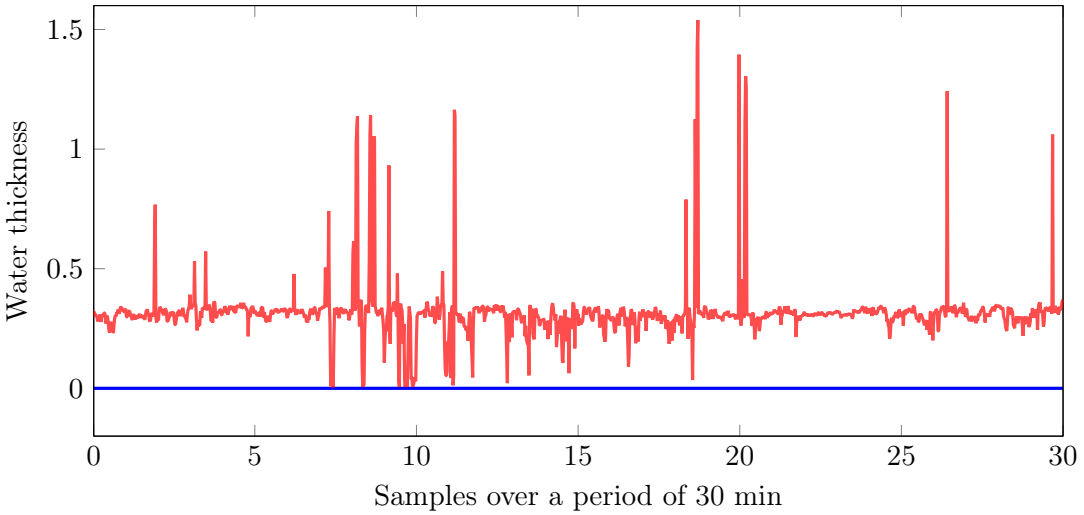


Figure 15.8: A plot of the water thickness for RCM411 (blue) and MARWIS (red) on snowy road conditions. The x-axis shows the samples along the road segment measured every second, and the y-axis is the water thickness in mm.

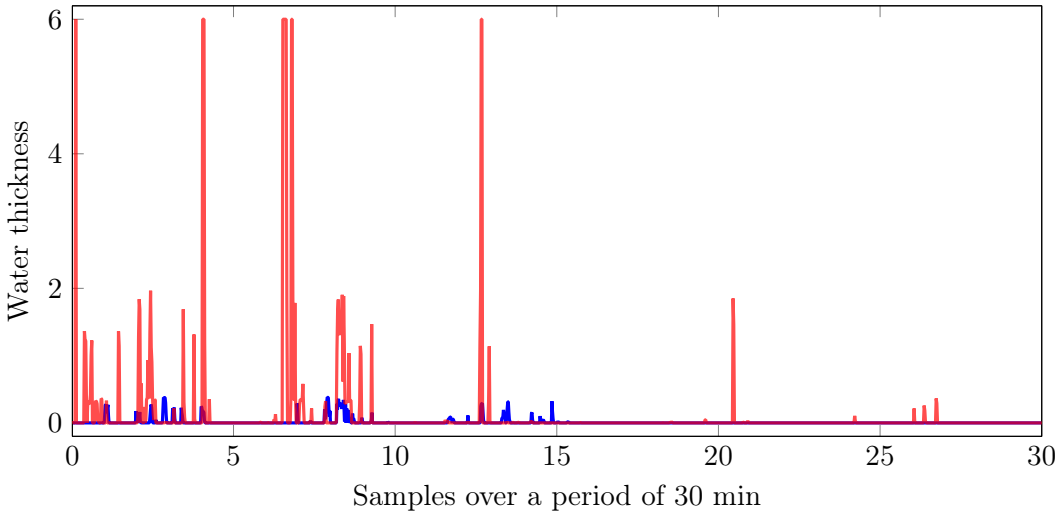


Figure 15.9: A plot of the water thickness for RCM411 (blue) and MARWIS (red) on dry and wet road conditions interchangeably. The x-axis shows the samples along the road segment measured every second, and the y-axis is the water thickness in mm.

In Figure 15.10 with road conditions being mostly icy the water thickness is dissimilar in both behaviour and value when comparing the sensors. RCM411 values are constantly zero while MARWIS values are fluctuating between 0 and 0.25. In the laboratory experiments in [10] the water thickness was not deliberately tested on ice, however, for one of the test cases the water froze almost instantly when applying it to the surface, and in these tests RCM411 did not detect water on the ice, while MARWIS did. This means that the behaviour of the water thickness for both sensors in the lab had an equal trend as the water thickness on snow. Comparing the field tests with the lab tests show that they behave similarly, and that ice is a little more unpredictable than snow, reflected in the fluctuating values of MARWIS.

In Figure 15.11 with mixed road conditions the water thickness is present on $\frac{2}{3}$ of the segment for both sensors. By studying the conditions on the road segment an observation is made that RCM411 shows water thickness above zero on road parts where it measures icy conditions.

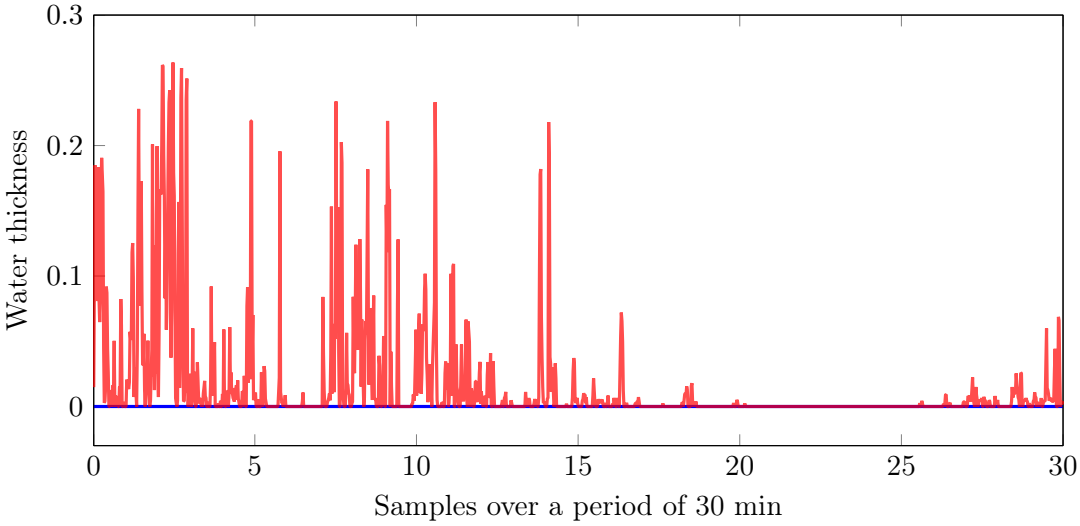


Figure 15.10: A plot of the water thickness for RCM411 (blue) and MARWIS (red) on icy road conditions. The x-axis shows the samples along the road segment measured every second, and the y-axis is the water thickness in mm.

This was not the case in Figure 15.10, where RCM411 showed zero water on ice. For MARWIS the results are as expected since this sensor shows water for all previous conditions except dry asphalt. The only anomaly in the MARWIS data is that it often shows max value. Taking into consideration the lab results in [10] the tests show that icy conditions with a thin film of water is detected by both sensors. This means that the field results for RCM411 is probably correct if we conclude that the icy parts on this segment had water on it while the previous experiment on ice did not. There are no indications from the lab tests as to why MARWIS would show max value when RCM411 is not even close to its max value.

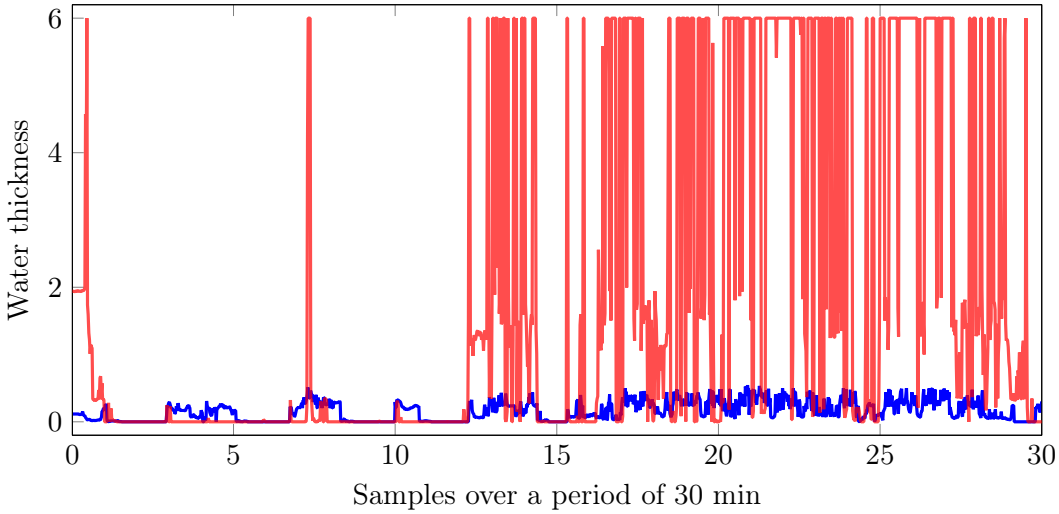


Figure 15.11: A plot of the water thickness for RCM411 (blue) and MARWIS (red) on mixed road conditions. The x-axis shows the samples along the road segment measured every second, and the y-axis is the water thickness in mm.

As a general conclusion we note that the field results for water thickness match the lab results. Some minor dissimilarities are detected, with the most prominent being the max value measurements (6 mm) of MARWIS. Possible reasons for this, as with the spikes in MARWIS friction data, could be loss of data (where the value is just set to 6 mm), placement and angle of the

sensor, or vibrations while driving.

Road temperature

The road temperature measurements, from RCM411 and MARWIS, for the four road segments, in Table 15.1, can be seen in figures 15.12, 15.13, 15.14, and 15.15, respectively. The overall impression of the measurements in all figures is that the trend of the temperature is alike for both sensors. When it comes to the values it seems that RCM411 (blue line in the graphs) has higher values for some road conditions than MARWIS (red line in the graphs), and for other conditions the values are almost equal.

In Figure 15.12 with road conditions being mostly snowy the behaviour of the road temperature is similar when comparing the two sensors, however, the values of RCM411 is constantly about 2 degrees Celsius higher than MARWIS, for the entire road segment. In the laboratory test in [10] the road temperature measurements were showing large differences between RCM411 and MARWIS. The reason for these differences was problems regarding calibration of the MARWIS sensor, which caused MARWIS to display temperatures that were on average 7 degrees Celsius colder than RCM411 on snowy conditions. In our field test it can be seen from the graph that we also obtained lower values for MARWIS, but only 2 degrees Celsius lower on average, so in that sense the calibration was probably more accurate for our tests.

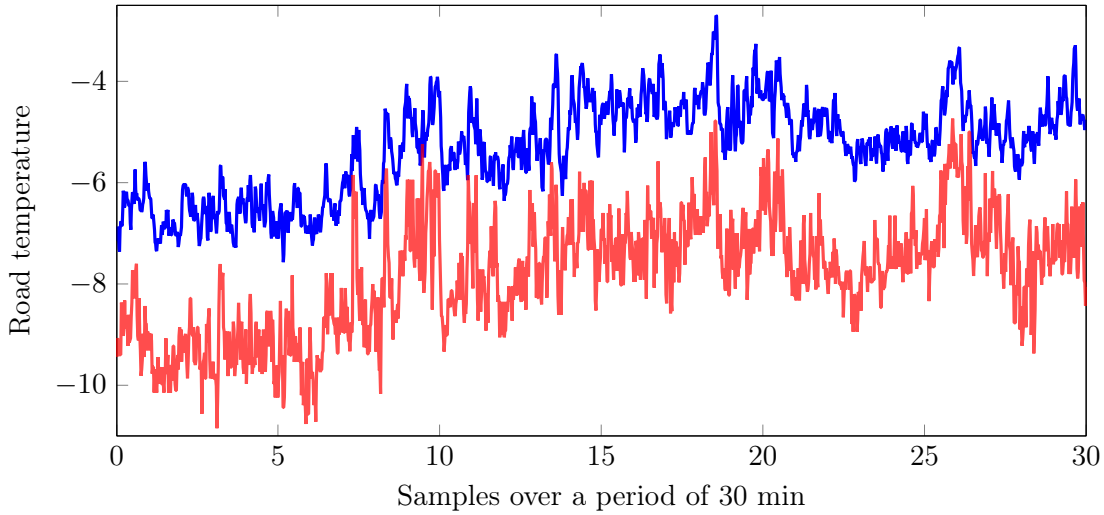


Figure 15.12: A plot of the road temperature for RCM411 (blue) and MARWIS (red) on snowy road conditions. The x-axis shows the samples along the road segment measured every second, and the y-axis is the road temperature in Celsius.

In Figure 15.13 with road conditions being dry and wet interchangeably the road temperature is almost equal in both behaviour and value for the sensors. In some areas MARWIS has some lower values compared to RCM411, but the overall temperature is mainly alike on the road segment. The laboratory tests in [10] did not test road temperatures on dry or wet roads, so a comparison against lab results are not made. However, based on the results from the field tests, RCM411 and MARWIS seem stable in regard to each other when measuring road temperature on dry and wet surfaces.

In Figure 15.14 with road conditions being mostly icy the road temperature for the sensors follows the same pattern as on snowy conditions. The behaviour is alike, but the values of RCM411 is on average about 2 degrees Celsius higher than MARWIS. The laboratory tests in [10] did not test road temperatures on icy roads, so a comparison against lab results are not

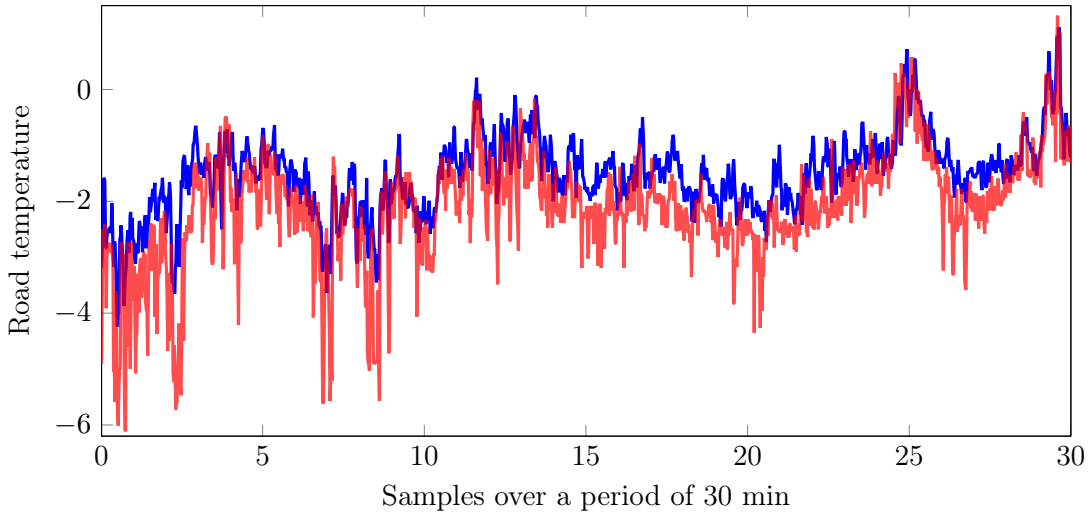


Figure 15.13: A plot of the road temperature for RCM411 (blue) and MARWIS (red) on dry and wet road conditions interchangeably. The x-axis shows the samples along the road segment measured every second, and the y-axis is the road temperature in Celsius.

made here either. From the field experiments the comparison in the graph looks a lot like the comparison in Figure 15.12, with slightly higher temperatures. From this, a conclusion that the road temperature from the sensors act equally on snowy and icy roads is plausible.

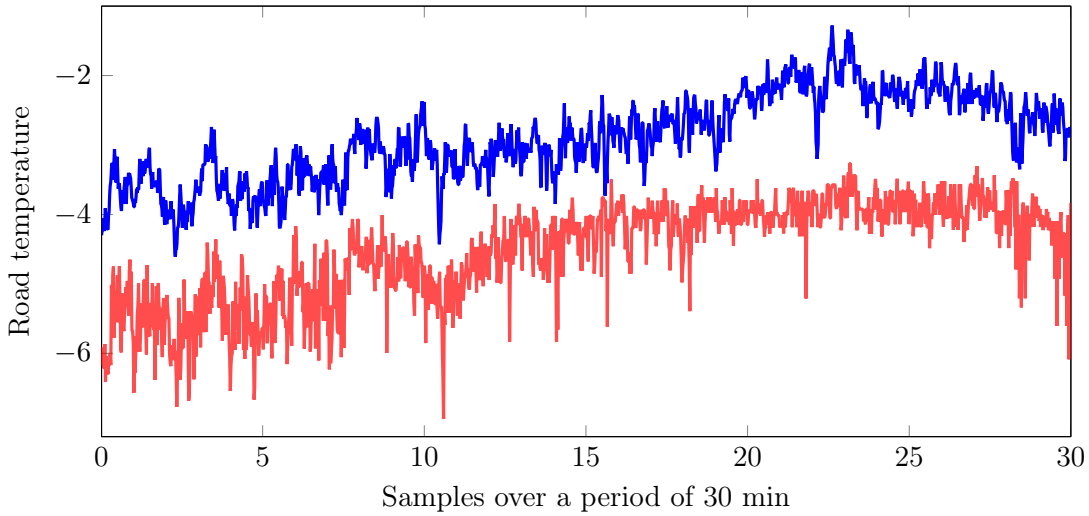


Figure 15.14: A plot of the road temperature for RCM411 (blue) and MARWIS (red) on icy road conditions. The x-axis shows the samples along the road segment measured every second, and the y-axis is the road temperature in Celsius.

In Figure 15.15 with mixed road conditions it appears that the road temperature follows the same trend as the three previous tests. In some areas the values are similar (here we expect dry or wet asphalt) and in other areas MARWIS shows lower temperatures than RCM411 (here we expect snow or ice covered asphalt). As a whole the sensors seem consistent when measuring road temperature, and we notice that the analogous spikes in the MARWIS data from the friction and water thickness experiments are not present in the road temperature measurements.

The general conclusion to the road temperature field tests is that for dry and wet roads the temperature is mainly equal for the sensors, while on snowy and icy roads the temperature from

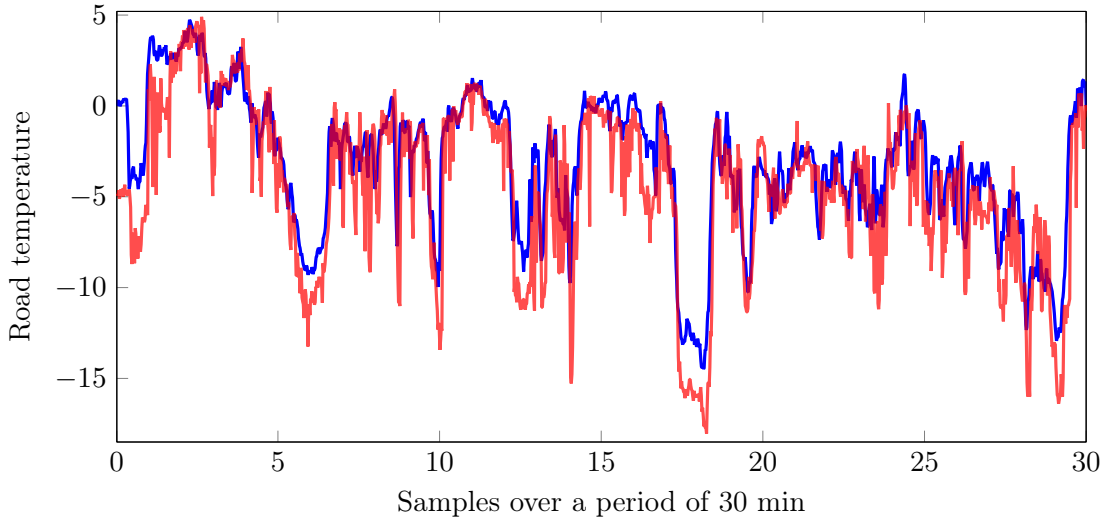


Figure 15.15: A plot of the road temperature for RCM411 (blue) and MARWIS (red) on mixed road conditions. The x-axis shows the samples along the road segment measured every second, and the y-axis is the road temperature in Celsius.

MARWIS is on average, in our case, 2 degrees Celsius lower than the temperature from RCM411. An explanation for the deviation between the RCM411 and MARWIS values on snow and ice could be the difference in placement of the sensors, or calibration of the sensors.

15.4.2 OBD-II

The data collected from the OBD-II is in itself not directly suitable for winter road assessment. However, the data can be used as a supplementary tool to other sensors in the assessment. From previous research OBD-II has been utilized together with a smartphone to detect vehicle skidding [4], it has been modified and expanded on to create a device that allows customization by the end-user to detect road conditions [33], and it has been used together with a smartphone, and road and air temperature to detect slippery roads [34]. The common factor in these three examples is that the OBD-II data is only one part in a larger integrated system, relying on algorithms with multiple input to evaluate road conditions. OBD-II can give information regarding the vehicle that other sensors know nothing about, for instance, vehicle wheel speed, vehicle acceleration, steering wheel angle, and brake pressure, to mention some relevant variables. The information collected from the vehicle can be a great addition in winter road assessment, both in order to increase useful data and maybe reduce the need for other sensors, hence saving costs. A relevant example that should be mentioned within this topic is [35] and their idea of a virtual sensor. Although they do not directly use the OBD-II, they still collect various data from the vehicle and run that through an algorithm to estimate the friction between road and tire, and the tire pressure. In our field experiments the OBD-II data has not been used to evaluate road conditions momentarily. Data has been collected sequentially for the whole test route, so future work may include tests involving OBD-II measurements.

15.4.3 Walabot

The Walabot sensor is an experimental radar sensor which we have decided to test as a tool in winter road assessment. The reason for choosing the Walabot sensor is its small size and low cost, as well as it being a radar. This is interesting since radar technology has not been tested extensively for winter road assessment, which in theory it should be suited for given that radar is not dependent on sight, being less affected by snowy or foggy conditions where, for instance,

optical sensors have a problem. The sensor is tested against itself on three segments with different road conditions (snow, ice, and dry/wet). Figure 15.16 shows the comparison of the three measurements; snowy conditions in blue, icy conditions in green, and dry/wet conditions in red. From the figure we see that the measurements are very similar, and it is hard to draw a conclusion based on this one result alone. Research on the use of radar in winter road assessment has been done before: two examples are [6] and [7]. In these two papers, the backscattering properties of various surfaces are studied to estimate road conditions. The conclusion drawn in both papers is that the radar is suitable to assess the road under certain conditions. Two important conditions are the incidence angle of the radar and the surface roughness. In our field experiments the incidence angle was zero in all tests, whereas in the two papers [6, 7] a large angle (50-80 degrees) was shown to highlight different conditions to a greater extent than small angles. In addition, small changes in the surface roughness was proven to cause large variation in the measurements. Based on these results and Figure 15.16 it can be concluded that more tests are required to say something constructive about the abilities of the Walabot to determine road conditions, however, we have opted for including these measurements to show some preliminary results concerning the setup and use of the Walabot.

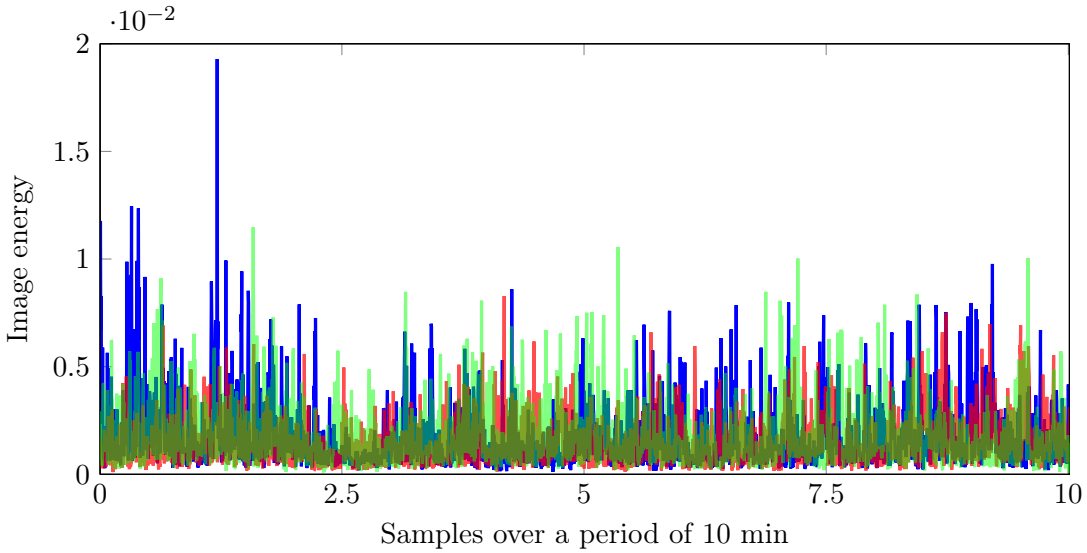


Figure 15.16: A plot of Walabot energy for snowy (blue), icy (green), and dry/wet (red) road conditions. The x-axis shows the samples along the road segment measured between 4 and 5 times every second, and the y-axis is the image energy.

When analysing the Walabot data, after the experiments were completed, we noticed two phenomena caused by outside forces which were affecting the Walabot measurements. The first phenomenon is that the Walabot sensor was sensitive to water and ice covering the sensor's casing. In Figure 15.17 the Walabot image energy (red) can be seen along a road segment, where in the middle of the segment RCM411 measures water (blue) on the asphalt. The figure shows that a little while after driving through water the image energy values drop significantly, as a consequence of water, or possibly ice (air temperatures were around -1 degree Celsius), covering the sensor. About 5 min after this incident the sensor was cleaned and values returned back to normal. The example given in Figure 15.17 was one of the most prominent examples of water or ice affecting the sensor in this way. No other Walabot data showed the same decrease in value, so a conclusion here is that conditions can disrupt the Walabot measurements if the setting is just right, however, this does not happen often and can be prevented by considering the placement and containment of the sensor. As mentioned, the reasoning in using Walabot in the experiments was that radar could be a useful addition to optical sensing as it does not depend on sight. But as the field experiment show, snow/slush build-up on the sensor is detrimental to its sensing ca-

pabilities, and it has some of the same problems as optical sensors in harsh conditions. However, a better configuration, both mechanical and electronic, would be interesting to investigate.

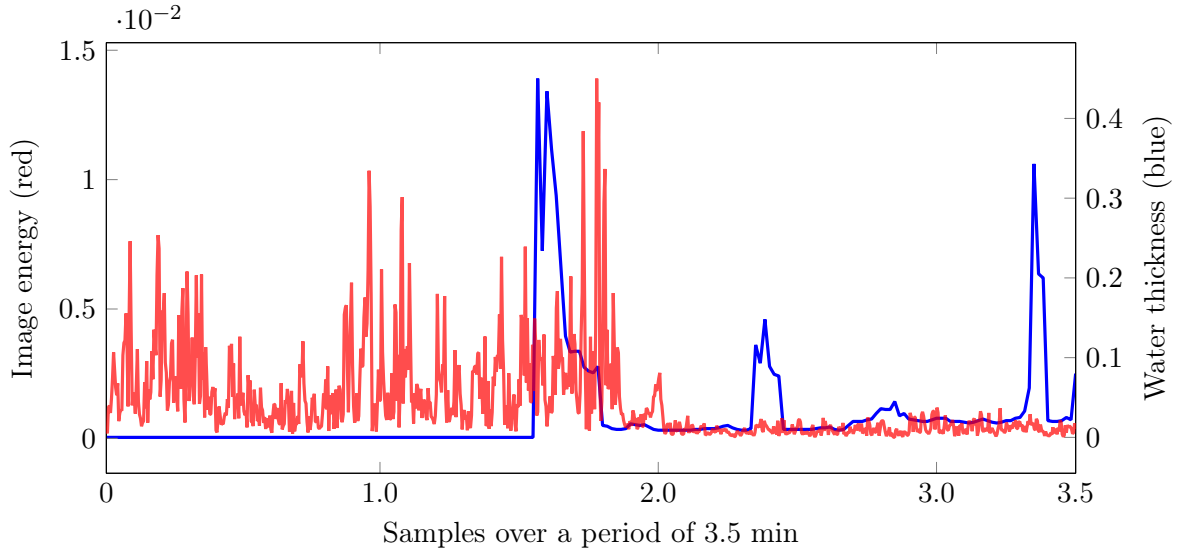


Figure 15.17: The Walabot energy (red) being affected by RCM411 water thickness (blue). The x-axis shows the number of samples during the measured period, between 4 and 5 samples every second for the Walabot and every second for the RCM411. The y-axis is the image energy for the red graph (Walabot) and water thickness (mm) for the blue graph (RCM411).

The second phenomenon of outside forces affecting the Walabot measurements was the sensor's sensitivity to vibrations of the car. An example can be seen in Figure 15.18 where the values of the Walabot data (red) approaches zero when the vehicle is at a standstill (represented by the blue line which is km driven). The small ripples in the area where the car is parked shows people getting out of the vehicle or in to the vehicle, and the door is opened and closed. In

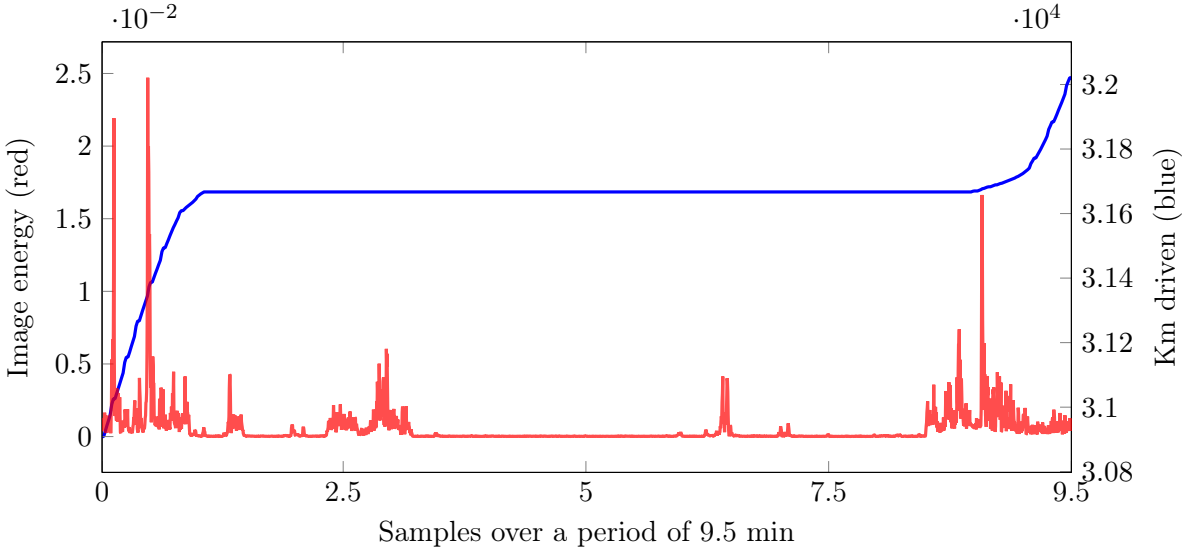


Figure 15.18: The Walabot energy (red) being affected by car vibration. The x-axis shows the number of samples during the measured period, between 4 and 5 samples every second for the Walabot and every second for the RCM411. The y-axis shows image energy for the red graph and km driven (from RCM411) for the blue graph, where the flat area indicates that the car is at a standstill.

addition to this example, all other stops along the test track showed similar behaviour from the Walabot. As the Walabot was configured to measure total image energy, with fast capture rate, the Doppler component of the signal is strongly connected to the effects of vehicle speed and vibrations. Change in surface conditions, should give a change in the backscatter component of the signal, but preliminary analysis show that it is so small compared to the effect of vibrations that it may be impossible to isolate it in this data set. Further investigations should consider other configurations of the radar, and the possible use of multi-resolution signal analysis (wavelet/mra) for extracting relevant parts of the radar signal.

The conclusion on using Walabot as a tool in winter road assessment is that, with the settings and placement of the sensor, and the fact that we collected image energy, the sensor did not give unambiguous answers. Reading previous research on the topic gave us reason to believe that more experiments with other parameters can give different, and possibly better, results. In addition, the behaviour of the sensor data led to a hypothesis that the sensor only return vehicle vibrations, and nothing else. If other results would have appeared given that, for instance, raw data was collected instead of image energy or the incidence angle was changed is a question for future work, which would be interesting to investigate.

15.5 Conclusion

In this paper we performed field experiments with four sensors, RCM411, MARWIS, OBD-II, and Walabot, for use in winter road assessment. RCM411 and MARWIS were tested against each other, and compared to laboratory experiments. OBD-II data was collected for future us as a supplementary tool in the road assessment. And initial tests for the experimental Walabot sensor were carried out.

The comparisons between RCM411 and MARWIS showed that the field tests followed the same pattern as the laboratory experiments, for all three variables (friction, water thickness, and the limited data on road temperature). Regarding the values of the data, RCM411 returned predictable values (as expected from the knowledge of the test route, and the lab tests) in all experiments, while MARWIS had predictable behaviour but experienced spikes in the values for friction and water thickness, possibly explained by loss of data, placement and angle of the sensor, or vibrations. Overall, the results from the RCM411/MARWIS experiments were consistent with the lab results, and the sensors showed similarities between themselves, although MARWIS had some incorrect readings and seemed to be more sensitive to small changes in the road conditions, especially for water thickness.

In the OBD-II analysis previous research showed that OBD-II data could be used as a supplementary tool in winter road assessment, mainly to detect slippery road surface. The research combined data from the OBD-II with data from other sensors, running them through an algorithm to evaluate road conditions. Results were positive and interesting, and gave good arguments for the use of OBD-II data in winter road assessment. In our project OBD-II has not been applied in the assessments yet, however, data from the test route exists and may be considered for future work.

Lastly, the Walabot tests, and previous research, indicated that more experiments should be conducted before giving a conclusion on the suitability of the sensor in winter road assessment. From the results in the field experiments different surface conditions did not appear considerably different in the Walabot data, and it was also discovered that the sensor was sensitive to water/ice covering the casing, and to vibrations caused by the vehicle. In future work controlled experiments should be performed where several parameters are changed, in order to increase the

test set and give a more accurate measurement.

15.6 Data Availability

The data used to support the findings of this study is available from the corresponding author upon request.

15.7 Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

15.8 Funding Statement

This research was in part financed by the Interreg-Nord project "WiRMa" [project ID 20201092].

References

[1] P. Jonsson. Classification of road conditions: From camera images and weather data. In *2011 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications (CIMSA) Proceedings*, pages 1–6, Sep. 2011.

[2] S. Kawai, K. Takeuchi, K. Shibata, and Y. Horita. A method to distinguish road surface conditions for car-mounted camera images at night-time. In *2012 12th International Conference on ITS Telecommunications*, pages 668–672, Nov 2012.

[3] R. Omer and L. Fu. An automatic image recognition system for winter road surface condition classification. In *13th International IEEE Conference on Intelligent Transportation Systems*, pages 1375–1379, Sep. 2010.

[4] Y. Hou, A. Gupta, T. Guan, S. Hu, L. Su, and C. Qiao. Vehsense: Slippery road detection using smartphones. In *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, pages 1–5, June 2017.

[5] Patrik Jonsson and Mats Riehm. Infrared thermometry in winter road maintenance. *Journal of Atmospheric and Oceanic Technology*, 29(6):846–856, 2012.

[6] J. Häkli, J. Säily, P. Koivisto, I. Huhtinen, T. Dufva, A. Rautiainen, H. Toivanen, and K. Nummila. Road surface condition detection using 24 GHz automotive radar technology. In *2013 14th International Radar Symposium (IRS)*, volume 2, pages 702–707, June 2013.

[7] V. V. Viikari, T. Varpula, and M. Kantanen. Road-condition recognition using 24-GHz automotive radar. *IEEE Transactions on Intelligent Transportation Systems*, 10(4):639–648, Dec 2009.

[8] Taisto Haavasoja, Juhani Nylander, and Pauli Nylander. Experiences of mobile road condition monitoring. In *Proceedings of SIRWEC 2012, Helsinki, Finland*, May 2012.

[9] Lufft. *MARWIS specifications*, 2019 (Accessed 2019-09-05). `https://www.lufft-marwis.com/en_US/specifications`.

[10] Laura Fay, Michelle Akin, Anburaj Muthumani, and Xianming Shi. Quantifying Salt Concentration on Pavement: Phase II – MARWIS and Teconer Laboratory Sensor Evaluation. Technical report, Western Transportation Institute, Montana State University, June 2018.

[11] Johan Wåhlin. Laboratory tests of optical road condition sensors, 2015 and 2016. Technical report, Statens Vegvesen, Norway; Trafikverket, Sweden; Vejdirektoratet, Denmark; Tiehallinto, Finland, Vegagerdin, Iceland and Landsverk Faroe Islands, November 2017.

[12] Alfredas Laurinavičius, Donatas Čygas, Audrius Vaitkus, Tomas Ratkevičius, Matas Bulevičius, Darjušas Mučinis, and Andrius Baltrušaitis. Research of snow melting materials performance efficiency for road winter maintenance. *Transport*, 31:322–332, 07 2016.

[13] Timo Sukuvaara, Kari Mäenpää, Riika Ylitalo, Heikki Konttaniemi, Juha Petäjäjärvi, Juho Veskoniemi, and Matti Autioniemi. Vehicular networking road weather information system tailored for arctic winter conditions. *International Journal of Communication Networks and Information Security (IJCNIS)*, 7:60–68, 04 2015.

[14] A. Ryguła, M. Honkisz, A. Maczyński, K. Brzozowski, and A. Konior. Mobile system for recording road weather conditions. *Archives of Transport System Telematics*, Vol. 9, iss. 3:42–45, 2016.

[15] Bernhard Schmiedel, Frank Gauterin, and Hans-Joachim Unrau. Study of system layouts for road wetness quantification via tire spray. *Automotive and Engine Technology*, 4(1):63–73, Jun 2019.

[16] Lauri Lovén, Virve Karsisto, Heikki Järvinen, Mikko J. Sillanpää, Teemu Leppänen, Ella Peltonen, Susanna Pirttikangas, and Jukka Riekki. Mobile road weather sensor calibration by sensor fusion and linear mixed models. *PLOS ONE*, 14(2):1–17, 02 2019.

[17] Armando Piccardi and Lorenzo Colace. Optical detection of dangerous road conditions. *Sensors*, 19:1360, 03 2019.

[18] Jakob Döring, Lakshan Tharmakularajah, Jakob Happel, and Karl-Ludwig Krieger. A novel approach for road surface wetness detection with planar capacitive sensors. *Journal of Sensors and Sensor Systems*, 8:57–66, 01 2019.

[19] Giovanni Pepe, Leonardo Gabrielli, Livio Ambrosini, Stefano Squartini, and Luca Cattani. Detecting road surface wetness using microphones and convolutional neural networks. In *Audio Engineering Society Convention 146*, March 2019.

[20] Kang Gui, Lin Ye, Junfeng Ge, Faouzi Alaya Cheikh, and Lizhen Huang. Road surface condition detection utilizing resonance frequency and optical technologies. *Sensors and Actuators A: Physical*, 297:111540, 2019.

[21] Choongheon Yang, Dukgeun Yun, Jinguk Kim, and Sungpil Shin. Estimation road surface temperature variation using commercial vehicle ambient sensor. *IOP Conference Series: Materials Science and Engineering*, 603:022027, 09 2019.

[22] G. Pan, L. Fu, R. Yu, and M. Muresan. Evaluation of alternative pre-trained convolutional neural networks for winter road surface condition monitoring. In *2019 5th International Conference on Transportation Information and Safety (ICTIS)*, pages 614–620, July 2019.

[23] J. Autioniemi, M. Autioniemi, J. Casselgren, H. Konttaniemi, T. Sukuvaara, and R. Ylitalo. Intelligent road. Technical report, Lapland university of Applied Sciences, Luleå University of Technology, Finnish Meteorological institute, 2015.

[24] Teconer. *Winter maintenance of roads and runways*, 2015 (Accessed 2019-09-05). `http://www.teconer.fi/en/winter.html`.

[25] Lufft. *USER Manual MARWIS/StaRWIS*, 2018 (Accessed 2019-10-01). `https://www.lufft.com/download/manual-lufft-marwis-starwis-en/`.

[26] Peter Rau. *User Manual - MARWIS App*, 2017 (Accessed 2019-10-02). `https://www.lufft.com/download/manual-marwis-ios-app-en/`.

[27] Lufft. *ViewMondo User's Manual*, 2018 (Accessed 2019-10-02). `https://www.lufft.com/download/manual-lufft-viewmondo-en/`.

[28] OBD Solutions. *STN1100 Family Reference and Programming Manual*, (Accessed 2019-10-21). `https://www.scantool.net/downloads/98/stn1100-frpm.pdf`.

[29] OBD Solutions. What is obd? Online, (Accessed 2019-10-21). `https://www.obdsol.com/knowledgebase/on-board-diagnostics/what-is-obd/`.

[30] Walabot. *Walabot - Technical Brief*, (Accessed 2019-09-05). `https://walabot.com/docs/walabot-tech-brief-416?type=pdf`.

[31] SparkFun. Getting started with walabot. Online, 2018 (Accessed 2019-10-02). https://learn.sparkfun.com/tutorials/getting-started-with-walabot?_ga=2.140213947.929731503.1570004485-127684519.1570004485.

[32] Youtube. Walabot Programmable 3D Sensor. Online videos, (Accessed 2019-10-03). https://www.youtube.com/channel/UCdQcnXMiEGizQvPODZIwvBA/videos.

[33] D. J. Enriquez, A. Bautista, P. Field, S. Kim, S. Jensen, M. Ali, and J. Miller. Canopnr: Can-obd programmable-expandable network-enabled reader for real-time tracking of slippery road conditions using vehicular parameters. In *2012 15th International IEEE Conference on Intelligent Transportation Systems*, pages 260–264, Sep. 2012.

[34] Ralph Robinson. Slippery road detection and evaluation. Technical report, University of Michigan, Transportation Research Institute, 2901 Baxter Road, Ann Arbor, MI 48109, May 2012.

[35] Fredrik Gustafsson, Niclas Persson, Markus Drevö, Urban Forssell, Henrik Quicklund, and Mats Löfgren. Virtual sensors of tire pressure and road friction. Technical Report 2376, Linköping University, Automatic Control, 2001.