# Comparative Analysis of Fixed-Length and Dynamic Segmentation for Feature Extraction from Non-Stationary Spatial Data

Abel Ayele, Kayla Hamlin

**Abstract**

This research presents a comparative analysis of non-stationary spatial data segmentation techniques such as fixed-length and dynamic segmentation based feature extraction efficiency. The study utilizes 5 miles of railway track geometry data, a non-stationary spatial dataset, to assess the effectiveness of both segmentation approaches. The profile (vertical alignment) of the track geometry is used for this purpose. For fixed-length segmentation, the track data is divided into segments of 264 feet (1/20th of a mile), resulting in about 102 segments. Dynamic segmentation is performed using an "l2" model-based change point detection algorithm, which adapts to natural variations in the signal. Key features such as standard deviation, kurtosis, and energy are extracted from both segmentation methods. Performance is evaluated based on multiple criteria, including the discriminative power of the features for classifying track safety and ride-quality conditions using statistical tests such as the f-test and Fisher score, consistency or signal quality across segments, measured using the variance of the signal-to-noise ratio (SNR), computational efficiency in terms of run-time and memory usage. Results indicate that, features from fixed-length segments have demonstrated better discriminative power between safety and ride quality classes, with higher Fisher scores and f-values showing strong statistical significance ($p < 0.05$). Additionally, fixed-length segmentation has shown a better performance with lower run-time and stable signal power across segments.

DOI: 10.5195/pur.2025.102

## Introduction

**Background**

Data segmentation refers to the process of dividing large data in to small segments for targeted analysis. Breaking the data into smaller sections facilitates analysis by isolating specific events, anomalies, or trends that contribute to better decision-making and more targeted results. This is particularly important in the analysis of non-stationary, spatial data for localized analysis and more accurate feature extraction. The non-stationary characteristics of a spatial data such as railway track geometry signal where patterns and structural changes occur unpredictably pose unique challenges in data processing and feature extraction. Non-stationarity refers to the variability of statistical properties of the data along the track length. The deterioration of track conditions, often exacerbated by frequent train passage and track bed deformation, further contributes to the non- stationary characteristics of the data (Chen et al. 2018). Applying different segmentation techniques to such data, helps explore and validate methods for effectively capturing critical features that reflect the actual condition of the track, leading to more accurate diagnostics and predictive maintenance strategies. This makes railway track geometry a compelling case study for advancing segmentation methods and improving analytical approaches in handling non-stationary spatial data across various applications.

Studies highlight the influence of analytical segment length on the assessment of track quality. Longer analytical segments can improve precision in track quality assessments through a Zero-crossings segmentation strategy (Dawod and Terdik 2024). This approach allows for a more refined analysis of track conditions, suggesting that while shorter segments may enhance resolution, longer segments can provide a broader context for understanding track behavior. This duality indicates that the choice of segment length should be tailored to the specific goals of the assessment, balancing detail with overall track performance analysis. Li et al. propose a model that divides continuous track lines into adjacent segments of equal length for health evaluation. Their work illustrates that the segmentation approach can significantly influence the derived health indices, which are crucial for resource allocation and maintenance scheduling (Li et al. 2019). This aligns with the findings of Offenbacher et al., who discuss the development of various Track Quality Indices (TQIs) that integrate multiple geometry parameters, suggesting that the mathematical modeling of these indices is sensitive to the segment lengths used in the analysis (Offenbacher et al. 2020). Another study also reinforces the idea that TQIs, which rely on the standard deviations of measured parameters, are influenced by the segment length (Karunianingrum and Widyastuti 2020). Their findings indicate that a comprehensive assessment of track quality should consider the segment length to ensure accurate performance indicators.

The two widely applicable types of segmentation are fixed-length and dynamic (or adaptive) segmentations. Fixed-length divides the data into segments of equal size, while dynamic segmentation adjusts the segmentation boundaries to accommodate the natural variations within the data. Dynamic segmentation offers flexibility that can be advantageous in handling non- stationary data. Studies propose a method that adapts segment lengths based on local signal characteristics, which can lead to more accurate decompositions of signals (Ghoraani and Krishnan 2012). Jiang et al. propose a new filtering and smoothing algorithm for railway track surveying that utilizes inertial measurement units (IMUs) and odometer data. Their findings suggest that the dynamic nature of railway track conditions necessitates a flexible segmentation approach to accurately monitor track deformations and irregularities (Jiang et al. 2017). Fixed-length segmentation has also shown to provide robust performance in various applications such as voice activity detection (VAD), automatic speech recognition (ASR) and speech translation tasks (Gaido et al. 2021; Kamper et al. 2014). Famili et al. argues, the use of fixed segments simplifies the complexity of data processing by facilitating direct comparisons over time while maintaining spatial integrity (Famili et al. 2019). Fixed-length segments provided consistent results when evaluating track quality indices, suggesting that a uniform segment length can enhance the reliability of feature extraction in spatial datasets (Majstorović et al. 2022). The findings by Zarembski et al. also suggested that consistent data segmentation can lead to better predictive modeling of track geometry exceptions (Zarembski, Palese, and Euston 2017). Gao et al. argue that fixed-length segmentation allows for easier comparison and analysis of mechanical motion curves, thereby facilitating more accurate fault diagnosis in the intelligent diagnosis for railway turnout switches (Gao et al. 2022). Putra et al. also highlighted the necessity of fixed-length segments for feature extraction in accelerometer-based fall detection systems. They argue that fixed-length overlapping or non-overlapping sliding windows are essential for effective data segmentation, which can be extrapolated to other domains, including railway track data analysis (Putra et al. 2018).

The choice between fixed length and dynamic segmentation depends on the type of data to be analyzed and the specific goals of the analysis. In the context of railway track geometry data, evaluating these segmentation techniques is crucial for understanding their performance, particularly in tasks related to feature extraction. Previous studies have primarily focused on evaluating data segmentation to assess track quality indices and modeling geometry exceptions. This research aims to systematically compare the effectiveness of these two methods in extracting features that are essential for assessing the overall condition of the track, such as standard deviation and energy, as well as identifying defects through extreme values (minimum and maximum amplitudes) and kurtosis from the geometry data. The scope of this study is limited to employing equal-length, non-overlapping segments for the fixed-length method and utilizing change point detection based on the “l2” model for dynamic segmentation.

**Track Geometry**

The geometry of the railroad track in space is mainly described by parameters such as surface/profile, alignment, gage, cross-level, twist, curvature, superelevation and warp. Profile describes the change in elevation of the rail surface in the vertical plane over a track length. Alinement refers to the layout of the track in the horizontal plane. Gage refers to the distance between the face of the rails at a distance of 5/8” below the top of the rail. The 5/8” value is based on typical contact location of the wheel flange. Cross-level refers to the difference in height between the two adjacent rails at any given point along the track. Twist refers to the algebraic difference of two cross level measurements a defined distance apart, typically 31' or 62'. Warp refers to the algebraic difference of the maximum and minimum cross level measurements within a defined window, typically 31' or 62'.

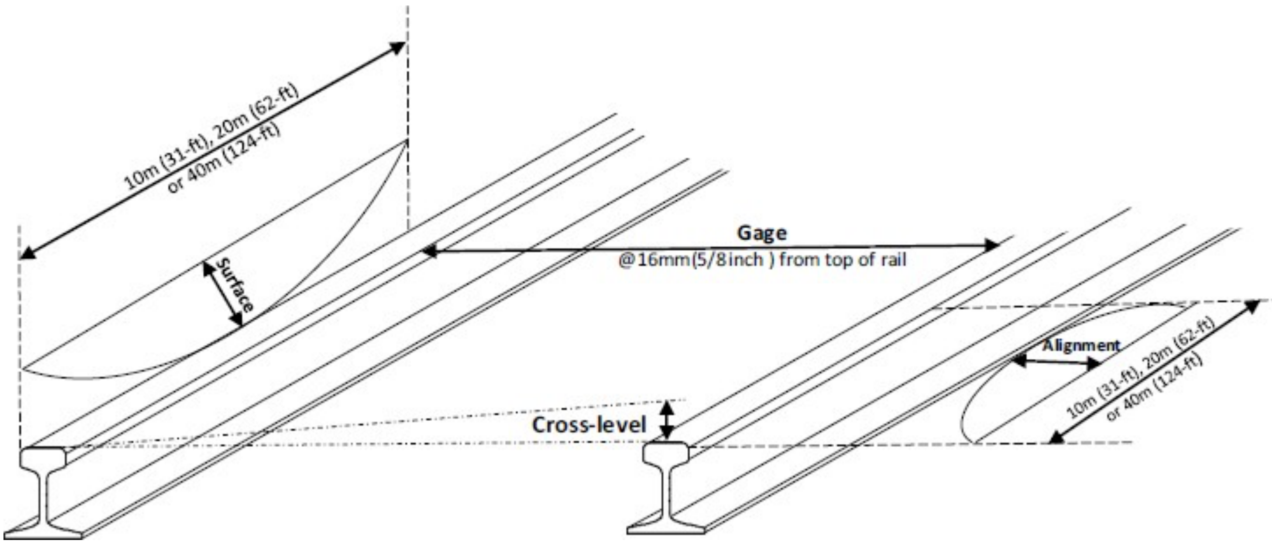


Figure 1: Schematic diagram of track geometry parameters (Lasisi & Attoh-Okine, 2019)

Effective analysis of railroad geometry data is crucial for identifying defects, enabling proactive maintenance planning and ensuring the safety of railway operations. Studies show that geometry defects and broken rail are the leading causes of freight derailments for major US railroads (Liu 2017). According to the train accident statistics by the Federal Railroad Administration (FRA) over the past 10 years, track failure is one of the leading causes of train derailments in the United States. Figure 2 shows the FRA's number of reported track related derailments in the US in the last ten years. Several studies have demonstrated that railway track irregularities, influenced by both vehicle dynamics and natural degradation, directly affect the operational safety of trains (Zhang, Cui, and Huang 2021; Lv et al. 2024). Furthermore, the statistical analysis of train derailment incidents by Liu et al. demonstrates a correlation between deteriorating track conditions and increased accident rates, noting that variations in track characteristics contribute significantly to accidents (Liu, Saat, and Barkan 2012).

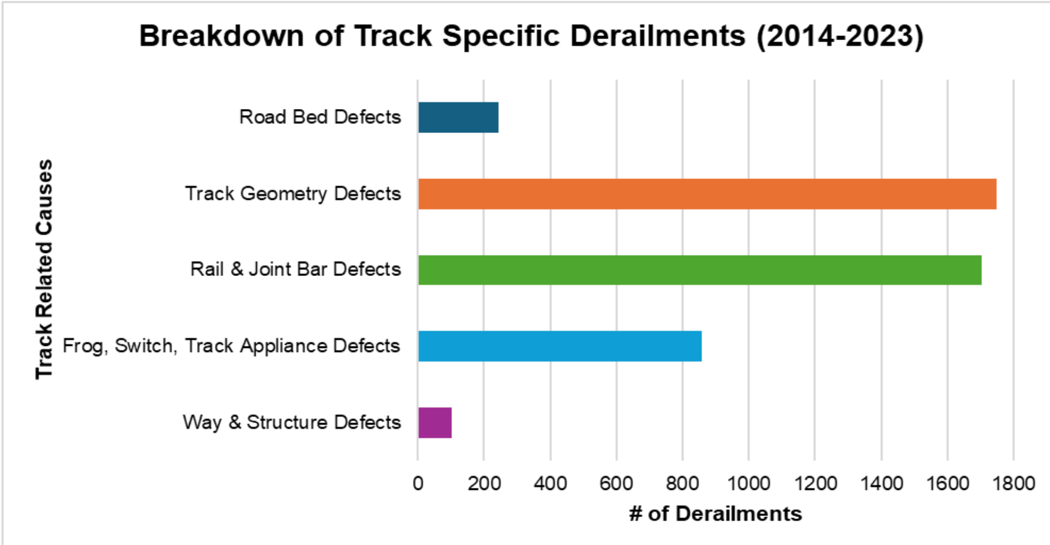


Figure 2: Number of derailments by track related causes

## Methodology

### About The Data

This study utilizes five miles of track geometry data obtained from a mainline heavy axle load, Class 4 freight track. The geometry data includes the vertical alignment (profile), the horizontal alignment, gauge, cross-level and twist of the track. Track geometry data serves as an ideal subject for studying segmentation of non-stationary spatial data due to its intrinsic variability and complex nature. In addition to track geometry data, ten years (2014–2023) of train accident report obtained from the FRA's Office of Safety Analysis was utilized in this research. The FRA train accident database documents track related accidents under five categories. These are roadbed, defects, track geometry defects, rail and joint defects, frog, switch and track appliance defects, and way and structure defects. Analyzing this accident data is crucial for understanding how irregularities in track geometry contribute to train accidents, highlighting the critical role that precise data segmentation plays in identifying and mitigating these risks.

### Data Segmentation

*Fixed-Length Segmentation*

Fixed-length segmentation involves dividing a continuous stream of data into uniform non-overlapping sections for further analysis. This approach ensures that each segment contains a similar number of observations, simplifying the comparison of track conditions across different sections (Famili et al. 2019). This method is particularly effective when there is a need to apply machine learning algorithms, which benefits from consistent input sizes across training samples. Mathematically, fixed-length segmentation can be represented as follows: Let the non-stationary spatial signal be represented as, $X = \{x1, x2, . . . , xn\}$, where $n$ is the total number of data points. The signal is divided into $m$ segments of equal length, $l$, such that: Segment 1: $\{x1, x2, . . . , xl\}$, Segment 2: $\{xl+1, xl+2, . . . , x2l\}$ ... Segment $m$: $\{x(m-1)l+1, x(m-1)l+2, . . . , xn\}$ where $m = n/l$, and $l$ is the predefined segment length.

*Dynamic Segmentation*

Dynamic segmentation creates segments of varying lengths to accommodate the data's inherent fluctuation. The "$l2$" model-based change point detection algorithm is a data-driven approach that adapts the segment lengths based on the local characteristics of the signal (Jaramillo, Nielsen, and Christensen 2021). Let the non-stationary spatial signal be represented as a time series, $X = \{x1, x2, . . . , xn\}$, where $n$ is the total number of data points. The algorithm seeks to find the optimal set of change points, $k = \{k1, k2, . . . , km\}$, where $m$ is the number of change points, such that the following cost function is minimized:

$$C(K) = \sum_{k=1}^{K} \sum_{t=t_{k-1}+1}^{t_k} (x_t - \hat{x}_k)^2 \tag{1}$$

where:

- $K$ is the number of change points.
- $\{t_1, t_2, \ldots, t_K\}$ are the change point positions.
- $\hat{x}_k$ is the mean of the signal in segment $k$, which is typically estimated as the average of $x_t$values within the segment $k$.
- $t_{k-1}$ and $t_k$ define the boundaries of each segment.

The algorithm iteratively searches for the optimal set of change points that minimize the overall cost function, adapting the segment lengths based on the local signal characteristics. The "l2" model, which is based on detecting changes in the mean of the signal, is a practical and efficient choice for many applications, particularly when working with large datasets such as track geometry data. This model primarily focuses on detecting mean shifts within segments and does not explicitly account for other statistical properties such as variance or assume full stationarity within a segment. This means that while it effectively identifies sudden mean changes, it may overlook variations in dispersion or higher-order statistical properties, limiting its ability to capture more complex signal behaviors. It struggles to detect gradual trends rather than high-frequency fluctuations. Since it is designed to identify abrupt changes in the mean, slow, continuous shifts in the signal might not be recognized as distinct change points. In contrast, high-frequency fluctuations, if they cause noticeable mean shifts, are more likely to be detected by the algorithm.

Unlike more complex non-linear models such as "rbf," which require substantial computational resources and can run into memory issues, the "l2" model is computationally lightweight. It provides a good trade-off by capturing important changes in the signal without the need for expensive kernel computations or high memory usage. While the "l2" model may not detect highly non-linear changes with the same precision as the "rbf" model, it is often sufficient for detecting significant transitions, making it ideal for real-time applications or large-scale analyses. By leveraging the advantages of dynamic segmentation while maintaining a low computational footprint, the "l2" model strikes a balance between accuracy and efficiency, offering a robust solution without requiring costly resources.

**Feature Extraction**

This research attempts to explore whether segmentation techniques offer more dependable and discriminative features for assessing track quality and safety by concentrating on feature extraction from both fixed-length and variable-length segments. Several assets of the signal's unpredictability, dispersion, and strength can be captured by the chosen characteristics, which include energy, kurtosis, and standard deviation. These characteristics are examined to determine how well they capture the track's underlying conditions and how well they serve to highlight locations that might need maintenance or provide safety hazards. Extreme values (min/max) help identify the most severe deviations in track geometry, highlighting critical defects such as sudden dips or peaks that could compromise train stability. Standard deviation measures the variability in track geometry while energy quantifies the overall intensity of track irregularities. Both standard deviation and energy help assess the general trends in the track geometry data. Skewness helps assess the asymmetry of track deviations, identifying whether track irregularities are biased toward one direction. Kurtosis detects transient defects by highlighting segments with extreme outliers, such as sudden bumps or cracks.

*Extreme Values*

Extreme values (minimum and maximum) are useful statistical features that can be extracted from non-stationary spatial data, such as railroad track geometry. The minimum value, denoted as $x\_min$, is the smallest value in the dataset. Minimum value can be expressed as: $xmin = min\{x1, x2, . . . , xn\}$. It represents the lowest point or the most extreme low value in the spatial signal. The maximum value, denoted as $xmax$ , is the largest value in the dataset. The maximum value can be expressed as: $xmax = max\{x1, x2, . . . , xn\}$. It represents the highest point or the most extreme high value in the spatial signal.

*Standard Deviation*

Standard deviation is a statistical feature that quantifies the amount of variability or dispersion in a set of data values. It indicates how much the individual data points in a dataset deviate from the mean (average) value of the dataset. A higher standard deviation indicates more spread out data, while a lower standard deviation indicates that the data points are closer to the mean.

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \quad (2)$$

Where $n$ represent the total number of data points in the dataset, $x_i$ represents each individual data point and $\bar{x}$ represents the mean of the dataset, calculated as $\bar{x} = \sum_{i=1}^{n} \frac{(x_i - \bar{x})}{n}$

*Energy*

Energy is the measure that quantifies the total magnitude of a signal or a dataset. It provides a sense of how "strong" or "intense" the signal is over its duration/distance or over the dataset's range.

$$Energy\ (E) = \sum_{n=0}^{N-1} |x[n]|^2 \tag{3}$$

$x_i$ represents the value of the signal at the $i^{th}$ point. $N$ is the total number of points in the signal or dataset. The absolute value squared $|x_i|^2$ ensures that all contributions are positive, emphasizing the magnitude of each point without regard to its direction.

*Skewness*

Skewness is a statistical measure that describes the asymmetry of the distribution of values in a dataset relative to its mean. It quantifies how much the distribution of the dataset leans to the left or right, providing an indication of the deviation from a normal distribution.

$$Skewness = \frac{N \sum_{i=1}^{N} (x_i - \bar{x})^3}{(N-1)(N-2)s^3} \tag{4}$$

$x_i$ represents the individual values in the dataset, $\bar{x}$ is the mean of the dataset, $s$ is the standard deviation, $N$ is the number of data points.

*Kurtosis*

Kurtosis measures the shape and spread of the decomposed signal. It is significant for detecting transient defects because it identifies sharp, isolated deviations from the normal track condition. High kurtosis might indicate the presence of transient events or defects. It can also be an indicator for a greater prevalence of outliers, which can be important for detecting transient events in the signal.

$$Kurtosis = \frac{N}{(N-1)(N-2)} \sum_{i=1}^{N} \left(\frac{x_i - \mu}{\sigma}\right)^4 - 3 \tag{5}$$

where $x_i$ represents each individual data point in the dataset, $\mu$ is the mean of the data points and $\sigma$ is the standard deviation of the data points.

The extracted features are evaluated within individual segments rather than across multiple segments. Since segmentation techniques divide the track into distinct sections, statistical features are calculated per segment to capture localized track conditions. This segmentation-based evaluation ensures that track maintenance efforts can be targeted at specific problematic areas rather than averaging out anomalies across the entire dataset.

## Scope and Limitation of the Study

In this research, for the fixed-length method, a predefined segment length of 264 $ft$ ($1/20th\ of\ a\ mile$) is utilized, a standard measurement frequently employed in railway track condition assessments. The choice of this specific length aims to align the study with industry practices, facilitating relevant and actionable insights. On the other hand, the dynamic segmentation approach employs a change point detection algorithm based on the "$l2$" model. This model was selected for its computational efficiency, providing a robust framework for identifying significant changes in the data. While this study provides a thorough comparison of fixed-length and dynamic segmentation methods, it acknowledges certain limitations. The use of the "$l2$" model for dynamic segmentation, while efficient, might not capture as complex changes in data as other models, such as the radial basis function ("$rbf$"). Therefore, the results might not generalize to scenarios where more sophisticated change point detection is required. Future research could extend this work by exploring the "$rbf$" model or other advanced dynamic segmentation algorithms. Such investigations could offer a broader understanding of how different segmentation strategies perform across a wider range of non-stationary spatial data scenarios.

# Exploratory Data Analysis (EDA)

### A Line Plot of Track Irregularities

The line plot indicated in Figure 3 shows the amplitude of the profile irregularities for the five miles track under study. Visual inspection of this plot reveals that the first two miles of this track contains multiple sharp spikes with large amplitudes which indicates that this section of the track is very rough compared to the last three miles which looks fairly smooth. This variation in the amplitude of the profile irregularities between track sections will have an implication in the dynamic segmentation of the track. The rougher sections with high amplitude and frequent changes might be segmented into shorter lengths to capture detailed variations and pinpoint specific issues. In smoother sections, longer segments could suffice as these areas show less variability and might not require as granular an analysis, leading to more efficient data processing and analysis.

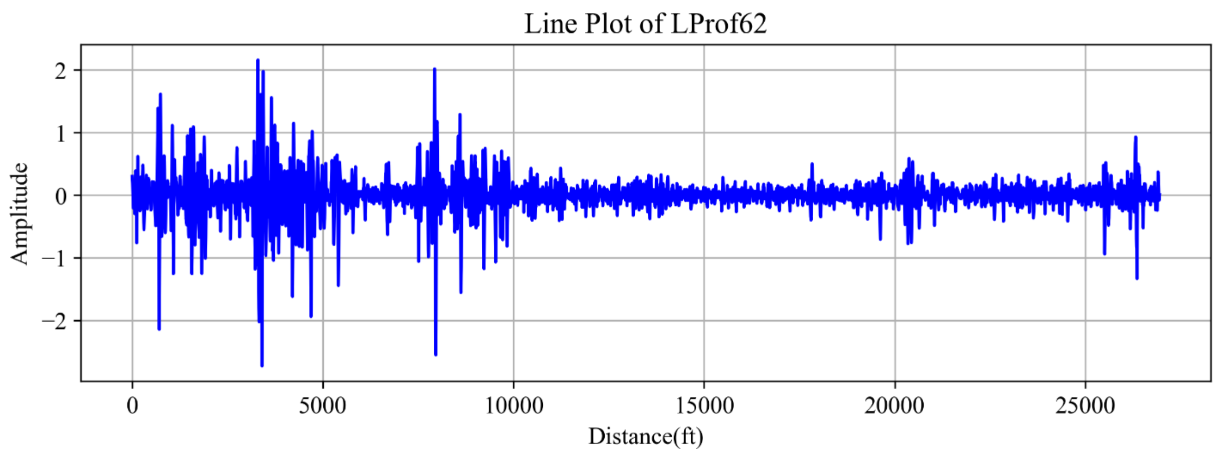


Figure 3: Line graph for Left Profile Signal

### Distribution of the Profile Signal

The histogram in Figure 4 illustrates the distribution of "LProf62" values, which is centered around the mean of 0. The distribution is markedly symmetric and exhibits a bell-shaped curve, typical of a normal distribution. Both tails extend towards the -2 and 2 extremes but with significantly lower frequencies, suggesting that extreme values are relatively rare. This pattern highlights the general stability of the "LProf62" values, with most deviations being minor and suggesting that any significant deviations might be anomalies or indicative of specific areas of concern along the profile being analyzed. With the majority of data points falling close to zero and showing minimal variability, the thresholds for defining segments in dynamic segmentation can be set based on deviations from the mean.

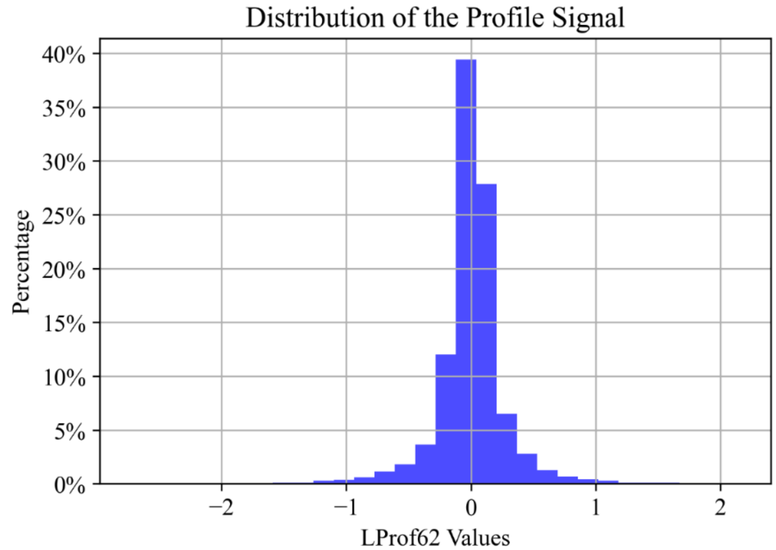


Figure 4: Distribution of the Left Profile Signal

## Results and Discussion

In the case of fixed-length segmentation, the left profile signal (LPof62') for the 5 miles (26,944 ft) track was divided into a predefined length of 1/20th of a mile (264 feet) which generates 102 segments with varying signal characteristics. Each segment contains 264 points, except for the last segment, which contains 248 data points due to the track length not being perfectly divisible by the segment length. In the case of dynamic segmentation, "l2" model based change point detection algorithm was utilized to generate about 48 segments from the same track. Figure 5 and Figure 6 show the fixed length and dynamic (variable length) segments of sample 1mile track.

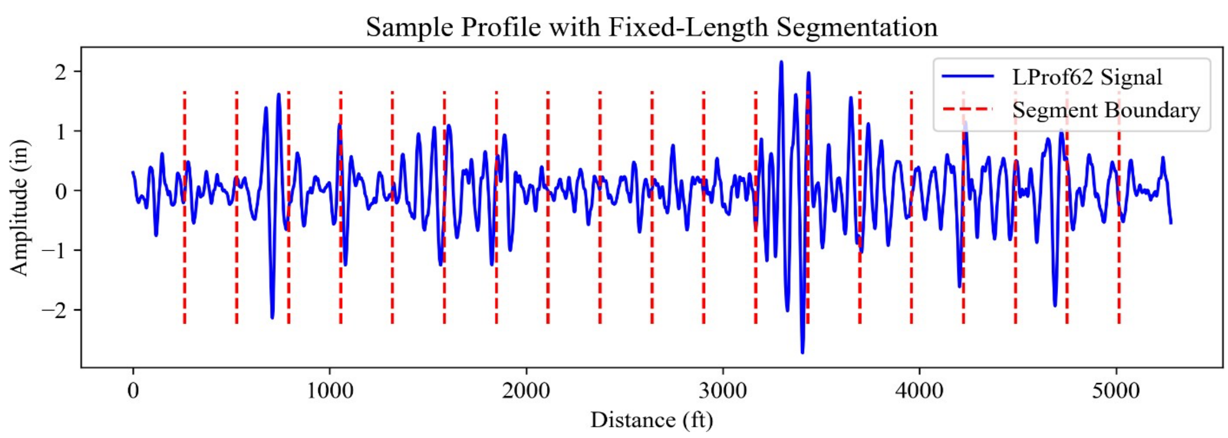


Figure 5: Fixed-length segmentation

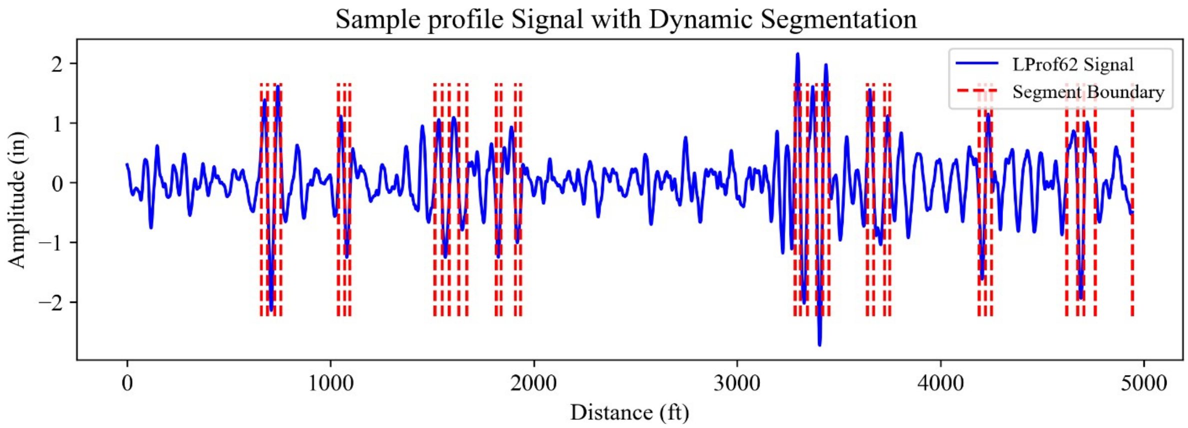

Figure 6: Dynamic (variable-length) Segmentation

The performance of each segmentation methods was evaluated based on (1) computational efficiency of the method, (2) signal/data stability across segments and (3) discriminability of each method. Computational efficiency was measured based on metrics such as run-time and memory usage of each segmentation method. The stability or consistency of the signals across the segments during the segmentation process was measured using the signal-to-noise ratio of each segments. The discriminative power of features extracted from each segmentation method was measured using statistical tests such as Fisher-score and f-test.

**Correlation Analysis**

The correlation matrices depicted in Figures 7 and 8 illustrate the relationships between features extracted from fixed-length and dynamic (variable-length) segments, respectively. Notably, certain features in the fixed-length segments display stronger correlations. For example, the standard deviation, a key statistical measure often utilized as a Track Quality Index (TQI) for assessing track conditions in the railroad industry (Offenbacher et al. 2020), shows a pronounced correlation with both the energy and the extreme values (minimum and maximum) of the geometry variable under analysis. This reflects expected behavior in practical track data scenarios, where these features often vary together indicating similar influences from track conditions.

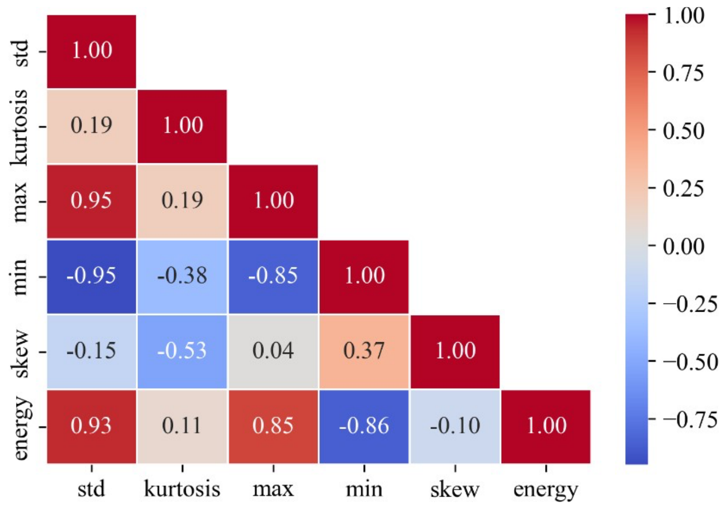


Figure 7: Correlation between Features Generated by Fixed-Length Method

The analysis of features from dynamic segmentation, as indicated by the correlation matrix in Figure 8, shows unexpected patterns compared to fixed-length segmentation. Notably, the moderate correlation of 0.62 between energy and kurtosis isn't considered significant, contrasting with typically stronger correlations seen in fixed methods, such as between standard deviation, energy, and amplitude extremes. The lower correlations involving standard deviation in dynamic segmentation suggest it may capture localized track conditions more distinctly, differing from the global trends highlighted in fixed-length methods. These findings may prompt a reevaluation of how metrics from dynamic segmentation are used in track maintenance and assessment.

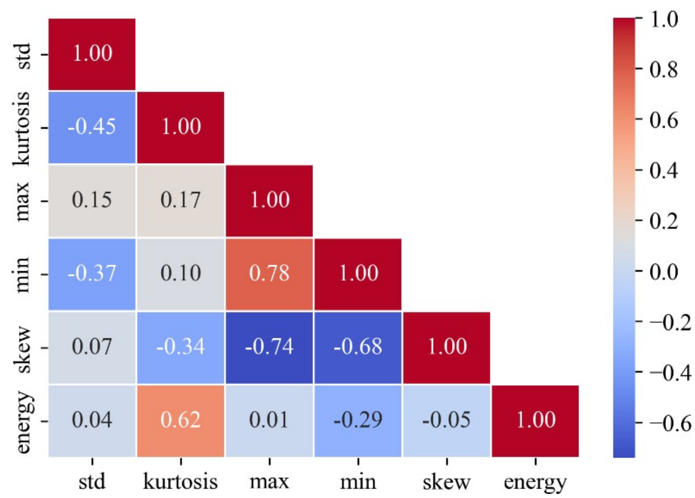


Figure 8: Correlation between Features Generated by Dynamic Method

**Discriminative Power (statistical Test)**

All the segments in the five track mile data were classified into two categories (ride quality and safety) based on the FRA's safety standard for class 6 track. The safety threshold stated in this standard stipulates the amplitude exceedance of 0.75 inches for a profile (vertical alignment). Segment with the amplitude of the profile signal exceeding 0.75 *inches* were labelled as a 'safety' and those with amplitude values under 0.75 *inches* were labelled as a 'ride quality' class. Metrics such as Fisher Score and F-test were used to evaluate how well the extracted features from each segmentation method differentiate between ride quality and safety segments.

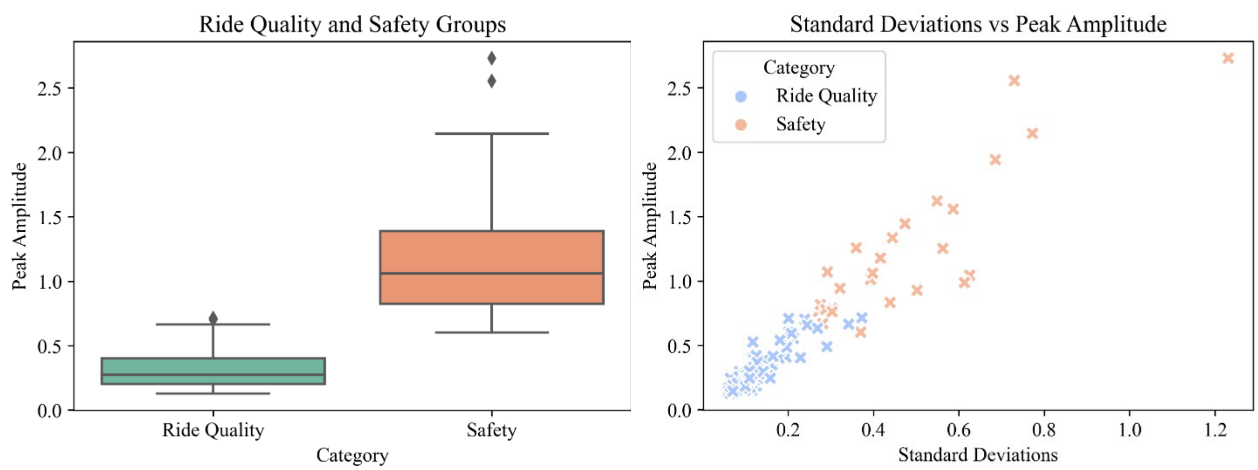


Figure 9: Ride Quality and Safety classes of the track segments

The minimum and maximum values, standard deviation, energy, kurtosis and skewness were extracted from both fixed-length and dynamic segments. The f-value, derived from ANOVA (Analysis of Variance), quantifies the degree to which these features vary between safety and ride quality groups compared to the variation within each group. It provides a ratio of between-group variance to within-group variance, where higher values indicate better discriminability. Similarly, Fisher-score measures the separation between classes (safety and ride quality) by calculating the ratio of the variance between the classes to the variance within the classes for each feature. Table 1 shows the summary of these statistical tests for both segmentation techniques.

Table 1: Feature Discriminability

| | Fixed-Length | | | Dynamic (Variable-Length) | | |
|---|---|---|---|---|---|---|
| | **Fisher-score** | **f-value** | **p-value** | **Fisher-score** | **f-value** | **p-value** |
| Min | 1.68 | 168.36 | 3.66E-23 | 0.01 | 0.4875 | 0.4886 |
| Max | 1.66 | 165.70 | 6.04E-23 | 0.005 | 0.0241 | 0.8772 |
| Std. dev | 1.66 | 165.62 | 6.13E-23 | 0.06 | 2.76 | 0.10 |
| Energy | 0.59 | 58.53 | 1.27E-11 | 0.02 | 0.83 | 0.37 |
| Kurtosis | 0.11 | 10.70 | 0.001471 | 0.003 | 0.13 | 0.72 |
| Skewness | 0.04 | 3.72 | 0.06 | 0.004 | 0.18 | 0.68 |

This result shows across all features, the fixed-length method constantly exhibits a larger Fisher score and F-value, highlighting its greater discriminatory potential. Features from the dynamic segmentation method on the other hand, resulted in far lower Fisher-score and F-values. These low values imply that the dynamic method's feature contribution is not sufficiently variable, which reduces its capacity to discriminate between the two classes. Considering the confidence interval of 95%, in the case of the fixed-length segmentation method, the $p - values$ associated to the f-test came under 0.05 indicating a strong statistical significance. However, in the case of dynamic segmentation method the $p - values$ for all the features are above 0.05, indicating a lack of statistical significance in distinguishing between safety and ride quality classes.

In fixed-length segmentation, when calculating the Fisher score, which assesses the ratio of between-class variance to within-class variance for each feature, consistent segment sizes can enhance the ability to distinguish between the ride quality and safety classes. Variations within segments are minimized relative to variations between segments, enhancing class discriminability. In dynamic segmentation, if change points coincide with or are triggered by outliers, the resulting segments might not accurately reflect the general condition of the track. This can affect the calculated features, leading them to capture these extremes rather than the typical conditions, potentially reducing their discriminative power as measured by Fisher scores. The other explanation is the fact that the "l2" model in dynamic segmentation is sensitive to noise in data, which might lead to identifying too many change points in noisy data scenarios. This can fragment the data excessively, reducing the effectiveness of the features extracted from these segments. The Fisher score and F-test (ANOVA) are appropriate for evaluating feature discriminability but may not fully capture the advantages of adaptive segmentation. These tests focus on global statistical differences, whereas dynamic segmentation is designed to detect localized changes. In real-world track condition monitoring, a combination of segmentation methods might be optimal fixed-length for consistent statistical evaluation and dynamic segmentation for identifying localized defects.

### Computational efficiency

Table 2 shows the summary of results for the for the performance metrics that evaluates computational efficiency and signal stability of the fixed-length and dynamic segmentation techniques.

Table 2: Computational Efficiency and Signal Stability

| | Fixed-length | Dynamic (variable-length) |
|---|---|---|
| Run Time (sec) | 0.2 | 306 |
| Memory Usage (MB) | 1.66 | 1.65 |
| SNR (std. dev) | 0.09 | 2.2 |

The change point detection algorithm ("l2" model), used for dynamic segmentation, exhibits poor computational efficiency, with a total run-time of 306 seconds to process five miles track compared to just 0.2 seconds for the fixed-length approach. The dynamic segmentation's complexity, which could include extra calculations, iterative procedures, or adaptive mechanisms that lengthen processing times, could be the cause of the significant discrepancy. Both methods used almost the same amount of memory, being 1.66 MB for the fixed system and 1.65 MB for the dynamic segmentation processes. Despite the large difference in run time, the small difference in memory usage indicates that both methods are efficiently tuned, with no visible memory consumption tradeoff. This finding suggests that higher memory needs are not the cause of the dynamic system's longer run time. The significant difference in run time underscores the high computational complexity of the change point detection method, which requires more processing time to adaptively identify segments based on signal changes. This makes it less practical for large datasets or real-time applications.

**Signal Stability**

With a significantly lower signal-to-noise ratio (SNR) standard deviation of 0.09, the fixed-length system exhibits higher signal reliability and consistency across segments. On the other hand, the dynamic segmentation method exhibits greater inconsistency in signal performance, as evidenced by its higher deviation of 2.2 shown in (Table 2). This instability in signal power across segments using the "$l2$" model based change point detection algorithm suggests that the dynamically divided segments may not consistently capture meaningful signal patterns, leading to increased noise and variability. In contrast, the fixed-length method shows much more stable signal power, with a lower SNR variation across segments.

The results discussed above suggest that the fixed-length segmentation method is more effective and statistically reliable compared to the dynamic method, which employs the 'l2' model-based change point detection, for applications that require feature extraction and discrimination between different classes for this type of data. Fixed-length segmentation produces stable and comparable segments, leading to statistically significant differences between track conditions. This method ensures that extracted features remain consistent, improving discriminability between safety-critical and ride quality sections. Dynamic segmentation, despite its lower Fisher scores and F-values, remains useful in scenarios where feature variation is more meaningful for identifying defects, such as capturing localized anomalies that might not be detected in uniform- length segments. However, its computational complexity and segment inconsistency make it less practical for large-scale applications.

## Conclusion and Future Directions

Fixed-length data segmentation has proven to be a more efficient and reliable method, providing better discriminative power, data stability and faster computation time compared to the dynamic method that utilizes “l2” model based change point detection algorithm. For the non-stationary spatial signals, such as railway track geometry data, the use of "l2" model based change point detection algorithm for feature extraction has shown poor performance. While this method theoretically offers flexibility by adapting to natural signal variations, it struggles to maintain consistent signal power across segments and fails to effectively distinguish between safety and ride quality features. Furthermore, its high computational cost and inconsistency in signal segmentation severely limit its practical utility in railroad applications, in particular feature extraction tasks. However, caution should be exercised not to overly generalize these results, due to the sensitivity to the type of data used and the specific model employed in the change point detection algorithm in dynamic segmentation. Future research should focus on exploring alternative adaptive models, such as Bayesian or machine learning-based approaches, to improve the accuracy and efficiency of dynamic segmentation for non-stationary spatial da

**References**

Chen, Qijin, Xiaoji Niu, Lili Zuo, Tisheng Zhang, Fuqin Xiao, Yi Liu, and Jingnan Liu. 2018. "A Railway Track Geometry Measuring Trolley System Based on Aided INS." Sensors (Switzerland) 18 (2). https://doi.org/10.3390/s18020538.

Dawod, Abdaljbbar, and György Terdik. 2024. "Use of Zero-Crossings Segmentation for Track Quality Assessment." Acta Technica Jaurinensis 17 (1): 8–21. https://doi.org/10.14513/actatechjaur.00726.

Famili, Afshin, Wayne Sarasua, Adika Mammadrahimli Iqbal, Devesh Kumar, and Jennifer Harper Ogle. 2019. "Short Segment Statewide Screening of Midblock Crashes in South Carolina." Transportation Research Record 2673 (10): 696–707. https://doi.org/10.1177/0361198119851068.

Gaido, Marco, Matteo Negri, Mauro Cettolo, and Marco Turchi. 2021. "Beyond Voice Activity Detection: Hybrid Audio Segmentation for Direct Speech Translation," April. http://arxiv.org/abs/2104.11710.

Gao, Yuxin, Yong Yang, Yuan Ma, and Weixiang Xu. 2022. "Study on Intelligent Diagnosis of Railway Turnout Switch Based on Improved FastDTW and Time Series Segmentation under Big Data Monitoring." Mathematical Problems in Engineering 2022. https://doi.org/10.1155/2022/7048813.

Ghoraani, Behnaz, and Sridhar Krishnan. 2012. "Discriminant Non-Stationary Signal Features' Clustering Using Hard and Fuzzy Cluster Labeling." EURASIP Journal on Advances in Signal Processing. Vol. 2012. http://asp.eurasipjournals.com/content/2012/1/250.

Jaramillo, Alfredo Esquivel, Jesper Kjær Nielsen, and Mads Græsbøll Christensen. 2021. "Speech Decomposition Based on a Hybrid Speech Model and Optimal Segmentation." In Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 4:3111–15. International Speech Communication Association. https://doi.org/10.21437/Interspeech.2021-47.

Jiang, Qingan, Wenqi Wu, Mingming Jiang, and Yun Li. 2017. "A New Filtering and Smoothing Algorithm for Railway Track Surveying Based on Landmark and IMU/Odometer." Sensors (Switzerland) 17 (6): 2–20. https://doi.org/10.3390/s17061438.

Kamper, Herman, Aren Jansen, Simon King, and Sharon Goldwater. 2014. Unsupervised Lexical Clustering Of Speech Segments Using Fixed-Dimensional Acoustic Embeddings. IEEE.

Karunianingrum, Dyni Indar, and Hera Widyastuti. 2020. "Penilaian Track Quality Index (TQI) Berdasarkan Standar Perkeretaapian Indonesia (Studi Kasus : Cirebon-Cikampek)." Jurnal Aplikasi Teknik Sipil. Vol. 18.

Lasisi, Ahmed, and Nii Attoh-Okine. 2019. "Machine Learning Ensembles and Rail Defects Prediction: Multilayer Stacking Methodology." ASCE-ASME Journal of Risk and

Uncertainty in Engineering Systems, Part A: Civil Engineering 5 (4). https://doi.org/10.1061/ajrua6.0001024.

Li, Qing, Qiyuan Peng, Rengkui Liu, Ling Liu, and Lei Bai. 2019. "Track Grid Health Index for Grid-Based, Data-Driven Railway Track Health Evaluation." Advances in Mechanical Engineering 11 (11). https://doi.org/10.1177/1687814019889768.

Liu, Xiang. 2017. "Statistical Causal Analysis of Freight-Train Derailments in the United States." Journal of Transportation Engineering 143 (2). https://doi.org/10.1061/jtepbs.0000014.

Liu, Xiang, M. Rapik Saat, and Christopher P.L. Barkan. 2012. "Analysis of Causes of Major Train Derailment and Their Effect on Accident Rates." Transportation Research Record, no. 2289 (January), 154–63. https://doi.org/10.3141/2289-20.

Lv, Shiwang, Yu Yin, Zhipeng Chen, Qingquan Li, Wei Tu, Dejin Zhang, Bochen Zhang, and Zijian Zhang. 2024. "Rapid and Precise Measurement Method for Track Static Geometry of Large-Span Bridge Railway." Structural Control and Health Monitoring 2024. https://doi.org/10.1155/2024/1224481.

Majstorović, Igor, Maja Ahac, Janusz Madejski, and Stjepan Lakušić. 2022. "Influence of the Analytical Segment Length on the Tram Track Quality Assessment." Applied Sciences (Switzerland) 12 (19). https://doi.org/10.3390/app121910036.

Offenbacher, Stefan, Johannes Neuhold, Peter Veit, and Matthias Landgraf. 2020. "Analyzing Major Track Quality Indices and Introducing a Universally Applicable Tqi." Applied Sciences (Switzerland) 10 (23): 1–17. https://doi.org/10.3390/app10238490.

Putra, I. Putu Edy Suardiyana, James Brusey, Elena Gaura, and Rein Vesilo. 2018. "An Event-Triggered Machine Learning Approach for Accelerometer-Based Fall Detection." Sensors (Switzerland) 18 (1). https://doi.org/10.3390/s18010020.

Zarembski, Allan M., Joseph W. Palese, and Todd L. Euston. 2017. "Correlating Ballast Volume Deficit with the Development of Track Geometry Exceptions Utilizing Data Science Algorithm." Transportation Infrastructure Geotechnology 4 (2–3): 37–51. https://doi.org/10.1007/s40515-017-0039-5.

Zhang, Xinchun, Ximin Cui, and Bo Huang. 2021. "The Design and Implementation of an Inertial Gnss Odometer Integrated Navigation System Based on a Federated Kalman Filter for High-Speed Railway Track Inspection." Applied Sciences (Switzerland) 11 (11). https://doi.org/10.3390/app11115244.