Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

Sivert Selnes

Railway condition monitoring applying the continuous and discrete wavelet transforms

A time-frequency analysis of wear and maintenance effects in the track geometry of railway turnouts

Master's thesis in Applied Physics and Mathematics
Supervisor: Ingelin Steinsland
June 2020

Arne Hückelheim/Wikimedia Commons

NTNU
Kunnskap for en bedre verden

Sivert Selnes

Railway condition monitoring applying the continuous and discrete wavelet transforms

A time-frequency analysis of wear and maintenance effects in the track geometry of railway turnouts

Master's thesis in Applied Physics and Mathematics
Supervisor: Ingelin Steinsland
June 2020

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

Preface

Written in Lyngby, Denmark, during the spring of 2020, this thesis concludes my Master of Science in Applied Physics and Mathematics, specialized in Industrial Mathematics, at the Norwegian University of Science and Technology. I had to apply for this programme because it was one of the few studies for which I could envision the content - mechanical engineering was named something with product development, and the cybernetics programme felt like rolling the dice with a dictionary in hand. Besides, the picture on the home page of Applied Physics and Mathematics looked cool (a man staring at a bright light, I thought he was welding). It was a no-brainer. Five years later, I am quite happy to make it through and conclude that the name of Applied Physics and Mathematics was a real promise, while the picture was not.

I sincerely thank my supervisor, professor Ingelin Steinsland (NTNU), for her input and light spirit. She has been of great support. I would also like to express my gratitude towards my co-supervisor, associate professor Line Clemmensen at the Technical University of Denmark (DTU), for accepting my visit and for introducing the topic of railways.

Finally, I thank my fellow students and friends for their excellent company in Trondheim, for their open-heartedness, colourful minds, and all our laughs together.

Lyngby, Danmark
June 2020

Sivert Selnes

©Geir Mogen/NTNU

Summary

In this thesis, wavelet transformations are considered as a method for condition monitoring of railway turnouts, utilizing track geometry data collected by a track inspection vehicle. The track geometry data are measurements of track width (gauge), rail height difference (cant), and vertical and horizontal variations of each rail individually, sampled every 25 cm. Three turnouts with known maintenance history are considered, each observed nine times by the track inspection vehicle for 2-4 years. The continuous wavelet transform (CWT) is applied to the data, investigating how the track geometry degrades with time (and tonnage) and evaluating the effect of track adjustments (ballast tamping, restoring the geometry). Both cases are compared to the current method for assessing track quality, typically more straightforward summary statistics. Furthermore, the discrete wavelet transform (DWT) is tested on a turnout subject to gradually increasing wear, comparing the response of the various levels of DWT coefficients to the results of the CWT, and the reference statistics. For the discrete *symlet 7*, the coarsest levels are found to correlate most with the measurements. The turnouts under analysis showed similar characteristics in terms of CWT frequency modes and peak positions in the geometry data. The study of track adjustments revealed large variations in how successful track corrections are, some lasting less than six months, others several years.

Sammendrag

Den diskrete og den kontinuerlige bølgetransformasjonen er anvendt på geometrimålinger som en metode for å overvåke sporgeometri i den danske jernbanens sporvekslinger. Geometrimålinger samles regelmessig for hele jernbanenettet av dedikerte måletog, og er mål av sporvidde, -helling, og skinnenes individuelle avvik i vertikal og horisontal retning, målt hver 25. cm. Studien tar utgangspunkt i tre sporvekslinger med kjent vedlikeholdshistorikk, hvis sporgeometri er målt ni ganger over en periode på 2-4 år. Den kontinuerlige bølgetransformasjonen (CWT) er anvendt på dataene for å visualisere eventuell forverring av sporgeometrien over tid. I tillegg vurderes effekten av maskinell sporjustering (tamping), da to av sporvekslingene gjennomgår sporjustering i løpet av perioden. Observasjonene er sammenlignet med skinnenes gjennomsnittlige standardavvik i horisontal og vertikal retning, hvilket er den nåværende metode Banedanmark (ansvarlige for jernbaneinfrastrukturen i Danmark) benytter for å evaluere sporkvalitet (dog primært for rette strekninger og i mindre grad sporvekslinger, på grunn av deres krevende utforming og varierende geometri). I tillegg er den diskrete bølgetransformasjonen (DWT) anvendt på en enkelt skinnes vertikale avvik i en sporveksel som utviser tegn på gradvis forringelse i både CWT og standardavvik, der de grovere detaljnivåene (i den diskrete transformasjonen) responderte mest på den økte slitasjen. Den kontinuerlige transformasjonen avdekket endringer i sporgeometri som forklarer sporvekslingens forringelse med et høyere detaljnivå enn den nåværende praksis tillater, hvilket er særlig nyttig for evalueringen av maskinell sporjustering. Undersøkelsene avdekket store variasjoner i justeringenes varighet; noen sporjusteringer gav forbedret geometri i kun seks måneder før sporet var tilbake til gammel oppførsel, mens andre eksempler viste forbedringer med varighet over 2 år.

Contents

Preface i

Summary i

Table of Contents iv

List of Tables v

List of Figures viii

1 Introduction 1

2 Turnouts & the Railway Universe 5
2.1 Turnouts . . . 5
2.2 Track Structure . . . 6
2.3 Track Life Cycle . . . 7

3 Data 9
3.1 Track Geometry Data . . . 9
3.2 Repair & Maintenance Records . . . 12
3.3 Cases . . . 12
3.4 Techniques for Track Geometry Monitoring . . . 13

4 Wavelets 15
4.1 Time-Frequency Analysis . . . 15
4.2 The Continuous Wavelet Transform . . . 17
4.2.1 Wavelets . . . 18
4.2.2 Transform . . . 19
4.2.3 Scale and Characteristic Frequency . . . 20
4.2.4 Scalograms . . . 24
4.3 The Discrete Wavelet Transform and Multiresolution Analysis . . . 25

4.3.1 Discrete Wavelets 27
4.3.2 Multiresolution Analysis 28
4.3.3 Vanishing Moments, Regularity, and Support 31

5 Method **33**
5.1 General Turnout Signature 34
5.2 Geometric Degradation 34
5.3 Track Adjustments 35

6 Results **37**
6.1 General Turnout Signature 38
6.2 Geometric Degradation 40
6.3 Track Adjustments 42
6.3.1 Turnout B 43
6.3.2 Turnout C 44

7 Discussion & Conclusion **49**

Bibliography **51**

A The Fourier transform, the short-time Fourier transform, and proof of the CWT **55**

B Scalograms **57**

C Code **63**
C.1 Geometry data pre-processing 63
C.2 Figures for theory 66
C.3 The continuous wavelet transform 71
C.4 The discrete wavelet transform 74

List of Tables

3.1 Track geometry measurements 10
3.2 Case turnout summary 12
3.3 Track geometry sampling 13

4.1 Scale and corresponding central frequency for the complex Morlet wavelet 23
4.2 High-pass/low-pass decomposition filter coefficients for the Daubechies *Symlet-7* 32

6.1 Mean std. dev. of `Level` and `Alignment`, turnout A (reference statistic) 41
6.2 DWT detail coefficients of levels 1-4, `Alignment(r)`, turnout A 43
6.3 Mean std. dev. of `Level` and `Alignment`, turnout B (reference statistic) 44
6.4 Mean std. dev. of `Level` and `Alignment`, turnout C (reference statistic) 45

List of Figures

2.1 Sketch of a turnout 6
2.2 The track superstructure: Rails, sleepers (in wood), fasteners securing the rails to the sleepers, and the ballast stones. The railpads are not used in this configuration, but are more often used with concrete sleepers. 7
2.3 The tamping machine 8

3.1 The track inspection car (TIC) 10
3.2 Principle of track measurements: `Level` and `Alignment` 11
3.3 Principle of track measurements: `Cant` and `Gauge` 11
3.4 Track geometry measurement example: `Level` and `Alignment`, turnout B 13
3.5 Track geometry measurement example: `Gauge` and `Cant`, turnout B . . 14

4.1 Time series analysis and Fourier analysis 17
4.2 Short-time Fourier transform and Wavelet analysis 17
4.3 A sine wave in time and frequency 18
4.4 A chirp signal in time and frequency 18
4.5 The complex Morlet wavelet and it's Fourier transform 22
4.6 The Mexican hat wavelet and it's Fourier transform 23
4.7 The real Morlet wavelet and it's Fourier transform 24
4.8 Track geometry measurement for wavelet comparison: `Alignment(l)`, turnout A 25
4.9 The real and imaginary coefficients of the CWT with the complex Morlet 26
4.10 The coefficient amplitude and energy of the CWT with the complex Morlet 26
4.11 The coefficient amplitude and energy of the CWT with the Mexican hat . 27
4.12 The coefficient amplitude and energy of the CWT with the real Morlet . . 27
4.13 The discrete wavelet transform (DWT) implemented as a multiresolution analysis (MRA) 31

6.1 Track geometry comparison (CWT), turnout A, recorded 2016-05-08 . . . 38
6.2 Track geometry comparison (CWT), turnout B, recorded 2016-05-08 . . . 39

6.3 Track geometry comparison (CWT), turnout C, recorded 2015-05-06 . . . 39
6.4 Alignment(r) of turnout A (CWT) 40
6.5 Gauge of turnout A (CWT) 41
6.6 Level(l) of turnout A (CWT) 42
6.7 Level(l) of turnout B (CWT) 44
6.8 Level(r) of turnout B (CWT) 45
6.9 Level(l) of turnout C (CWT) 46
6.10 Alignment(r) of turnout C (CWT) 47

B.1 Level(r) of turnout A, recorded from 2014-02-22 to 2016-11-05 . . . 57
B.2 Cant of turnout A, recorded from 2014-02-22 to 2016-11-05 58
B.3 Alignment(l) of turnout A, recorded from 2014-02-22 to 2016-11-05 58
B.4 Alignment(r) of turnout B, recorded from 2013-03-09 to 2016-11-05 59
B.5 Gauge of turnout B, recorded from 2013-03-09 to 2016-11-05 59
B.6 Level(r) of turnout C, recorded from 2012-06-23 to 2016-11-05 . . . 60
B.7 Alignment(l) of turnout C, recorded from 2012-06-23 to 2016-11-05 60
B.8 Gauge of turnout C, recorded from 2012-06-23 to 2016-11-05 61

Chapter 1

Introduction

Turnouts, often referred to as switches and crossings (S&Cs), connect the railway network, allowing trains to change tracks. They are vital pieces of the railway infrastructure and must be kept in good working condition at all times, as faulty turnouts will affect all adjoining tracks and potentially stall large parts of the railway network. Unlike ordinary tracks, turnouts have movable parts and small discontinuities where the rails cross, but still need to handle the large stresses of several hundred tons trains. This demands a relatively complex structure that makes them susceptible to a broad range of wear, increasing the forces in play when a train passes, and ultimately posing a risk of hazardous failure. S&Cs are particularly frequent in areas where tracks are numerous, and traffic is high, complicating the maintenance planning even further. However, initiating service and maintenance will often stall large parts of the infrastructure too, and both the railway down-time and the maintenance operations themselves are very costly. To maintain the approximately 3500 turnouts in Denmark takes about a third of the total maintenance budget, according to Barkhordari et al. (2017), some 900 M. DKK every year. The decision to maintain, renew, or pass, is, therefore, a critical and difficult one.

Since maintaining the turnouts in a timely manner has such a high priority, precise condition monitoring is key. The complex structure, moving parts, and precision components of switches and crossings make inspections mainly a manual job; however, there exist several sources of various data types for assisting the monitoring. The data treated in this work is track geometry measurements, provided by Banedanmark, managing the Danish railway infrastructure. These measurements are not currently used in a systematic way to determine the condition of turnouts specifically, only for longer track segments (wear thresholds exist for the whole track, including turnouts, but the data for turnouts is not used for maintenance planning, Jøndrup (2019)). It is, therefore, the ambition of this thesis to explore a methodology that can utilize the collected track geometry data in better ways, leading to knowledge enabling the prediction of maintenance demand.

In this work, I propose a method for condition monitoring of track geometry in turnouts

by applying wavelet transformations, as an alternative to the traditional condition monitoring using summary statistics (typically the standard deviation of the measurements as an indication of quality). The work is a continuation of my project thesis written in autumn 2019, treating the same data but using less sophisticated feature extraction to review the geometry. Wavelets are mathematical tools intended for signal processing and time-frequency analysis, in particular. In contrast to Fourier analysis where the analyzing functions are complex exponentials (sines and cosines) of infinite support and energy, wavelets are analyzing functions of finite support and finite energy. This allows for the localization of frequency content in time (or equivalently, in space) within a signal, useful for working with non-stationary signals where the location of certain frequency matters. Developed in the late 80s and early 90s, the discrete wavelet transform (DWT) quickly became the international standard for image representation and compression, known as JPEG2000, Taubman and Marcellin (2002). The DWT is also how the FBI can digitally store and compress their fingerprint bank, consisting of more than 100 million fast searchable fingerprints, (Bradley et al., 1993; Babb and Moore, 2007). The continuous wavelet transform (CWT) is instrumental in medicine, for studying abnormalities in biosignals (for example electrocardiograms/ECGs), Addison (2017), and even used to analyze rudder responses of test aircraft like NASA's F18s, Brenner (2003). The ability of wavelets to represent signals is universal, in the sense of decomposition, compression, reconstruction, or feature extraction.

Importantly, wavelets have also found a wide application in wear and degradation modelling, in particular for the analysis of the vibration signature from gearboxes and engines (Lin and Zuo, 2003; Wu and Liu, 2008), wind-turbines (Jiang et al. (2011)), and roller bearings (Ocak et al. (2007)). Many researchers have applied wavelet analysis to acceleration box data from trains (measuring the train vibrations as it travels the track), identifying rail surface imperfections like rail cracks, wheel flats, and other high-frequency rail irregularities. Caprioli et al. (2007) did a comparative analysis of wavelet and Fourier transform-based techniques, while Jia and Dhanasekar (2007) used wavelets to identify wheel flats. Hopkins and Taheri (2010), Molodova et al. (2013), Molodova et al. (2014), and Cantero and Basu (2015) all considered classification of surface defects in ordinary tracks, from acceleration box data. A more recent study also considered acceleration data for turnouts, Sysyn et al. (2019), and there are a handful of studies that consider other types of data searching for rail defects, for example, magnetic field data of Hall sensors, Toliyat et al. (2003).

Interestingly, there are few or no examples of analysis of track geometry data using wavelets, neither for turnouts nor straight tracks. The combination of the track geometry monitoring as such a major data source for condition monitoring, and that railway infrastructure managers like Banedanmark are still just partially able to utilize this information for their switches and crossings, motivates a study of wavelets as an analysis tool for track geometry data in turnouts. The goal of this work is, therefore, to investigate the usefulness of wavelets for monitoring of track geometry data. The analysis is unsupervised in nature as there are no labels available, although the repair and maintenance records are used to control for major maintenance interventions.

The thesis is written with the following structure: Chapter 2 gives an overview of rail-

way turnouts, track structure, and life cycle, and tamping as a track adjustment technique. The data is presented in Chapter 3, describing the track geometry data and the maintenance records, as well as the practice of the Danish railway infrastructure manager for assessing geometry data. The theory of time-frequency analysis, the continuous wavelet transform, and the discrete wavelet transform is given in Chapter 4. The method used to select turnouts and combine the track geometry data and wavelets to produce the results is described in Chapter 5. The results themselves are given and discussed in Chapter 6, comparing the CWT and the DWT to the current practice of analyzing the wear and geometric deterioration, and track adjustment effects. In Chapter 7, the method in general is discussed and some thoughts for future work are given.

Chapter 2

Turnouts & the Railway Universe

This chapter presents the topic of railway turnouts, track structure, and maintenance. Oftentimes, tracks, turnouts, rails, and S&Cs can be a confusing language for the uninitiated reader. By remembering that it takes two rails and some sleepers to make a track, and a couple of adjoining tracks to make a turnout (S&C), you will be fine. The first part introduces turnouts and some appurtenant terminology in brief. Then the general track structure is presented in depth (quite literally), finally bringing us to the track life cycle and maintenance techniques.

2.1 Turnouts

The turnout type treated in this thesis is illustrated in Figure 2.1. It consists of three main regions and two transition zones.

- *The switch panel.* The first part of the turnout. It consists of two mechanically movable switch blades, directing the train to the correct set of rails. The position of the blades decides whether the passing train continues on the main track or change to the diverging track.
- *The closure panel.* This region lies between the switch blades and the crossing panel. It contains a set of rails for the tracks in the main direction and a set of rails for tracks in the divergent direction.
- *The crossing panel.* This is where the rail in the main direction and the rail in the divergent direction cross each other, meeting at the V-shaped crossing nose.
- *The transition zones.* Before the switch blade (the beginning of the turnout itself) and after the crossing panel. The transition zones are not technically part of the

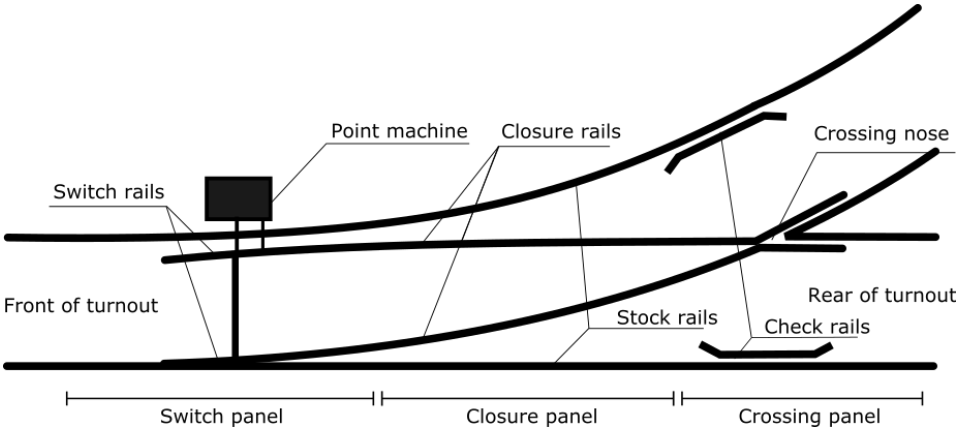


Figure 2.1: Sketch of a turnout with the regions and important parts indicated.

turnout, but they are considered here since the measurements begin and end in these zones to avoid the loss of potentially important information.

2.2 Track Structure

One distinguishes the parts of a track into two main categories: track *superstructure* and track *substructure.* The track superstructure consists of the *rails, fasteners, rail pads, sleepers,* and the *ballast*, see Figure 2.2 (some authors regard the ballast layer as substructure, Tzanakakis (2013) and Jensen (2016)). Starting from the top, the rails are resting on the sleepers, fixed by the fasteners. To dampen some of the high-frequency forces occurring when a train passes, thin rail pads (1-2 cm) are commonly fitted between the rail and the sleeper. The ballast is a layer consisting of crushed stones, on which the sleepers rest. The ballast is central to the following discussion, and serves several purposes:

- Transfer the loads from the train through the rails and sleepers to the substructure, such that the forces acting are as evenly distributed as possible.
- Fixing the track. By ensuring as little movement as possible under each train passing, the strain on rails and sleepers is kept to a minimum, and the gradual deterioration of the track geometry is reduced.
- Drain away water from the track.
- Serve as a reshapable layer between the track and the substructure, such that the track geometry can be adjusted and maintained, even with varying ground conditions and after many years in service.

The track *substructure* consists of the subballast and the subgrade/formation. The subballast is a layer of finer material gravel, resting on top of the subgrade, which is typically leveled earth or rock formation. The substructure is designed to be a steady platform for

Figure 2.2: The track superstructure: Rails, sleepers (in wood), fasteners securing the rails to the sleepers, and the ballast stones. The railpads are not used in this configuration, but are more often used with concrete sleepers.

the superstructure to rest. As it is only maintained or changed in major track maintenance campaigns and may stay untouched for more than 40 years (Tzanakakis (2013)), the substructure is included for completeness, but not relevant for the following discussion.

2.3 Track Life Cycle

When a turnout is new, it rests on the ballast consisting of clean, new skeletal grains. To make sure the turnout has the perfect, projected geometry, it undergoes a series of track adjustments, Jensen (2016). Track adjustments are usually performed by *tamping*, although it sometimes refers to minor manual adjustments (in this context, track adjustment refers to tamping unless otherwise specified). Tamping the tracks reshapes and compacts the ballast, making the tracks rest evenly on the substructure. A tamping machine, Figure 2.3, lifts the tracks slightly, while hydraulically controlled arms are pushed into the ballast, compacting the ballast stones. The machine smooths out track imperfections and ensures that the ballast is evenly distributed without voids.

The reason for tamping new tracks multiple times is two-fold. Firstly, tamping smooths out imperfect geometry, but larger deviations may need several tampings to be within specifications, Jøndrup (2019). Secondly, the new and clean gravel in the ballast layer has a lot of small, sharp edges that quickly wear down when the track is exposed to heavy train loads. As a consequence, the track geometry tends to deteriorate quickly in the initial period, before it settles as the ballast stabilizes. This is called the break-in period, and full stability of the track is typically anticipated after a trainload of approximately 100

Figure 2.3: Tamping machine. ©Plasser and Theurer.

000 tons, Jensen (2016). Tracks are therefore frequently tamped in the break-in period, to ensure that the tracks settle (close) to optimal geometry.

In addition to setting new tracks, tamping is generally the response when the measured track geometry exceeds some pre-defined limits at a later stage. Tracks are typically re-adjusted every 1-5 years depending on the accumulated tonnage and the geometric conditions. Furthermore, as the ballast slowly degrade, the forces between wheels and rails increase (interaction forces due to irregularities in tracks or wheels are known as dynamic forces, whereas the weight of the train itself is called a static force). The increased dynamical forces accelerate structural fatigue in the rails (cracks) and surface wear (abrasive wear, plastic flow, corrugation, and creep, to mention a few, Jensen (2016). These, in turn, increase the contact forces further driving the geometric deterioration. It is, therefore, crucial to limit the dynamic forces to a minimum, and the presence of this vicious circle is well known to all companies responsible for maintaining railway infrastructure. That brings the decision problem of when to schedule track maintenance because the actions, on the one hand, are themselves costly and cause delays, but on the other, keep the track geometric nice and crisp and reduce the dynamical forces and thus the track wear and maintenance needed in the future.

Chapter 3

Data

The data sets considered are provided by Banedanmark, the Danish railway infrastructure manager, and they are made anonymous to comply with their data policy. Three data sets are used: The track geometry recordings, the tamping records, and the manual service history, including routine inspections, smaller repairs like track welding, and complete turnout renewals.

3.1 Track Geometry Data

Track geometry data is recorded by a track inspection car (TIC), Figure 3.1. The TIC regularly inspects the Danish railways, travelling the network at up to a speed of 120 km/h. A railway segment (straight tracks or turnouts) is recorded 1-6 times a year, depending on the admissible track speed (the quality class) and the traffic volume. The TIC is equipped with a range of sensors, including inertial measurement systems, lasers, and cameras, monitoring various aspects of the rail surface, track geometry, and the surrounding vegetation. In particular, the TIC collects the track geometry measurements, which are a range of measures of the vertical and lateral rail movement, sampled every 25 cm. The resulting data series resembles the various aspects of track evenness. For every sample taken, the following geometric features are measured:

- *Level.* Sometimes also called vertical profile. The vertical deviations, from a reference line (see below). Left and right rail are measured independently. Figure 3.2.
- *Alignment.* The horizontal deviations (lateral/sideways movement) from a reference line. Left and right rail are measured independently. Figure 3.2.
- *Gauge.* The perpendicular distance between the left and right rail, given as the deviation from the optimal rail spacing. Figure 3.3.
- *Cant.* The height difference between the left and right rail. Figure 3.3.

Figure 3.1: Track inspection car (TIC), the Eurailscout UFM 120. ©Fisher (2014).

The reference line is computed as the moving average of the surrounding 65 meters for each rail, see Jensen (2016). Some variables, like `gauge`, are easily defined as the distance between the rails. The `level` and `alignment`, however, are measured independently on each rail, relative to this moving average reference line.

By European standard EN 13848-1, the measurements of `level` and `alignment` are filtered and divided into three signals with different frequency content: "short wavelength" $\lambda \in (3, 25)$, "long wavelength" $\lambda \in (25, 70)$, and finally $\lambda \in (70, 150)$ (typically only used to monitor dams and other geo-technical structures, Jensen (2016)). Wavelengths shorter than 3 m are filtered out and not available from the TIC. The data set has a large proportion of missing values for `alignment_D2` (long wavelength), and this part of the signal is therefore discarded for the analysis. The variables are summarized in Table 3.1.

Variable	Description
`Level_R_D1`	Vertical deviation, right rail, short wave period
`Level_L_D1`	Vertical deviation, left rail, short wave period
`Level_R_D2`	Vertical deviation, right rail, long wave period
`Level_L_D2`	Vertical deviation, left rail, long wave period
`Alignment_R_D1`	Horizontal deviation, right rail, short wave period
`Alignment_L_D1`	Horizontal deviation, left rail, short wave period
`Gauge`	Track gauge - the spacing between the rails
`Cant`	Track banking - height difference between the rails

Table 3.1: Summary of the track geometry measurements used in the analysis.

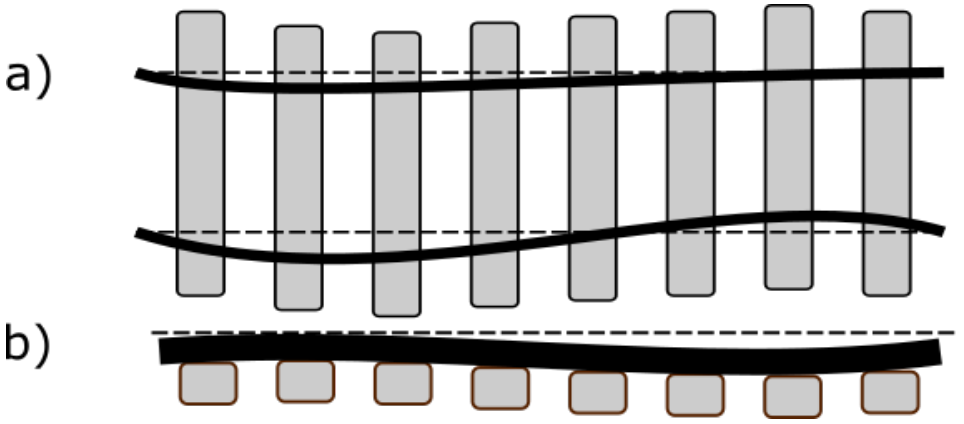


Figure 3.2: Illustration of `alignment` (a) and `level` (b), measured separately for each rail. The dashed lines illustrate the horizontal and vertical reference lines.

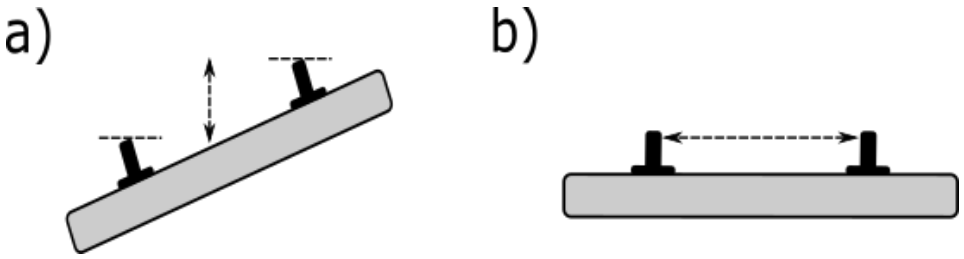


Figure 3.3: Illustration of `cant` (a) and `gauge` (b).

Transforming the Driving Direction: A Small Warning

The driving direction of the measurement car has been standardized in a data preparation procedure by Hovad et al. (2019). The data is initially not well aligned spatially, nor recorded with the same driving direction every time. Hence some turnouts are measured while in a right turn, in a left turn, facing direction or opposite direction. The combination of all these configurations means the data is initially not comparable. Hovad et al. (2019) applied a series of variable transformations to make the data uniform, ensuring that the measurement is the same standard movement (travelling direction and cornering direction). The data preparation comes with a small precaution, as the authors found that the driving direction had some influence on the measurements of `alignment_L_D1`, `gauge`, and `cant` (i.e. driving direction of the TIC matters for the measured geometry, which introduces problems when the measurement series are ”flipped” for standardization). This is the price to pay for using the track geometry data for turnouts as the measurements need to be of uniform direction, and that introduces a bit more noisy results than if all were collected the same way. That means that one should be cautious with results that alternate between two clear patterns, as this might indicate that the TIC sometimes travels the turnout in the opposite direction than usual. One such possible event is noted in the

results.

3.2 Repair & Maintenance Records

Two data sets of maintenance records are considered. The first is the tamping history of each turnout. The other data set contains both inspection and maintenance orders, except for the tamping interventions. The maintenance records do not play a large role in this work but are used to make sure the chosen cases are not ”contaminated” with geometry altering repairs (such as the complete renewal of the turnout, for example). The maintenance action for track adjustment, tamping, is included in the analysis to investigate the effect it has on the track geometry. The motivation for including tamping and neglecting minor repairs, like tightening bolts and surface grinding, is the following:

Track geometry data is commonly used to assess the need for tamping the ballast (Tzanakakis (2013) and Jensen (2016)). The data is not, however, well suited to detect minor faults like loose bolts or perhaps even surface cracks. This is partly due to the sampling rate of track geometry measurements (every 25 cm, while surface defects are typically 1-2 cm long, Molodova et al. (2013)), the movement range measured (the vertical deviations of a typical surface defect are around 0.02 mm, while geometry variables can easily vary between $\pm$ 10 mm over a turnout, Molodova et al. (2014)), and the international standards of storage, discarding the signal content of wavelength less than 3 m, Jensen (2016). Simply put, track geometry measurements capture smoother and larger deviations better than small abrupt faults leading to minor high-frequency fluctuations (such fault are still important to capture, but typically found using acceleration box data, see Chapter 1). Track adjustment campaigns are therefore the primary maintenance type to take into account for track geometry data.

3.3 Cases

Three turnouts in the data material are analyzed, presented in Table 3.2. Periods of observations are chosen so that each turnout is recorded nine times, during which periods two of the turnouts underwent track adjustments, while one did not. Furthermore, there are not reported any (other) major geometry-altering repairs or renewals of the case turnouts. However, it must be emphasized that there may be maintenance events missing from the records, according to Banedanmark officials.

	Turnout A	Turnout B	Turnout C
First Observation	2014-02-22	2013-03-09	2012-06-23
Last Observation	2016-11-05	2016-11-05	2016-11-05
No. of observations in Period	9	9	9
Days Since Renewal (1st obs.)	3614	6642	5652
Track Adjustments in Period	0	1	1

Table 3.2: Observation period and count, age, and track adjustments of the case turnouts.

Each turnout is measured as follows: The measurements of the turnouts start 10 meters in front of the switch blades and extend 10 meters behind the crossing nose, to ensure that no important information is missed. The specific switch treated in this analysis is approximately 43 meters long and measurements are taken every 25 cm, for every variable. One full observation of the whole segment consists of eight variables à 250 samples, Table 3.3. Figures 3.4 and 3.5 illustrate a track record, plotting the geometric measurements of turnout B, recorded 2016-11-05.

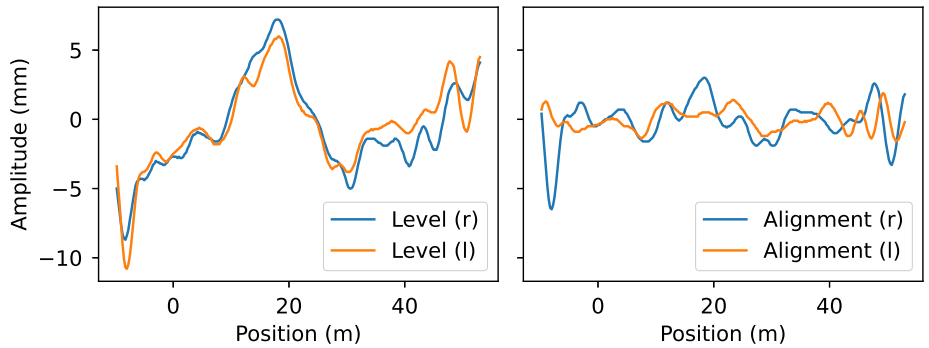


Figure 3.4: Track geometry measurements of turnout B, recorded 2016-11-05.

Sampling frequency	4 (every 0.25 m)
Total no. of samples	250
Turnout length	42.5 m
Measurement length	62.5 m

Table 3.3: Sampling of the turnout track geometry.

3.4 Techniques for Track Geometry Monitoring

The practice of most railway infrastructure managers in the world, including Norway, Sweden, and Denmark, is to assess track quality by setting thresholds for point failures (deviations in a single track geometry measurement) and unevenness (standard deviation used for longer track segments), Soleimanmeigouni et al. (2018), Jøndrup (2019) and EN 13848-1. Although a well-proven method for monitoring the condition of straight tracks, it is not so easily applied to turnouts, because S&Cs generally cause much larger variations in the recordings than seen on straight tracks and even a great variability in geometry between different turnout types. This entails limited usage of track geometry data for S&Cs, necessitating frequent and extensive manual inspections as a consequence.

The Danish national standards for track geometry intervention limits (concerning data gather from the TIC, not manual measurements) are ”Banenormerne BN1-38: Sporbeliggenhedskontrol og sporkvalitetsnormer”, Jøndrup (2019). The geometry is assessed in

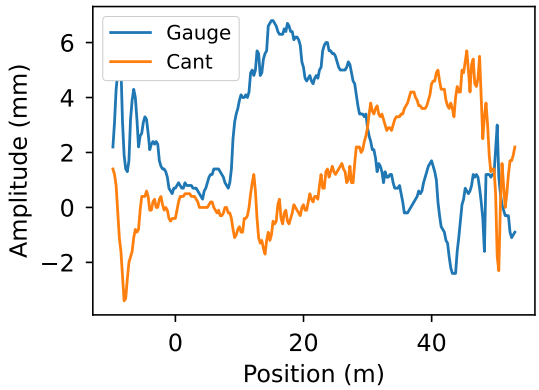


Figure 3.5: Track geometry measurements of turnout B, recorded 2016-11-05.

two ways:

- *Point faults.* If the deviations of the measurements exceed a defined threshold, the track will have to be adjusted within a short period of time, and the allowed speed limit will be lowered temporarily. A point fault found in a turnout will initiate a manual inspection to localize the potential fault.
- *Unevenness.* The track unevenness is assessed by the mean of the standard deviations of `level(r)` and `level(l)`, and `alignment(r)` and `alignment(l)` over segments of specified length, typically 200 m, Jensen (2016). These statistics are used to plan to tamp in long stretches, but not used to schedule tamping in turnouts.

To keep the discussion of geometric degradation informed when testing the continuous and the discrete wavelet transform, the standard deviations of the vertical and horizontal are added to the results for comparison. Point faults are not considered, as the work is focused on measures of the gradually worsening geometry, rather than the function or malfunction of specific parts. Besides, the point failures are defined by subjectively set thresholds, and even though the standard deviations are also used with this kind of heuristics, the development of the standard deviations themselves are interesting for the results and discussion to compare track unevenness.

Chapter 4

Wavelets

The first section of this chapter introduces the essential concepts of time-frequency analysis and the general idea of wavelet transformations. The second section treats the continuous wavelet transform (CWT), and demonstrates a handful of different wavelets applied to the geometry measurements. The third section walks through the foundation for the discrete wavelet transform (DWT) and briefly discusses the discrete wavelet analyzing properties.

4.1 Time-Frequency Analysis

The goal of time-frequency analysis is to obtain information on a signal in terms of both it's time and frequency content. A continuous signal $f(t)$ gives the perfect resolution in time, meaning that, for any given time t_0, the signal value $f(t_0)$ is determined exactly. In many cases, however, one is more interested in the frequencies found in the signal around that time t_0. Playing the piano, for instance, requires not only to know the intensity (like a crescendo), but also what key to play - the frequency. The problem of playing the piano by just reading the pressure wave amplitude $f(t)$ of a song is that it has no frequency *resolution* (not listening, however, as it turns out that up to a normalizing constant, the inner ear actually performs a wavelet transformation itself, see Daubechies (1992)).

The Fourier transform, on the other hand, maps the signal in it's entirety to the frequency domain. Effectively, this is a change of basis of the signal to sines and cosines, and the signal may now be assessed as a function of frequency, $\hat{f}(\xi)$. The transform and it's inverse is defined in Appendix A. This very general ability of mapping functions between time and frequency domains constitute the foundation for everything from radar detection technology, representing images and music, and fast computations and algorithms in computers. The Fast Fourier Transform (FFT) is considered the backbone of modern electronics, and Gilbert Strang named it "the most important numerical algorithm of our lifetime" Strang (1994). The Fourier transform of a signal has perfect frequency resolution and no time

resolution. Figure 4.1 displays (generically) the time and frequency resolution of a signal in a time series and a Fourier representation, and illustrates how signals can be chopped up in either arbitrary fine pieces in time, or in arbitrary fine pieces of frequency.

Since one can represent a signal perfect in either time or frequency, it is tempting to ask if it can be represented in both time and frequency simultaneously. To see why that is a bit of a problem, consider the following.

A natural way to think of frequencies is by their inverse, the wavelengths. Think of sea waves that pass your stationary boat, floating in the open ocean. You might have an idea about when the wave arrives and when it has passed, but what is it's location, exactly? Simply because the wave spans several meters (or taking a small portion of time to pass the boat), knowing the exact *position* is impossible. A shorter wave takes up less space and passes the boat faster, and may be more precisely localized than the large ones. However, pinpointing the wave exactly is yet not possible, as it still has a certain extent. Thus, something observed in the frequency domain (the waves in the water) may be localized in time or space only to a certain degree of precision (even if we are staring really, really hard at it). This result is what is formalized in the famous uncertainty principle of Werner Heisenberg (often associated with quantum mechanics, as the wave properties of very small particles lead to a trade-off between determining the position and momentum).

While the idea of the Fourier transform is to apply an integral convolution to the signal and a basis of harmonics with infinite support (or equivalently, infinite energy), the basic idea of the wavelet transform is to use a basis of small waves, the so-called *wavelets*. Wavelets have finite energy, and thus good localization properties in both time and frequency. By *scaling* and *shifting* the wavelet, one can control it's frequency content and position relative to the signal, in turn affecting what it will pull out from the signal convolution (the transform). Figure 4.2 shows the generic picture of varying the resolution of time and frequency, the trade-off imposed by the uncertainty principle. The plot to the right of Figure 4.2 shows the wavelet transform, performed by iterative re-scaling of the wavelet to match various parts of the signal frequency content, while shifting it along the signal for each "level" (scale).

The left plot shows the Short-time Fourier transform (STFT), noted for comparison and defined in Appendix A. The STFT divides the signal into pieces of constant length, and then Fourier transforms the content within each piece, obtaining some time localization to the cost of losing the wavelengths that exceed the window length. We can understand even more of the basic drawings that explain the idea of scaling and translation of wavelets with the Heisenberg uncertainty in mind. A single box in these plots is often referred to as an "atom" or "Heisenberg box" in the literature since their minimal area is bounded by the product of the uncertainty in time and frequency, see for example Mallat (2009). The important point is to note that the area of the boxes in Figure 4.2 is constant (although the artistic skills of the author may not do the Heisenberg uncertainty principle proper justice).

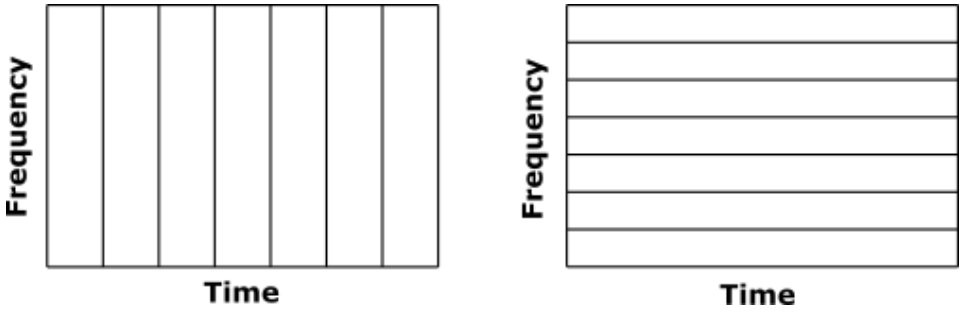


Figure 4.1: Generic picture of a time series analysis (left) and a Fourier analysis (right).

4.2 The Continuous Wavelet Transform

The continuous wavelet transform (CWT) is an integral transformation of type

$$(Tf)(\omega) = \int dx\, K(x,\omega)\, f(x), \tag{4.1}$$

transforming some function $f(x)$ to a new function $(Tf)(\omega)$ of different parameters. The kernel K (also called analyzing function), a wavelet, is stretched and dilated in order to pick up different frequency content as tightly localized as possible. The transform results in a coefficient for every particular choice of scale and translation, and repeating the transform for different scales and translations of the wavelet results in a map of coefficients, describing the time and frequency content of the signal.

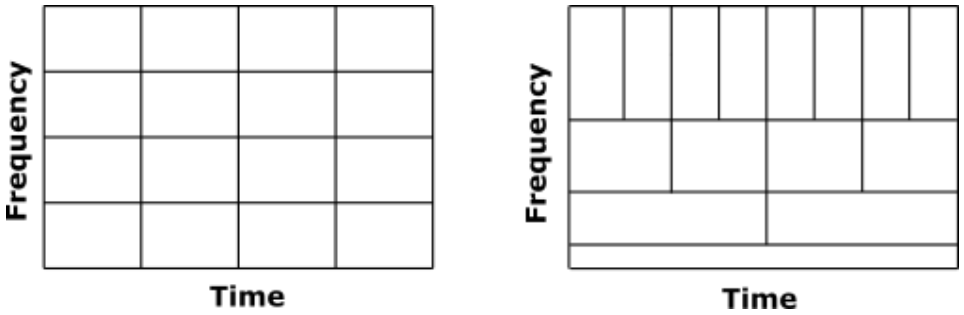
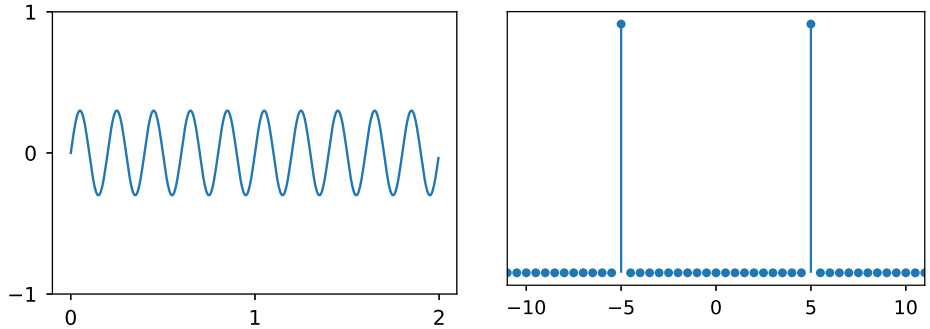


Figure 4.2: Time-frequency resolution of the Short-time Fourier transform (left), and the wavelet transform (right).

One could restrict the analyzing function to obtain a better time localization. Figure 4.4 shows a short impulse, very compact in time. As a result, bandwidth in the Fourier domain is extremely wide, showing the trade-off between the ability to localize of time and frequency content tightly.

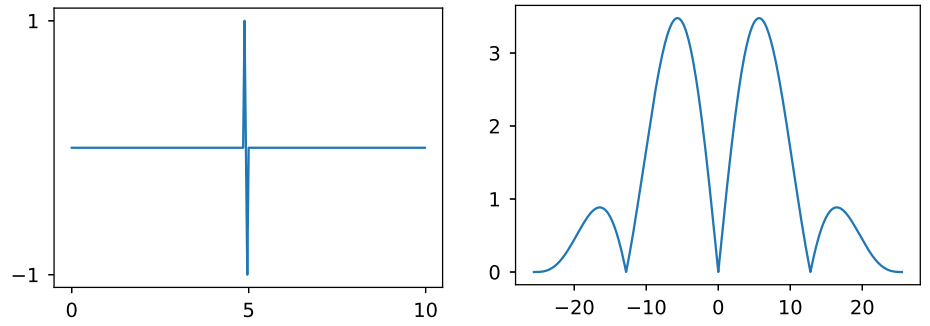


(a) $f(t)$ **(b)** $\hat{f}(\xi)$

Figure 4.3: A sine wave, frequency 5, in time domain (left) and Fourier domain (right).

4.2.1 Wavelets

The criterion for an analyzing function to be a wavelet is called the *admissibility condition*, defined as follows.



(a) $f(t)$ **(b)** $\hat{f}(\xi)$

Figure 4.4: A short and abrupt chirp signal and it's corresponding Fourier transform. The impulse is tightly localized in the time domain, but the frequency content of the signal is wide spread .

Definition 4.2.1. ***The admissibility condition.*** *A wavelet $\psi(t)$ is admissible if $\psi \in L^2(\mathbb{R})$, and it satisfies*

$$C_\psi = 2\pi \int d\xi \, |\xi|^{-1} \, |\hat{\psi}(\xi)|^2 < \infty, \tag{4.2}$$

where $\hat{\psi}(\xi)$ denotes the Fourier transform of $\psi(x)$.

The admissibility condition must be satisfied by the wavelet for the CWT to be well defined, as it will be clear that a wavelet that results in an unbounded C_ψ will just blow up

the transform. For practical purposes, the admissibility condition is equivalent to

$$\int dx\, \psi(x) = 0, \tag{4.3}$$

Daubechies (1992), since wavelets used in analysis are usually imposed a more "strict" decay than needed to satisfy the admissibility condition.

Furthermore, let different wavelets $\psi(x)$ be generated from a *mother wavelet*, indexed by a scale parameter a, setting the dilation of the wavelet, and by a translation parameter b, deciding the position of $\psi_{a,b}(x)$ relative to the signal $f(x)$.

Definition 4.2.2. ***The mother wavelet.*** *Wavelets can be generated from a mother wavelet,*

$$\psi_{a,b}(x) = |a|^{-1/2}\, \psi\left(\frac{x-b}{a}\right), \tag{4.4}$$

where $a, b \in \mathbb{R}$, $a \neq 0$ determines the scaling and translation.

4.2.2 Transform

The CWT is parameterized with the scale a and translation b of the mother wavelet and defined as follows.

Definition 4.2.3. ***The continuous wavelet transform.***

$$(Tf)(a,b) = \int dx\, f(x)\, |a|^{-1/2}\, \overline{\psi\left(\frac{t-b}{a}\right)} \tag{4.5}$$

$$= \langle f,\, \psi_{a,b} \rangle, \tag{4.6}$$

where $(Tf)(a,b)$ is a new function of the two parameters scale and translation, $\overline{\psi(\cdot)}$ denotes the complex conjugate, and $\langle\ ,\ \rangle$ the L^2-inner product.

The transform is linear,

$$\alpha(Tf)(a,b) + \beta(Tg)(a,b) = (T(\alpha f + \beta g))(a,b), \tag{4.7}$$

which means the sum of the transform of each variable is the same as the transform of the sum of all variables, which will allow us to compute the scalogram directly from the aggregated signal.

Furthermore, the original signal can be recovered analytically by the so-called *resolution of the identity*, yielding

$$f = C_\psi^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{da\, db}{a^2} \langle f,\, \psi_{a,b} \rangle\, \psi_{a,b}, \tag{4.8}$$

where $\langle f,\, \psi_{a,b} \rangle$ is the coefficients for the particular scaled and dilated wavelet $\psi_{a,b}$, and C_ψ the wavelet admissibility constant. This is the reason why the mother wavelet must be bounded, as the inverse transform would not exist otherwise.

Proposition 4.2.1. ***The inverse wavelet transform.*** *For all* $f, g \in L^2(\mathbb{R})$*,*

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{da\,db}{a^2}(Tf)(a,b)\overline{(Tg)(a,b)} = C_\psi \langle f, g\rangle. \tag{4.9}$$

The proof as given by Daubechies (1992) is written in appendix.

When choosing wavelets for a given problem, it is sometimes required to have a fine localization, sometimes required to have a fine frequency resolution, depending on the practical application. Formal measures of wavelets' spread in time and frequency are for example the wavelet variance in space (or time) and frequency:

$$\sigma_x^2 = \int (x - \langle x\rangle)^2 \, |\psi(x)|^2 \, dx, \tag{4.10}$$

and

$$\sigma_\xi^2 = \frac{1}{2\pi} \int (\xi - \langle \xi\rangle)^2 \, |\hat{\psi}(\xi)|^2 \, d\xi, \tag{4.11}$$

where $\langle x\rangle$ and $\langle \xi\rangle$ denotes the expectation in space and frequency, respectively. This leads to the formal result that limits the trade-off between resolution and localization:

Theorem 4.2.1. ***The Heisenberg Uncertainty principle.*** *The product of uncertainty (the variance) in time*

$$\sigma_x^2 = \int (x - \langle x\rangle)^2 \, |\psi(x)|^2 \, dx, \tag{4.12}$$

and in frequency

$$\sigma_\xi^2 = \frac{1}{2\pi} \int (\xi - \langle \xi\rangle)^2 \, |\hat{\psi}(\xi)|^2 \, d\xi, \tag{4.13}$$

is always constrained by the inequality

$$\sigma_x^2 \, \sigma_\xi^2 \geq \frac{1}{2}, \tag{4.14}$$

see for example Kutz (2013), relating directly to the resolution trade-off in Figures 4.1 and 4.2.

4.2.3 Scale and Characteristic Frequency

To implement the continuous wavelet transform, a range of scales must be set to match the frequency content one would like to analyze. The signals in question are filtered (a least in the case of horizontal `level` and vertical `alignment`), with resulting period $\lambda \in (3, 75)$ m. To match this with the frequencies of the various mother wavelets, we must utilize their "representative" frequencies. Several representative frequencies are used to describe wavelets. However, since a wavelet actually contains a multitude of frequencies, not just one as for the Fourier analyzing functions, there are a couple of ways to choose this representative frequency:

- *Peak frequency* f_p. The frequency with the highest energy (the peak) in the wavelet energy spectrum.
- *Band-pass frequency* f_c. The second-order moment of the wavelet energy spectrum.
- *Central frequency* f_0. The frequency peak of the wavelet in the Fourier domain.

The wavelet energy spectrum is defined as

Definition 4.2.4. ***The Wavelet Energy Spectrum.*** *The energy spectrum of a wavelet is given as the squared modulus of the wavelet in Fourier domain,*

$$E(\xi) = |\hat{\psi}(\xi)|^2, \tag{4.15}$$

where ξ denotes frequency.

From the mother wavelet, we see that the choice of scale has an inverse relationship with the frequencies of the wavelet (controlled by the parameter a in the argument).

$$\psi_{a,b}(x) = |a|^{-1/2}\,\psi\left(\frac{x-b}{a}\right), \tag{4.16}$$

Both the peak, band-pass, and central frequency will serve as a representative frequency f and satisfy the inverse relationship with the scale a:

$$f \propto \frac{1}{a}, \tag{4.17}$$

Addison (2017). The central frequency f_0 is chosen here, so the representative frequency for a given scale is

$$f = \frac{f_0}{a}, \tag{4.18}$$

often known as the pseudo-frequency.

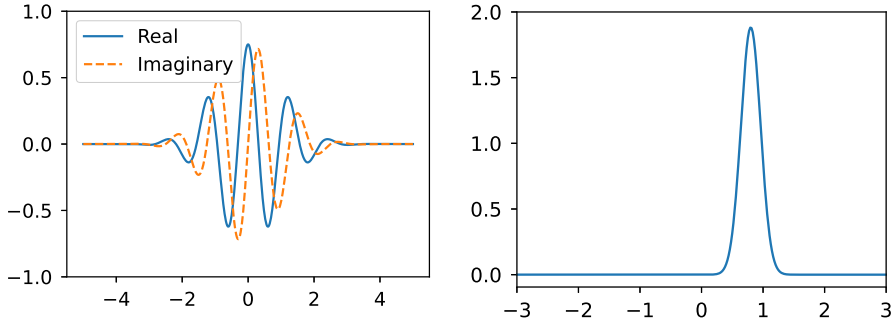
Setting the correct scales is illustrated with the complex Morlet wavelet (on of three wavelet candidates considered below). The choice of continuous wavelet is not changing the nature of the CWT results, but there are some that fit the specific demands of resolution better than others, as we will see. The complex Morlet wavelet, Figure 4.5, is a typical wavelet to use for analysis, and especially for signal from material responses, Addison (2017). It consists of a Gaussian envelope and a complex exponential, and being complex it captures information about both amplitude and phase. It is given as

$$\psi(x) = \pi^{-1/4}\,\exp\left(\frac{-x^2}{2}\right)\left(\exp(2\pi i f_0 x) - \exp\left(\frac{2\pi f_0)^2}{2}\right)\right), \tag{4.19}$$

where f_0 denotes the central frequency. The term $\exp((2\pi f_0)^2/2)$ is a correction for a nonzero mean of the complex exponential, but is usually dropped as the deviation from zero becomes negligible for frequencies $f_0 \gg 0$, Addison (2017). It then takes the nicer form of

$$\psi(x) = \pi^{-1/4}\,\exp\left(\frac{-x^2}{2}\right)\,\exp(2\pi i f_0 x), \tag{4.20}$$

where the first term is the normalization to ensure unit energy, the second is the Gaussian envelope, and the third the complex exponential with the real cosines and imaginary sines.



(a) $\psi(x)$ **(b)** $\hat{\psi}(\xi)$

Figure 4.5: The Complex morlet wavelet and it's Fourier transform. The solid and the dotted line in **(a)** represents the real and imaginary part, respectively. It consist of an complex exponential as signal carrier and a Gaussian envelope ensuring decay. Since the wavelet is complex it has only nonzero frequencies.

The Fourier transform of this wavelet is

$$\hat{\psi}(\xi) = \pi^{1/4}\sqrt{2}\exp\left(\frac{1}{2}\left(2\pi\xi - 2\pi f_0\right)\right). \tag{4.21}$$

The central frequency is set to $f_0 = 0.8$ (a value around $\sqrt{1/2\ln(2)} \approx 0.849$ is commonly used for the complex Morlet, see for example Addison (2017). With an central frequency of $f_0 = 0.8$ we get

$$f_{upper} = \frac{1}{3} = \frac{0.8}{a} \qquad a_{upper} = 2.4 \tag{4.22}$$

$$f_{lower} = \frac{1}{75} = \frac{0.8}{a} \qquad a_{lower} = 60. \tag{4.23}$$

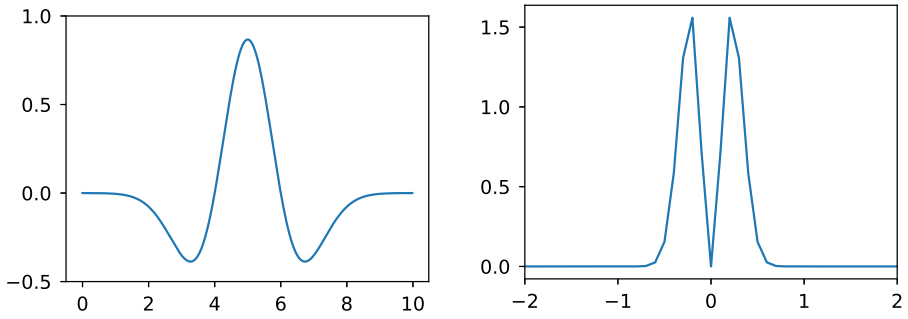
Table 4.1 displays the corresponding scales, periods and frequencies. Setting the scales for the other wavelets can be done in similar fashion.

The Mexican hat wavelet, Figure 4.6, and the real Morlet wavelet, Figure 4.7, are classical wavelets and therefore also natural candidates, Mallat (2009). The Mexican hat has it's name from the characteristic shape, and is constructed by taking the second derivative of a Gaussian:

$$\psi(x) = \frac{2}{\pi^{1/4}\sqrt{3\sigma}}\left(\frac{x^2}{\sigma^2} - 1\right)\exp\left(\frac{-x^2}{2\sigma^2}\right). \tag{4.24}$$

Scale a	Period (m)	Freq. (m^{-1})
2.4	3	0.33
3.9	4	0.25
6.4	8	0.13
12.7	16	0.06
25.5	32	0.03
50.9	64	0.02

Table 4.1: The corresponding period, frequency and scales for a complex Morlet with central frequency $f_0 = 5/2\pi \approx 0.8$.



(a) $\psi(x)$ **(b)** $\hat{\psi}(\xi)$

Figure 4.6: The Mexican hat wavelet and it's Fourier transform

The Fourier transform also share the Gaussian shape,

$$\hat{\psi}(\xi) = \frac{-\pi^{1/4}\sqrt{8}\,\sigma^{5/2}}{\sqrt{3}}\,\xi^2\,\exp\left(\frac{-\sigma^2\,\xi^2}{2}\right). \tag{4.25}$$

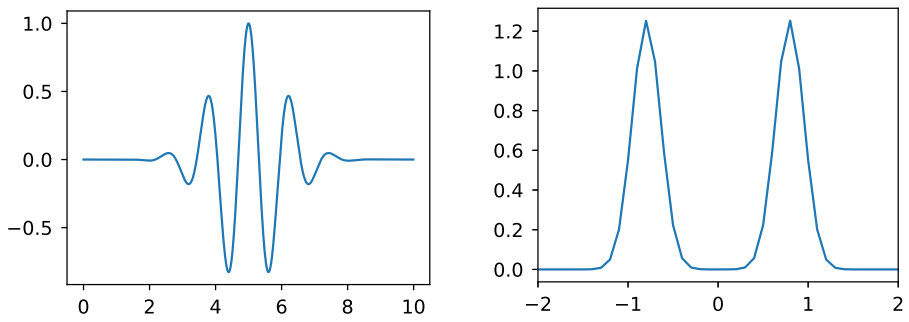
The real Morlet wavelet is often advocated as a suitable wavelet for engineering purposes, see for example Lin and Qu (2000). The real Morlet wavelet, often just called the Morlet, is only the real part of the complex version:

$$\psi(x) = \exp\left(\frac{-t^2}{2}\right)\cos(5t). \tag{4.26}$$

In the Fourier domain, it is given by

$$\hat{\psi}(s\xi) = \pi^{-1/4}\,\exp\left(\frac{-(s\xi-\xi_0)^2}{2}\right)U(s\xi), \tag{4.27}$$

where $U(\cdot)$ is the Heaviside function. Figure 4.7 displays the Morlet and it's corresponding Fourier transform, showing only the real part of the signal.



(a) $\psi(x)$ **(b)** $\hat{\psi}(\xi)$

Figure 4.7: The real Morlet wavelet and it's Fourier transform

4.2.4 Scalograms

The continuous wavelet transform is typically plotted in a heat map as function of scale a and translation b (so-called *scalograms*), giving either the transform coefficients directly, $T(a, b)$, or the two-dimensional wavelet energy distribution $E(a, b) = |T(a, b)|^2$. Practice differs in the literature of whether one focuses on the coefficients themselves or the energy spectrum. There are several advantages to the energy representation, as the sum of the coefficients can be interpreted as the total energy in the signal. For computational purposed one have to bear in mind that the CWT utilizes a redundant basis, so it does not preserve energy. However, transformations with the same wavelet and the same set of scales/translations will still compare relative to each other.

Finally, one may, of course, normalize the plots to have either unit energy (coefficients sum to one) or peak energy 1 (divide by peak energy), for example, since the visualization is dependent on setting the colour map levels to some minimum and maximum values. Torrence and Compo (1998) took a more statistical point of view and normalized the coefficients by their variance. By plotting $T(a, b)/\sigma^2$, $\sigma^2 = \mathrm{Var}[T(a, b)]$, they obtained a better understanding of whether the coefficients were statistically significant. However, normalizing the transform result in this way has the adverse effect that the scalograms become difficult to compare to each other, destroying the trends when observing long term behaviour. Unit variance would not make sense as the development of the geometry will also give development in variance. The same argument goes with unit energy, as the energy level in the signal is changing over time, perhaps even being the primary statistic of interest. The approach when comparing scalograms developing over time, therefore, is to set the extreme colour level to the maximum peak energy found in the period one is considering, as this allows the development to be easily followed, visually. To illustrate the properties of the different wavelets, a signal sample from the left rail `alignment`, turnout A, is chosen for the CWT, Figure 4.8.

Figure 4.9 shows the complex Morlet applied to the illustration signal, with one scalogram for the real coefficients representing the amplitude and one scalogram for the imaginary

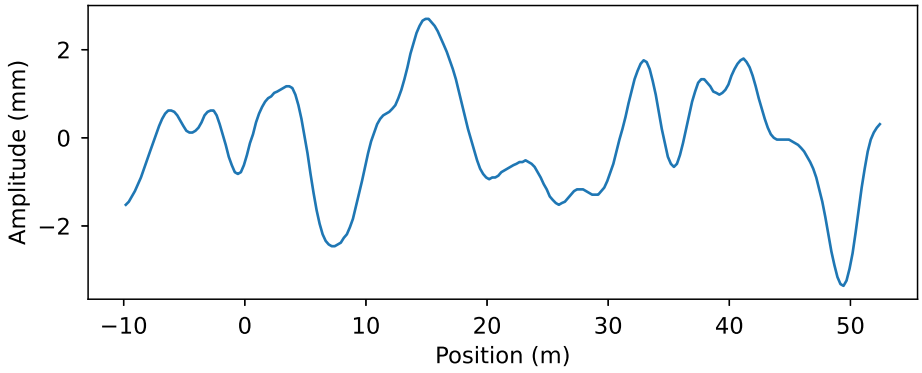


Figure 4.8: Track geometry measurement for wavelet comparison: `Alignment(l)`, turnout A, recorded 2016-05-08.

part representing the phase. The complex wavelet is practical in the settings where the signal phase changes are important, as this is captured by the complex exponential lying in the core of the complex Morlet. This work first and foremost explores the power of the various frequency components at different locations along the turnout; Thus it is interesting to assess the amplitude and the squared amplitude (energy) of the signal. For the complex Morlet, this is shown in Figure 4.10. The energy content is smeared due to the combination of the two components. Figure 4.11 shows the CWT using the Mexican hat wavelet, applied to the example signal. The Mexican hat has a characteristic frequency of $f_c = \sqrt{5/2}/2\pi \approx 0.251$, Addison (2017). The left plot shows the absolute value (the magnitude) of the transform. The observed features/nodes seem to be well separated in the spatial direction but at the cost of somewhat blurry frequency content. The right plot of Figure 4.11 displays the squared absolute value of the coefficients, $|Tf|^2$, the "energy spectrum". Plotting the energy spectrum, it is still easy to determine the high energy concentrations in the picture and, at the same time, there is possible to see some other details and get an idea of their relative strength and significance. Finally, Figure 4.12 shows the resulting CWT using the real Morlet. This wavelet information about the magnitude only, similar to Figure 4.11. Again, the reasonable visualization seems to be the energy spectrum, which both enable the identification of the major modes and locality, and still provide a reasonable amount of detail.

4.3 The Discrete Wavelet Transform and Multiresolution Analysis

Resembling the information stored in images, voice signals, or other types of data by wavelet transform coefficients, we may not want just to visualize the information in nice scalograms. More importantly, we often want the signal to be represented uniquely (so that different information remains different), but also as compactly as possible (using few

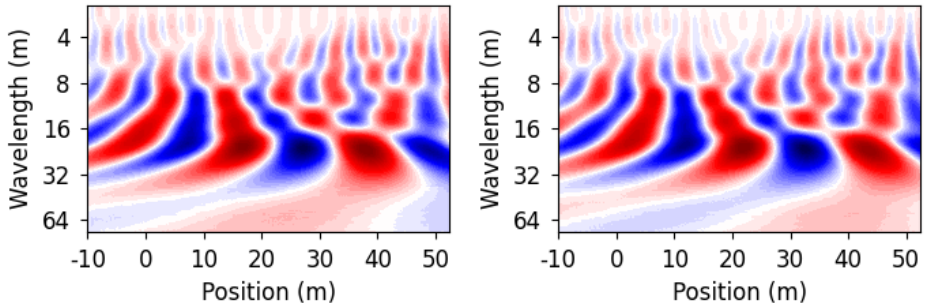


Figure 4.9: The real coefficients (left) and the imaginary coefficients (right) of the CWT with the complex Morlet wavelet (bandwidth 4, centre frequency 0.82), applied to the example measurement.

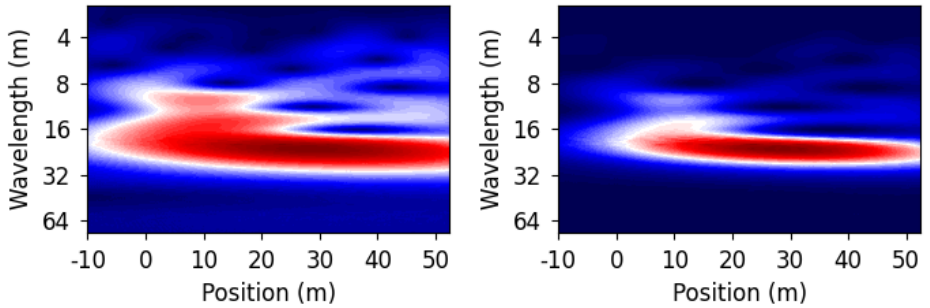


Figure 4.10: The coefficient amplitude (left) and the amplitude squared (right) of the CWT with the complex Morlet wavelet (bandwidth 4, centre frequency 0.82), applied to the example measurement.

coefficients).

Mathematically speaking, this depends on the choice of basis functions used for decomposing the signal. If they share properties (i.e. their inner product is nonzero), the contributions from the signal may be picked up by several basis functions, leading to redundancy in the coefficients, meaning more than one coefficient carries the same piece of information. On the other hand, if the signal spans a space that the basis functions do not fully reach, we will fail to represent all information in the signal, leading, ultimately, to a disappointed Netflix-user when half of the image was lost in the decomposition. And so, the next step of the chapter to build a framework where bases of mutually orthogonal wavelets can exist and have a good time. This framework is called Multiresolution Analysis (MRA), one of the 1990s greatest mathematical inventions.

One can extend the idea of the wavelet transform to Multiresolution analysis and complete/orthogonal wavelet bases as follows. By only allowing the scaling and translation to be done in discrete steps m and n, it is possible to find a wavelet family $\psi_{m,n}$ that is complete, or in other words constitutes an orthogonal basis, as this allows the exact expansion

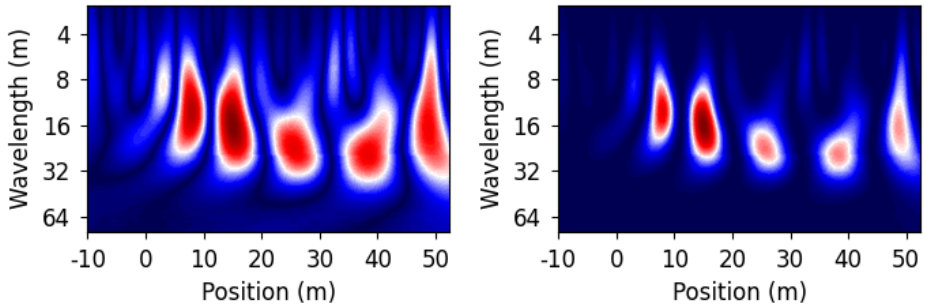


Figure 4.11: The coefficient amplitude (left) and the amplitude squared (right) of the CWT with the Mexican hat wavelet (centre frequency 0.25), applied to the example measurement.

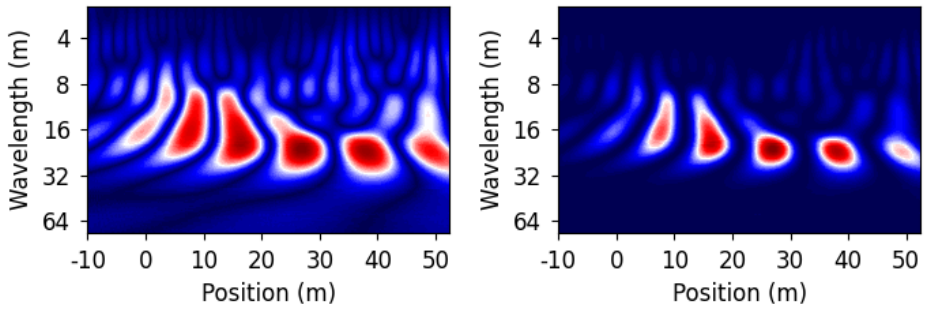


Figure 4.12: The coefficient amplitude (left) and the amplitude squared (right) of the CWT with the real Morlet wavelet (centre frequency 0.80), applied to the example measurement.

of the signal $f(x)$:

$$f(x) = \sum_{m,n} \langle f,\, \psi_{m,n} \rangle \, \psi_{m,n}. \tag{4.28}$$

This allows the expansion of a signal into orthogonal wavelet bases, just as one would expand a signal into its Fourier modes/basis functions.

4.3.1 Discrete Wavelets

Consider again the generic construction of the discretized mother wavelet, using integers $j, k \in \mathbb{Z}$ to scale and translate the wavelet. By discretizing the parameters,

$$a = a_0^m, \qquad b = n b_0 a_0^m, \tag{4.29}$$

one obtains discretized version of the mother wavelet:

Definition 4.3.1. ***The discrete mother wavelet.*** *Discrete wavelets, indexed by m and n*

are generated by

$$\psi_{m,n}(x) = a_0^{-m/2}\,\psi\left(\frac{x - nb_0a_0^m}{a_0^m}\right), \tag{4.30}$$

where $a_0, b_0 \in \mathbb{R}$, $a_0 \neq 0$ *as before, and* $m, n \in \mathbb{Z}$ *scale and translate the wavelet.*

Analyses in MRA are done in $\log_2$-scale and the translation in steps of one. Set $a_0 = 2$ and $b_0 = 1$ to obtain

$$\psi_{m,n}(x) = 2^{-m/2}\,\psi(2^{-m}x - n). \tag{4.31}$$

Intuitively, this results in halving the scale of the wavelet for every increment of m (dyadic scaling), whereas n slides the wavelet around. Next, define the orthogonality of wavelets, as this will be needed to have an informed discussion of basis functions.

Definition 4.3.2. ***Orthogonality.*** *Two wavelets* $\psi_{m,n}$ *and* $\psi_{j,k}$ *are said to be orthogonal if their inner product satisfies*

$$\langle \psi_{m,n},\, \psi_{j,k} \rangle = \int \psi_{m,n}(x)\,\psi_{j,k}(x)\,dx \tag{4.32}$$

$$= \delta_{mj}\,\delta_{nk}, \tag{4.33}$$

where δ_{ij} *is the Kronecker-delta*

$$\delta_{ij} = \begin{cases} 0, & \textit{for} \quad i \neq j, \\ 1, & \textit{for} \quad i = j. \end{cases} \tag{4.34}$$

Thus, the inner product of orthogonal wavelets are nonzero if and only if they have equal level of dilatation ($m = j$) and position ($n = k$), i.e. $\psi_{m,n} = \psi_{j,k}$, in which case the inner product is exactly 1.

4.3.2 Multiresolution Analysis

The following presentation is based on Daubechies (1992), and Kutz (2013). The notation of authors on the topic differs in the convention of using 2^m or 2^{-m} as the scaling factor. The notation here is adapted to using 2^{-m}, making it align with the (somewhat untypical) view fronted by Ingrid Daubechies.

Consider a sequence of subspaces V_j, embedded/nested within each other

$$\cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \cdots, \tag{4.35}$$

with five specific properties determining how the subspaces are related, scales, and spans $L^2(\mathbb{R})$:

(i) $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$,

(ii) $\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R})$,

(iii) $f \in V_j \iff f(2^j\cdot) \in V_0$,

(iv) $f \in V_0 \Longrightarrow f(\cdot - n) \in V_0,$ for all $n \in \mathbb{Z}$, and finally,

(v) A *scaling function* $\phi \in V_0$, such that the sequence of functions $\{\phi_{0,n} = \phi(x - n)\}$ consitute an orthogonal basis for V_0, and, for all $j, n \in \mathbb{Z}$, $\phi_{j,n} = 2^{-j/2}\,\phi(2^{-j/2}x - n)$.

Property (i) states that the intersection of all subspaces V_j is the empty set, while (ii) states that the span of the union of all subspaces is equal to the space of square-integrable functions, $L^2(\mathbb{R})$. This ensures that the limit of orthogonal projections P_j onto the subspace at level j converge to the actual function

$$\lim_{j \to -\infty} P_j f = f \qquad \forall f \in L^2(\mathbb{R}), \tag{4.36}$$

as one move up the ladder of "resolution".

The concept of multiresoultion stems from (iii), which determines how the subspaces scale. The subspaces V_j can be seen as just scaled versions of the reference space, V_0. If a function f is contained in V_j, then it is also contained in V_0 if one multiply it's argument by 2^j (effectively scaling the signal by j octaves, for the *do-re-mi* enthusiast). Property (iv) ensures invariance under integer translation, meaning that if a function f is in V_j, so does the n-translated version of it. Combining (iii) and (iv) yields

$$f \in V_0 \Longrightarrow f(\cdot - n) \longrightarrow f(\cdot - 2^j n) \in V_j \quad \forall n \in \mathbb{Z}. \tag{4.37}$$

Because the sequence of subspaces V_j are nested,

$$\cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \cdots, \tag{4.38}$$

one can decompose subspace V_j into the subspace V_{j+1} and compliment of V_{j+1} in V_j:

$$V_j = V_{j+1} \oplus W_{j+1}. \tag{4.39}$$

The subset W_{j+1} is the *orthogonal complement* of V_{j+1} inside the subspace of V_j, in other words what is "missing" of V_j on the lower level of resolution V_{j+1}:

$$V_{j+1} \perp W_{j+1}. \tag{4.40}$$

Orthogonal compliments on different levels of resolution are therefore themselves orthogonal,

$$W_j \perp W_{j'} \quad \text{for} \quad j \neq j'. \tag{4.41}$$

Recursively decomposing the subspaces V_j, one obtains

$$\begin{aligned} V_j &= V_{j+1} \oplus W_{j+1} \\ &= V_{j+2} \oplus W_{j+2} \oplus W_{j+1} \\ &= V_{j+3} \oplus W_{j+3} \oplus W_{j+2} \oplus W_{j+1} \\ &\vdots \\ V_j &= V_J \oplus \bigoplus_{k=1}^{J-j} W_{J-k}, \end{aligned} \tag{4.42}$$

which showes the increasing level of signal detail captured by the orthogonal compliments W_j. Finally, expanding all the way to $j = \pm\infty$, the sum of the contributions span the space of $L^2(\mathbb{R})$ (properties of (i) and (ii)):

$$L^2(\mathbb{R}) = \bigoplus_{j\in\mathbb{Z}} W_j. \tag{4.43}$$

Note that the subspaces W_j also have the scaling property as V_j (iv), $f \in W_j \implies f(\cdot 2^j) \in W_0,$ for all $n \in \mathbb{Z}$.

What makes the multiresolution approach work is the fact that, whenever we have a construction satisfying properties (i-vi), there exists an orthonormal wavelet basis of $L(R^2)$, $\{\psi_{j,k},\, j,k \in \mathbb{Z}\}$, generated by $\psi_{j,k} = 2^{-j/2}\,\psi(2^{-j/2-k})$ such that for all $f \in L(R^2)$,

$$P_j f = P_{j+1} f \,+ \sum_{k\in\mathbb{Z}} \langle f,\, \psi_{j+1,k}\rangle\, \psi_{j+1,k}(x) \tag{4.44}$$

where P_j denotes the orthogonal projection operator onto V_j. By including all levels of resolution j we get

$$f = \sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} \langle f,\, \psi_{j,k}\rangle\, \psi_{j,k}(x). \tag{4.45}$$

The result is summarized in this theorem, as given in Daubechies (1992):

Theorem 4.3.1. ***Existence of an orthogonal, discretized wavelet basis.*** *Suppose the collection $\{V_j\}$ is a multiresolution of $L^2(\mathbb{R})$ satisfying properties (i)-(v), there exists a corresponding wavelet basis $\{\psi_{j,k};\, j,k \in \mathbb{Z}\}$ such that*

$$P_j = P_{j+1} \,+ \sum_{k\in\mathbb{Z}} \langle \cdot,\, \psi_{j+1,k}\rangle\, \psi_{j+1,k}(x). \tag{4.46}$$

The wavelets ψ may be constructed by the relation

$$\psi(x) = \sqrt{2} \sum (-1)^{n-1}\, \bar{c}_{-n-1}\, \phi(2x-n), \tag{4.47}$$

where

$$c_n = \langle \phi,\, \phi_{1,n}\rangle = \sqrt{2} \int dx\, \phi(x)\, \overline{\phi(2x-1)}. \tag{4.48}$$

There is certainly more to be said to make this a more complete introduction, however, this will suffice to have an intuition and some formal knowledge about how a signal is represented in a wavelet basis. Figure 4.13 illustrates how the MRA is implemented, repeatedly splitting the signal into approximation (V_j) and detail (W_j) levels. The detail and approximation at each level may be interpreted as a high- and low-pass filters, respectively, which leave coarser and coarser signal approximation left as the finer details are pulled out.

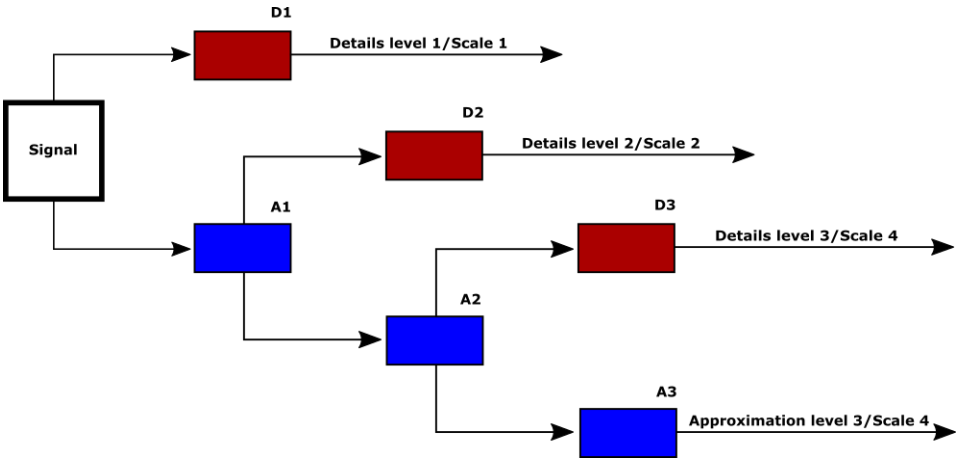


Figure 4.13: The discrete wavelet transform implemented as a multiresolution analysis, recursively decomposing the signal into finer and finer detail coefficients.

4.3.3 Vanishing Moments, Regularity, and Support

Discrete wavelets function as discrete filters, with some length and some coefficient value for every entry. Discrete wavelets are often described by their number of vanishing moments, where a moment is defined as

$$m_k = \int_{-\infty}^{\infty} f(x)\, x^k \, dx. \tag{4.49}$$

The vanishing moment of a wavelet is the polynomial order k a signal f may take, and still integrate to zero with the wavelet (zero correlation), thus the moment m_k "vanishes". A wavelet with a high number of vanishing moments will be able to correlate with signals of high regularity, or smoothness, and is, therefore, suitable to use for analysis with complex signals (complex as in smooth and heavily fluctuating, not imaginary). The downside to a high number of vanishing moments is that the wavelet has to be very wide, in order to incorporate the sought after complexity. This means they have wider support and can be scaled and translated to fewer levels and positions than more basic wavelets of a lower order, although these cannot replicate a high order signal. As a consequence, a large part of (discrete) wavelet design is about achieving a high number of vanishing moments for the most compact support possible, for which the Daubechies wavelets have proven to

Low-pass coefficients	High-pass coefficients
h0 = 0.0026818146	g0 = -0.0102681767
h1 = -0.0010473849	g1 = 0.0040102449
h2 = -0.0126363034	g2 = 0.1078082377
h3 = 0.0305155132	g3 = -0.1400472404
h4 = 0.0678926935	g4 = -0.2886296318
h5 = -0.0495528349	g5 = 0.7677643170
h6 = 0.0174412551	g6 = -0.5361019171
h7 = 0.5361019171	g7 = 0.0174412551
h8 = 0.7677643170	g8 = 0.0495528349
h9 = 0.2886296318	g9 = 0.0678926935
h10 = -0.1400472404	g10 = -0.0305155132
h11 = -0.1078082377	g11 = -0.0126363034
h12 = 0.0040102449	g12 = 0.0010473849
h13 = 0.0102681767	g13 = 0.0026818146

Table 4.2: High-pass/low-pass decomposition filter coefficients for the Daubechies *Symlet-7*.

be optimal, Daubechies (1992). For the analysis, the Daubechies *symlet-7* (7 vanishing moments) is chosen, a wavelet which is based on the original Daubechies *db* wavelets, but optimized with respect to symmetry, Mallat (2009). Table 4.2 shows the decomposition filter coefficients for the symlet-7, Lee et al. (2019) (the reconstruction coefficients are not used in the analysis and therefore not included).

Chapter 5

Method

Three turnouts from the data material have been selected for analysis, summarized in Table 3.2. They have been selected based on the following criteria:

- They have all been observed by the track inspection vehicle at least nine times, recording track geometry data. This makes them suitable cases for studying long-term track development.
- Two out of the three turnouts have undergone track adjustment (tamping) in the period.
- One out of the three turnouts has not undergone track adjustment in the period.
- Except for the major tamping campaigns, the turnouts are only subject to minor repairs (according to the repairs history) not believed to have a major effect on track geometry (welding, tightening bolts, etc.).
- The turnout is not completely renewed in the period.

As a preliminary treatment of the track geometry data, the `level_D1` and `level_D2` signals (wavelength periods 3-25 m and 25-75 m) are added together, to analyze the signals as complete as possible, concerning frequency content. Due to a large proportion of missing observations in `alignment_D2` for both left and right rail, only `alignment_D1` is kept (wavelength period 3-25 m), and simply referred to as `alignment`. Furthermore, the data is not rescaled or centred by e.g unit variance or unit energy, as this would defy the goal of monitoring wear development and track adjustment effect over time.

The CWT and the DWT transformations are applied using *Pywavelets*, Lee et al. (2019), a package for wavelet analysis in Python. The real Morlet wavelet is chosen for the continuous wavelet transform. The properties of the various wavelet candidates are displayed in Chapter 4, for which the real Morlet has favourable properties (although the qualitative result does not depend on the wavelet choice). For the discrete wavelet transform, the *symlet*

7 is applied. The choice of the discrete wavelet, similar to the choice of the continuous wavelet, is not affecting the qualitative results and comparisons, although the coefficient values would change if applied with different wavelets, Addison (2017).

5.1 General Turnout Signature

First, a single track-recording campaign is considered for each of the turnouts, and the CWT is applied to each measurement variable, to compare their relative strength. For a given turnout, one measurement campaign is arbitrarily chosen from the nine campaigns possible. Then each variable is transformed with the CWT, and the squared magnitude of the coefficients are displayed in a scalogram, similar to the example in Chapter 4, Figure 4.12, right plot. The six scalograms, one for each variable, are normalized by setting the extreme colour levels to the highest energy peak of the six transforms so that the relative signal strengths can be compared. The procedure is repeated for the two remaining turnouts so that the three ”signatures” can be compared to each other.

5.2 Geometric Degradation

The second step is the analysis of geometry development over time. For this, the turnout without any associated major maintenance is investigated, exploring the ”natural” degradation rate. Since each turnout observation consists of multiple geometry variables, the development is considered by following one variable at a time. The CWT is applied to every observation of a variable from the nine track-recording campaigns, resulting in a sequence of nine scalograms for each variable. The sequences of scalograms are plotted as the squared magnitude of the CWT coefficients, and normalized by setting the extreme colour levels to the highest energy peak of the sequence. This allows for easy comparison between the turnout observed at different times, similar to the normalization approach described above.

The standard deviations for the vertical and horizontal deviations are computed for comparison, since rules and regulations typically determine track quality in terms of the permitted standard variation in the vertical and horizontal directions, see Chapter 3. These are computed by taking the standard deviation of the `alignment` and `level` measurements over the turnout segment in full length, averaging the left and right rail. The summary statistics are given in tables with one row per track-recording campaign, for comparison to the CWT-sequences.

Also, the discrete wavelet transform (DWT) is applied to a single variable of the turnout, which has the signs of a monotonous wear path, judging from the CWT-sequence and the standard deviations. The summed, squared coefficients of each detail level and each measurement campaign are computed, to assess which detail levels that correspond to the visually determined wear found with the CWT. The intention is to explore how the DWT picks up on changes in track geometry, as this is less easily visualized than with the CWT.

5.3 Track Adjustments

The third and final part is the analysis of track adjustment effects. Here, the two turnouts subject to tamping are investigated with the CWT, and the plots are sequences of scalograms made similarly as for the geometry degradation. The only difference is that the events of track adjustment campaigns are indicated in the plots, so it is easier to look for changes in the geometry. Also in this part, the standard deviations for each campaign are listed in a table for each turnout, for comparison to the CWT-sequences.

Chapter 6

Results

The results are organized into three sections, with an approach as described in Chapter 5. The first illustrates the general turnout signature, comparing the behaviour of the geometric variables to each other, recorded in the same campaign. The second part investigates how the track geometry deteriorates over time, monitoring the track geometry recorded over several campaigns. The last part concerns the effect of track adjustments, assessing the changes in the turnout geometry that occurs after a track adjustment, also monitored over several campaigns.

The analysis of wear and tamping is done with the CWT. The results are compared to the mean, standard deviation of vertical and horizontal directions, the current practice of Banedanmark to assess track quality (but mainly used to assess long straight track segments). The various detail levels of the DWT are computed on a single example to investigate what levels respond to track development. The DWT detail level coefficients are compared to the development seen from visual inspections (CWT) and the standard deviations.

The CWT scalograms are visualized by plotting the squared magnitude of the coefficients, interpreted as signal energy. The y-axes (the scale, indicating frequency) are inverted and plotted in $\log_2$-scale (octaves), with units given as the wave period (wavelength in meters) instead of frequency. The x-axes range from -10 to 50 meters, where 0 meters marks the entry of the turnout (tip of the switch blades) and 40 meters the turnout exit, just behind the crossing nose. The reader is encouraged to revisit the turnout sketch in Figure 2.1 and the illustrations of geometry measurements in Figures 3.2 and 3.3 if some terms are unfamiliar.

6.1 General Turnout Signature

Figure 6.1 shows turnout A, recorded 2016-05-08. Starting with the horizontal deviations, we see that `level(l)` and `level(r)` are dominating. The right and left side are equal both in terms of frequency and node position along the x-axis (so the general signature in terms of frequency content and localization is equal), but there is clearly more energy in the left side movement. The oscillations are of relatively low frequency, however, with a period around 40 m. There are weak oscillations in `alignment(r)` and `alignment(l)`, and there is a weak, but noteworthy high frequency occurring in `alignment(r)` and `level(l)` just over the crossing nose (30-40 m). The `gauge` is quite pronounced, and reflects the weak oscillations with period 20 m in `alignment(l)`. It shows a concentration of energy around the turnout entrance which is not seen in `alignment` right or left, however. A possible explanation for this is the missing `alignment_D2` measurements with period 25-70 m. Finally, the `cant` mostly consists of very low frequencies, with two peaks at 0 and 30 meters in addition to some low energy high frequencies over the crossing nose.

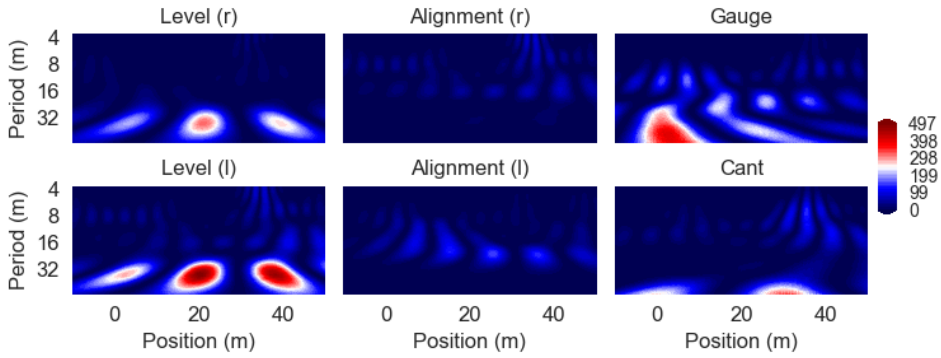


Figure 6.1: Track geometry comparison (CWT), turnout A, recorded 2016-05-08. The scalograms plot the squared magnitude of the CWT coefficients, for each measured geometric variable.

Next, consider the composition of turnout B, recorded 2016-05-08, Figure 6.2. `Level` is again prominent, this time for the right rail. The energy is concentrated around low frequencies, and there are similar node positions/peaks as for turnout A. Again, we observe a subtle, high-frequency oscillation in `alignment(r)` and `level(l)` over the crossing nose, similar to turnout A. Furthermore, there is a weak pattern seen in `alignment(l)`, appearing somewhat stronger in `gauge`. As for turnout A, the `gauge` deviations are centred around the turnout entrance for B, too, and the longwave signature in this variable indicates that some longwave information is indeed missing from `alignment`.

Figure 6.3 plots turnout C, recorded 2015-05-06. Here, the right and left rail `levels` are similar, in both energy and time/frequency peak positions. Both oscillates at a frequency quite a lot higher than seen in turnout A, Figure 6.1, and turnout B, Figure 6.2, have a period around 10 m. Turnout C also have the delicate, high-frequency oscillation in `alignment(r)` and `level(l)` over the crossing nose, seen in both A and B. It seems

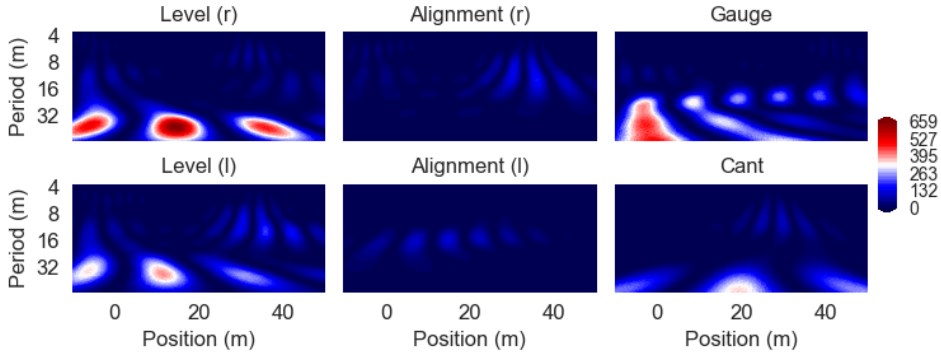


Figure 6.2: Track geometry comparison (CWT), turnout B, recorded 2016-05-08. The scalograms plot the squared magnitude of the CWT coefficients, for each measured geometric variable.

that the deviations in `alignment(r)` and `alignment(l)` cancels in `gauge`, as the same pattern cannot be observed there. Finally, the `cant`-signature is again deep down the frequency spectrum, with peaks about 30 m apart (period 60 m).

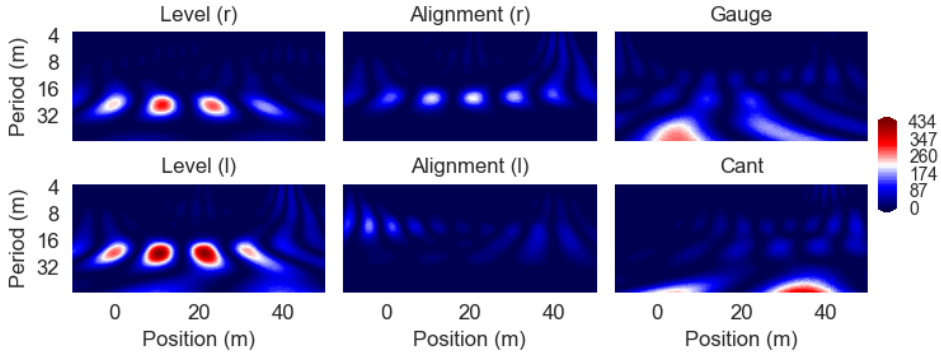


Figure 6.3: Track geometry comparison (CWT), turnout C, recorded 2015-05-06. The scalograms plot the squared magnitude of the CWT coefficients, for each measured geometric variable.

The composition of the three case turnouts in terms of the CWT proves to be relatively equal. Track movement is generally dominated by `level(r)` and `level(l)`. The oscillations from `alignment` are generally modest in terms of energy, but typically with twice the frequency one would normally observe in the vertical direction. For some reason, all case turnouts have the property that the `level(l)` and `alignment(r)` share a high-frequency signature over the crossing nose. The `gauge` typically has a more pronounced oscillation than seen in `alignment`, left or right, which may indicate that the horizontal deviations of the left and right rail tend to amplify each other, instead of cancelling (or being in phase, rather).

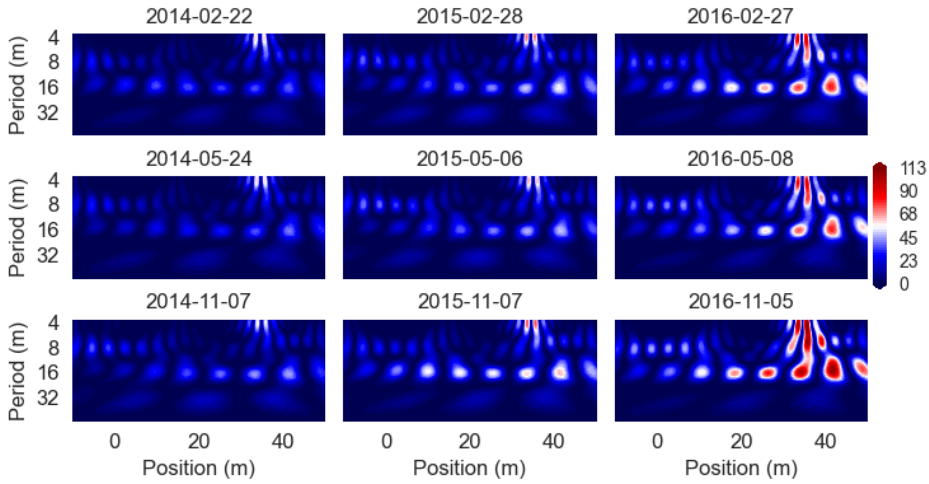


Figure 6.4: `Alignment(r)` of turnout A, recorded from 2014-02-22 to 2016-11-05. The scalograms shows the squared magnitude of the CWT coefficients using the real Morlet wavelet.

6.2 Geometric Degradation

Next, we consider the wear and geometric degradation of the turnouts as time progresses. The wear is evaluated with the CWT and compared to tables of the currently used quality indices for `level` and `alignment`. Case turnout A is followed from early spring 2014 to fall 2016, with each geometric measurement plotted separately, such that potential developments in the turnout geometry can be monitored as detailed as possible.

The scalograms of `alignment(l)`, Figure B.3, show a clear trend toward higher wear in the lateral alignment of the rails resulting in more energetic sideways movement as time passes. The oscillation grows most rapidly between 2014-11 to 2015-05, before it remains at approximately the same level. Figure 6.4 shows the `alignment(r)`, and again there is a clear pattern developing toward worsening lateral geometry, but on this side, it is more concentrated between 30 and 40 meters into the turnout. Two separate frequencies occur over the crossing nose, one with period 10 m and the other with higher frequency, period < 4 m.

Figure 6.5 shows the `gauge` develop from summer 2014 to winter 2016. It appears to have some kind of seasonal changes to it, as the long-wavelength oscillations seem to increase around November and February, and then decrease again in the summer mounts. This might also be an effect of the driving direction transformations done outside this project, Hovad et al. (2019). The vertical deviations have a more dominant and steady signature, see Figure 6.6. However, the energy seems to build up a little during 2014, before being constant throughout 2015, and then somehow dying off in 2016. Interestingly, there have been no major track adjustments in this period. Meanwhile, `level(r)` is fairly constant, and the combination results in a build-up of `cant` (but at a very low frequency), see

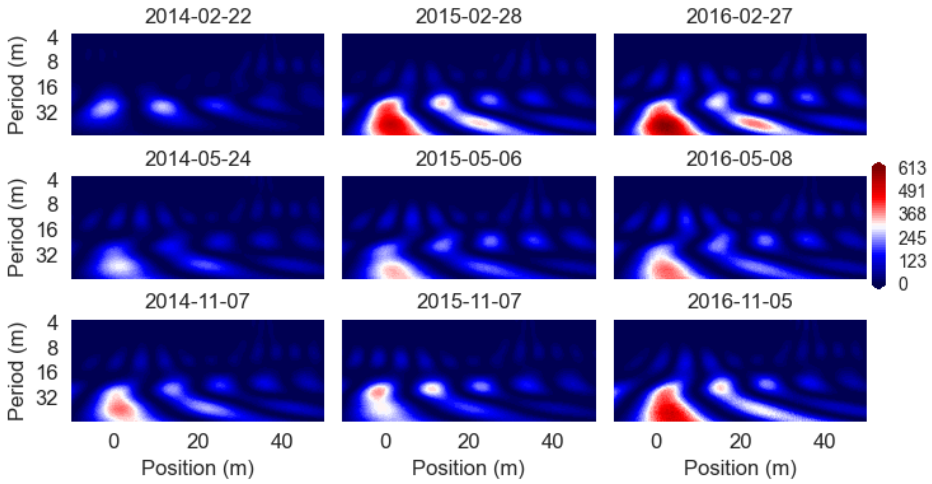


Figure 6.5: `Gauge` of turnout A, recorded from 2014-02-22 to 2016-11-05. The scalograms shows the squared magnitude of the CWT coefficients using the real Morlet wavelet.

Figures B.1 and B.2.

Table 6.1 shows the standard deviation for the vertical and horizontal direction, used by the railway manager to evaluate the condition of the track. Comparing Figures B.3 and 6.4 to $\sigma_{\texttt{alignment}}$, it is not obvious that there is an oscillation building up in the horizontal alignment, although a slight increase is seen in the standard deviation. Furthermore, the vertical deviation $\sigma_{\texttt{level}}$ is relatively stable, whereas the result of the CWT seems to be that the oscillations are losing energy in 2016-05-08 and 2016-11-05, Figure 6.6.

	$\sigma_{\texttt{level}}$	$\sigma_{\texttt{alignment}}$
2014-02-22	1.67	1.10
2014-05-24	1.80	1.16
2014-11-07	1.84	1.12
2015-02-28	1.84	1.13
2015-05-06	1.82	1.17
2015-11-07	1.81	1.16
2016-02-27	1.81	1.26
2016-05-08	1.82	1.28
2016-11-05	1.83	1.35

Table 6.1: Mean standard deviations of `level` (left and right) and `alignment` (left and right), turnout A.

To explore the DWT based features, the right-side horizontal alignment `alignment(r)` of turnout A is chosen as a reference, Figure 6.4. The `alignment(r)` has a clear de-

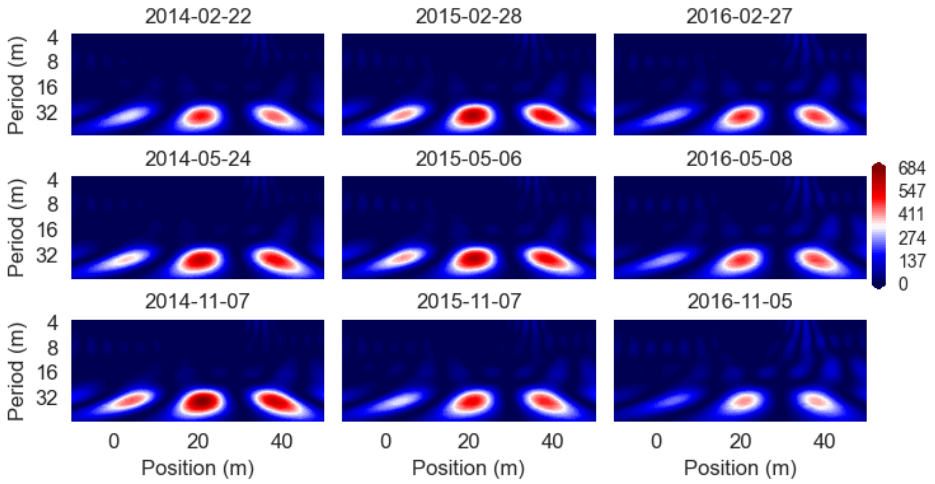


Figure 6.6: `Level(l)` of turnout A, recorded from 2014-02-22 to 2016-11-05. The scalograms shows the squared magnitude of the CWT coefficients using the real Morlet wavelet.

velopment towards increased geometry deviations as time progresses, (although relatively weak compared to the vertical directions) and serves as a good example signal to evaluate the DWT. Table 6.2 displays the squared, summed DWT detail coefficient for each level 1 to 4. (Four detail levels is the maximum number of levels possible for the *symlet-7*, applied to a signal of length 250. This is due to the repeated signal downsampling with factor 2 and the required wavelet support.) The finest levels have small (squared) coefficients, and level 1 is constant at 0.02 except for a jump to 0.28 (2015-02-28), unexplained by the standard deviations, Table 6.1, or the CWT, Figure 6.4. The level 2 coefficients fluctuate several times in the period and show no clear sign of monotonously increasing frequency content. The coarser levels (3 and 4) store larger parts of the signal, and this is also where the greatest development is seen. The summed squared details of level 3 increase from 60.16 to 93.93 and level 4 increase from 69.98 to 122.85 in the period. The mean standard deviation is also slightly increasing in the period, Table 6.1, confirming that there is a greater variability developing, also indicated by the CWT in Figure 6.4.

6.3 Track Adjustments

Turnout B and C were subject to track adjustment (tamping) 2015-10-03 and 2014-08-25, respectively. By following the development of the transformed geometry data with the CWT, the effect of the tamping campaigns may be evaluated and compared to the behaviour of the turnout before and after the campaigns.

	Level 1	Level 2	Level 3	Level 4
2014-02-22	0.02	0.21	60.16	69.98
2014-05-24	0.02	0.26	68.12	64.57
2014-11-07	0.02	0.26	57.33	78.13
2015-02-28	0.28	0.38	68.84	89.19
2015-05-06	0.02	0.30	58.78	98.11
2015-11-07	0.02	0.30	76.06	81.80
2016-02-27	0.02	0.31	71.37	110.37
2016-05-08	0.02	0.23	72.33	115.54
2016-11-05	0.02	0.31	93.93	122.85

Table 6.2: Squared and summed DWT detail coefficients of levels 1-4, `alignment(r)`, turnout A.

6.3.1 Turnout B

Figure 6.7 displays the `level(l)` before and after the track adjustment, done 2015-10. The tamping has a clear effect at the next observation one month later, but it then falls back, more or less, to the same geometry as before the tamping: the high-frequency oscillations over the crossing nose (seen 2014-05-24) are no longer prominent, but the low frequencies reoccur six months later to keep on the ”steady-state”.

Another event standing out is the peculiar recording taken 2014-05-24. This sudden deviation has no explanation in the maintenance data, and manifests itself as weird measurements in the other variables as well (Figures 6.8, B.4). It is possible that this is the result of a measurement train (TIC) driving in the opposite direction of usual, and that we observe a potential problem of the driving direction transformations.

Also the `level(r)`, Figure 6.8, is an example that tamping does not always cure the unwanted geometry deterioration. It is clear that the peaks of the vertical fluctuations actually amplified between 2015-05 and 2015-11, shifting a bit to the front of the turnout. The tamping 2015-10 did not introduce any high-frequency faults, but it is clear that a strong, low-frequency component emerged after the campaign. It stays relatively constant for as long as the observations extend, one year after. But, on the other hand, the lateral movement in the right rail (`alignment(r)` shows some positive effects of the tamping, muting the ongoing oscillation (Figure B.4). However, the `gauge`, Figure B.5, indicates that the change is merely temporary. One month after the track adjustments the energy is considerably reduced, but it goes back to the regular pattern six months later and onward.

Table 6.3 displays the mean standard deviations of `level` and `alignment`. Observe the unusually high levels in both $\sigma_{\texttt{level}}$ and $\sigma_{\texttt{alignment}}$ 2014-05-24, agreeing with Figures 6.7 and B.4. Furthermore, there is a very modest drop in both variables after the tamping, much like the scalograms indicate.

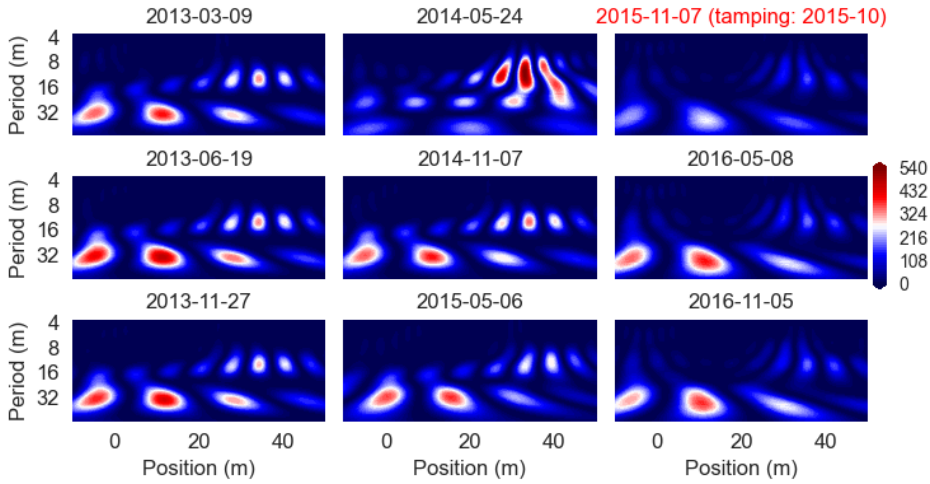


Figure 6.7: `Level(l)` of turnout B, recorded from 2013-03-09 to 2016-11-05. The scalograms shows the squared magnitude of the CWT coefficients using the real Morlet wavelet.

6.3.2 Turnout C

Moving on to turnout C, tamped 2014-08-25, there effects of track adjustment are more positive in the vertical direction (`level`), Figures 6.9 and B.6. The campaign clearly smooth out the deviations in both rails, and the effect seem quite permanent within the window of observations. The lateral directions are split, as `alignment(l)`, Figure B.7, show nearly no change at all, whereas the `alignment(r)`, Figure 6.10, shows a good and lasting effect. The ”total” effect of the sideways movement seen in `gauge`, Figure B.8, seems questionable, as there is little evidence for geometric improvement of this variable.

	$\sigma_{\texttt{level}}$	$\sigma_{\texttt{alignment}}$
2013-03-09	2.06	1.24
2013-06-19	2.14	1.20
2013-11-27	2.08	1.27
2014-05-24	2.46	2.41
2014-11-07	2.03	1.26
2015-05-06	1.96	1.21
2015-11-07	1.90	1.10
2016-05-08	2.13	1.13
2016-11-05	2.07	1.10

Table 6.3: Mean standard deviations of `level` (left and right) and `alignment` (left and right), turnout B.

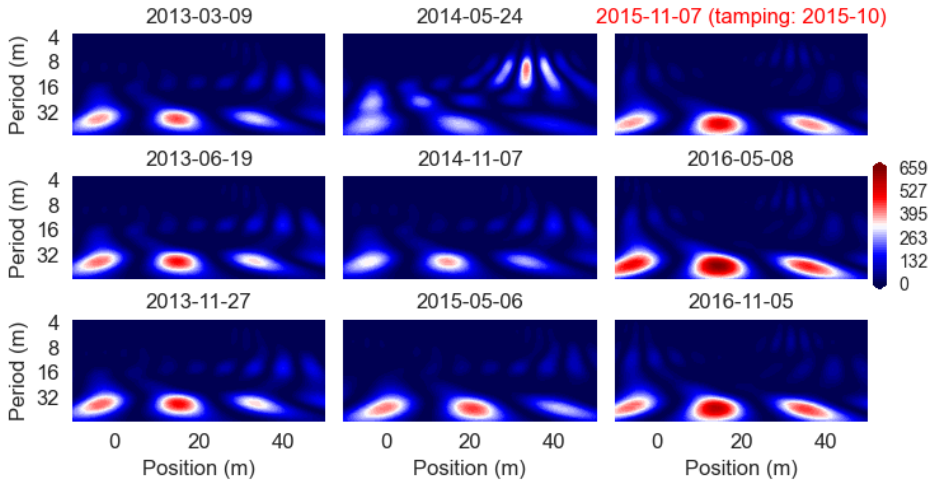


Figure 6.8: `Level(r)` of turnout B, recorded from 2013-03-09 to 2016-11-05. The scalograms shows the squared magnitude of the CWT coefficients using the real Morlet wavelet.

Figure B.7 shows an example of how the CWT may be more insightful than the standard deviation of segments, presented in Table 6.4. The standard deviation of left side `alignment` was more or less the same in 2012-06 as it was in 2016-11, but from Figure B.7 we know that the deterioration in the ballast has actually moved from the front of the rear of the turnout. In addition, there appear to be significant (loosely speaking) seasonal changes in both vertical and horizontal direction for `level` and `gauge` in the CWT, unexplained by the repair records. This is potentially interesting for exploring ground conditions, causing different wear trends.

	$\sigma_{\texttt{level}}$	$\sigma_{\texttt{alignment}}$
2012-06-23	2.96	1.37
2013-06-19	2.84	1.55
2013-11-27	2.94	1.54
2014-05-24	2.93	1.63
2014-11-07	1.58	1.28
2015-05-06	1.80	1.37
2015-11-07	1.89	1.44
2016-05-08	1.43	1.30
2016-11-05	1.63	1.41

Table 6.4: Mean standard deviations of `level` (left and right) and `alignment` (left and right), turnout C.

To summarize, most variables showed a clear improvement in the first months after tamping. For turnout B, the observations suggest that the improvements are merely temporary

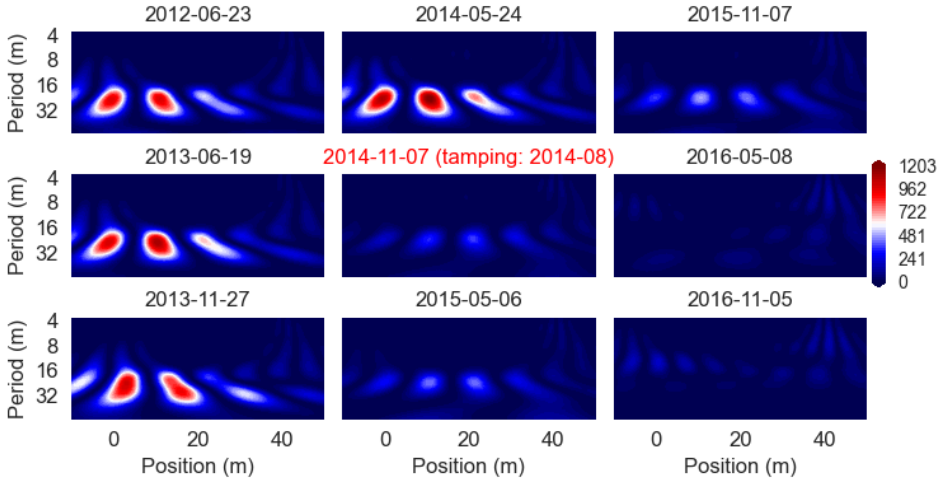


Figure 6.9: `Level (l)` of turnout C, recorded from 2012-06-23 to 2016-11-05. The scalograms shows the squared magnitude of the CWT coefficients using the real Morlet wavelet.

changes (around six months) before the track settles to ”as good as old” geometry. The vertical movements were just shifted, for example, but maintained the same state as before tamping. On the other hand, turnout C serves as an example of more successful tamping. The track adjustments had a very good effect on `level` for both rails, although the effect on `alignment` and `gauge` were more questionable.

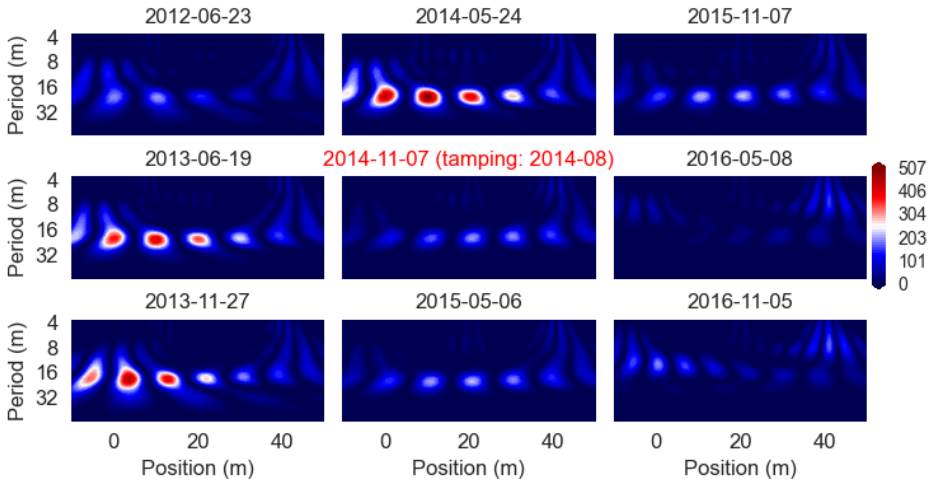


Figure 6.10: `Alignment(r)` of turnout C, recorded from 2012-06-23 to 2016-11-05. The scalograms shows the squared magnitude of the CWT coefficients using the real Morlet wavelet.

Chapter 7

Discussion & Conclusion

Investigation of turnout A, which did not undergo a track adjustment in the observation period, reveals that one can, indeed, observe the turnout geometry degrade using the geometry data and the CWT. The transformation of each separate variable allows the overall geometric condition of the turnout can be followed, improving the understanding of the need for maintenance and the out-of-spec severity. Hence, it is also a valuable tool to assess the effect of major maintenance actions affecting geometry, such as tamping. If combined with previous tamping history, it is an easy task to judge when the turnout has reached the former levels of geometry degradation. In total, the application of the continuous wavelet transform gives several advantages over the traditional way of evaluating track geometry data using the standard deviations.

The discrete wavelet transform may be used to generate an alternative track quality index, however, the current use of vertical and horizontal standard deviation as a quality indication seems to capture the overall condition and trend. The DWT coefficients are difficult to interpret (at least more difficult than using scalograms or standard deviations) and offer little understanding in an unsupervised context. Due to the orthogonality and thereby non-redundancy of discrete wavelets, they may be better suited as features in a supervised analysis, given that some type of sensible labelling is available together with the track geometry.

For future work, two paths should be considered. For the companies dealing with railway infrastructure, it is quite clear from the literature that acceleration box data is the way to go for more detailed supervised analyses of discrete faults. They are small devices that may be mounted on the bogies of trains in regular service, practically providing real-time information of the track condition. This type of vibration data is sensitive even to small irregularities, and enough research exists to implement decision rules based wavelet analysis. The track geometry data still serves an important purpose in deciding on track adjustment actions, and the other approach for predictive maintenance should further investigate how long it takes for particular segments to degrade to the CWT profiles seen

before maintenance, as this provides a helpful time estimate for planning track adjustment campaigns.

Bibliography

Addison, P.S., 2017. The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance. CRC press.

Babb, B., Moore, F., 2007. The best fingerprint compression standard yet, in: 2007 IEEE International Conference on Systems, Man and Cybernetics, IEEE. pp. 2911–2916.

Barkhordari, P., Galeazzi, R., de Miguel Tejada, A., Santos, I., 2017. Low-complexity behavioral model for predictive maintenance of railway turnouts, in: 2017 Annual conference of the prognostics and health management society, Prognostics and Health Management Society.

Bradley, J.N., Brislawn, C.M., Hopper, T., 1993. Fbi wavelet/scalar quantization standard for gray-scale fingerprint image compression, in: Visual Information Processing II, International Society for Optics and Photonics. pp. 293–304.

Brenner, M.J., 2003. Non-stationary dynamics data analysis with wavelet-svd filtering. Mechanical systems and signal processing 17, 765–786.

Cantero, D., Basu, B., 2015. Railway infrastructure damage detection using wavelet transformed acceleration response of traversing vehicle. Structural Control and Health Monitoring 22, 62–70.

Caprioli, A., Cigada, A., Raveglia, D., 2007. Rail inspection in track maintenance: A benchmark between the wavelet approach and the more conventional fourier analysis. Mechanical Systems and Signal Processing 21, 631–652.

Daubechies, I., 1992. Ten lectures on wavelets. volume 61. Siam.

EN 13848-1, 2019. Railway applications - Track - Track geometry quality - Part 1: Characterization of track geometry. Standard. European Committee for Standardization (CEN).

Fisher, M., 2014. Track inspection car, wikimedia commons.

Hopkins, B.M., Taheri, S., 2010. Track health monitoring using wavelets, in: ASME 2010

Rail Transportation Division Fall Technical Conference, American Society of Mechanical Engineers Digital Collection. pp. 9–15.

Hovad, E., Thyregod, C., Andersen, J.F., Lyndgaard, C.B., Spooner, M.P., Rodrigues, A.F.D.S., Ersbøll, B.K., 2019. The effect of driving direction on spatially aligned track recording car measurements in turnouts. International Journal of Rail Transportation , 1–15.

Jensen, P.J., 2016. Sporteknik. Banedanmark and Teknik, LCC Spor.

Jia, S., Dhanasekar, M., 2007. Detection of rail wheel flats using wavelet approaches. Structural Health Monitoring 6, 121–131.

Jiang, Y., Tang, B., Qin, Y., Liu, W., 2011. Feature extraction method of wind turbine based on adaptive morlet wavelet and svd. Renewable energy 36, 2146–2153.

Jøndrup, P.M., 2019. Banenormerne BN1-38: Sporbeliggenhedskontrol og sporkvalitetsnormer. Banedanmark.

Kutz, J.N., 2013. Data-driven modeling & scientific computation: methods for complex systems & big data. Oxford University Press.

Lee, G., Gommers, R., Waselewski, F., Wohlfahrt, K., O'Leary, A., 2019. Pywavelets: A python package for wavelet analysis. Journal of Open Source Software 4, 1237.

Lin, J., Qu, L., 2000. Feature extraction based on morlet wavelet and its application for mechanical fault diagnosis. Journal of sound and vibration 234, 135–148.

Lin, J., Zuo, M., 2003. Gearbox fault diagnosis using adaptive wavelet filter. Mechanical systems and signal processing 17, 1259–1269.

Mallat, S., 2009. A wavelet tour of signal processing: The sparse way. Elsevier.

Molodova, M., Li, Z., Núñez, A., Dollevoet, R., 2013. Monitoring the railway infrastructure: Detection of surface defects using wavelets, in: 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013), IEEE. pp. 1316–1321.

Molodova, M., Li, Z., Núñez, A., Dollevoet, R., 2014. Automatic detection of squats in railway infrastructure. IEEE Transactions on Intelligent Transportation Systems 15, 1980–1990.

Ocak, H., Loparo, K.A., Discenzo, F.M., 2007. Online tracking of bearing wear using wavelet packet decomposition and probabilistic modeling: A method for bearing prognostics. Journal of sound and vibration 302, 951–961.

Soleimanmeigouni, I., Ahmadi, A., Kumar, U., 2018. Track geometry degradation and maintenance modelling: A review. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit 232, 73–102.

Strang, G., 1994. Wavelets. American Scientist 82, 250–255.

Sysyn, M., Gruen, D., Gerber, U., Nabochenko, O., Kovalchuk, V., 2019. Turnout monitoring with vehicle based inertial measurements of operational trains: a machine learning approach. Communications-Scientific letters of the University of Zilina 21, 42–48.

Taubman, D.S., Marcellin, M.W., 2002. Jpeg2000: Standard for interactive imaging. Proceedings of the IEEE 90, 1336–1357.

Toliyat, H.A., Abbaszadeh, K., Rahimian, M.M., Olson, L.E., 2003. Rail defect diagnosis using wavelet packet decomposition. IEEE Transactions on Industry Applications 39, 1454–1461.

Torrence, C., Compo, G.P., 1998. A practical guide to wavelet analysis. Bulletin of the American Meteorological society 79, 61–78.

Tzanakakis, K., 2013. The railway track and its long term behaviour: a handbook for a railway track of high quality. volume 2. Springer Science & Business Media.

Wu, J.D., Liu, C.H., 2008. Investigation of engine fault diagnosis using discrete wavelet transform and neural network. Expert Systems with Applications 35, 1200–1213.

Appendix A

The Fourier transform, the short-time Fourier transform, and proof of the CWT

The Fourier transform. *The Fourier transform is defined as*

$$(\mathcal{F}f)(\omega) = \hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)\, e^{-ix\omega}\, dx. \tag{A.1}$$

The inverse Fourier transform. *The inverse Fourier transform is then given by*

$$f(x) = \hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\mathcal{F}f)(\omega)\, e^{ix\omega}\, d\omega. \tag{A.2}$$

The short-time Fourier transform. *The STFT is given by*

$$(\mathcal{F}^{short}f)(\omega, t) = \int ds\, f(s)\, g(s-t)\, e^{-i\omega s}, \tag{A.3}$$

where $g(\cdot)$ is the window function, restricting the support of the harmonics.

Proof of the inverse continuous wavelet transform. *The proof of the inverse CWT, i.e. the resolution by identity, as given in Daubechies (1992).*

Proof.

$$\begin{aligned}
\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{da\,db}{a^2}(Tf)(a,b)\,\overline{(Tg)(a,b)} &= \iint\frac{da\,db}{a^2}\left[\int\hat{f}(\xi)\,|a|^{\frac{1}{2}}\,e^{-ib\xi}\,\overline{\hat{\psi}(a\xi)}\,d\xi\right]\\
&\quad\left[\int\overline{\hat{g}(\xi')}\,|a|^{\frac{1}{2}}\,e^{ib\xi'}\,\hat{\psi}(a\xi')\,d\xi'\right]\\
&= 2\pi\int da\,a^{-2}\int d\xi\,F_a(\xi)\,\overline{G_a(\xi)}\\
&= 2\pi\int da\,|a|^{-1}\int d\xi\,\hat{f}(\xi)\,\overline{\hat{g}(\xi)}\,|\hat{\psi}(a\xi)|^2\\
&= 2\pi\int d\xi\,\hat{f}(\xi)\,\overline{\hat{g}(\xi)}\int da\,|a|^{-1}|\hat{\psi}(a\xi)|^2\\
&= C_\psi\langle f,g\rangle,
\end{aligned}$$

last interchange is permitted by Fubini's theorem. □

Fubini's theorem. *If the double integral of the absolute value of a function f is finite,*

$$\int\int|f(x,y)|\,dy\,dx<\infty, \tag{A.4}$$

then the order of integration can be changed,

$$\int\int|f(x,y)|\,dy\,dx=\int\int|f(x,y)|\,dx\,dy. \tag{A.5}$$

The Plancherel theorem. *Let f be any square integrable function on the real line $\mathbb{R}$, and $\hat{f}$ it's frequency spectrum in Fourier domain. Then the square of the modulus of f is equal to the square of the modulus of the frequency spectrum of f,*

$$\int_{-\infty}^{\infty}|f(x)|^2\,dx=\int_{-\infty}^{\infty}|\hat{f}(\xi)|^2\,d\xi. \tag{A.6}$$

Appendix B

Scalograms

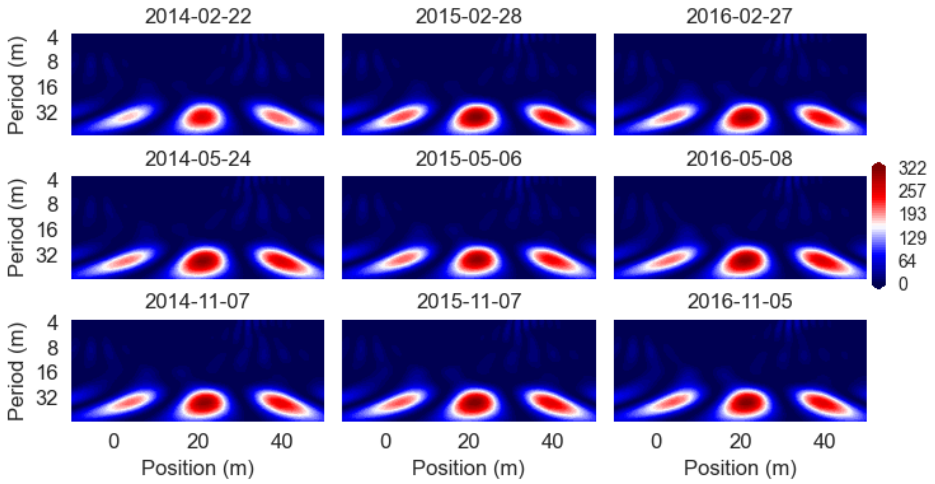


Figure B.1: `Level(r)` of turnout A, recorded from 2014-02-22 to 2016-11-05. The scalograms shows the squared magnitude of the CWT coefficients using the real Morlet wavelet.

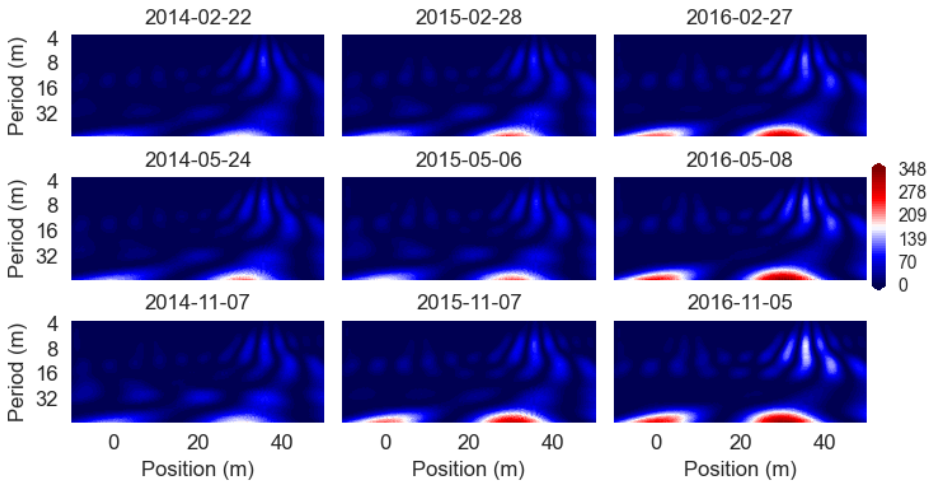


Figure B.2: `Cant` of turnout A, recorded from 2014-02-22 to 2016-11-05. The scalograms shows the squared magnitude of the CWT coefficients using the real Morlet wavelet.

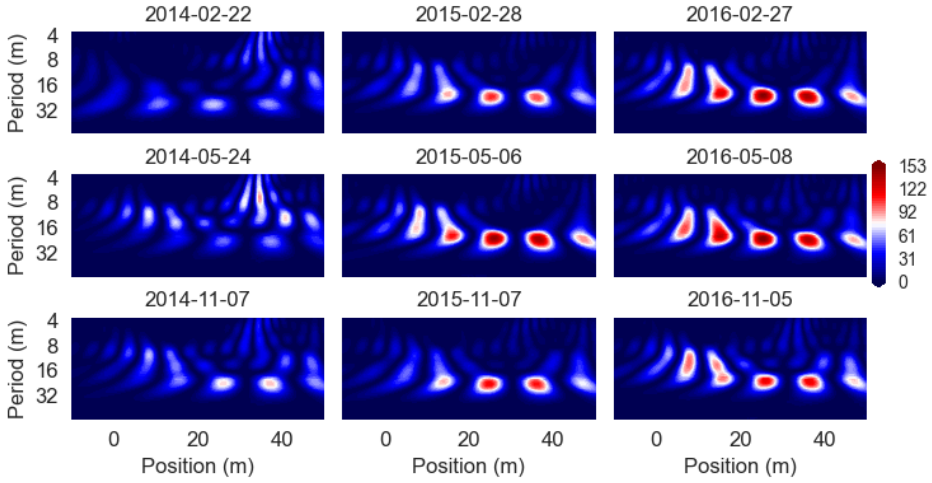


Figure B.3: `Alignment(l)` of turnout A, recorded from 2014-02-22 to 2016-11-05. The scalograms shows the squared magnitude of the CWT coefficients using the real Morlet wavelet.

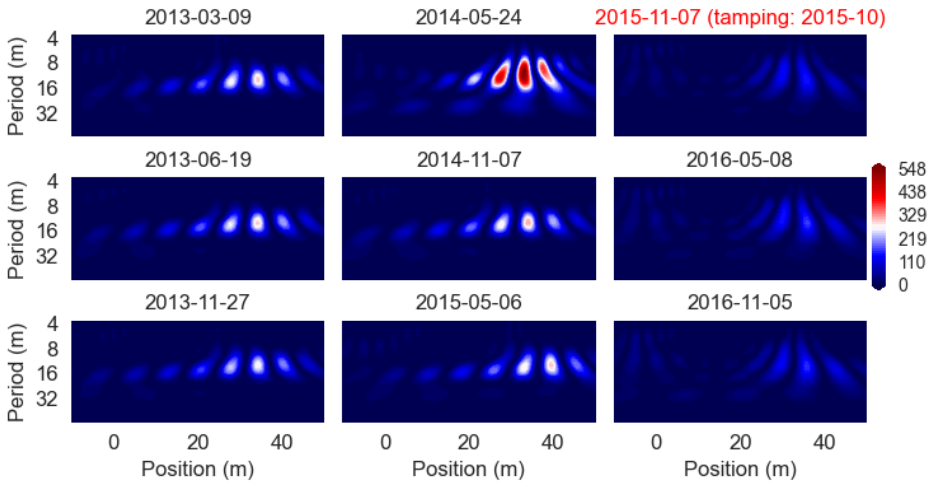


Figure B.4: `Alignment(r)` of turnout B, recorded from 2013-03-09 to 2016-11-05. The scalograms shows the squared magnitude of the CWT coefficients using the real Morlet wavelet.

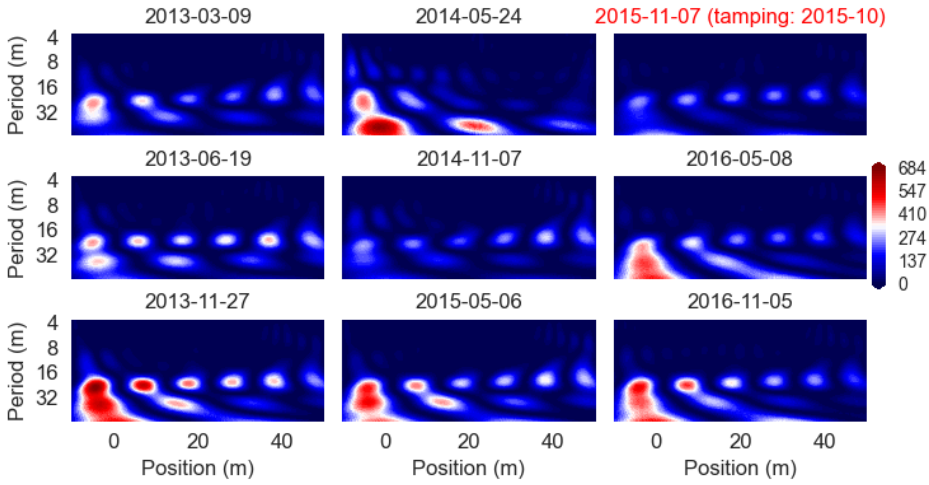


Figure B.5: `Gauge` of turnout B, recorded from 2013-03-09 to 2016-11-05. The scalograms shows the squared magnitude of the CWT coefficients using the real Morlet wavelet.

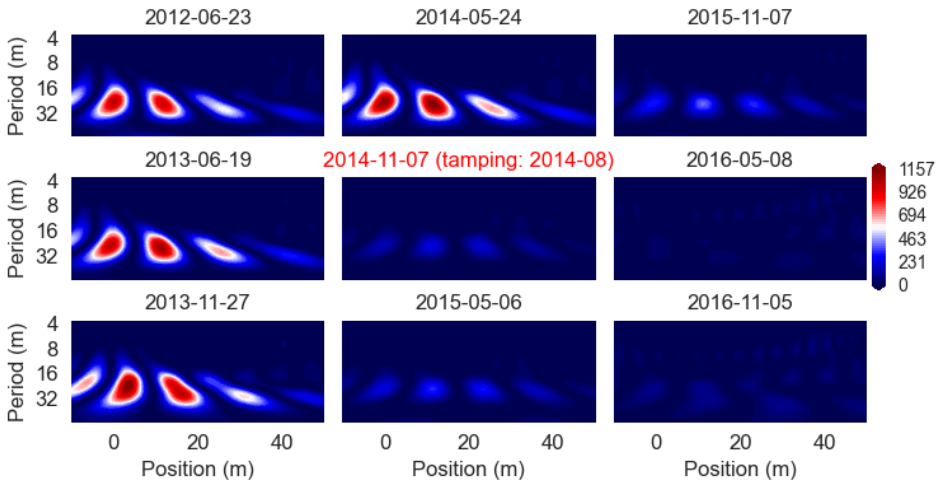


Figure B.6: `Level(r)` of turnout C, recorded from 2012-06-23 to 2016-11-05. The scalograms shows the squared magnitude of the CWT coefficients using the real Morlet wavelet.

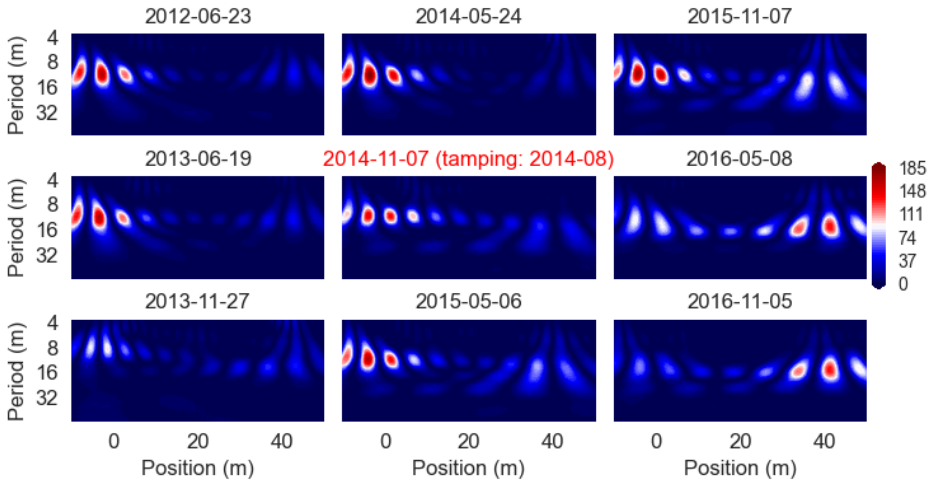


Figure B.7: `Alignment(l)` of turnout C, recorded from 2012-06-23 to 2016-11-05. The scalograms shows the squared magnitude of the CWT coefficients using the real Morlet wavelet.

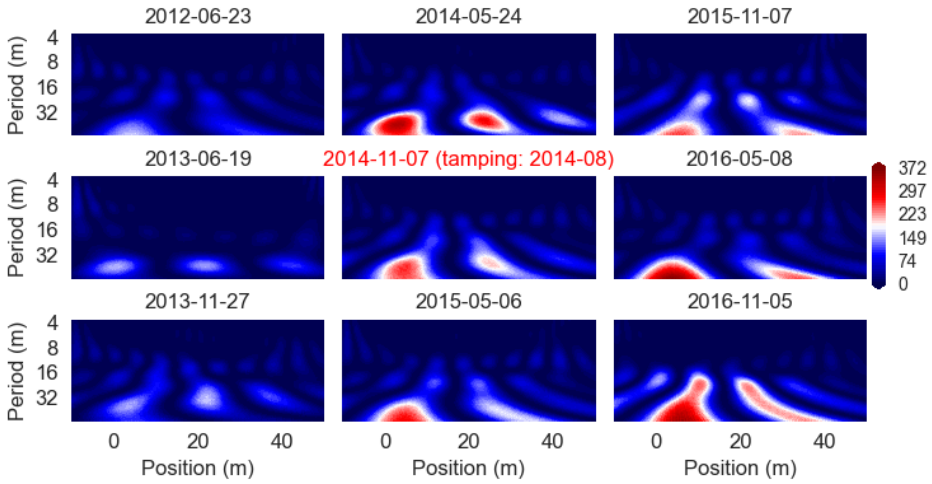


Figure B.8: `Gauge` of turnout C, recorded from 2012-06-23 to 2016-11-05. The scalograms shows the squared magnitude of the CWT coefficients using the real Morlet wavelet.

Appendix C

Code

C.1 Geometry data pre-processing

```
import pandas as pd
import numpy as np
import pickle


def prepare_matrix_from_txt(filename_txt):

    data = pd.read_csv("data/{}".format(filename_txt))

    # Add d1 and d2
    data["Level_L_D1"] += data["Level_L_D2"]
    data["Level_R_D1"] += data["Level_R_D2"]

    data = data.drop(columns=["Level_L_D2", "Level_R_D2"])

    # Change km to meters
    data["km"] *= 1e3  # convert to meters

    # Rename and make lower case
    data.rename(
        columns={
            "Level_L_D1": "Level_L",
            "Level_R_D1": "Level_R",
            "FinalCluster": "Class",
            "km": "Position",
        },
        inplace=True,
    )
    data.rename(columns=lambda x: x.lower(), inplace=True)

```

```
    # Change order
    turnout = data["turnout"]
    date = data["date"]
    pos = data["position"]
    reg = data["region"]
    days = data["days"]
    clas = data["class"]
    evt = data["eventid"]

    data = data.drop(
        columns=["turnout", "date", "position", "region", "days", "class",
        ↪  "eventid"]
    )

    data["position"] = pos
    data["date"] = date
    data["eventid"] = evt
    data["region"] = reg
    data["turnout"] = turnout
    data["days"] = days
    data["class"] = clas

    colname_list = data.columns.values.tolist()
    return data, colname_list


def matrix_to_cube(data):  # input: pandas object

    data = data.to_numpy()

    # Shape to cube
    distinct_switches = len(np.unique(data[:, 10]))
    N, M = data.shape
    i0 = 0
    cube = np.zeros(shape=(250, M, distinct_switches))
    j = 0  # frame number
    for i in range(N-1):  # iterate over all turnout-date obs.
        if data[i, 10] != data[i + 1, 10]:  # if event-id changes
            if i - i0 >= 249:
                signal = data[i0: i0 + 250, :8]
                meta = data[i0: i0 + 250, 8:]  # store frame
                cube[:, :, j] = np.concatenate((signal, meta), axis=1)
                j = j + 1  # increment frame number
            i0 = i + 1  # increment current index

    cube = cube[:, :, np.nonzero(cube[0, 0, :])[0]]  # remove zero-frames

    return cube


def fast_load(np_filename): # for loading processed npy-files
    with open("data/column_names.txt", "rb") as fp:
        column_names = pickle.load(fp)
    data = np.load("data/{}.npy".format(np_filename))
    return data, column_names
```

```
87  # -----------  Prepare data from scratch -----------
88
89  X, colnames = prepare_matrix_from_txt('my_data.txt')
90
91  with open("data/column_names.txt", "wb") as fp:  # need only to be done
    ↪  once
92      pickle.dump(colnames, fp)
93
94  cube = matrix_to_cube(X)
95  np.save("data/cube_spatial", cube)
```

C.2 Figures for theory

```
import numpy as np
import matplotlib.pyplot as plt
import pywt

from numpy.fft import fft, fftfreq, fftshift

import preprosessing as pre

plt.style.use('seaborn-paper')


def illustration_plots():
    n = 2 ** 10  # Make grid
    L = 10
    t0 = np.linspace(0, L, num=n + 1)
    t = t0[0:n]
    dt = t[1] - t[0]

    a = np.arange(0, n / 2)  # For plotting fft with the correct
    ↪  frequencies
    b = np.arange(-n / 2, 0)
    c = np.concatenate((a, b), axis=0)
    k = (1 / L) * c
    ks = np.fft.fftshift(k)

    # --- Sine signal ---
    f = 0.3 * np.sin(0.5 * (2 * np.pi) * t)
    ft = np.fft.fft(f) * dt

    fig, axes = plt.subplots(1, 2, figsize=(7.5, 3))

    for i, ax in enumerate(axes):
        if i == 0:
            ax.plot(t, f)
            ax.set_yticks([-1, 0, 1])
        elif i == 1:
            ax.stem(ks, fftshift(abs(ft)), basefmt='None',
            ↪  use_line_collection=True)
            ax.set_xlim([-1, 1])
        ax.tick_params(labelsize=12)
    plt.tight_layout()
    plt.show()
    # plt.savefig('fig/illustrations/03sine_5pi.eps')

    # --- Chrip signal  ----
    f = np.zeros((n))
    f[549:556] = np.array([0.5, 1, 0.5, 0, -0.5, -1, -0.5])
    ft = np.fft.fft(f)

    fig, ax = plt.subplots(figsize=(4, 3))
    ax.plot(t, f)
    ax.set_xticks([0, 5, 10])
    ax.set_yticks([-1, 1])
    ax.tick_params(labelsize=12)
```

```
    plt.tight_layout()
    plt.show()
    # fig.savefig('fig/illustrations/chirp.eps')

    fig, ax = plt.subplots(figsize=(4, 3))
    ax.plot(ks, np.fft.fftshift(abs(ft)))
    ax.tick_params(labelsize=12)
    plt.tight_layout()
    # fig.savefig('fig/illustrations/chirp_fft.eps')
    plt.show()

    # --- Mexican hat wavelet ---
    mexican = (2 / np.sqrt(3)) * np.pi ** (-1 / 4) * np.exp(-(t - 5) ** 2
    ↪  / 2) * (1 - (t - 5) ** 2)
    t_mexican = np.fft.fft(mexican) * dt

    fig, ax = plt.subplots(figsize=(4, 3))
    ax.plot(t, mexican)
    ax.set_yticks([-0.5, 0, 0.5, 1])
    # ax.set_yticklabels([-0.5, 0, 0.5, 1])
    ax.tick_params(labelsize=12)
    fig.tight_layout()
    plt.show()
    # plt.savefig('fig/illustrations/mexican.eps')

    fig, ax = plt.subplots(figsize=(4, 3))
    ax.plot(ks, fftshift(np.abs(t_mexican)))
    ax.set_xlim([-2, 2])
    ax.set_yticks([0, 0.5, 1, 1.5])
    ax.tick_params(labelsize=12)
    fig.tight_layout()
    plt.show()
    # plt.savefig('fig/illustrations/mexican_fft.eps')

    # --- Complex Morlet ---
    ksi = np.linspace(-5, 5, n)

    real = np.pi ** (-1 / 4) * np.exp(-ksi ** 2 / 2) * np.cos(2 * np.pi *
    ↪  0.8 * ksi)
    imag = np.pi ** (-1 / 4) * np.exp(-ksi ** 2 / 2) * np.sin(2 * np.pi *
    ↪  0.8 * ksi)

    complex_t = np.pi ** (1 / 4) * np.sqrt(2) * np.exp(-(1 / 2) * (2 *
    ↪  np.pi * ksi - 2 * np.pi * 0.8) ** 2)

    fig, ax = plt.subplots(figsize=(4, 3))
    ax.plot(t, real, linestyle='solid', label='Real')
    ax.plot(t, imag, linestyle='dashed', label='Imaginary')
    ax.set_yticks([-1, -0.5, 0, 0.5, 1])
    ax.legend(fontsize=12)
    ax.tick_params(labelsize=12)
    fig.tight_layout()
    plt.show()
    # fig.savefig('fig/illustrations/complex-morlet.eps')

    fig, ax = plt.subplots(figsize=(4, 3))
    ax.plot(ksi, complex_t, label='Fourier transform')
```

```
106     ax.set_yticks([0, 0.5, 1, 1.5, 2])
107     ax.set_xlim([-3, 3])
108     ax.tick_params(labelsize=12)
109     plt.show()
110     # fig.savefig('fig/illustrations/complex-morlet-fft.eps')
111 
112     # --- Real valued Morlet ---
113     morlet = np.exp(-(t - 5) ** 2 / 2) * np.cos(5 * (t - 5))
114     t_morlet = np.fft.fft(morlet) * dt
115 
116     fig, ax = plt.subplots(figsize=(4, 3))
117     ax.plot(t, morlet)
118     ax.tick_params(labelsize=12)
119     fig.tight_layout()
120     plt.show()
121     # plt.savefig('fig/illustrations/morlet.eps')
122 
123     fig, ax = plt.subplots(figsize=(4, 3))
124     ax.plot(ks, abs(fftshift(t_morlet)))
125     ax.set_xlim([-2, 2])
126     ax.tick_params(labelsize=12)
127     fig.tight_layout()
128     plt.show()
129     # plt.savefig('fig/illustrations/morlet-fft.eps')
130 
131 
132 def scalogram_theory(data, attr, wavelet):
133     x = data[:, 8]
134     dx = x[1] - x[0]
135     contour = np.linspace(0, 1, num=5, endpoint=True)
136     # scales = np.arange(2 * 4, 62.5 * 4)  # corresponds to wavelengths
        ↪  2-62.5 m
137     scales = np.arange(2 * 1.3, 62.5 * 1.3)  # for mexican hat
138     fig, axes = plt.subplots(1, 2, figsize=(7, 2.5))
139     axes = axes.ravel()
140 
141     coef, freq = pywt.cwt(data[:, attr], scales, wavelet, dx)
142     period = 1.0 / freq
143     period_p = np.log2(period)  # period logged for plotting purposes
144 
145     abs_coef = abs(coef)
146     real = np.real(coef)
147     imag = np.imag(coef)
148     power = abs_coef ** 2
149     unit_energy = power / np.max(power)  # coefficients for plotting
        ↪  purposes
150 
151     list_coef = [real, imag]
152     list_names = ['Real part', 'Imaginary part']
153 
154     list_coef = [abs_coef, unit_energy]
155     list_names = ['Transform Modulus', 'Energy Distribution']
156 
157     for i, (ax, coef_plot) in enumerate(zip(axes, list_coef)):
158         contour_fill = np.linspace(np.min(coef_plot), np.max(coef_plot),
            ↪  num=50, endpoint=True)
159         # axes[i, j].contour(x, period_p, coef_p, contour, cmap=cmap)
```

```
160         im = ax.contourf(x, period_p, coef_plot, contour_fill,
            ↪  extend="both", cmap='seismic')
161         yticks = 2 ** np.arange(np.ceil(period_p.min()),
            ↪  np.ceil(period_p.max()))
162         ax.set_yticks(np.log2(yticks))  # "log-truncate" y-direction
163         ax.set_yticklabels(yticks.astype(int))
164         ax.invert_yaxis()
165         ax.set_xticks([-10, 0, 10, 20, 30, 40, 50])
166         ax.set_xticklabels([])
167         ax.tick_params(labelsize=12)
168         # ax.set_title(list_names[i], fontsize=12)
169         ax.set_ylabel("Wavelength (m)", fontsize=12)  # fontsize=12
170         ax.set_xticklabels([-10, 0, 10, 20, 30, 40, 50])
171         ax.set_xlabel("Position (m)", fontsize=12)
172
173     fig.tight_layout()
174     fig.subplots_adjust(wspace=0.3)
175     # plt.savefig('fig/cmorl-demo.png')
176     # plt.savefig('fig/cmorl-squared-demo.png')
177     # plt.savefig('fig/mexican-demo.png')
178     # plt.savefig('fig/morlet-demo.png')
179     plt.show()
180
181
182 def plot_geometry_raw(data, column_names):
183     x = data[:, 8]
184     fig, axes = plt.subplots(nrows=1, ncols=2, figsize=(7.5, 3),
        ↪  sharey=True, sharex=True)
185     axes = axes.ravel()
186
187     for i, ax in enumerate(axes):
188         ax.plot(x, data[:, 2 * i], label=column_names[2 * i])
189         ax.plot(x, data[:, 2 * i + 1], label=column_names[2 * i + 1])
190         ax.tick_params(labelsize=12)
191         ax.legend(fontsize=12)
192         ax.set_xlabel("Position (m)", fontsize=12)
193         if i == 0:
194             ax.set_ylabel("Amplitude (mm)", fontsize=12)
195     plt.tight_layout()
196     plt.show()
197     # fig.savefig('fig/illustrations/level-align-B-105.eps')
198
199     fig, ax = plt.subplots(figsize=(4, 3))
200     ax.plot(x, data[:, 4], label=column_names[4])
201     ax.plot(x, data[:, 5], label=column_names[5])
202     ax.tick_params(labelsize=12)
203     ax.legend(fontsize=10)
204     ax.set_xlabel("Position (m)", fontsize=12)
205     ax.set_ylabel("Amplitude (mm)", fontsize=12)
206     plt.tight_layout()
207     plt.show()
208     # fig.savefig('fig/illustrations/gauge-cant-B-105.eps')
209
210
211 def line_plot(data, attr, list_names):
212     fig, ax = plt.subplots(figsize=(7, 3))
213     ax.plot(data[:, 8], data[:, attr], label=list_names[attr])
```

```
214     ax.set_ylabel("Amplitude (mm)", fontsize=12)
215     ax.set_xlabel("Position (m)", fontsize=12)
216     ax.tick_params(labelsize=12)
217     print(np.datetime64('1970-01-01') + np.timedelta64(data[0,
        ↪  -6].astype(int), 'D'))
218     plt.tight_layout()
219     # plt.show()
220     plt.savefig('fig/illustrations/A-align-l-2016-05-08.eps')
221
222
223 ############################################################
224 pretty_names = ['Level (r)', 'Level (l)', 'Alignment (r)', 'Alignment
    ↪  (l)',
225                 'Gauge', 'Cant', 'Twist pc1', 'Twist pc2', 'Position',
                    ↪  'Date',
226                 'Event-id', 'Region', 'Turnout', 'Days', 'Class']
227 data, names = pre.fast_load('cube_spatial')
228
229 illustration_plots()
230 plot_geometry_raw(data[:, :, 105], pretty_names)  # turnout 2840
231 line_plot(data[:, :, 367], 3, pretty_names)
232 scalogram_theory(data[:, :, 367], 3, 'mexh')  # 'morl', 'cmor4.0-0.82'
```

C.3 The continuous wavelet transform

```
import matplotlib.pyplot as plt

import numpy as np
import pywt

import preprosessing as pre


def plot_all_variables(turnout, list_names, power):
    transforms = []
    max_coef = 0
    min_coef = 0

    for i in range(6):
        coefs, freqs = transform(turnout[:, i])
        coefs = coefs ** power
        if np.max(coefs) > max_coef:
            max_coef = np.max(coefs)
        if np.min(coefs) < min_coef:
            min_coef = np.min(coefs)
        transforms.append((coefs, freqs))

    plot_multiple_transforms(transforms, min_coef, max_coef, list_names)
    return


def transform(signal, wave='morl'):
    coefficients, frequencies = pywt.cwt(signal,
                                         scales=np.arange(3 * 4, 50 * 4),
                                         wavelet=wave,
                                         sampling_period=0.25)
    return coefficients, frequencies


def plot_transform(coefficients, frequencies,
                   title=None,
                   colormap='seismic'):
    contour_fill = np.linspace(np.min(coefficients),
                               np.max(coefficients),
                               num=45, endpoint=True)
    period = 1.0 / frequencies
    t = np.linspace(-10, 50, num=250, endpoint=True)

    fig, axes = plt.subplots(figsize=(7, 5))  # 6,8 standard
    im = axes.contourf(t, np.log2(period), coefficients,
                       contour_fill,
                       extend="both", cmap=colormap)
    yticks = 2 ** np.arange(np.ceil(np.log2(period).min()),
                            np.ceil(np.log2(period).max()))
    axes.set_yticks(np.log2(yticks))  # log y-direction
    axes.set_yticklabels(yticks.astype(int))
    axes.set_xticks([-10, 0, 10, 20, 30, 40, 50])
    axes.invert_yaxis()
    axes.tick_params(labelsize=11)
```

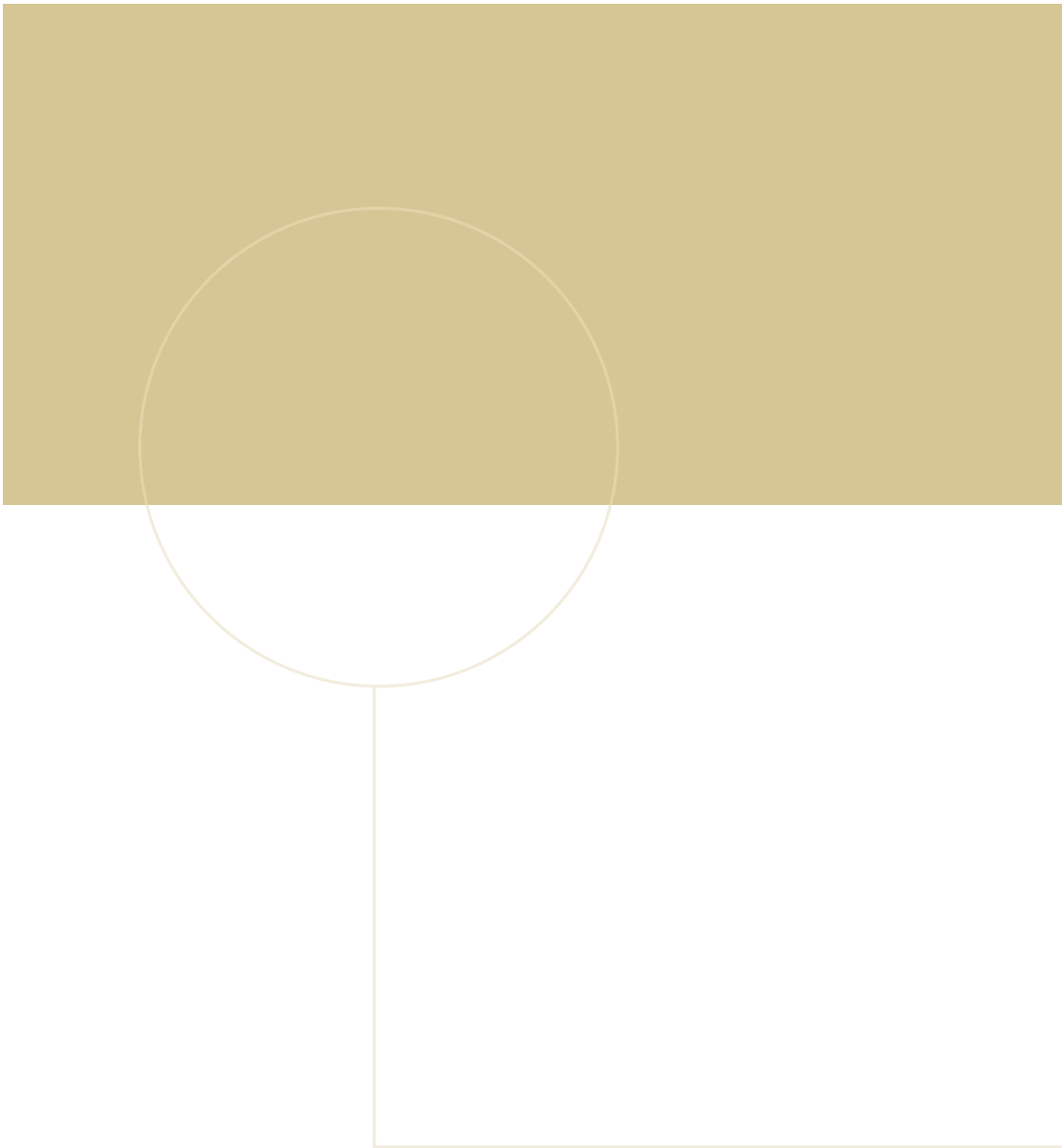
```
    axes.set_ylabel("Wavelength (m)", fontsize=11)
    axes.set_xlabel("Position (m)", fontsize=11)
    if title != None:
        axes.set_title(title, fontsize=12)
    fig.subplots_adjust(bottom=0.15, right=0.88, top=0.9)
    cbar_ax = fig.add_axes([0.91, 0.45, 0.02, 0.3])
    fig.colorbar(im, cax=cbar_ax, orientation="vertical", format='%.0f')
    plt.show()


def plot_multiple_transforms(list_transforms, low_val, high_val, titles):
    n = len(list_transforms)
    fig, axes = plt.subplots(3, 3, figsize=(7.5, 4))  #  3,3 fs=(7,4) or
    ↪  2, 3, figsize=(7, 3)
    axes = axes.ravel(order='F')

    for i, (ax, title, (coefficients, freq)) in enumerate(zip(axes,
    ↪  titles, list_transforms)):
        contour_fill = np.linspace(low_val, high_val, num=45,
        ↪  endpoint=True)
        period = 1.0 / freq
        t = np.linspace(-10, 50, num=250, endpoint=True)
        im = ax.contourf(t, np.log2(period), coefficients,
                         contour_fill,
                         extend="both",
                         cmap='seismic')
        yticks = 2 ** np.arange(np.ceil(np.log2(period).min()),
                                np.ceil(np.log2(period).max()))
        ax.invert_yaxis()
        ax.set_yticks([])
        ax.set_xticks([])
        ax.tick_params(labelsize=12)
        ax.set_title(title, fontsize=12)
        if (i == 0) or (i == 1) or (i==2):
            ax.set_ylabel("Period (m)", fontsize=12)  # fontsize=14
            ax.set_yticks(np.log2(yticks))  # log y-direction
            ax.set_yticklabels(yticks.astype(int))
        if i == 2 or (i == 5) or (i == 8):  #2-5-8 or 1-3-5
            ax.set_xticks([0, 20, 40])
            ax.set_xlabel("Position (m)", fontsize=12)
        # if i == 6: # 9910:4:2014-08, 2840:6:2015-10
            # ax.set_title('{} (tamping: 2015-10)'.format(title),
            ↪  fontsize=12, color='r')
    fig.tight_layout()
    # axes[-1].remove()
    cbar_ax = fig.add_axes([0.925, 0.41, 0.015, 0.25])
    fig.colorbar(im, cax=cbar_ax, orientation="vertical",
                 ticks=np.linspace(low_val,high_val, num=6),
                 ↪  format='%.0f')
    fig.subplots_adjust(right=0.92, hspace=0.40, wspace=0.07)


def cwt_multiple_turnouts(data, observation_range,
                          vars=[0, 1, 2, 3, 4, 5],
                          power=1):
    transformations = []
```

```
107     dates = []
108     max_coef = 0
109     min_coef = 0
110 
111     for i in observation_range:
112         turnout = data[:, :, i]
113         dates.append(np.datetime64('1970-01-01') +
            ↪  np.timedelta64(turnout[0, -6].astype(int), 'D'))
114         signal = np.sum(turnout[:, vars], axis=1)
115         coefs, freqs = transform(signal)
116 
117         coefs = coefs ** power
118 
119         if np.max(coefs) > max_coef:
120             max_coef = np.max(coefs)
121         if np.min(coefs) < min_coef:
122             min_coef = np.min(coefs)
123         transformations.append((coefs, freqs))
124 
125     return transformations, min_coef, max_coef, dates
126 
127 
128 # -----------------  Initialization  ---------------------
129 plt.style.use('seaborn')
130 morlet = pywt.ContinuousWavelet('morl')
131 mexican = pywt.ContinuousWavelet('mexh')
132 complex_morlet = pywt.ContinuousWavelet('cmor4.0-0.82')
133 
134 pretty_names = ['Level (r)', 'Level (l)', 'Alignment (r)', 'Alignment
    ↪  (l)',
135                 'Gauge', 'Cant', 'Twist pc1', 'Twist pc2', 'Position',
                    ↪  'Date',
136                 'Event-id', 'Region', 'Turnout', 'Days', 'Class']
137 fast_names = ['level_r', 'level_l', 'align_r', 'align_l', 'gauge', 'cant']
138 data, names = pre.fast_load('cube_spatial')  # load data cube
139 dict_turnouts = {1000:480, 2000:360, 3000:96}  # dummy turnout numbers for
    ↪  publication
140 
141 
142 # -------------  CWT of variable collections and of variable sequences
    ↪  --------------------
143 for k, v in dict_turnouts.items():
144     coef, freq = transform(data[:, 4, v])
145     plot_transform(coef, freq)  #  Plot a single transform
146     for variable, attribute in enumerate(fast_names):  #  Plot a range of
        ↪  turnouts for a specific variable
147         cwts, minimum, maximum, list_dates = cwt_multiple_turnouts(data,
            ↪  range(v, v + 9), vars=[variable], power=2)
148         plot_multiple_transforms(cwts, minimum, maximum, list_dates)
149         # plt.savefig('fig/case-{}/{}.png'.format(k, attribute))
150 
151 plot_all_variables(data[:, :, v+5], pretty_names, power=2)   # Plot a
    ↪  range of variables for a specific turnout
152 # plt.savefig('fig/case-{}/contri.png'.format(k))
153 
154 plt.show()
```

C.4 The discrete wavelet transform

```
import numpy as np
import pandas as pd
import scipy.stats as stats
import pywt

import preprosessing as pre

pd.set_option('display.width', 150)
pd.set_option('precision', 2)

data, names = pre.fast_load('cube_spatial')  # load cube spatial
dict_turnouts = {1000:430, 2000:360, 3000:96}  # dummy turnout numbers, for publication
wave = 'sym7'

def get_features_dwt(coefs):  # return SSE for each detail level of alignment
    levels = len(coefs) - 1
    sse = np.zeros(levels)
    coefs.reverse()  # get order cD1, cD2, etc.
    for i in range(levels):
        detail = coefs[i]
        sse[i] = np.sum(detail[3:-3]**2)
    return sse

def tqi_stddev_sse(data, turnouts):
    date_list = []
    var_list1 = ['Std. dev. Level' , 'Std. dev. Alignment']
    var_list2 = ['Level 1', 'Level 2', 'Level 3', 'Level 4']
    tqi_reference = np.zeros((len(turnouts), len(var_list1)))
    tqi = np.zeros((len(turnouts), len(var_list2)))

    for i, obs in enumerate(turnouts):
        date = (np.datetime64('1970-01-01') + np.timedelta64(data[0, -6, obs].astype(int), 'D'))
        date_list.append(date)
        tqi_reference[i, 0] = np.mean((np.std(data[:, 0, obs]), np.std(data[:, 1, obs])))  # level
        tqi_reference[i, 1] = np.mean((np.std(data[:, 2, obs]), np.std(data[:, 3, obs])))  # alignment

        coefficients = pywt.wavedec(data=data[:, 2, obs], wavelet=wave)  # alignment-r for case A
        tqi[i, :] = get_features_dwt(coefficients)

    df_reference = pd.DataFrame(tqi_reference, index=date_list, columns=var_list1)
    df = pd.DataFrame(tqi, index=date_list, columns=var_list2)

    return (df_reference, df)


for k, v in dict_turnouts.items():
    tqi_ref, tqi = tqi_stddev_sse(data, range(v, v+9))
```

```
48
49      print('Turnout {} \n Standard deviations'.format(k))
50      print(tqi_ref.to_latex(index=True))
51      print('Turnout {} \n SSE'.format(k))
52      print(tqi.to_latex(index=True))
```



NTNU

Kunnskap for en bedre verden