University of South-Eastern Norway

FMH606 Master's Thesis 2022

Master of Science, Industrial IT and Automation

Real-time monitoring of train wheels and track conditions based on the time series analysis and multivariate data analysis

Santosh Dangal

Faculty of Technology, Natural sciences and Maritime Sciences

Campus Porsgrunn

University of South-Eastern Norway

www.usn.no

Course: FMH606 Master's Thesis, 2022

Title: Real-time monitoring of train wheels and track conditions based on time series analysis and multivariate data analysis

Number of pages: 50

Keywords: Time-series analysis; train; Data Acquisition; fault

Student: Santosh Dangal

Supervisor: Maths Halstensen

External partner: Per Ivan Januschas, CEMIT Digital Porsgrunn, Norway

Summary:

As the railways are a significant section of transportation infrastructure, it is crucial to use good maintenance procedures for railway networks. Condition-monitoring of railway wheels and tracks is especially important where extremely unfortunate failure happens. So regular examination of railway tracks and wheels health is required to maintain safe and reliable train operations. The traditional method of manually inspecting rail tracks is inefficient and prone to human error and bias. This study aims to improve the aging railway system by using an automated solution to address these difficulties.

A series of experimental tests were carried out before data collection. A LabVIEW application was made for recording data and tested with a function generator and Data Acquisition (DAQ) device. The provided Cemit Data Collection (CDC) time-series data were converted into frequency domain through Fast Fourier Transform (FFT) and Wavelet Transform (WT) in MATLAB. Multivariate data analysis was done through FFT and WT data for fault detection using Principal Component Analysis (PCA) and Partial Least Square Regression (PLS-R).

During PCA, the score plot of FFT data depicted the important samples present in it. However, the score plot of WT data illustrated that important samples were absent. Hence, after further analysis of FFT data using PLSR, the calibration model was built which was validated from test set validation. From this, Y predicted was almost equal to the Y reference, hence the model performance can be significant.

From comparing the score plots of FFT and WT data, it is concluded that the WT data is not useful for discovering faults in the tracks using multivariate data analysis because no relevant samples for further investigation were detected. To get better results, it's also a good idea to analyze multi-channel or more sensor data.

Preface

The research for this master's thesis was carried out as a part of the master's degree program in Industrial IT and Automation at the University of South-Eastern Norway. This project is initiated by CEMIT Digital for real-time monitoring of railway track and wheel conditions. Manually inspecting rail tracks is inefficient and prone to human error and bias. This study aims to improve the aging railway system by using an automated method to address these difficulties.

I would like to express my heartfelt gratitude to Professor Maths Halstensen for his continuous encouragement during the project, for assisting me in overcoming problems, and for providing valuable suggestions, direction, and comments throughout the project.

I would also like to thank the Project manager, Senior advisor of CEMIT Digital, Per Ivan Januschas for his invaluable direction, technical assistance, devotion, commitment, and feedback, all of which contributed to the successful completion of the project. I am so thankful to CEMIT team members for their support to run the project.

I also want to express my gratitude for my family and friends' unwavering support and affection during this project, which helped me to keep going and this project work would not have been accomplished without their input.

Porsgrunn, 15-05-2022

Santosh Dangal

Contents

List of figures

Nomenclature

CMPE	Composite Multiscale Permutation Entropy
DAQ	Data Acquisition
DFT	Discrete Fourier Transform
FFT	Fast Fourier Transform
GPS	Global Positioning System
GUI	Graphical User Interface
HF	High Frequency
LED	Light Emitting Diode
LF	Low Frequency
MRA	Multi-Resolution Analysis
NI	National Instrument
PCA	Principal Component Analysis
PIR	Passive Infrared
PLS-R	Partial Least Square Regression
RMSEC	Root Mean Square Error of Calibration
RMSEP	Root Mean Square Error of Prediction
TDMS	Technical Data Management Streaming
TRV	Track Recording Vehicle
WPT	The Wavelet Packet Transform
WT	Wavelet Transform

1 Introduction

In most countries, railways make up a significant section of transportation infrastructure. To avoid service disruptions and preserve the system's safety, it's crucial to use good maintenance procedures for railway networks. For example, the monitoring of railway wheels and track conditions must be regularly done before the failure happens [1]. Condition-monitoring systems are especially important in essential machinery, where extremely unfortunate failures must be minimized due to the possible costs and threats to human life [2].

The railway industry still performs regular preventative maintenance at intervals determined by the distance traveled [2]. Between non-destructive testing examinations, which are currently the industry standard, no information regarding the machine's status is available. Condition-monitoring systems would be a huge benefit to this business in terms of boosting comfort and reliability [3].

For railway track monitoring, several types of sensor technology have been used, including laser technology [4], camera technology [5],[6], and inertial sensors (accelerometers, gyrometers, and so on) [7]. Depending on the main goal of monitoring, some of these technologies can be utilized jointly or independently. Laser optical sensors are typically thought to be expensive, and they can also be costly to maintain. Camera-based systems are less expensive to implement, but they require costly image processing technologies and deliver limited data. Because of their simplicity, low cost, and effectiveness, inertial sensors have been widely employed. The responses of inertial sensors positioned on the train axle box, bogie, or within the cabin are continuously measured [1].

The control of certain features that change when a fault develops is often used in condition monitoring. Choosing features is difficult, especially when vibration signals are employed; because they are complicated and include a lot of information. Signal processing tasks are necessary to make the accessible information in a raw signal easier to manage while ensuring that the defect information is preserved. There are a variety of signal processing tools that can be used [2]. The current study describes a system for monitoring the condition of railway tracks and wheels in real time. Experimental vibration signals were recorded on a freight train and placed a sensor on a cabin. The feature selection was performed, resulting in the identification of a frequency band that appears to provide credible information regarding the fracture. Reliable multivariate data analysis was chosen using the knowledge of the features picked.

1.1 Problem description

Maintaining safe and reliable train operations is necessary for regular examination of railway track health. Cracks, rail discontinuities, burnt wheels, and misalignment generated on the rails as a result of non-maintenance, pre-emptive investigations, and delayed identification pose a serious threat to rail transport safety. The old method of manually examining rail tracks using a railway cart is inefficient and subject to human mistakes and bias. This research intends to improve the old railway system to overcome these concerns by implementing an automated method.

1.2 Aims and objectives

The main objective of this thesis is to predict the fault in railway tracks and give an indication or alarm to the driver that there is a breakage in the railway line. For this, vibrating sensors are used to detect breakage in the railway track. Specific objectives are outlined as follows:

- Developing a Data Acquisition (DAQ) system including software for acquisition (stream to disk) of continuous-time series from multiple sensors.
- Developing software for signal processing of the recorded time series data involving identification and implementation of relevant signal processing methods for the analysis of train vibrations.
- Calibrating multivariate regression models using multivariate data analysis.

1.3 Structure of the report

This thesis contains six main chapters. The first chapter gives an overview of the background, research objectives, and problem description, which gives a proper understanding of the overall project. The second chapter includes the literature review of previous research and some theories related to streaming to disk and time series analysis. The third chapter is about the method and equipment used in this project. The fourth chapter gives results obtained during the project with proper graphs and a discussion about the results. The fifth chapter gives the overall conclusions and findings of this study. And the last chapter presents the future works for making this project more efficient and economical.

2 Literature review

This chapter includes the previous research, streaming to disk, time series analysis, and multivariate data analysis.

2.1 Previous research

Several types of research have been done on the safety and security of railways. In a study [1], an algorithm for detecting railway track faults using data from an in-service train's bogie was proposed. A standard Irish Rail passenger train was instrumented with an accelerometer sensor and a Global Positioning System (GPS). The GPS technology was used to record the train's location with approximate forward speed. Sixty train passes were used to inspect six kilometers of the Dublin-Belfast railway route. The exact amplitudes (acceleration signals) were obtained using the Hilbert transform. A new method for determining the energy levels of amplitudes in the space domain was proposed. The passes with low energy due to low train forward speeds were removed using a data cleaning procedure. The cleaned data was then used and scaled using a new scaling procedure that emphasizes the importance of the vehicle's forward speed in determining the energy level. The scaled energy signals had a high level of consistency, which was suitable for inspecting railway tracks regularly. A persistently high amount of scaled energy across numerous passes was thought to signal a likely track fault. After the train measurements, Track Recording Vehicle (TRV) data was used to verify the track condition in two regions with high energy levels. The high-energy areas of the track were shown to correspond to the defects found in the TRV data.

Figure 2.1 (a) presents an overview of the trailer bogie (i.e., non-powered) of an Irish Rail Hyundai Rotem InterCity fleet car, which was instrumented utilizing inertial sensors, and was put on the trailer (non-powered) bogie [1]. To reduce noise contamination from the power train, sensors were mounted on the trailer bogie. Figure 2.1 (b) shows a tri-axial accelerometer as close to the bogie's center of mass as possible. The data was collected at a 500 Hz sampling rate. A GPS antenna was used to capture the train's location at a frequency of 5 Hz.
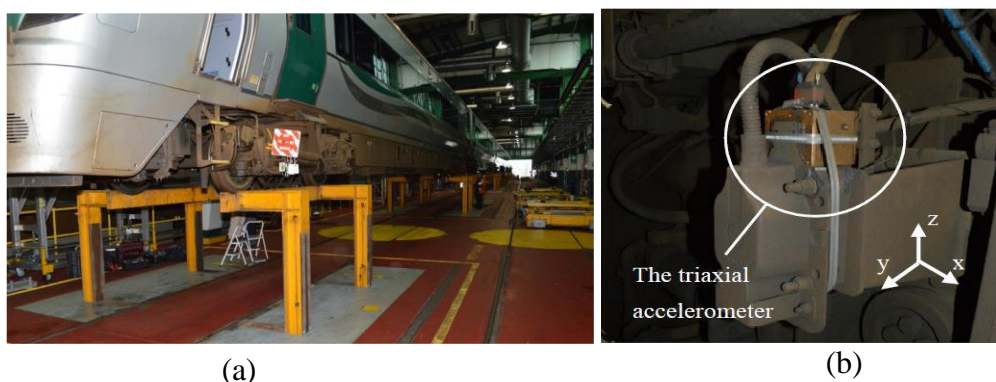

(a)   (b)

Figure 2.1: (a) Instrumented train (b) triaxial sensor [1].

To avoid people, goods, and asset losses, real-time train track detection and adaptability are critical [8]. To track the problems, Yongzhi Min et al. [9] created a real-time visual portable machine vision inspection system. It had an acoustic emission sensor (for track breaks) as well as a Passive Infrared (PIR) sensor for detecting animal dead bodies on the track. The system, which was written in LabVIEW, obtained high-quality photographs from a light source environment by first adding hoods and a Light Emitting Diode (LED) auxiliary light source to the image acquisition equipment. The H value of the color image was immediately used to extract the original image in the second stage, which can reduce image preparation time and was ideal for a target area with a small range. The interference of a significant amount of redundant information was removed via morphological processing, and the direction chain code was employed to swiftly extract the defect's shape features for defect type detection. The rail area rapid locating method detects track problems in real-time, and the system has tight detection time constraints. A track's latitude and longitude coordinates were communicated to a nearby base station if a crack was discovered.

The Wavelet Transform (WT) has been frequently employed for crack diagnosis in recent years. The WT is particularly well suited to this type of research since it operates in both the time and frequency domains, which is necessary when dealing with malfunctioning shaft signals that exhibit nonlinear and nonstationary effects [2]. The Wavelet Packet Transform (WPT energy) in particular, has been successfully exploited as a characteristic for crack identification in shafts [10],[11], [12]. The WPT energy's key advantage is that it turns a raw signal's complicated structure into a simple energy structure in the frequency domain, significantly lowering the amount of information while preserving the information linked to the fracture [2]. This feature has been successfully applied to crack detection in railway axles [13]. The WPT energy has also been utilized to assess the integrity of a rail fastening system in other railway applications [14]. In [15] they use the WPT energy to choose the frequency band with the most defect information for axle bearing diagnosis. Later, the time-domain signal is reconstructed using Composite Multiscale Permutation Entropy (CMPE), and the frequency band chosen is employed as a feature.

2.2 Streaming to disk

A Technical Data Management Streaming (TDMS) file is used to store the time-series data. National Instruments created the TDMS open-source file format. It's a binary format designed for data streaming, thus it can manage dynamically growing files [16]. It contains two types of data, Meta data (Names and properties) and Raw data (Measurement data in binary format). The purposes of using the TDMS files are as follows;

- To store test or measurement data.
- To create structures for grouping the data.
- To store headers/properties about your data.
- To read and write at a high speed.

2.3 Time-series analysis

Time-series analysis is a method of studying a collection of data points over a period of time. The ability to depict how variables change over time distinguishes time-series data from other types of data. In other words, time is an important variable since it reveals how the data changes through time as well as the outcomes. It provides an additional source of data as well as a predetermined order of data dependencies. To maintain consistency and dependability, time series analysis often requires a high number of data points. A large data collection ensures that the sample size is representative and that analysis can cut through noisy data. Time series data can also be used for forecasting or anticipating future data based on previous data [17].

2.3.1 Fast Fourier Transform (FFT)

W. T. Cochran et al. [18] describe FFT as it is a computational method that allows digital computers to do signal analysis such as filter simulation and power spectrum analysis. The FFT is a significant measurement method that breaks down a signal into its various spectral components and hence offers frequency information. The FFTs are employed in machine or system defect analysis, quality control, and condition monitoring. The FFT is an optimized algorithm for performing the "Discrete Fourier Transform" (DFT). A signal (for example time series) is sampled and separated into its frequency components over a time period. Single sinusoidal oscillations at different frequencies having their amplitude and phase are depicted in Figure 2.2. The signal has three separate dominating frequencies across the period.
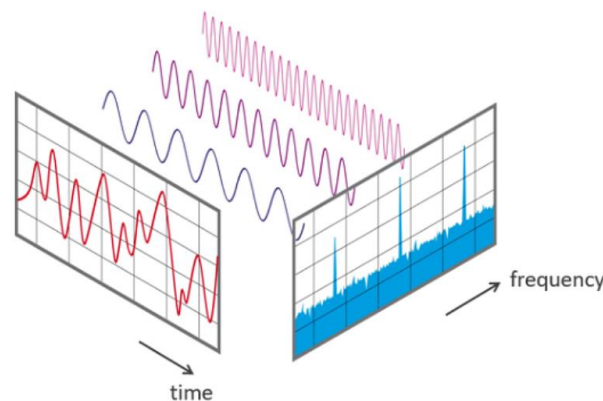

Figure 2.2: A signal in frequency and time domain [19].

2.3.2 Wavelet transforms (WT)

Using the discrete wavelet transform, wavelet-based Multi-Resolution Analysis (MRA) decomposes a signal into several components at different resolutions. As shown in Figure 2.3 a signal is decomposed into some detailed (high-frequency) components and approximation (low-frequency) components [20].
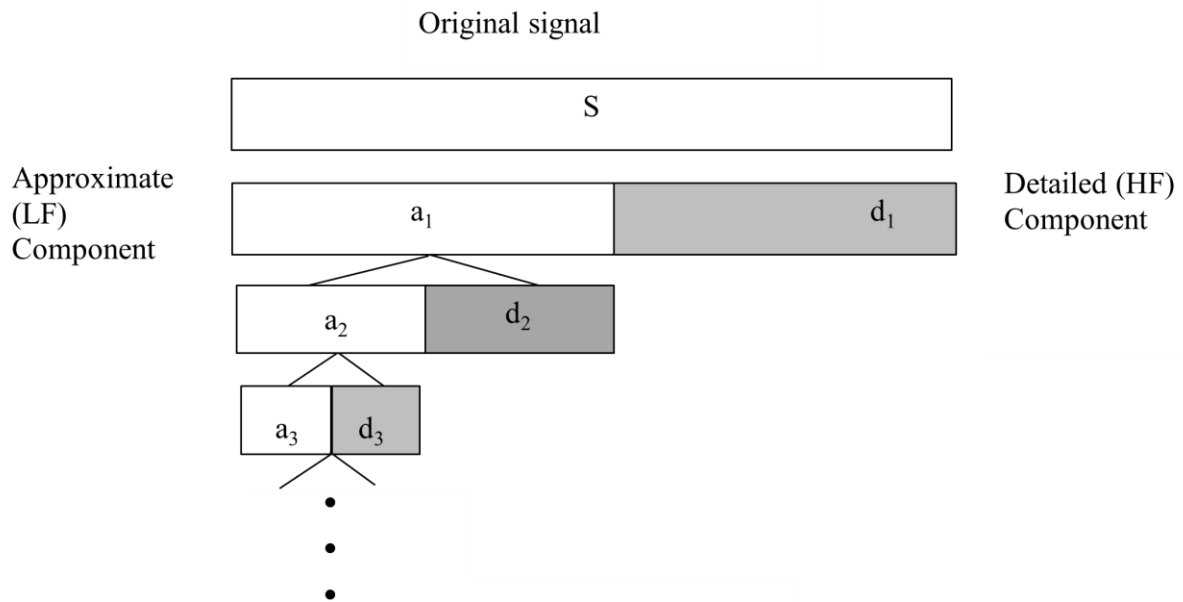
Original signal

Figure 2.3: Decomposition of signal into low frequency (LF) and high frequency (HF) components [20].

where,

a_1, a_2, a_3 = Approximate component (Low frequency)

d_1, d_2, d_3 = Detailed component

$$S = \text{original signal} = d1+d2+d3+ \ ..... \ + dn \ +an \qquad\qquad 2\text{-}1$$

2.4 Multivariate data analysis

Multivariate data analysis is a kind of statistical analysis in which more than two dependent variables are used to produce a single result [21]. There is a different type of methods that can be used for data analysis but in this study Principal Component Analysis (PCA) and Partial Least Square methods (PLS) are used for data analysis.

2.4.1 Principal Component Analysis (PCA)

This method decomposes one data matrix into 'information' and 'noise' [21]. It visualizes the relation between samples and the variables. The score plot depicts objects/samples interrelationships and the loading plot presents variable interrelationships. From this method, outlier can be detected which was used as important samples in this study. The background information of this method is described in the following section.

Scaling or weighting

Scaling or weighting means the preprocessing of data before analysis. The goal of preprocessing is to turn the data into the most useful format for analysis. The methods use for scaling or weighting are standardization and autoscaling. In standardization, the inverse of standard deviation (SDev) is used as a scaling factor. Every element of the X matrix is multiplied by 1/SDev.Whereas in autoscaling, centering and scaling both are used [22].

Outliers

Outliers are observations discovered experimentally which are different from the original data structure. Figure 2.4 shows the outlier in a mild and extreme position. Outliers should be deleted if they are the outcome of incorrect measurements because they will affect the model. In some cases, the outlier plays an important role so it should not be removed otherwise the model will be incorrect [22]. In this study, the outlier is taken as an important sample for making the model.
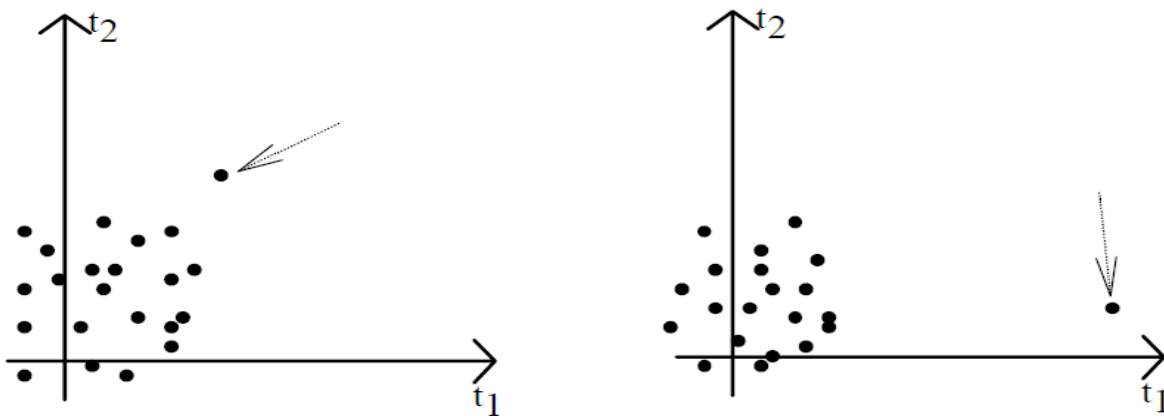
Figure 2.4: Mild and extreme outlier [22]

2.4.2 Partial least Square Regression (PLSR) method

K. H. Esbensen et al. [22], describe PLS as a method that was created to handle completely multivariate regression scenarios (both X- and Y-spaces are multivariate). PLS's main accomplishment is that it addresses MLR's shortcomings while also improving PCR in terms of prediction ability with fewer, and more interpretable components.

2.4.3 Calibration stage

It is the first stage of multivariate modeling (X, Y). The matrices X and Y are involved in multivariate calibration. The dependent variable(s) is represented by the Y matrix, while the independent variables are represented by the X matrix. Using known X and Y data a regression model is built.

The calculation of calibration variance or modeling error is described in the following steps.

A model based on X_{cal} and Y_{cal} is made which means the multivariate model (X, Y) is calibrated as shown in Figure 2.5 [22].

Figure 2.5 Calibration of the multivariate model (X, Y) [22].

The X_{cal} is used to feed the model to predict $\hat{y}_{cal}$
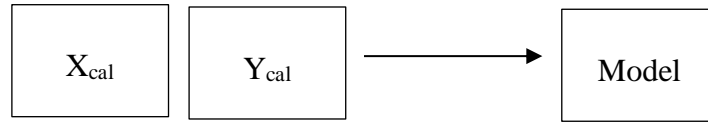
$$X_{cal} + Model \rightarrow \hat{y}_{cal}$$
<div align="right">2-2</div>

By comparing the predicted and measured y_{cal} model error is calculated as given in the equation

$$Modelling\ error = \hat{y}_{cal}\text{-}y_{cal}$$
<div align="right">2-3</div>

Finally, Root Mean Square Error of Calibration (RMSEC) is calculated by,

$$RMSEC = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_{i,cal}-y_{i,cal})^2}{n}}$$
<div align="right">2-4</div>

2.4.4 Validation stage

The testing of the performance of the model by a predetermined set of test result specifications is called validation. Validation testing, for example, in prediction model validation, is concerned with the model's capacity to predict on a new data set that was not used in its creation. The test set is the name given to this new data set [22]. The calculation of validation variance is described in the following steps.

First, the model is applied to the X_{val} values (the test set) and it predicts the proper $\hat{y}_{val.}$

The predicted $\hat{y}_{val}$ calculated as [22]:

$$X_{val} + Model \rightarrow \hat{y}_{val}$$
<div align="right">2-5</div>

Now, predicted ($\hat{y}_{val}$) and measured y_{val} are compared to find the Prediction error, which is given below,

$$Prediction\ error = \hat{y}_{val} - y_{val}$$
<div align="right">2-6</div>

Finally, Root Mean Square Error of Prediction (RMSEP) is given by

<div align="right">2-7</div>

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_{i,val} - y_{i,val})^2}{n}}$$

3 Materials and methods

This section describes the proposed methodology and equipment used in this study.

3.1 Methods

Figure 3.1 depicts a detailed block diagram of the methods employed in this study. The core methods include sensing elements, Data acquisition, signal processing, and multivariate data analysis.
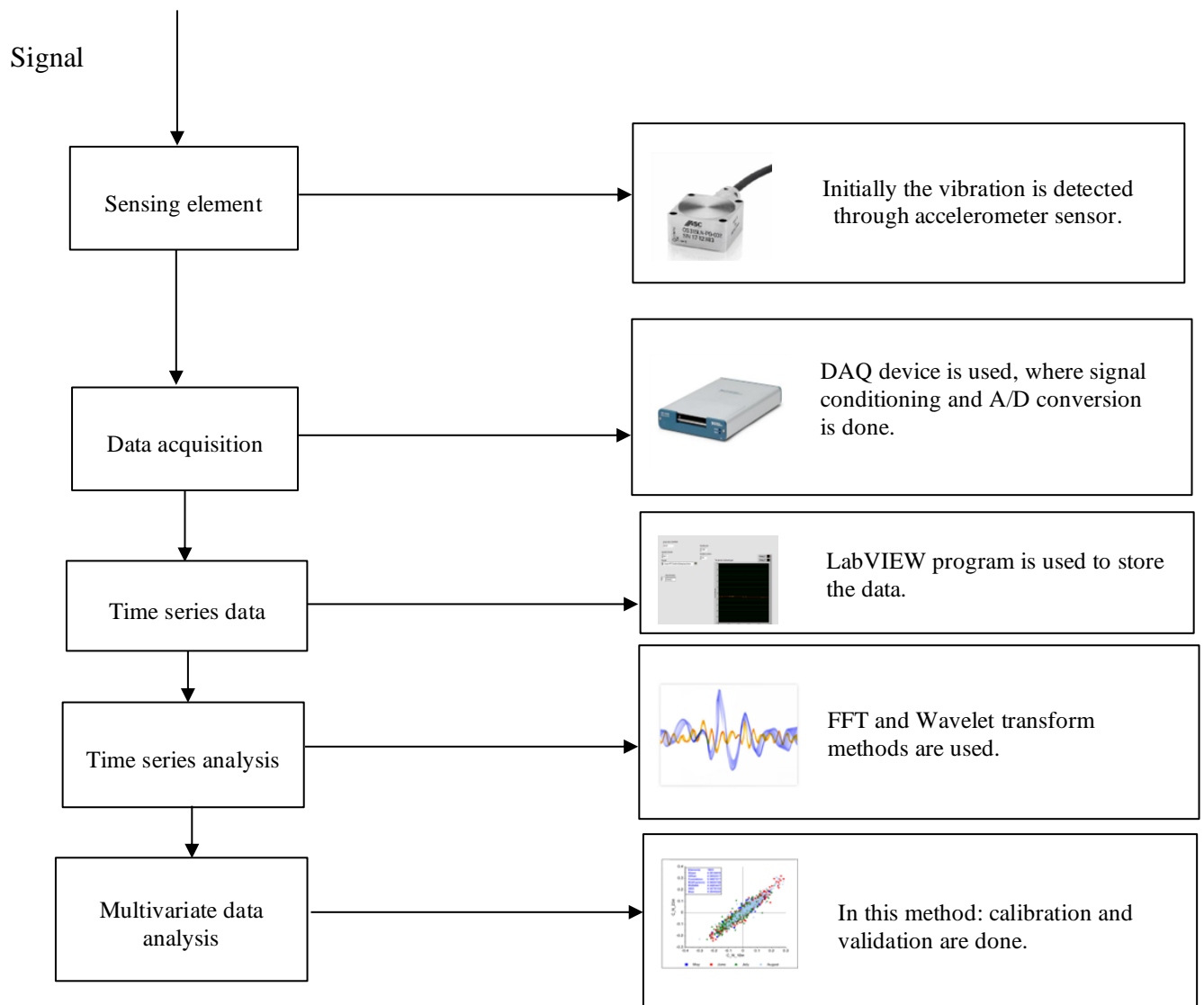

Figure 3.1: Overview of proposed methodology.

Sensing element:

In this research, accelerometer sensors are used to assess the vibration of the railway structure and to monitor the readings in real-time for interesting frequencies. Triaxial capacitive accelerometer sensors were planned as sensing devices to measure the signals. In the next stage, the measured signal is sent to Data Acquisition.

Data Acquisition:

The USB X series NI USB-6363 device by National Instruments (NI-DAQ) was planned to use to digitize the analog signals. A LabVIEW-based Graphical user interface (GUI) has been created for storing the data on disk in TDMS format. Since LabVIEW Programming Language is also by NI, the DAQ hardware was nicely integrated with the application software.

Time series data:

Time series data, which also means time-stamped data, are data points recorded in time order. A LabVIEW program was used to store time-series data on disk in TDMS format.

3.2 Experiment setup

This section details the setup's working method as well as the hardware and software applications that were planned to employ during the project.

3.2.1 Locomotive

The locomotive that was planned to use for the data collection was a freight train. This train is used for carrying raw materials to cement factories. The train run with approximately a weight of 896 tons with loads and 316 tons (information from train operator) when unloaded. The train operates between Porsgrunn and Brevik with an approximate distance of 9.73 kilometers. Figure 3.2 (a) shows the proposed locomotive and (b) shows the map where the locomotive runs.
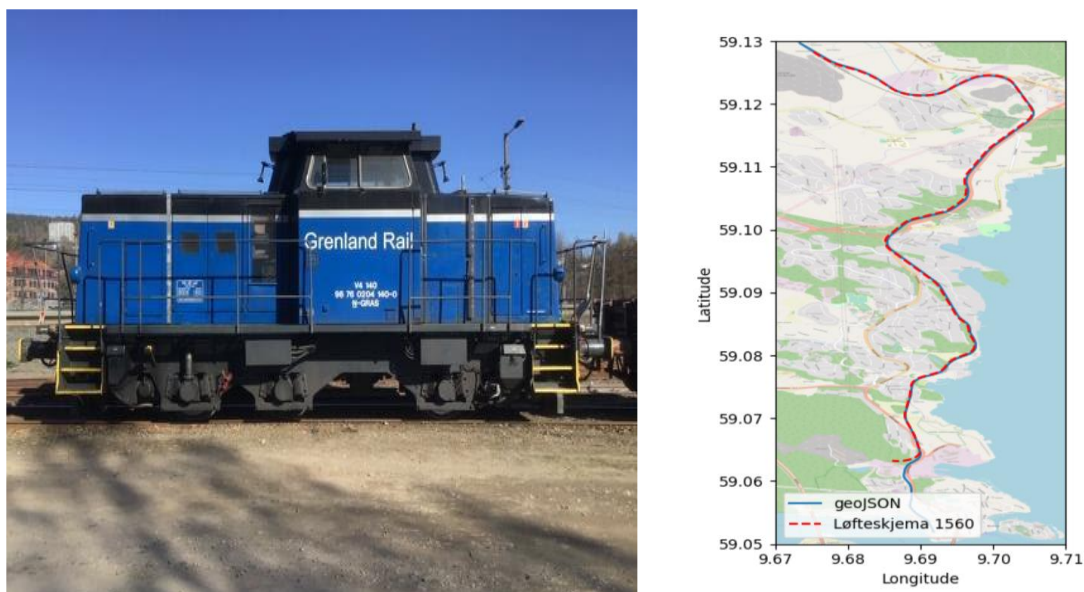

Figure 3.2 (a) Diesel freight locomotive for experiment (b) Train route from Porsgrunn to Brevik

3.2.2 Sensors

The sensors purposed for the data collection are two MEMS capacitive triaxial accelerometers. The triaxial sensors are chosen as the vibration from all the axis was needed. These were planned to place near the wheel. The locations of these sensors are chosen concerning the high vibrational area. The sensors were tested in the lab before going to mount on the train. They were tested by connection with NI USB-6363 DAQ device and LabVIEW program. Figure 3.3 shows the triaxial capacitive accelerometer sensor. And Table 3.1 shows the details of the sensor.

Figure 3.3: Capacitive triaxial sensor [23].

Table 3.1: Detail specification of sensor [23].

Sensor Name	Measurement range	Axis	Frequency range	Operating Temperature range	Supply voltage	Differential analog output voltage
Triaxial MEMS capacitive	±2 to ±400g	x, y, z	2kHz (±5)	-40 to +100 °C	6 to 40 V	±4 V

3.2.3 Data Acquisition

The USB series NI USB-6363 device by National Instruments (NI-DAQ) was planned to use for data acquisition which is shown in Figure 3.4 a) and b) shows the channels of DAQ. It processes the physical data from the real world and converts it into digital numeric values that can be manipulated by Computer. The connection to the DAQ channel from the sensor output during the testing of sensors can be seen in Table 3.2.
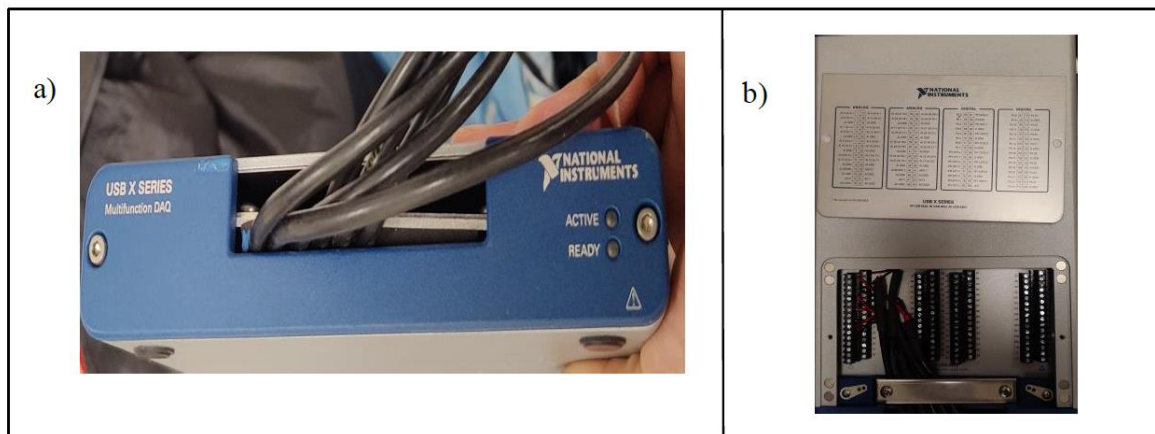
Figure 3.4 : (a) NI DAQ module (b) Channels of DAQ

Table 3.2: Connections of the sensor with DAQ during testing of sensor [23].

Sensor cable number	Pin	DAQ channel connected	LabVIEW channel
1	Power	-	-
2	GND	GND	-
3	Signal+ (X-Axis)	1	Channel 0
4	Signal- (X-Axis)	2	
5	Signal+ (Y-Axis)	4	Channel 1
6	Signal- (Y-Axis)	5	
7	Signal+ (Z-Axis)	7	Channel 2
8	Signal- (Z-Axis)	8	

3.2.4 Application

A LabVIEW application was developed for collecting time-series data in TDMS format. Figure 3.5 shows the front panel of the application. Both the LabVIEW application and the DAQ device have been developed under NI, hence they worked quite well without any technical complications.
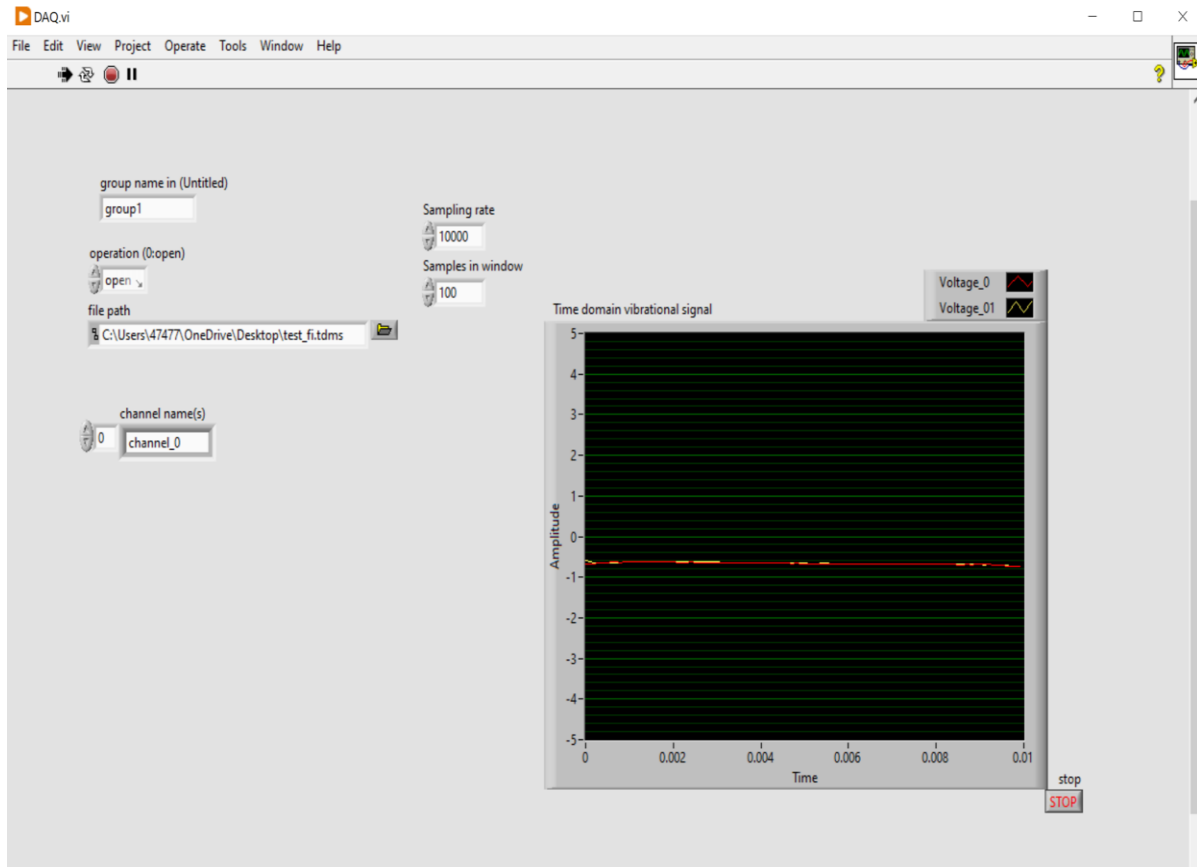
Figure 3.5: Front panel of LabVIEW program visualizing time-series signal during testing of sensor

3.2.5 Testing of Application

Testing of the LabVIEW application was done through the NI USB-6251 DAQ device which is shown in Figure 3.6 (a). And for the input, the function generator was used with different frequencies and voltages which is shown in Figure 3.6 (b). Firstly, the application was made for single-channel and then it was updated for more channels which means for more sensors or other axes of the tri-axial sensor. For this, two function generators were used which can be seen in Figure 3.7. During the testing time, the time series data were recorded in TDMS format, which is shown in Figure 3.8.
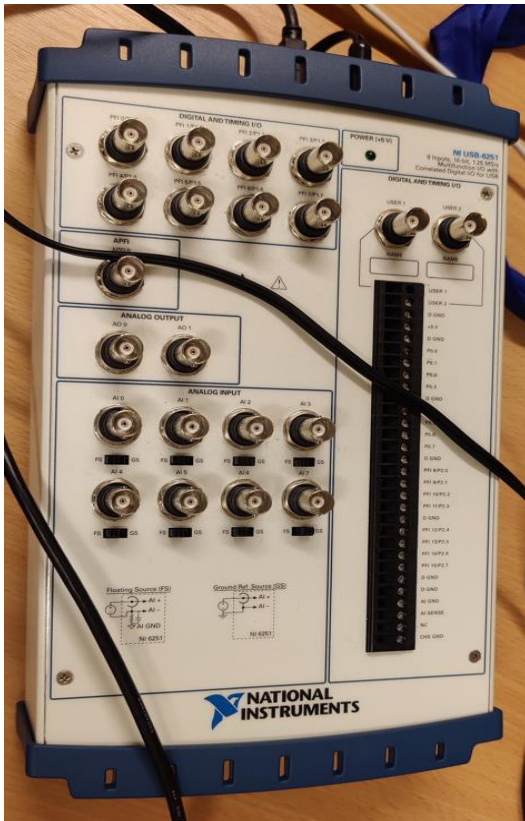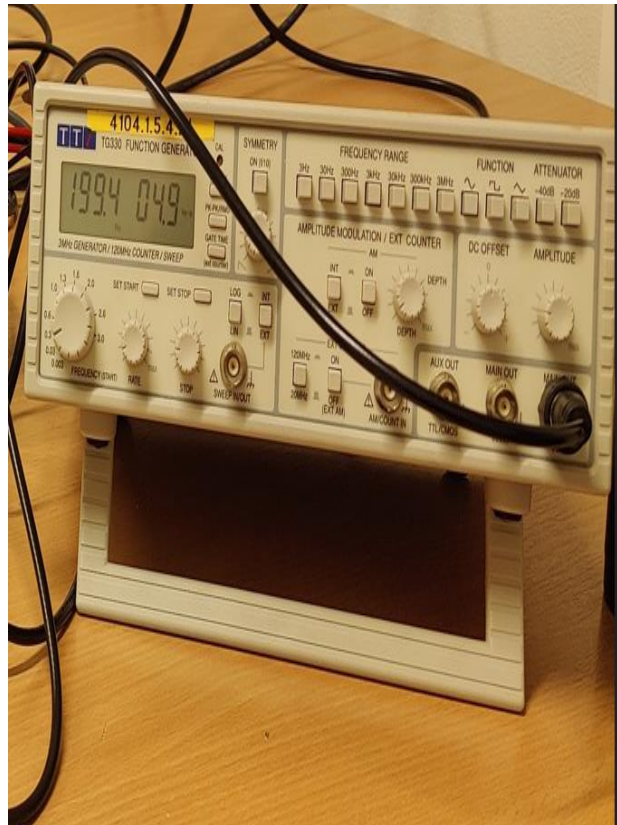
(a)   (b)

Figure 3.6 : (a) NI USB-6251 DAQ device and (b) Function or signal generator
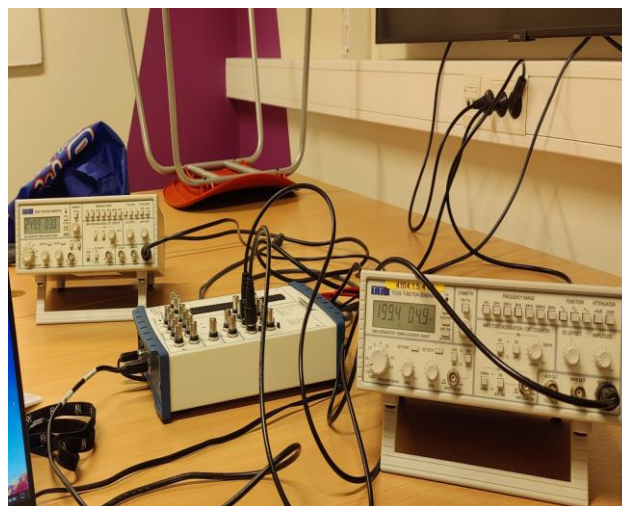
Figure 3.7: Testing LabVIEW Application with two function generators.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Root Name	Title	Author	Date/Time	Groups	Description							
2	first_test				1								
3													
4	Group	Channels	Description										
5	Group1	1											
6													
7	Group1												
8	Channel	Datatype	Unit	Length	Minimum	Maximum	Description	NI_ChannelName	NI_UnitDescription	wf_increment	wf_samples	wf_start_offset	wf_start_time
9	channel_00	DT_DOUBLE	Volts	35400				Voltage	Volts	0.001	100	0	2022-02-28 03:03:50.
10													
11													
12													

Figure 3.8: TDMS file with detail about time-series data.

Figure 3.8 shows the details of the time series data with starting time, time step, amplitude (voltage), and length of data. Figure 3.9 shows the plot of time series data recorded from a signal generator with 41Hz of 4 peak to peak (p-p) voltage and Figure 3.10 shows the frequency domain plot of the same time series data which shows the almost 41Hz signal so the LabVIEW program tested successfully.
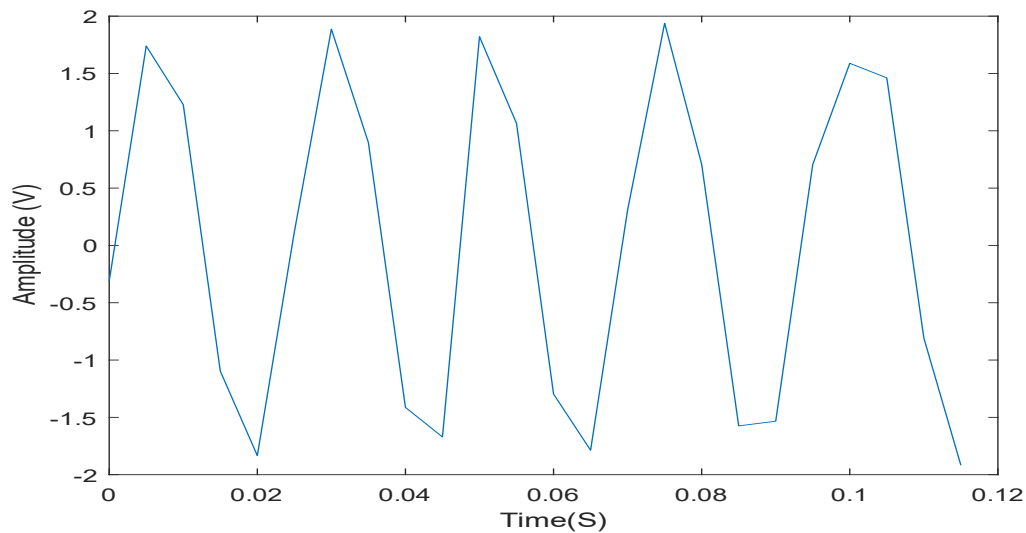
Figure 3.9: Time-domain plot of time series data for 0.115 sec.
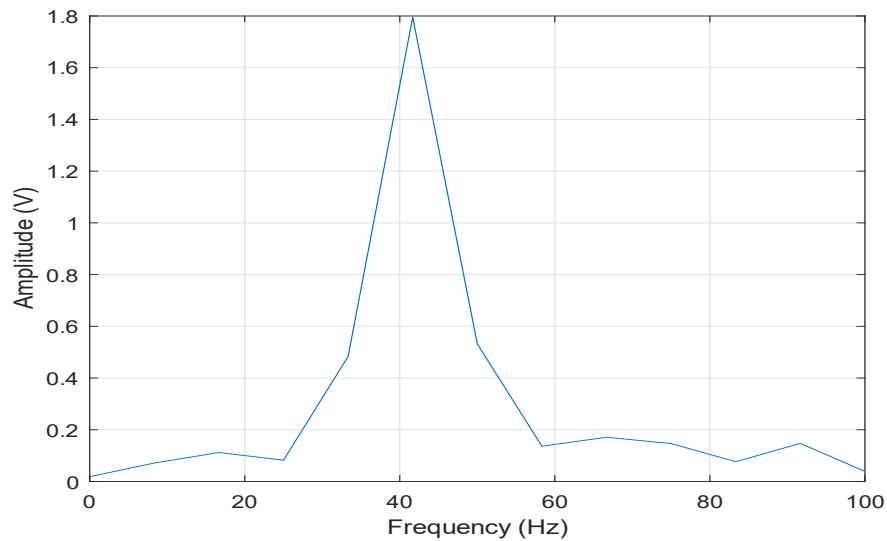
Figure 3.10: Frequency domain plot of 41Hz signal.

3.2.6 Sensor testing

The capacitive triaxial sensor was tested before going to install it in the train which can be shown in Figure 3.11 and Figure 3.12. The senor wires connected to DAQ as a reference to Table 3.2. The sensor was tested successfully by manually vibrating the sensor on the x-axis, the y-axis and z-axis respectively which can be visualized in LabVIEW as shown in Figure 3.12.

(a)

(b)
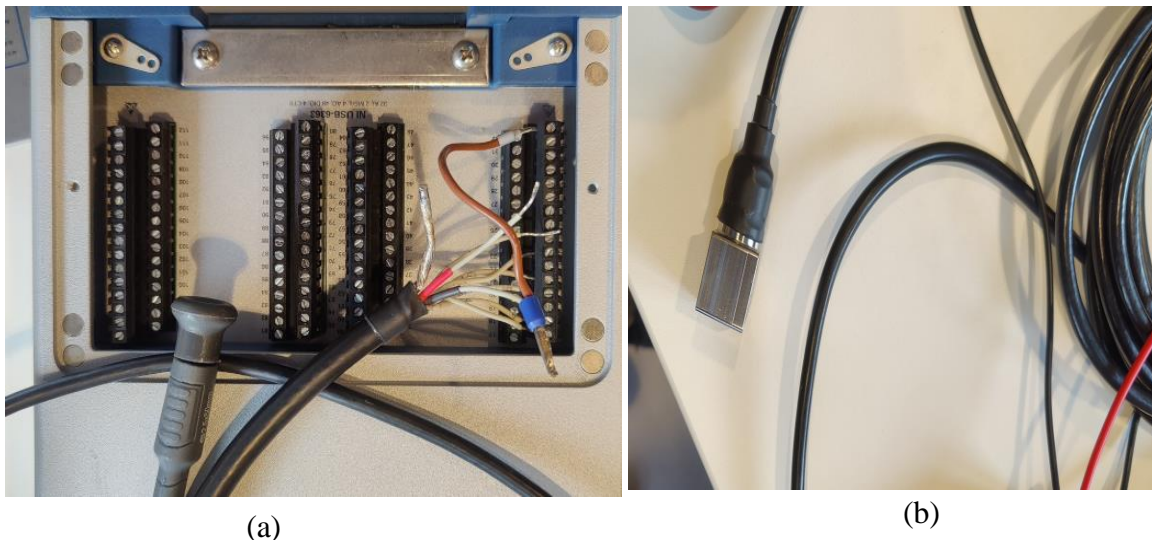
Figure 3.11: (a) Sensor wire connected in DAQ device (b) Triaxial capacitive sensor
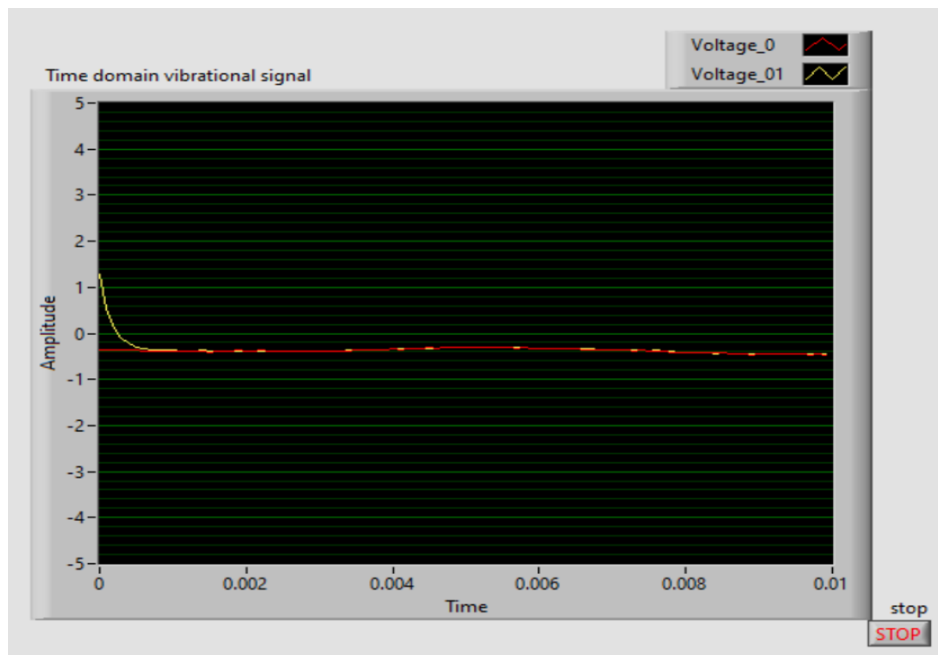
Figure 3.12: Amplitude variation of sensor data during testing.

3.3 Process description

First, the LabVIEW application was made for storing the time series data and tested successfully. The application was tested by giving input as a sine wave signal from a function generator with different voltage (p-p) as shown in Figure 3.13.
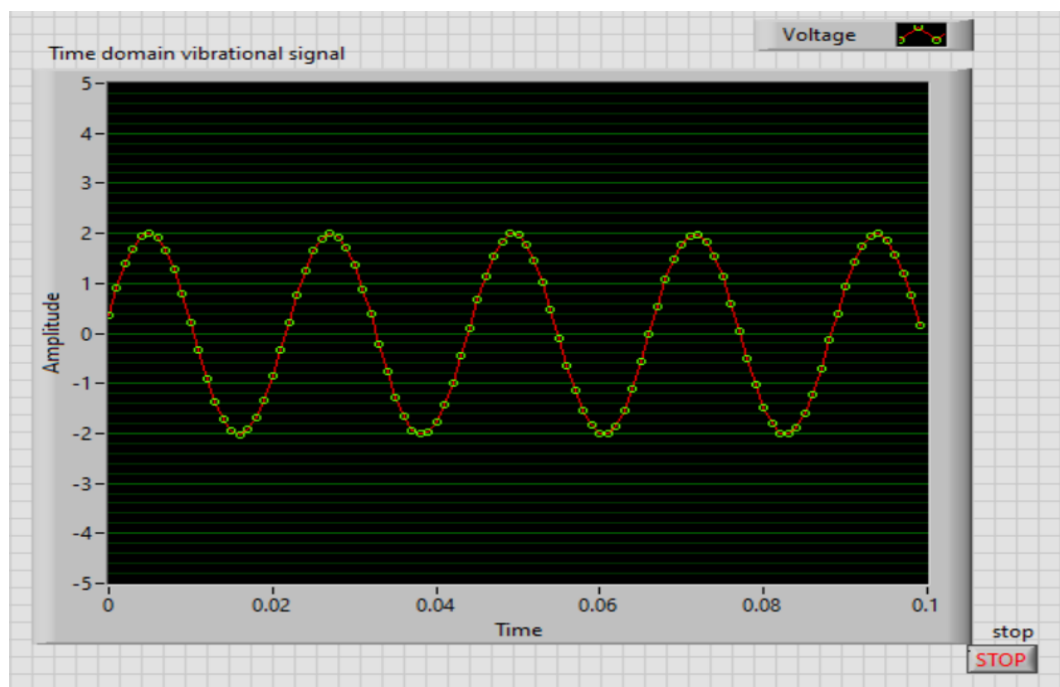
Figure 3.13: Sine wave generated with 4V (p-p) from function generator visualized through LabVIEW.

Due to the busy schedule of train operations, the data could not be collected, so the analysis was conducted through the Cemit Data Collector (CDC) time-series data provided by CEMIT. These data were collected by the accelerometer sensor placed on the cabin of the train as information given by CEMIT.

These time-series data are converted into the frequency domain by FFT through the MATLAB program with favorable window size (128) and windowing transform which will be discussed further in the result chapter. The result from the FFT was stacked in a row and made a matrix (5102*65) for multivariate data analysis. In multivariate data analysis, the PCA was used for detecting the most important samples and the Matrix was made with output (fault) for the PLSR method for making the model. The other method of time series analysis, WT also used to break down the frequencies into small frequencies for feature extraction.

The fault detection part was done from the multivariate data analysis which will be discussed further in the result section. For the detection of the exact position of fault in the track, the comparison of the timing for the fault locations with the respective distances in the data provided by CEMIT was done. The details will be discussed in section 4.3 of the results chapter.

4 Results and discussion

All the results described here are outputs from the previously described methodology. Here, MATLAB and Unscrambler software were used for the data analysis. For the time series analysis, time-series data recorded from the tri-axial sensor with a sampling frequency of 500 Hz, time step 0.002, and total time period of 1271.905 seconds was used as shown in Figure 4.1.
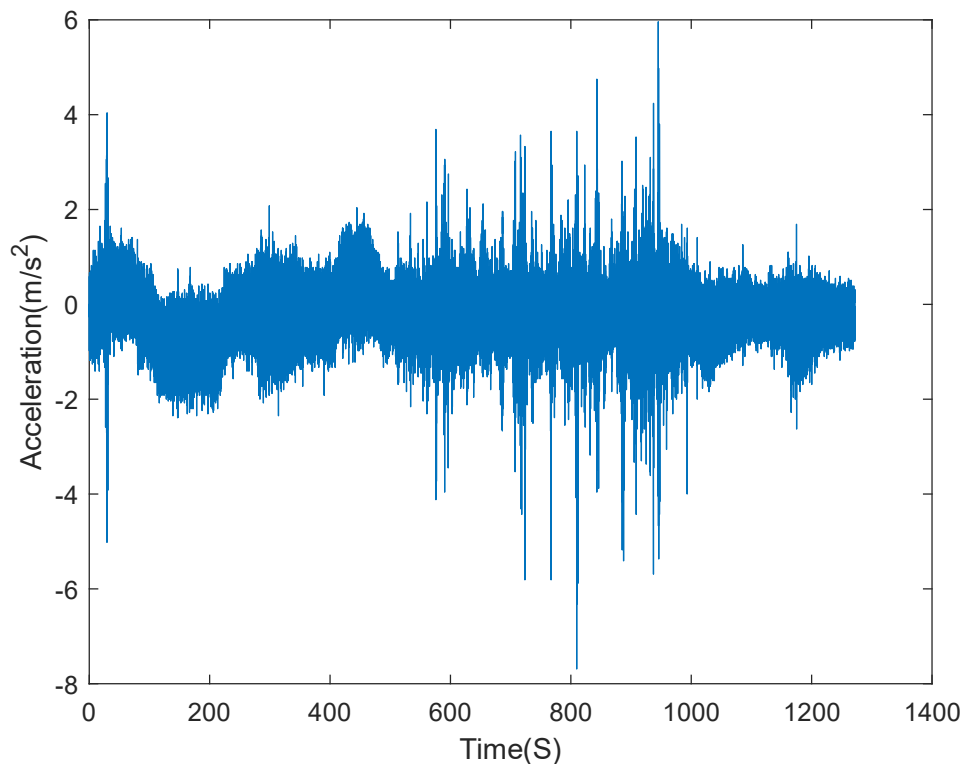
Figure 4.1: Time-domain plot of x-axis (channel_00) data for 1271.905 seconds

Figure 4.1 shows the times domain plot for the tri-axial sensor x-axis (channel_00) which possibly depicts more vibration (acceleration) between 600 and 1000 seconds as the amplitudes are at a high position.

And the time series plot for the y-axis (channel_01) is shown in Figure 4.2. From this also, it can be assumed that more vibration occurs between 600 and 1000 seconds.

Figure 4.2: Time-domain plot of y-axis (channel_01) data for 1271.905 seconds.

Now plotting the z-axis (channel_02) time-series data, the plot obtained is shown in Figure 4.3. This plot also shows that more vibration is around 600 to 1000 seconds, but the highest vibration is nearly at 900 seconds.
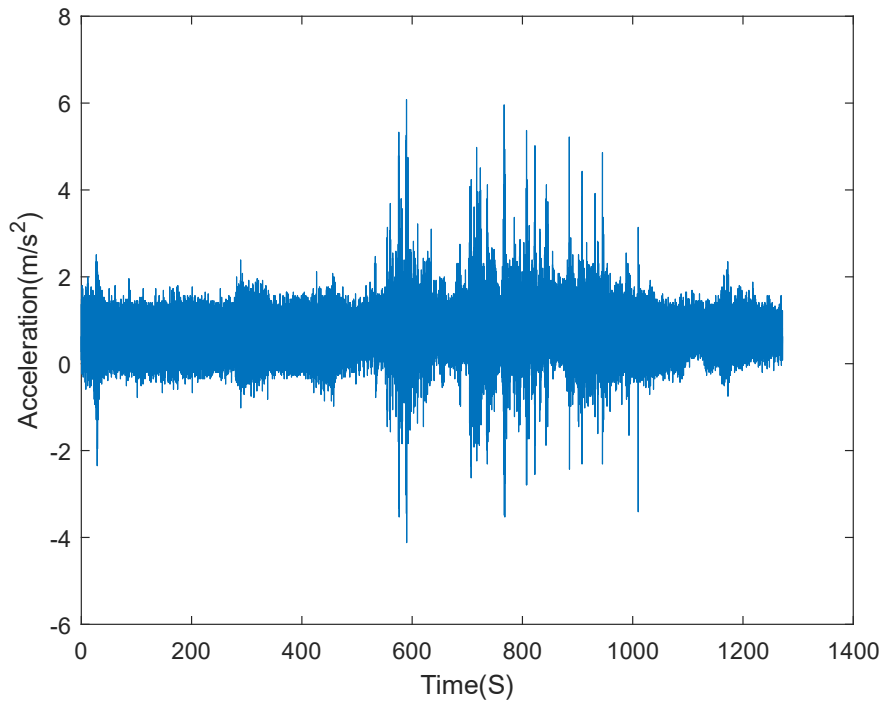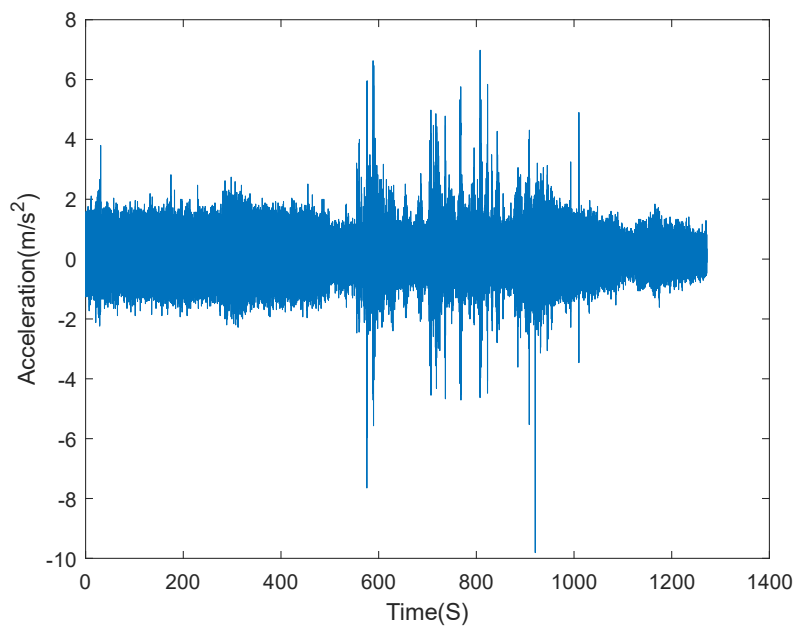

Figure 4.3: Time-domain plot of z-axis (channel_02) data for 1271.905 seconds.

From this time series data, time series analysis is done through FFT and Wavelet Transform. After that, multivariate data analysis was performed which will be discussed in the next topics.

4.1 Analysis from Fast Fourier Transform (FFT)

For the FFT, some calculations were done before doing the analysis,

Calculation of time step:

$$\text{Time step} = \frac{window\ size}{sampling\ frequency} = \frac{128}{500} = 0.256 \text{ Hz} \qquad \text{4-1}$$

Calculation of frequency resolution:

Here the frequency resolution (Δf) is calculated by the formula,

$$\Delta f = \frac{Sampling\ frequency}{window\ size} = \frac{500}{128} = 3.90 \text{ Hz} \qquad \text{4-2}$$

Now, 128 samples (window size) with a time of 0.256 seconds (see equation 4-1) from x-axis time series data (see Figure 4.1) were chosen, and FFT was done which can be shown in Figure 4.4. The frequency-domain plot shows the signal of 150 Hz. For Windowing transform, Blackman Harris was used with window size 128 which can be seen in MATLAB code in appendix B.

The FFT of 128 samples gave the 65 frequencies data which were stacked in rows one after another and made a matrix of 5102*65 (for one channel) with every time step 0.256 (for calculation see equation 4-1) for multivariate data analysis. This will be explained in next section.

Similarly, FFT of other channels (y-axis and z-axis) was done which is shown in Figure 4.5 and Figure 4.6.
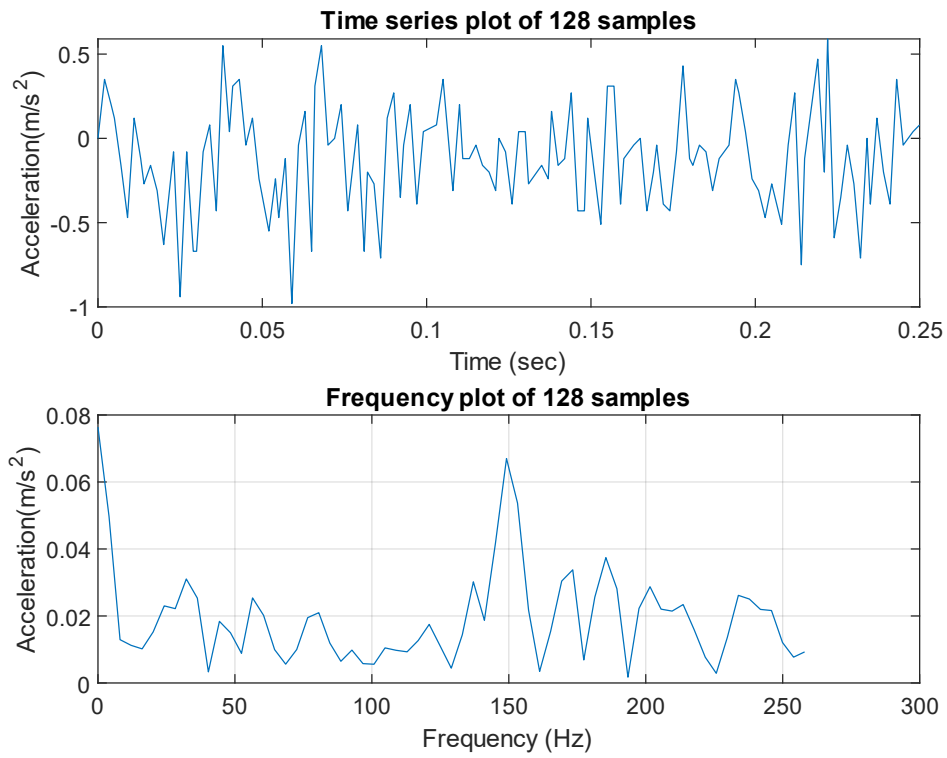
Figure 4.4: Time domain and frequency domain plot of 128 samples of channel_00 (x-axis).
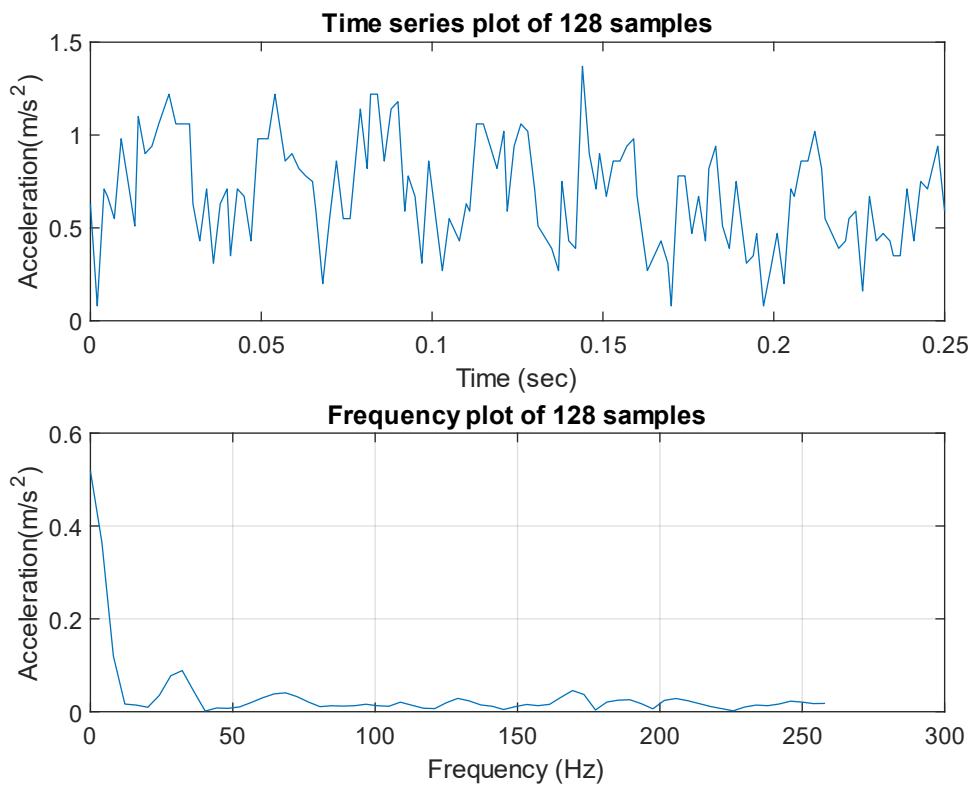
Figure 4.5: Time domain and frequency domain plot of 128 samples of channel_01 (y-axis).

In Figure 4.5 the frequency domain plot shows the signal containing frequency below 50Hz with low amplitudes whereas, in Figure 4.6, it seems that the signal of 150Hz with high amplitudes is presented. Now for multivariate data analysis, only one channel (x-axis) data is used which will be discussed in the next topic.

The frequency-domain plot in Figure 4.5 reveals that the signal contains frequencies below 50Hz with low amplitudes, but in Figure 4.6, the signal appears to be 150Hz with significant amplitudes. Only one channel (x-axis) of data is used for multivariate data analysis, which will be described in the next topic.
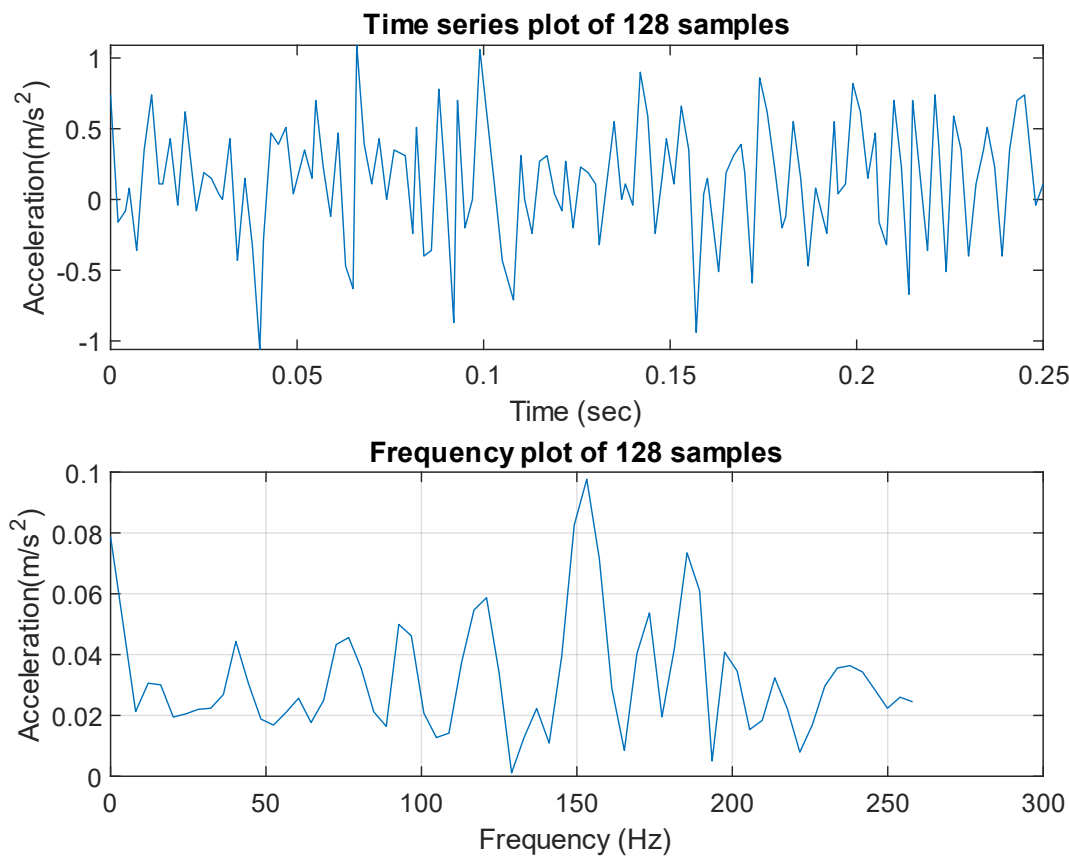
Figure 4.6: Time domain and frequency domain plot of 128 samples of channel_02 (z-axis).

4.2 Multivariate data analysis of FFT data

The variance of variables was first checked using the matrix produced from the FFT, as illustrated in Figure 4.7. In comparison to other variables, the variables from approximately 1 to 11 have a considerable variance. As a result, the data should be scaled to improve the analysis results. Then PCA was used to select important samples from the score plot for fault detection, as shown in Figure 4.8.
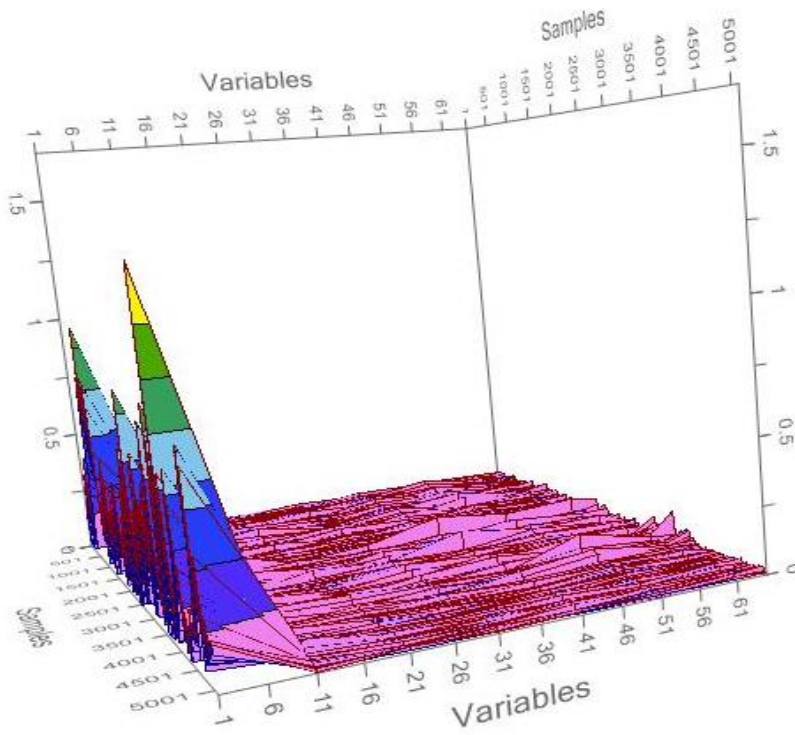
Figure 4.7: Matrix plot of FFT data.
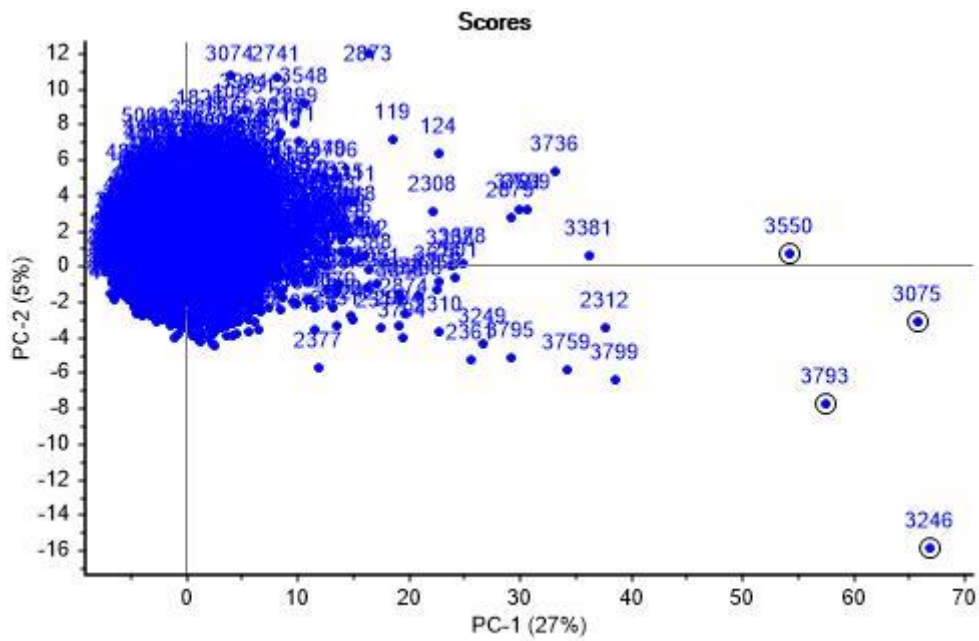
Figure 4.8: Score plot of samples in 2D scatter form.

The score plot, shown in Figure 4.8, depicts how the objects are related to one another. The sample index numbers 3075, 3246, 3550, and 3793 were chosen as crucial samples since they are far apart from other objects, making them the model with output as a fault (1) and non-fault (0) in the track. The model was then created using PLS using input (X) and output (Y). Figure 4.9 depicts the model's calibration, with a predicted vs reference plot.
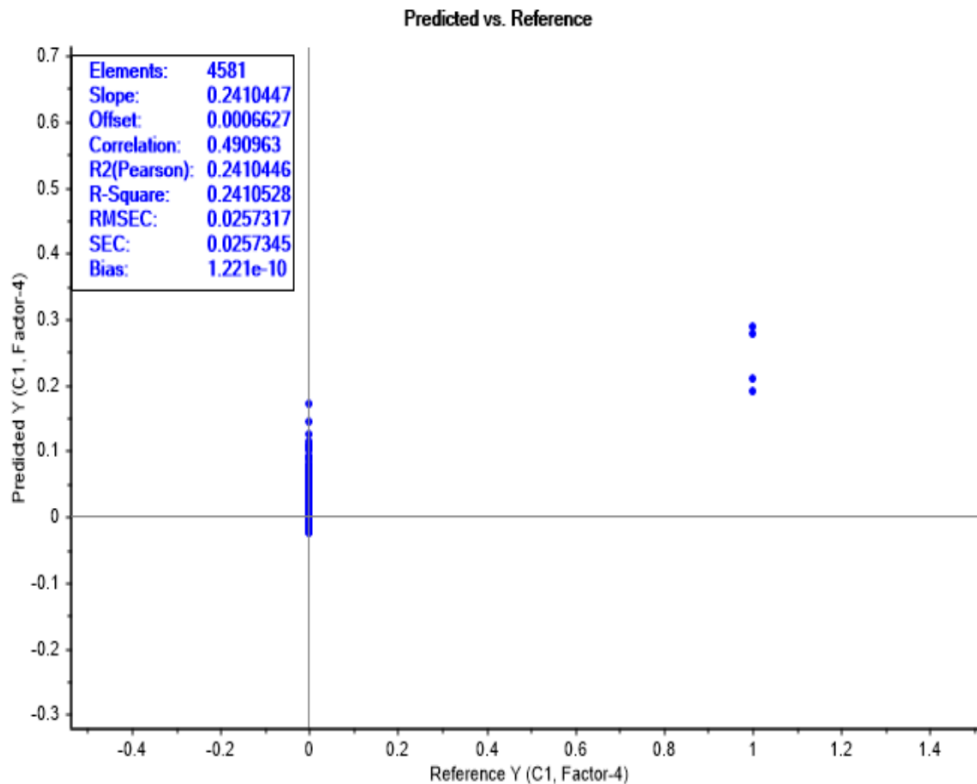
Figure 4.9: Predicted versus reference plot of calibration.

At the calibration stage, the predicted value was compared to the reference value (0 or 1) as in Figure 4.9. Root Mean Square Error of Calibration (RMSEC) of roughly 0.02573 is used to fit the given data and calibration model. The slope is just 0.2410 in this case, which could be due to the reference only having a 0 or 1 value.

A test set with a different data matrix was used to validate the model's performance. The PCA was used to select important samples, as shown in Figure 4.10. The sample numbers 3465, 3683, and 3684 were selected to create an output matrix for the validation test set matrix. Figure 4.11 shows the model validation with the predicted versus reference.

Figure 4.10: Score plot of test set data.

Figure 4.11: Predicted versus reference plot of validation.

Figure 4.11 shows that the Root Mean Square Error of Prediction (RMSEP) is 0.02469 indicating that the model's performance is significant.

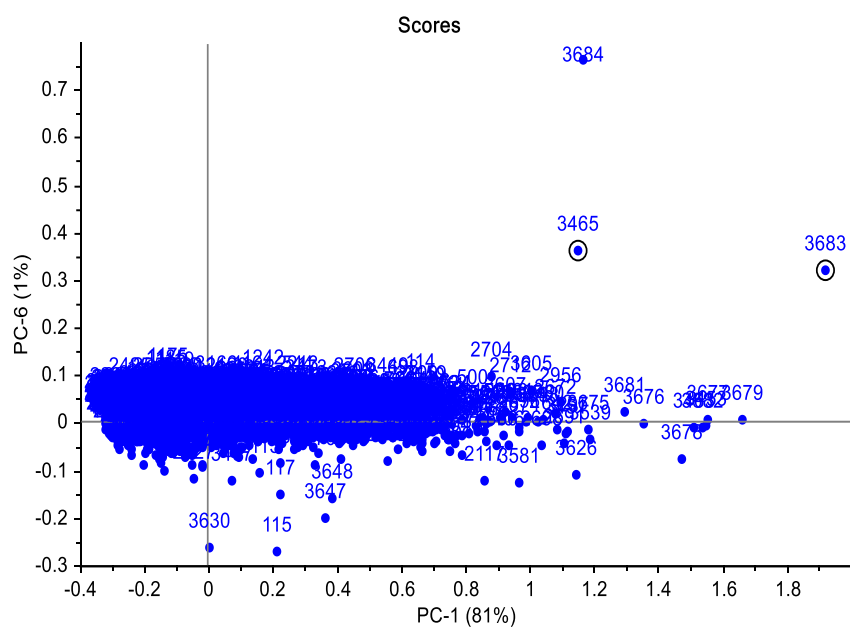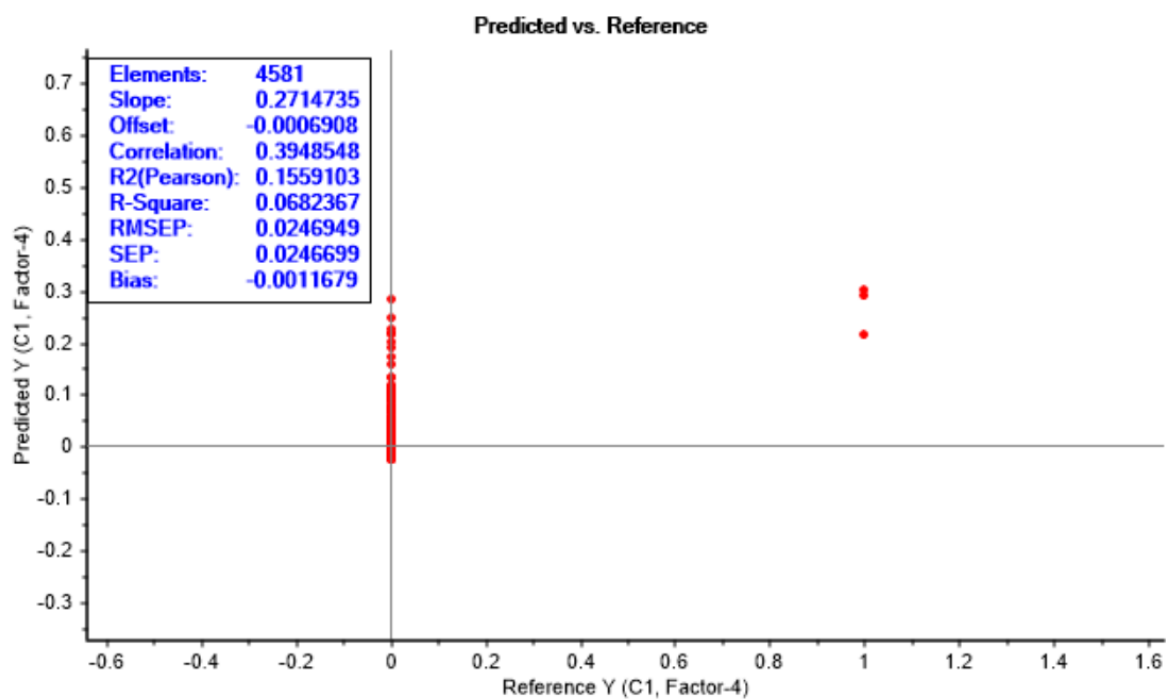|	Y Reference	Y Predicted		Y Reference	Y Predicted
	1	2		1	2
3452	0.0000	0.0900	3674	0.0000	0.0705
3453	0.0000	0.0298	3675	0.0000	0.0089
3454	0.0000	0.0522	3676	0.0000	0.0515
3455	0.0000	0.0172	3677	0.0000	-0.0013
3456	0.0000	0.0103	3678	0.0000	0.0745
3457	0.0000	0.0447	3679	0.0000	-0.0139
3458	0.0000	0.0169	3680	0.0000	0.1073
3459	0.0000	0.0489	3681	0.0000	0.0098
3460	0.0000	0.0157	3682	0.0000	0.0286
3461	0.0000	0.0135	3683	1.0000	0.2915
3462	0.0000	0.0202	3684	1.0000	0.2174
3463	0.0000	0.0041	3685	0.0000	0.0296
3464	0.0000	0.1321	3686	0.0000	0.2008
3465	1.0000	0.3035	3687	0.0000	0.1343
3466	0.0000	0.0080	3688	0.0000	0.0453
3467	0.0000	0.0996	3689	0.0000	0.0029
3468	0.0000	0.0172	3690	0.0000	-0.0146
3469	0.0000	0.0173	3691	0.0000	0.0172
3470	0.0000	0.0993	3692	0.0000	-0.0070
3471	0.0000	0.0207	3693	0.0000	0.0083

Figure 4.12 : Comparison between reference and prediction of validation

Figure 4.12, Y predicted is almost equal to the Y reference, hence the model performance can be significant. The model can be made more accurate by analyzing more data sets and making a better output matrix.

4.3 Location of fault in the track

The train operates from Porsgrunn to Brevik. The distance between the start and stop of the train is 9.73 km. Based on the score plot of the samples shown in Figure 4.13, it is of interest to evaluate the fault locations between Porsgrunn and Brevik.

The fault location can be evaluated by knowing the timing of the train for the samples given by

$$Time = time\ step * index\ of\ the\ sample \qquad \text{4-3}$$

The timing for different outliers as shown in Figure 4.13 can, then, be calculated as

Timing of sample number 3075 = 0.256 sec * 3075 = 787.2 sec

Timing of sample number 3246 = 0.256 sec * 3246 = 830.976 sec

Timing of sample number 3550 = 0.256 sec * 3550 = 908.8 sec

Timing of sample number 3793 = 0.256 sec * 3793 = 971.008 sec

Next, comparing the timing for the fault locations with the respective distances in the data provided by CEMIT, the distances of the fault locations from the starting point (Porsgrunn) are given as,

Distance of sample number 3075 = 5.91 Km

Distance of sample number 3246 = 6.49 Km

Distance of sample number 3550 = 7.40 Km

Distance of sample number 3793 = 7.99 Km

The fault locations can be shown in
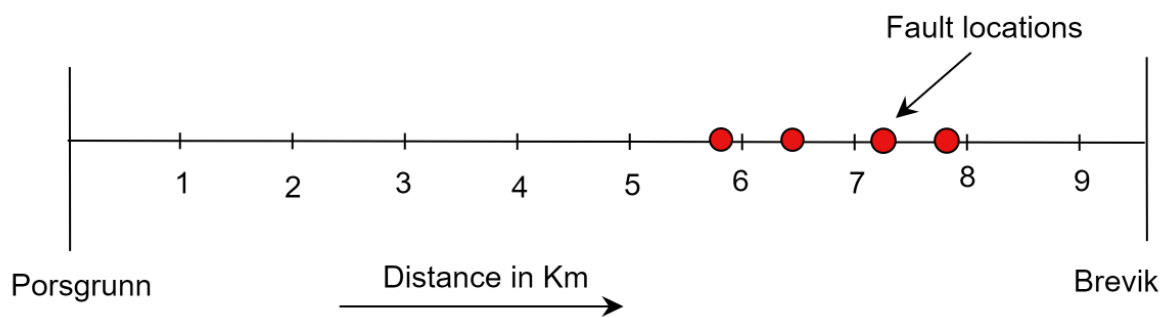
Figure 4.13: The location of faults in the train track.

4.4 Wavelet transform (WT)

Another method of time-series analysis was chosen as WT which decomposed the signal into detailed (HF) and approximation (LF) coefficients as shown in Figure 4.14. These decomposed frequencies were used to make a matrix (5102*160) for further analysis through multivariate data analysis. The detail coefficient from level 1 to level 5 shows the high-frequency components and the approximation at level 5 is the low-frequency component. These frequencies are chosen as the features of data and used for analysis which will be discussed in the next topic.
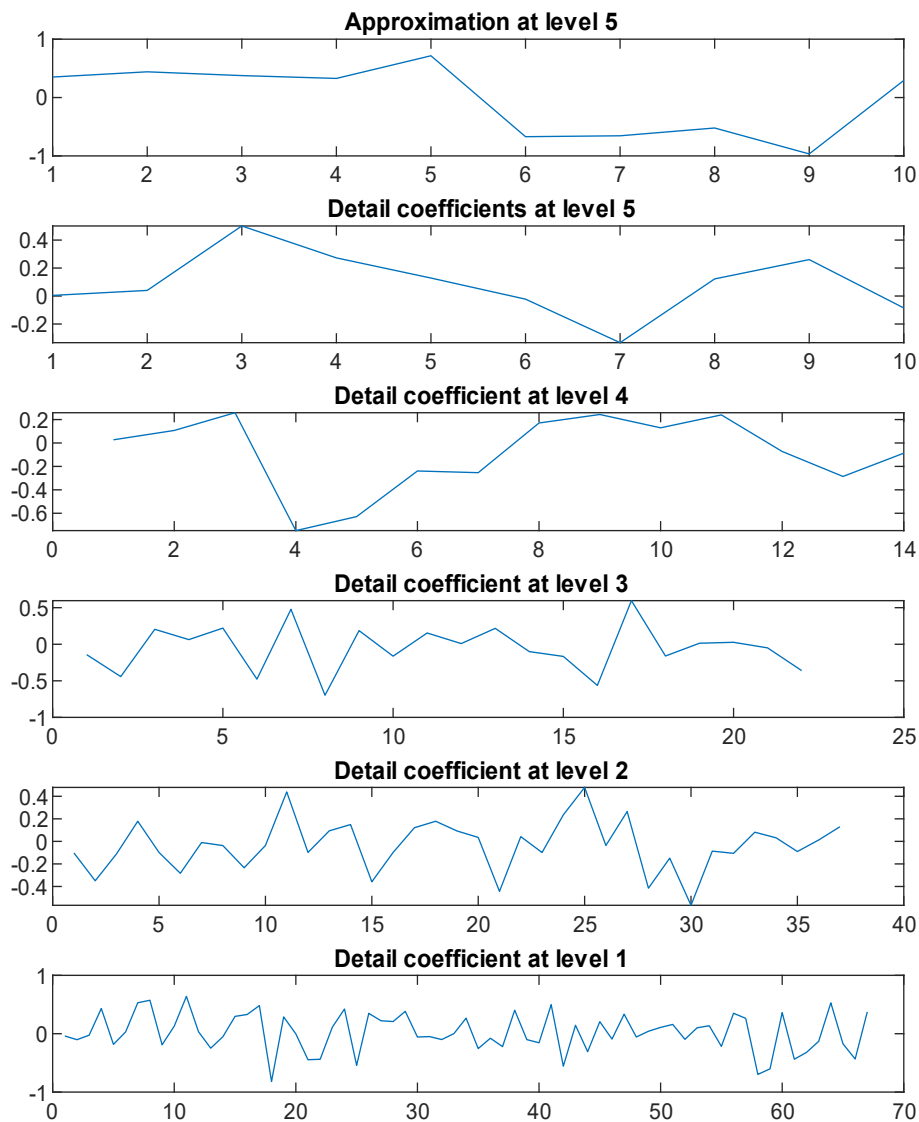
Figure 4.14: Decomposition of time series signal (128 samples) into low frequency (LF) and high frequency (HF) components.

4.5 Multivariate data analysis of WT data

To discover relevant samples, the PCA of WT data (5102*160) was used. The score plot of WT data is shown in Figure 4.15 which depicts objects are spread all over and there are no significant samples to choose from for further analysis.
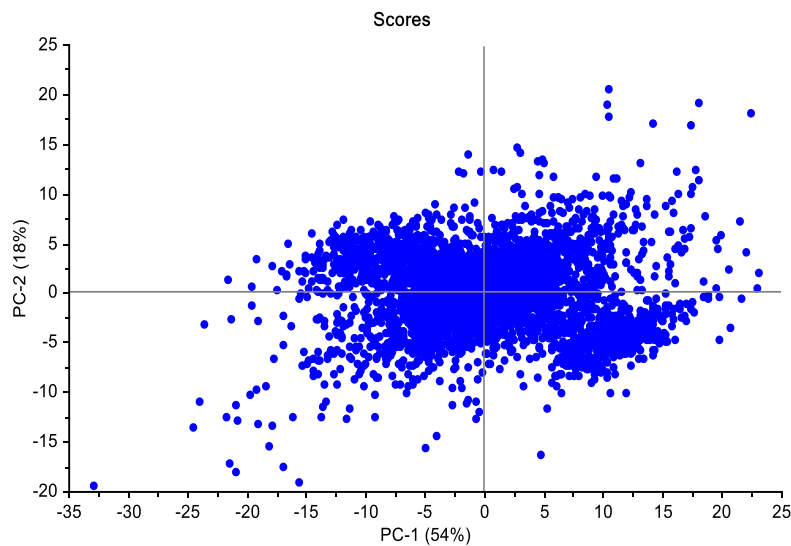
Figure 4.15: The score plot of WT data.

4.6 Discussion on the number of channels

For multivariate data analysis, only one channel (channel_00/ x-axis) was used in this research. Only three Principal Components (PC1, PC2, PC3) were found in the PCA of one channel, implying that only three components were available to accommodate the enormous number of samples. Other channels (y-axis, z-axis) can be used to see the outcomes of the analysis. Additional information about the data was accessible if many sensors (more channels) were positioned in different places of the train. This can aid in the creation of a useful output matrix for subsequent analysis. As a result, multi-channel or more sensor data are recommended.

4.7 Discussion on calibration and validation stage

For the calibration, the input matrix ($X_{5102*65}$) and output matrix (Y_{5102*1}) with four faults were used. From Figure 4.9, by analyzing the RMSEC value alone which is 0.02458. It means the calibration model is working fine but when looking at the slope (0.2284) which is very low (for a perfect model, the slope should be nearly equal to 1). This may arise because of output matrix contains only 0 or 1 binary numbers. Similarly, for validation, a test set data was used which was taken from another data set with the same direction running of the train. A test set matrix with ($X_{4581*65}$) and output matrix (Y_{4581*1}) with three faults. From Figure 4.11 the RMSEP is 0.0239 and the slope is 0.2481, which is also very low. So, it is also a topic for further analysis.

4.8 Discussion on multivariate data analysis of FFT and WT data

From the multivariate data analysis of FFT data, some interesting results were detected to analyze as discussed in section 4.7. However, from WT data after plotting the score plot (see Figure 4.15), no interesting samples were detected for further analysis. This means WT data are not useful for detecting faults in tracks using multivariate data analysis.

5 Conclusion

It is critical to implement good railway network maintenance methods to avoid service disruptions and maintain system safety. The purpose of the project is to explore if railway track and wheel conditions can be analyzed using time series and multivariate data analysis to time series data.

The FFT and WT were used to convert the time series data (from CDC) into the frequency domain. PCA and PLSR were used to do multivariate data analysis on FFT and WT data for fault detection. By comparing the score plot of FFT and WT data, it is concluded that the WT data is not beneficial for finding faults in the tracks using multivariate data analysis, because no interesting samples were discovered for further examination. Furthermore, it is advised to analyze multi-channel data to obtain better results.

6 Future works

In this study, the time-series and multivariate data analysis were done to detect whether there is a fault or not in the track. This can be further proceeded to make a real-time monitoring system. Some work can be done to improve further as in the following points.

➢ In this study, the analysis was done to detect the train tracks only. So, for the wheels fault detection, its geometry should be known, which can be the next step of the analysis.

➢ GPS with more frequency range can be used for tracking at a small time stamp, where only 5Hz GPS was used in this research.

➢ The IoT-based system can be implemented for real-time monitoring.

References

[1] A. Malekjafarian, E. OBrien, P. Quirke, and C. Bowe, 'Railway Track Monitoring Using Train Measurements: An Experimental Case Study', *Appl. Sci.*, vol. 9, no. 22, p. 4859, Nov. 2019, doi: 10.3390/app9224859.

[2] M. J. Gómez, C. Castejón, E. Corral, and J. C. García-Prada, 'Railway Axle Condition Monitoring Technique Based on Wavelet Packet Transform Features and Support Vector Machines', *Sensors*, vol. 20, no. 12, p. 3575, Jun. 2020, doi: 10.3390/s20123575.

[3] E. M. Vinberg, M. Martin, Alfi Hadi Firdaus, Yihao Tang, and Alireza Qazizadeh, 'Railway Applications of Condition Monitoring', 2018, doi: 10.13140/RG.2.2.35912.62729.

[4] S. Fallah Nafari, M. Gül, A. Roghani, M. T. Hendry, and J. R. Cheng, 'Evaluating the potential of a rolling deflection measurement system to estimate track modulus', *Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit*, vol. 232, no. 1, pp. 14–24, Jan. 2018, doi: 10.1177/0954409716646404.

[5] M. Bar-Am and Z. Yaakov, '(54) ON-TRAIN RAIL TRACK MONITORING', p. 10.

[6] B. A. Fosburgh, '(54) RAILWAY TRACK MONITORING', p. 19.

[7] E. J. OBrien, P. Quirke, C. Bowe, and D. Cantero, 'Determination of railway track longitudinal profile using measured inertial response of an in-service railway vehicle', *Struct. Health Monit.*, vol. 17, no. 6, pp. 1425–1440, Nov. 2018, doi: 10.1177/1475921717744479.

[8] R. Shafique *et al.*, 'A Novel Approach to Railway Track Faults Detection Using Acoustic Analysis', *Sensors*, vol. 21, no. 18, p. 6221, Sep. 2021, doi: 10.3390/s21186221.

[9] Y. Min, B. Xiao, J. Dang, B. Yue, and T. Cheng, 'Real time detection system for rail surface defects based on machine vision', *EURASIP J. Image Video Process.*, vol. 2018, no. 1, p. 3, Dec. 2018, doi: 10.1186/s13640-017-0241-y.

[10] M. J. Gómez, C. Castejón, and J. C. García-Prada, 'Crack detection in rotating shafts based on 3× energy: Analytical and experimental analyses', *Mech. Mach. Theory*, vol. 96, pp. 94–106, Feb. 2016, doi: 10.1016/j.mechmachtheory.2015.09.009.

[11] M. J. Gómez, C. Castejón, E. Corral, and J. C. García-Prada, 'Analysis of the influence of crack location for diagnosis in rotating shafts based on 3 x energy', *Mech. Mach. Theory*, vol. 103, pp. 167–173, Sep. 2016, doi: 10.1016/j.mechmachtheory.2016.05.006.

[12] M. J. Gómez, C. Castejón, and J. C. García-Prada, 'Automatic condition monitoring system for crack detection in rotating machinery', *Reliab. Eng. Syst. Saf.*, vol. 152, pp. 239–247, Aug. 2016, doi: 10.1016/j.ress.2016.03.013.

[13] M. Gómez, E. Corral, C. Castejón, and J. García-Prada, 'Effective Crack Detection in Railway Axles Using Vibration Signals and WPT Energy', *Sensors*, vol. 18, no. 5, p. 1603, May 2018, doi: 10.3390/s18051603.

[14] G. Guohua, Z. Yu, D. Guanghuang, and Z. Yongzhong, 'Intelligent Fault Identification Based On Wavelet Packet Energy Analysis and SVM', in *2006 9th*

International Conference on Control, Automation, Robotics and Vision, Singapore, 2006, pp. 1–5. doi: 10.1109/ICARCV.2006.345306.

[15]   X. Wang, Z. Lu, J. Wei, and Y. Zhang, 'Fault Diagnosis for Rail Vehicle Axle-Box Bearings Based on Energy Feature Reconstruction and Composite Multiscale Permutation Entropy', *Entropy*, vol. 21, no. 9, p. 865, Sep. 2019, doi: 10.3390/e21090865.

[16]   F. Czerwinski and L. B. Oddershede, 'TimeSeriesStreaming.vi: LabVIEW program for reliable data streaming of large analog time series', *Comput. Phys. Commun.*, vol. 182, no. 2, pp. 485–489, Feb. 2011, doi: 10.1016/j.cpc.2010.10.019.

[17]   'Time Series Analysis: Definition, Types, Techniques, and When It's Used', *Tableau*. https://www.tableau.com/learn/articles/time-series-analysis (accessed May 02, 2022).

[18]   W. T. Cochran *et al.*, 'What is the fast Fourier transform?', *Proc. IEEE*, vol. 55, no. 10, pp. 1664–1674, 1967, doi: 10.1109/PROC.1967.5957.

[19]   'FFT'. https://www.nti-audio.com/en/support/know-how/fast-fourier-transform-fft (accessed May 03, 2022).

[20]   H. Tsunashima, Y. Naganuma, A. Matsumoto, T. Mizuma, and H. Mori, 'Condition Monitoring of Railway Track Using In-Service Vehicle', *Reliab. Saf. Railw.*, p. 25.

[21]   B. Dutta, 'What is Multivariate Data Analysis? | Analytics Steps'. https://www.analyticssteps.com/blogs/what-multivariate-data-analysis (accessed May 13, 2022).

[22]   K. H. Esbensen, *Multivariate data analysis - in practice: an introduction to multivariate data analysis and experimental design*, 5th ed., Repr. Oslo: CAMO, 2012.

[23]   'ASC OS-315LN-PG - ASC Sensors'. https://www.asc-sensors.de/sensoren/asc-os-315ln-pg/ (accessed May 16, 2022).

Appendices

Appendix A Master thesis description.

University of South-Eastern Norway

Faculty of Technology, Natural Sciences and Maritime Sciences, Campus Porsgrunn

FMH606 Master's Thesis

Title: Real time monitoring of train wheels and track conditions based on time series analysis and multivariate data analysis

USN supervisor: Maths Halstensen

External partner: Per Ivan Januschas, CEMIT Digital Porsgrunn, Norway

Task background:

CEMIT Research/Product Development Project

Broken rails

Bane Nor intends to roll out the European Rail Transport Management System (ERTMS) in Norway during the period of 2021-2034. This, in short, means a transition to a digital signal system with no further use of the existing signaling system. The information will be sent directly to the train drivers console inside the train via a dedicated GSM-R system (GSM-R is a mobile phone system reserved for communication within the rail industry).
This means that the principle of a closed circuit, which used in today's existing signaling system will no longer be available to analyze the track for breakage in the rails.
CEMIT wants to look at whether it is acceptable to investigate this using acoustic chemometrics. We believe that this is a good potential for this technology to detect and predict the distance between a train and a rail break, and where acoustic chemometrics will be used to investigate if there is actually a break in the track by analyzing the actual sound emitted as a result of a train wheel rolling on the track.

«Screaming rails»

CEMIT also has proposed a theory that acoustic chemometrics can be used to look at causal relationships that cause rolling stock to create "wailing rails" in some cases. In tight turns, this will be "sort of normal" since the friction between the flange of the wheel and the rail increases at the entrance to a sharp curve. But what could be the reason why the train makes such noise even in wider curves? CEMIT wants to look at the total data that such a soundscape creates to see if something stands out, and whether acoustic chemometrics are suitable for analyzing these relationships. We will therefore look at whether we can isolate deviations in rail/train wheels that together produce this undesired soundscape. This can have a number of different causes, but e.g. bogie geometry that is not 100% aligned with the rail aisle may cause the wheels to constantly be left at the wrong angle to flange, thereby increasing wear and tear, or the wheel axle may be subjected to geometric travel, which in addition will cause vibrations in the train that will cause negative effects for the infrastructure, rolling stock and passengers.

Rolling stock wheels wear & tear

Furthermore, CEMIT recognizes that a train wheel in use wears out. Therefore, over time, the circumference of the wheel decreases. There are set criteria for the permissible minimum and maximum wheel diameter. At CEMIT we have therefore formulated a hypothesis that when the wheel wears down, the frequency and amplitude height will change quite noticeably. If this entails any level of correctness, CEMIT also hypothesizes that one will be able to real-time analyze this sound, and thus have a good data basis for predicting the remaining wheel life up to the next overhaul of the wheel, and likewise the remaining lifetime in total. The section on "screaming rails" mentions several of the causes of wear and tear, and also several of the consequences this may lead to.

Damage to/breakage in wheel/material fatigue/un-roundness/wheel profile/wear & tear og bearings and brakes

In the same way that a wheel that wears gradually will change its frequency, this change is likely to occur if the wheel is subjected to other changes over time, especially with sudden changes. CEMIT has formed a hypothesis that all these various influences on the wheels will have an acoustic "overall picture" and that by analyzing the sound these conditions will create, we will be able to distinguish them, isolating each individual condition. With this as a starting point, as a theory it will only be our imagination that limits what we can find. This will require an effort in the way of mathematical modelling, but not to the extreme in the context of engineering.

Other potential applications

- Listening to power loads drawn at the railway switching machines.
- Possible tools REW Room Equalizing Wizard. Real Time Analyzer.
- Potential source of information: GKN – They are using has predictive maintenance with the extraction of vibrations and acoustics for predictive analysis
- More information can be found by searching for "predictive analysis and maintenance"

It is aspired that good results based on this project will lead to peer-reviewed publication in an academic journal and the student group will be co-authors of the publication.

<u>**Task description:**</u>

1. **Data acquisition:** Develop a DAQ system including software for acquisition (stream to disk) of continuous time series from multiple sensors. The system should be based on National Instruments hardware and software.
2. **Feature extraction:** Develop software for signal processing of the recorded time series data involving identification and implementation of relevant signal processing methods for analysis on train vibrations.
3. **Multivariate Data Analysis and modelling:** Calibrate multivariate regression models for e.g. prediction, detection, "early warning", ... etc.

<u>**Student category**</u>: Reserved for IIA student Santosh Dangal

<u>**Is te task suitable for online students (not present at the campus)?**</u> No

Practical arrangements: N/A

Supervision:
As a general rule, the student is entitled to 15-20 hours of supervision. This includes necessary time for the supervisor to prepare for supervision meetings (reading material to be discussed, etc).

Signatures:

Supervisor (date and signature):   02.02 2022 Maths Halstensen

Student (write clearly in all capitalized letters): SANTOSH DANUJAL

Student (date and signature): 01.02.2022

Appendix B

MATLAB FILE

A list of all MATLAB scripts used to produce the results of this thesis is given in Table B.1.

Table B.1: A list of files and their description.

File name	Description
ff_window.m	Script used to do Fast Fourier transform and stacks results vertically.
ff_plot.m	Script used to plot the time domain and frequency domain of 128 samples.
wavelet_decom.m	Script used to do Wavelet transform and stacks results vertically.
wavelet_plot.m	Script used to plot the decomposed frequencies from wavelet transform.

ff_window.m

```matlab
close all;
clear all;
clc;
 % Load file
[d,s] = xlsread('cdc1.xlsx');

 % Time vector
t = d(:,1);

 % acceleration data in time domain
v = d(:,2);

% Number of samples
L = length(t);
a=1;
b=128;

%interval
N= 128;

c=1;
while b<L


    % sampling time interval
    Ts = mean(diff(t));

    % sampling frequency
    Fs = 1/Ts;

    % Nyquist frequency
    Fn = Fs/2;

    % chosing number of samples from acceleration data
    vc=v(a:b);

    % windowing with normalized Fast Fourier Transform
    FTv = fft(vc.*blackmanharris(length(vc)))/N;

    % Frequency vector
    Fv = linspace(0, 1, fix(N/2)+1)*Fn;

    % Index vector
    Iv = 1:length(Fv);
    % absolute value
    y= abs(FTv(Iv))*2;

    %figure(c)
    %plot(Fv,y)

    if c==1
        abc=y';
```

```
    else
        abc=cat(1,abc,y');
    end
    %grid
    %xlabel('Frequency (Hz)')
    %ylabel('Amplitude (V)')
    c=c+1;
    a=a+interval;
    b=b+interval;

end
```

ff_plot

```
close all
clear all
clc
[d,s] = xlsread('cdc1.xlsx');
t = d(1:130,1);
v = d(1:130,2);
L = length(t);
N=128;
a=1;
b=128;
Ts = mean(diff(t));
Fs = 1/Ts;
Fn = Fs/2;
vc=v(a:b);
FTv = fft(vc.*blackmanharris(length(vc)))/N;
Fv = linspace(0, 1, fix(N/2)+1)*Fn;
Iv = 1:length(Fv);
y= abs(FTv(Iv))*2;
subplot(2,1,1)
plot(t,v)
title('Time series plot of 128 samples');
xlabel('Time (sec)')
ylabel('Acceleration(m/s^2)')
subplot(2,1,2)
plot(Fv,y)
title('Frequency plot of 128 samples')
```

wavelet_decom.m

```matlab
clc;
clear all;
close all;
[d,s] = xlsread('cdc1.xlsx');
t = d(:,1);
v = d(:,2);
a=1;
b=128;
L =length(t);
while b<L
    vc=v(a:b,:);
    %wavelet decomposition at level 5
    [c,l]= wavedec(vc,5,'db4');
    % low-frequency component (approximation)
    approximation =appcoef(c,l,'db4');
    % high-frequency components (details)
    [d1,d2,d3,d4,d5]=detcoef(c,l,[1 2 3 4 5]);
    data = cat(1,d1,d2,d3,d4,d5,approximation)';
    if b==128
        data1=data;
    else
        data1=cat(1,data1,data);
    end
    a=b+1;
    b=b+128;

end
```

wave_plot.m

```matlab
clc;
clear all;
close all;
[d,s] = xlsread('cdc1.xlsx');
t = d(1:130,1);
v = d(1:130,2);
L = length(t);
a=1;
b=128;
while b<L
    vc=v(a:b,:);
    [c,l]= wavedec(vc,5,'db4');
    approximation =appcoef(c,l,'db4');
    [d1,d2,d3,d4,d5]=detcoef(c,l,[1 2 3 4 5]);
    subplot(6,1,1)
    %plot(approximation);
    title('Approximation at level 5');
```

```matlab
%subplot(6,1,2)
%plot(d5)
title('Detail coefficients at level 5');
subplot(6,1,3)
plot(d4)
title('Detail coefficient at level 4');
subplot(6,1,4)
plot(d3)
title('Detail coefficient at level 3');
subplot(6,1,5)
plot(d2)
title('Detail coefficient at level 2');
subplot(6,1,6)
plot(d1)
title('Detail coefficient at level 1');
data = cat(1,d1,d2,d3,d4,d5,approximation)';
if b==128
    data1=data;
else
    data1=cat(1,data1,data);
end
a=b+1;
b=b+128;

end
```