Wayside Railway Switch and Crossing Monitoring Using Isolation Forest Anomaly Scores

Yang Zuo *, Praneeth Chandran ©, Johan Odelius © and Matti Rantatalo ©

Abstract: Railway switch and crossing (S&C) systems have complicated moving structures compared with regular rail. They require multiple components that vary in complexity. The complexity of railway S&C, together with the fact that they are discontinuous points of the system, makes them vulnerable to defects such as squats. A squat on the switching rail could potentially cause rail breakage and lead to catastrophic results, such as derailment. In this study, a method based on anomaly scoring was investigated to estimate the status of an S&C system with respect to squat defects. The proposed method was tested in a real environment under controlled measurement sequences. The results show that the methods can differ between an S&C with squats and another one without them.

Keywords: railway; anomaly detection; anomaly score; rail squat; squat detection; machine learning; unsupervised learning; isolation forest

1. Introduction

Rail transportation is one of the most important modes of transportation, providing a safe and sustainable means of transporting goods and people across vast distances. It is also a vital component of the global economy, enabling the movement of goods and connecting businesses and consumers across different regions.

The safe and efficient operation of railway systems relies heavily on the condition of the track infrastructure, including the switches and crossings (S&Cs) that allow trains to change tracks or direction [1]. However, the failure of S&C components can lead to derailments, delays, and accidents, resulting in significant safety risks and economic losses [2]. Therefore, timely detection of S&C defects is crucial to ensure the safe and reliable operation of railway systems.

Several studies have demonstrated the importance of railways in the global economy and the transportation industry. For instance, a report by Meżyk et al. [3] highlighted the role of railways in reducing greenhouse gas emissions and promoting sustainable transportation. A study by Donaldson [4] highlighted the economic benefits of investing in railway infrastructure, such as increased trade, job creation, and reduced transport costs. Moreover, several studies have emphasised the importance of S&C in railway operations. For instance, a report by the Trafikverket in Sweden highlighted the role of S&C in ensuring the safe and efficient operation of railway systems [5]. A study by Naybour [6] emphasized the importance of S&C maintenance and inspection in reducing the risk of derailments and accidents.

Various techniques have been explored in the previous literature for S&C defect detection. Manual on-site inspections of rail and S&C are one of the oldest methods. They have several weaknesses, such as being time-consuming, labour-intensive, prone to human error, subjective, and challenging to detect defects in real-time. These weaknesses emphasize the need for more advanced and efficient methods of rail and S&C inspection to improve safety...
and reliability. The techniques include visual inspection [7–9], ultrasonic testing [10–12], use of the eddy current testing (ECT) system [13,14], and measuring vibrations [15–17]. However, all these current methods have limitations of one kind or another. The detection capability of visual inspection is constrained to solely identifying the defects on the surface of the rail that are not concealed by snow or any other obstructions. Furthermore, the accuracy of detection can be impeded by the presence of dust, dirt, or water on the lens. The method using an ultrasonic sensor is susceptible to technical difficulties arising from the operating or environmental conditions, which impairs its effectiveness. Moreover, this technology encounters obstacles in identifying defects that are located close to the surface [18]. The ECT method’s primary hurdle is the lift-off effect, which impacts the ECT signal and leads to inaccurate data interpretation [19]. The majority of the studies using vibration data focus on using axle box acceleration (ABA) or bogie vibration (BA). Rather than being suitable for monitoring an S&C, both of these methods are more appropriate for inspecting a significant length of the rail since the measuring point is on the train [20]. For example, if not all trains have measuring systems installed, defects with faster failure development could be missed while trains with no measurement system are occupying the track. By introducing sensors in the S&C, monitoring of the S&C for each train passage could be achieved.

Railway squat is a common type of rail defect that occurs in railway infrastructure and can lead to serious consequences, such as derailment and accidents. With the rapid development of technology, many researchers have explored various methods for automatic squat detection in the railway infrastructure. As Grosonni et al. [21] stated, approximately one-third of the crossing panel failures they documented were associated with squats. One popular means of squat detection is to use algorithms developed to detect such defects from videos/images captured during visual inspection. The Gabor filter bank with SVM classifier [22], fuzzy key performance indicators [23], and the Gaussian mixture model [24] are a few alternatives. The detection capability of the visual-based technique is constrained to identify only surface defects. Obstacles such as snow in the rail and dust, dirt, or water on the lens would impede the accuracy. Another approach for squat detection involves laser measurements. In one study, Lesiak et al. simplified a technique for detecting squat defects using laser scatterometry, which was able to detect both artificial and real squats [25]. De Becker et al. [26] implemented the covariance method and the normal intersection method to identify the defects and their locations. Nonetheless, the laser transmitter experiences comparable issues to those encountered by the camera lens. Another method involves the use of lock-in thermography to detect squats on the rail steel head [27,28]. The infrared camera encounters similar issues to those experienced by the camera lens and laser transmitter.

Another popular approach for squat detection is the use of accelerometers installed on trains to measure the ABA. An automatic squat detection algorithm was proposed by Molodova et al. [29]. They also performed a parametric study to understand the relationships of different parameters with the measured ABA signal [30]. Li et al. improved the algorithm to enable the detection of light squats [15]. Yin et al. [31] highlighted several limitations associated with ABA measurements. These drawbacks include the susceptibility of the ABA signal to various factors, such as the specific type of axle box, the condition of the wheel axle bearings, and even variations in the wheel profile. Moreover, it is worth noting that vibrations originating from both the bearing itself and any potential bearing defects have the potential to contaminate the ABA measurements. Wei et al. [17] proposed using an empirical mode decomposition method to detect squats and corrugation by analysing the bogie acceleration (BA). The drawback is that if not all trains have the ABA measurement systems installed, defects with faster failure development time could be missed while the trains with no measurement system occupy the track. On the other hand, by introducing vibration sensors mounted in the S&C, monitoring of the S&C for each train passage could be performed [32].
Timely detection of S&C squats is crucial for ensuring the safe and reliable operation of railway systems. Several studies have been performed to explore the feasibility of measuring the vibrations from an accelerometer mounted on the rod of a point machine [20,32]. These studies provide promising results. This study aims to further verify the results by applying the method with a locomotive at normal speed in a semi-controlled real environment. To verify previous laboratory experiments, a track section with two S&Cs, one old and one new, was instrumented with an accelerometer, and passages of a single locomotive were analysed by applying the previously developed method [20] for squat detection. Applying such a monitoring method on a large scale could potentially reduce the possibility of derailment accidents at the S&C and also lower the costs for S&C-related maintenance costs. In other words, this method could make the operation and maintenance of railways more sustainable.

Following the Introduction, Section 2 gives an overview of the used method and the working process, while Section 3 presents the results obtained from analysing the data and provides discussion of the results. Finally, Section 4 summarizes the findings of the study and proposes possible future studies.

2. Methods

2.1. Test Setup

The experiment was carried out in a real environment with two S&Cs that were 40 and 3 years old, respectively. The locomotive used was a Z67 manufactured by Hägglunds and the distance between the centers of the two axles was 4 m. The locomotive used in the field test is shown in Figure 1 and the field test site is shown in Figure 2. The old S&C is illustrated in Figure 3 and the new one in Figure 4. In these two figures, one side of the S&C is named \( S_0 \) and the other side \( S_1 \), as illustrated. The squats in the old S&C are illustrated in Figure 5. The dimensions of the squats are listed in detail in Table 1.

![Figure 1. Locomotive Z67 used for the field test.](image)

<table>
<thead>
<tr>
<th>Squat Name</th>
<th>Max Width (x-Direction) (mm)</th>
<th>Max Width (y-Direction) (mm)</th>
<th>Max Depth (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>22</td>
<td>11</td>
<td>0.6</td>
</tr>
<tr>
<td>B</td>
<td>18</td>
<td>12</td>
<td>1.3</td>
</tr>
<tr>
<td>C</td>
<td>10</td>
<td>15</td>
<td>0.8</td>
</tr>
<tr>
<td>D</td>
<td>44</td>
<td>19</td>
<td>1.6</td>
</tr>
<tr>
<td>E</td>
<td>27</td>
<td>34</td>
<td>0.8</td>
</tr>
</tbody>
</table>
The accelerometer used in the experiment was of the PCB 608A11 type. It was mounted onto the rail at a distance of 16 cm to the right of the center of the point machine using a magnet. Figure 6 shows how the accelerometer was mounted in the field test. The accelerometer was positioned to measure the vibrations in the z-direction. The main parameters of the accelerometer are listed in Table 2. For each S&C, the locomotive was driven from $S_0$ to $S_1$ via the through direction of the S&C, twice for each test scenario. The vibration data were recorded through a traditional data acquisition system (DAQ) DAQ9174. The captured data were directly stored on a computer. The measuring system operated with a sampling frequency of 10,240 Hz.
Figure 5. The rail squats on the 40-year-old S&C. The locations of the squats A-E are pointed out using red arrows.

Figure 6. PCB 608A11 sensor montage and positioning.

Table 2. Parameters of the accelerometer used.

<table>
<thead>
<tr>
<th>Name</th>
<th>Range (Hz)</th>
<th>Sensitivity (mV/g)</th>
<th>Destruction Limit (g)</th>
<th>Resonant Frequency (kHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>608A11</td>
<td>0.5–10,000</td>
<td>10.2</td>
<td>50</td>
<td>22</td>
</tr>
</tbody>
</table>

2.2. Signal Processing

The signal-processing methods employed in this study are depicted in Figure 7. A third-order Butterworth low-pass filter with a 5 Hz cutoff frequency was applied to the acceleration signal to extract the impacts from each wheel passage, called the axle pattern. By comparing the axle pattern in the time domain together with the measured axle distance of 4 m, the speed of the locomotive was estimated for each test run.
To preprocess the vibration signals, a third-order Butterworth band-pass filter, setting the cutoff frequencies at 50 Hz for the lower end and 2.5 kHz for the upper end, was applied to the captured signal. This processed signal and the estimated speed of the locomotive were used together to convert the acceleration signal into the spatial domain. Since the estimated speed for each test case was not the same, all the signals were resampled to achieve the same sampling frequency as the test case with the highest sampling frequency in the spatial domain. This frequency was calculated to be 1.56 kHz. After re-sampling, the signals were aligned and truncated to have the same length with the help of the achieved axle pattern. The length of the processed vibration signal was truncated to 128.32 m in the spatial domain. The wavelet denoising was then applied. The denoising procedure involved a multi-faceted approach, consisting of a level 9 decomposition using the Symlet 4 wavelet, along with the application of the empirical Bayesian denoising method. Additionally, median thresholding and a level-dependent noise estimator were applied to optimize the denoising process. The denoised signal was further converted into scale-averaged wavelet power (SAWP). Both the spatial domain signal and the SAWP signal were segmented and a group of 11 features was then extracted for each segment from both the spatial domain and the SAWP. These steps follow the same procedure presented in a previous study by Zuo et al. [20] with data collected from a testbed. In the final step of the procedure, the six most important features according to the corresponding Laplacian feature importance score were fed into the isolation forest algorithm to calculate the anomaly scores.

![System diagram of the signal-processing procedure.](image)

2.3. Cutoff Frequency

In Figure 8, the original and the processed vibration signals are presented for a train passage. To facilitate the comparison of the signal, the processed signal was amplified. A low-pass filter with a 5 Hz cutoff frequency was chosen to filter the original time domain signal. The processed signals were then used to further extract the axle passage pattern.

2.4. Axle Pattern Detection and Speed Estimation

The axle pattern can be determined from Figure 8. The two biggest negative amplitude peaks of each processed signal represent the moment when the wheels passed over the sensor. When the speed information was available, by measuring the time difference between the two negative peaks, one could calculate the distance between the wheels and vice versa. In this experiment, the distance between the front and back wheels was 4 m; hence, the speed of the locomotive could be calculated. The speeds for the four tests were estimated to be 7.59 m/s, 7.87 m/s, 6.57 m/s and 7.05 m/s, respectively.
2.5. Time Domain to Spacial Domain Conversion and Synchronisation

By utilising the estimated speed information, the signals were converted into the spatial domain. The low-pass filtered signals were aligned by calculating the corresponding cross-correlation between the different tests. The same shift values were later used to align the original signals. The aligned signals are visualised in Figure 9. Here, the 0 m point represents when the front wheel of the locomotive passes through the accelerometer.

![Figure 8](image1.png)

**Figure 8.** The original time domain signal and low-pass Butterworth filter processed signal with 5 Hz cutoff frequency.

![Figure 9](image2.png)

**Figure 9.** Four test signals in spatial domain aligned.
2.6. Wavelet Denoising and SAWP

Wavelet denoising serves as a sophisticated signal-processing method designed to effectively eliminate noise from a signal. Its approach involves decomposing the signal into distinct frequency bands utilising wavelet transforms. The core principle underlying wavelet denoising is to identify and isolate noise residing in the high-frequency sub-bands, subsequently applying targeted suppression to remove this interference. Crucially, the method ensures that essential signal information within the low-frequency sub-bands remains preserved, maintaining the integrity and utility of the processed signal. The details of the wavelet denoising procedure and the equations were described in detail in a previous study [33].

SAWP is a measure of the energy content of a signal at different scales using wavelet transforms. It is defined as the average power of the wavelet coefficients over a certain range of scales; it provides information about the overall strength of the signal across different frequency bands. The detailed formula for calculating SAWP was presented in a previous study [20].

The SAWP enables the thorough exploration of power variations across a spectrum of temporal scales, offering a precise and effective means of identifying and characterising power bursts within vibration signals. It creates a power time series that serves as a comprehensive record of how the power fluctuations unfold over time. This SAWP time series serves as a valuable foundation for subsequent analysis. The two features extracted from it are instrumental in quantifying and characterising the intensity and nature of the events in question, offering critical insights into the severance of the squat.

2.7. Anomaly Detection

Anomaly detection is a process of identifying data points or patterns that deviate from the expected or normal behaviour of a system. It is a key technique in various scientific fields, including data science, machine learning, and cybersecurity, among others. By nature, squat defects are rare and abnormal compared with rails without defects. Therefore, anomaly detection could be utilised to detect squats.

Anomaly detection techniques vary depending on the type of data being analyzed and the specific domain or application. Some common approaches include classification-based methods, clustering-based methods, nearest-neighbor-based methods, statistical methods, information-theoretic methods and spectral methods [34]. Isolation forest is a popular algorithm for anomaly detection and was used in this study. The idea of this method is to use random forests to isolate anomalies in the data. The algorithm works by constructing a number of binary trees, where each tree is built by randomly selecting a subset of features and then splitting the data along random values within the range of the selected feature. The process of randomly selecting features and values to split the data continues until each data point is isolated in its own leaf node. A more detailed description of the implementation was introduced by Zuo [20] in a previous study.

3. Results and Discussions

3.1. Feature Selection

Table 3 provides a summary of the 11 extracted features from the time domain signal and the SAWP. These features are, therefore, naturally categorised into two groups. The first group includes all the time-domain statistical features and the second group comprises two features extracted from the SAWP.

The method used for feature ranking was calculation of the corresponding Laplacian feature importance score for each feature. For the features that were similar to each other, only the ones with the highest score were kept by evaluating the cross-correlation values between the features. Here, a higher feature importance score implies the corresponding feature is of higher importance. With MATLAB, the Laplace feature importance scores were calculated and ranked. The achieved ranking is visualised in Figure 10. To address the correlations between the features, the following procedure was adopted: First, the most
significant feature was selected and its cross-correlation with the remaining features was computed. Any features whose cross-correlation value with the most important feature exceeded 0.9 were considered to be redundant and were removed. This process was repeated for the second most important feature remaining, and so on. This procedure terminated when none of the features remaining had a cross-correlation value greater than 0.9. Consequently, the remaining features were 10, 11, 9, 8, 5, and 6.

Table 3. All 11 features extracted from both domains.

<table>
<thead>
<tr>
<th>Feature Type</th>
<th>Features</th>
<th>Feature Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>time domain features</td>
<td>RMS 1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>standard deviation 2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>shape factor 3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>kurtosis 4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>skewness 5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>peak to peak amplitude 6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>impulse factor 7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>crest factor 8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>clearance factor 9</td>
<td>9</td>
</tr>
<tr>
<td>SAWP time series</td>
<td>number of peaks 10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>total peak power 11</td>
<td>11</td>
</tr>
</tbody>
</table>

![Feature ranking using the Laplacian score.](image)

3.2. Anomaly Scores

The anomaly scores were calculated using an isolation forest algorithm and the selected features. Figure 11 shows the anomaly scores for a test case with the old S&C for each segment. The whole signal was divided into 400 segments. The length of the signal corresponded to around 128.3 m; therefore, each segment represented a distance of around 0.32 m in the spatial domain. The y-axis was set to range from $-6$ to $6$ so that the details of the rest of the signal could be visualised and compared with the achieved anomaly scores. This setting cuts some of the biggest amplitude peaks but the absolute values of these peaks are not of interest in this Figure.
3.3. Threshold Selection

The threshold to decide whether a segment was an anomaly was calculated using the 95% percentile of the whole data size according to a previous study performed by Zuo et al. [20]. The histogram of the anomaly score and the achieved threshold are plotted in Figure 12. The threshold is reasonable and cuts the long tail away from the main distribution.

Figure 11. Anomaly score generated using an isolation forest algorithm and the selected 6 features.

Figure 12. Histogram of the anomaly scores and the threshold.

3.4. Anomaly Indicators for the S&C

Figure 13 displays all four test cases and their respective anomaly scores which exceed the specified threshold. Since there are two different S&Cs, the achieved results show it is possible to apply the proposed method for comparing the squat status of different S&Cs of similar types. The older S&C with squats had more anomalies identified compared with the newer one.

The information presented in Figure 8 shows that the two biggest negative amplitude peaks of each processed signal represent the moment when the front and rear wheels
passed over the sensors’ installation positions. Therefore, it is possible to further divide the achieved anomaly indicators into three parts. The first part represents the time before the front wheels have passed over the accelerometers and is denoted part 1. The second part represents the time the front wheels have passed over the accelerometers but the back wheels have not passed over the accelerometers and is denoted part 2. The third part represents the time that the back wheels have passed over the accelerometers and is denoted part 3. The lengths of these three parts were truncated to be 32 m, 6 m and 32 m, respectively. For the second part, the noise level was high because the wheel–rail contact was nearby the accelerometers. To make sure the anomaly scores and counts represented the status of the S&C, only the first and the third part were used for squat status monitoring. The accumulated anomaly score of the anomalies and the number of anomalies for each part from the four test runs are listed in Table 4.

![Figure 13. Anomaly scores over the threshold using the isolation forest algorithm.](image)

Compared with a previous study performed by Zuo et al. [20], the results are consistent. In both the low-speed condition before, and when the locomotive was travelling at around 30 km/h, the occurrence of squats increased both the number of anomalies detected and the accumulated anomaly score. When combining part 1 and part 3, the average accumulated scores of the new and old S&C were 3.75 and 13.58, and the average numbers of anomalies were 6 and 20, respectively.

<table>
<thead>
<tr>
<th>Test Run</th>
<th>Part</th>
<th>Accumulated Anomaly Score</th>
<th>Number of Anomalies</th>
</tr>
</thead>
<tbody>
<tr>
<td>old S&amp;C run 1</td>
<td>1</td>
<td>9.13</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5.62</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>5.03</td>
<td>8</td>
</tr>
<tr>
<td>old S&amp;C run 2</td>
<td>1</td>
<td>7.09</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>6.29</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>5.91</td>
<td>9</td>
</tr>
<tr>
<td>new S&amp;C run 1</td>
<td>1</td>
<td>1.35</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.09</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.85</td>
<td>3</td>
</tr>
<tr>
<td>new S&amp;C run 2</td>
<td>1</td>
<td>2.53</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.71</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.76</td>
<td>3</td>
</tr>
</tbody>
</table>
Through the automated detection of squat defects and monitoring of the S&C status, major accidents like derailments can be proactively prevented. Furthermore, incorporating the presented method into the maintenance process can result in substantial cost savings by eliminating unnecessary scheduled maintenance tasks. These advantages not only enhance the safety of railway transportation but also contribute to its sustainability, making it an even more appealing choice for meeting future transportation needs.

4. Conclusions and Future Work

This study was carried out on two real S&Cs to verify that the method of collecting vibration data near a point machine to detect rail squats is applicable in actual use cases. The achieved results show that this method not only works in a test bed when the speed of the boogie is low, but also works when using a locomotive with a reasonable speed of around 30 km/h. It also confirms that a low-pass filter can be used to identify when the wheels are passing over the sensor’s location. This part of the method can be further utilised either to identify the axle pattern when the speed is known or to calculate the speed when the axle pattern is available. The accumulated anomaly scores and the number of anomalies of the whole S&C can be utilised to estimate the status of squat defects. A new approach to split the whole signal into three parts was proposed. The results show that it is possible to use part 1 and part 3 together to calculate the ‘average accumulated anomaly score’ and the ‘average number of anomalies’ to estimate the health of the S&Cs related to squat defects. The value of the average accumulated anomaly score of the old S&C was around 3.6 times greater than the new one. The average number of anomalies of the old S&C was almost 3.4 times greater than the new one.

A number of new studies could be performed related to this work in the future. One possible study would be to collect more data from one S&C to perform a statistical analysis to further investigate and verify the results. Data could also be collected from real operating trains with more sophisticated axle patterns to demonstrate that the approach can be utilised in an actual situation when trains are in operation. A possible long-term study is to ascertain the feasibility of combining the power of deep learning and federated learning to propose a national solution for monitoring the status of S&C, including squats, wear, and other defects, without data exchange.

Author Contributions: Methodology, Y.Z. and M.R.; software, Y.Z.; validation, Y.Z.; formal analysis, Y.Z.; data curation, J.O.; writing—original draft, Y.Z.; writing—review & editing, Y.Z., P.C. and M.R.; visualization, Y.Z.; supervision, M.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: Data presented in the study are available on request with the corresponding author

Acknowledgments: This research received assistance from Alstom Luleå to access the facilities and collect the data. Really grateful!

Conflicts of Interest: The authors declare no conflict of interest.

References


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.