

Gaussian Mixture Segmentation for Managing Deterioration of Large-Scale Road Networks

Pengcheng Zhang¹, Wen Yi, Yongze Song², *Senior Member, IEEE*, Peng Wu²,
Keith Hampson, and Ammar Shemery

Abstract—Transportation infrastructure significantly influences the development of sustainable transportation systems and the effective management of road networks. Accurate segmentation of road infrastructure provides valuable data and a structured approach for the effective management of road deterioration. However, existing methods are limited in segmenting road deterioration data due to the lack of probabilistic assessment and the inability to effectively handle outliers. To address this gap, the Gaussian mixture segmentation (GMS) model is introduced, applying a Gaussian mixture distribution to model the variability in road deterioration. Our approach utilizes spatial line segmentation to effectively segment large-scale road networks into meaningful segments. The GMS model was applied to road deterioration data from the South West region of Western Australia, identifying and segmenting roads based on their distribution characteristics using the Jensen-Shannon (JS) divergence. To assess the performance of the GMS model, metrics such as the number of segments, coefficient of variation (CV), Caliński-Harabasz Index, and Davies-Bouldin Index were evaluated. The results demonstrate that the GMS model outperformed existing segmentation methods, achieving an increase in the average percentage of segments with a CV lower than 0.25 by 23.5% to 98.8%, along with a reduction in the average Davies-Bouldin Index by 25.2% to 64.4% and improvements in the average Caliński-Harabasz Index, which increased by 23.0% to 131.1%. This approach enhances the understanding of the spatial distribution of road deterioration, informing maintenance strategies for large-scale road networks and addressing the complexities of road infrastructure management, including traffic dynamics and environmental impacts.

Index Terms—Gaussian mixture distributions, road network, road deterioration, transportation management, spatial line segmentation.

I. INTRODUCTION

TRANSPORTATION infrastructure supports smooth traffic flow, enhances road safety, promotes alternative transportation modes, and enables efficient resource management

Received 19 September 2024; revised 23 February 2025; accepted 2 April 2025. Date of publication 22 April 2025; date of current version 6 August 2025. This work was supported in part by the National Natural Science Foundation of China under Grant 72201229 and Grant 72361137006 and in part by the Centre for Infrastructure Delivery Research Funding under Grant P0055981. The Associate Editor for this article was M. Guo. (*Corresponding author: Yongze Song.*)

Pengcheng Zhang and Wen Yi are with the Department of Building and Real Estate, The Hong Kong Polytechnic University, Hong Kong (e-mail: pengcheng.zhang@connect.polyu.hk; wen.yi@polyu.edu.hk).

Yongze Song and Peng Wu are with the School of Design and the Built Environment, Curtin University, Perth, WA 6102, Australia (e-mail: yongze.song@curtin.edu.au; peng.wu@curtin.edu.au).

Keith Hampson and Ammar Shemery are with the Sustainable Built Environment National Research Centre, Perth, WA 6102, Australia (e-mail: k.hampson@sbenrc.com.au; a.shemery@sbenrc.com.au).

Digital Object Identifier 10.1109/TITS.2025.3560357

[1], [2], [3]. As the foundation of transportation infrastructure, road networks facilitate the movement of people and goods while enhancing connectivity and accessibility [4]. Large-scale road networks often span thousands of kilometers and encompass diverse traffic patterns, environmental conditions, and asset characteristics, presenting significant challenges for management and maintenance across metropolitan regions, states, or even entire countries. Well-maintained road networks can reduce congestion [5], [6] and mitigate negative environmental impacts [7].

However, road deterioration is a significant issue influenced by various factors such as climate and traffic [8], [9]. Effectively managing road deterioration presents two main challenges. First, acquiring high-quality data related to road deterioration is crucial, including collecting detailed information on factors such as surface conditions, environmental impacts, and traffic loads to accurately assess deterioration rates [10]. Second, developing suitable and effective management methods is essential. Current indicators, such as pavement service life, often show a low correlation with actual road conditions due to factors like design lifespan and fluctuations in traffic flow [11]. In addition, effective management methods must address the various movements, wear and tear, and structural damage that contribute to increasing maintenance demands [12].

Current methods for managing road deterioration can be categorized into data-driven analysis and multi-user demand analysis [13], [14]. The data-driven analysis utilizes advancements in sensor technology to gather extensive road condition data, which is increasingly applied in local construction and maintenance [15]. This approach improves network-level management by providing high accuracy and valuable information, mainly through large-scale spatial analysis [16]. However, it lacks effective segment-level indicators to quantify deterioration accurately [17]. In contrast, multi-user demand analysis focuses on the different needs of various road users. This approach highlights the significance of segment-level data, which is often more practical than raw observational data [18]. While it captures diverse user requirements, it may not always offer the detailed condition assessments that data-driven methods provide. Consequently, there is a need for precise segment-level methods to assess road deterioration, which would facilitate long-term maintenance decision-making for government agencies.

Road segmentation effectively divides large-scale road networks into segment-level units, allowing for detailed analysis

of road deterioration and providing valuable decision support for maintenance. The road segmentation model identifies homogeneous segments based on deterioration characteristics, facilitating targeted maintenance, proactive planning, performance monitoring, and integrated analysis [19], [20]. This method enables transportation planners and engineers to tackle issues more effectively by breaking down road networks into meaningful segments, which aids in developing customized plans and targeted interventions to enhance traffic flow, such as identifying hotspots for congestion and high traffic volumes [21], [22], [23]. Road segmentation could be expanded to both private and public transport, which are useful for evaluating the risk of accidents [24], [25]. Ultimately, these advancements in road segmentation contribute to safer and more efficient transportation systems.

Previous research focused on methods for defining road segments that can be classified into three categories [20]. Firstly, one common approach involves segmenting roads as sections between intersections, considering road numbers, names, directions, and intersection locations within the network [26]. This method simplifies the identification of individual road sections and helps road users with communication and navigation. Secondly, another approach involves identifying road segments with consistent properties, such as the number of lanes, width, pavement surfacing type, construction material, or soil type [12]. This method is particularly useful for managing and maintaining infrastructure, as it allows for grouping similar road sections to allocate resources efficiently and plan maintenance. Thirdly, road segments are often characterized by similar or approximate attributes, including traffic flows, road deterioration, and local environmental conditions [27], [28], [29]. Transportation practitioners can gain a better understanding of the factors affecting traffic operations by analyzing attributes within road segments, such as congestion patterns, road surface conditions, and environmental impacts [22], [30], [31]. Such analysis aids in optimizing road infrastructure investments, enhancing traffic flow, and developing targeted sustainable transportation strategies.

Despite the effectiveness of existing road segmentation methods, they exhibit notable limitations [32]. One significant issue is their sensitivity to sudden changes in observations, particularly the presence of outliers, which can adversely affect segmentation results. For example, outliers in road deterioration variables can severely impact the final segmentation outcomes [33], subsequently influencing road maintenance strategies. Furthermore, there needs to be more consideration of the probability distribution of road deterioration variables; effective segmentation requires that variables within the same road section exhibit similar probability distributions [22], as deviations can lead to misestimations of deterioration rates and hinder the development of effective environmental mitigation measures.

To overcome the existing limitations in road segmentation, the primary objective of this study is to develop a novel segmentation approach that addresses the challenges of managing road deterioration in large-scale road networks. The GMS model utilizes the probability distribution of road deterioration variables in terms of sensor monitoring data

for segmentation and employs Jensen-Shannon divergence to measure distribution distance, allowing for effective segmentation that is less sensitive to abrupt changes and outliers. Accurately identifying homogeneous road segments based on their deterioration characteristics enables the GMS model to distinguish segments with higher deterioration rates. These high-deterioration areas can then be prioritized for timely maintenance, enabling government agencies to make informed long-term maintenance plans. This approach enhances the management of pavement deterioration, resulting in more effective decision-making across road networks.

This study proposes a Gaussian mixture segmentation (GMS) model for spatial line segmentation to manage deterioration in large-scale road networks. The work is original in several key ways. First, the innovative use of probability distribution is introduced to segment road deterioration variables, enabling the identification of high-deterioration areas. The model can assist government agencies in making long-term maintenance plans and improving overall road conditions. Second, the method enhances road segmentation accuracy by addressing the sensitivity of existing methods to abrupt changes and outliers, which can distort segmentation results. This improvement leads to a more precise identification of high-deterioration road segments, ultimately facilitating more effective maintenance strategies. Finally, a case study was conducted in the South West region, employing three metrics to compare this method with traditional segmentation approaches. The results indicate that the GMS model outperforms existing methods. This study lays the groundwork for developing targeted maintenance strategies and contributes to advancing research in this field by integrating the probability distribution of road deterioration variables.

II. RELATED WORKS

Methods related to this study are reviewed in this section, including road segmentation methods, Gaussian mixture segmentation and deep learning-based segmentation methods.

A. Road Segmentation Methods

In recent years, various emerging methodologies have been developed for road segmentation, each offering unique advantages for analyzing road networks. The coefficients of variation (CV)-based methods focus on identifying segments with minimal coefficients of variation, indicating high homogeneity within the segment [34]. This approach is beneficial for traffic analysis and pavement condition assessment, as it highlights sections of the road that exhibit consistent characteristics [35]. On the other hand, the change-point detection algorithm (CDA) detects change points in road data by analyzing cumulative area functions and their slope functions [36]. This method effectively identifies abrupt changes in traffic patterns or road conditions, making it valuable for dynamic road environments [30], [37], [38]. The constant length (CL) method segments roads into fixed-length sections, providing a straightforward approach to road analysis [22]. Meanwhile, the coefficient of variation (MCV) method aims to identify

road segments with the minimum coefficient of variation, effectively detecting data groups with high homogeneity [35].

Considering the heterogeneous characteristics of road segmentation data, clustering methodologies have been extensively employed to reveal hidden relationships and complex patterns [39]. Clustering techniques effectively group road segments based on similarity assessments of their attributes and spatial distributions. For instance, Depaire et al. utilized latent class clustering to discern homogeneous collectives of road segments [40], while Luca et al. employed C-mean clustering to identify sets of road segments for subsequent modeling endeavors [41]. Kumar and Toshniwal adeptly used K-mode clustering and association rule mining to identify the circumstances underlying road segment characteristics [42], illuminating distinctive trends and latent patterns within different clusters. Ghadi et al. applied K-means clustering to stratify road segments into spatial clusters, facilitating the identification of high-risk segments [43]. These clustering techniques have proven invaluable in road segmentation by enhancing the understanding of segment characteristics, such as traffic flow and surface conditions while facilitating the identification of patterns that inform targeted maintenance strategies and optimize resource allocation [44]. In addition, road segmentation methods based on spatial heterogeneity have been developed to address the complex nature of road networks [45]. Song et al. proposed a spatial heterogeneity-based segmentation (SHS) model that achieves higher homogeneity and heterogeneity among road segments compared to other methods, offering improved management of infrastructure systems [46]. These methodologies collectively contribute to a more nuanced understanding of road segmentation, enabling more effective analysis and management of road networks.

B. Gaussian Mixture Segmentation

Gaussian Mixture Models (GMMs) have gained significant traction in segmentation tasks due to their ability to effectively model complex distributions [47]. In image segmentation, GMMs offer a probabilistic framework that adeptly handles multimodal data distributions, making them particularly suitable for scenarios with substantial variations in pixel intensity [48]. Early implementations of GMMs employed the Expectation-Maximization (EM) algorithm to model pixel intensity distributions, yielding robust results across various applications, including medical imaging and object recognition [49].

To enhance segmentation performance, researchers have increasingly integrated spatial information alongside color or intensity data. Techniques that combine GMMs with Markov Random Fields (MRFs) have demonstrated significant potential in enforcing spatial coherence among segmented regions, thereby improving segmentation accuracy [50]. Furthermore, hierarchical GMMs enable multi-scale representations that capture varying levels of detail, allowing models to adapt to local geometric characteristics. This advancement is particularly valuable for applications such as point cloud segmentation and 3D reconstruction, where capturing the intricacies of spatial relationships is crucial [51].

The integration of GMMs with deep learning frameworks has significantly enhanced segmentation capabilities [52]. Utilizing Convolutional Neural Networks (CNNs) for feature extraction alongside Gaussian Mixture Models (GMMs) for probabilistic modeling, researchers have developed hybrid approaches that leverage the strengths of both methodologies [53]. This combination has enhanced segmentation performance, especially in challenging visual conditions. Wang et al. demonstrated the potential of Gaussian mixture models in complex scenarios, highlighting their effectiveness in object detection and localization [54]. Modeling the distribution of object instances with Gaussian mixtures allows these methods to adapt to various scales and occlusions, thereby enhancing the robustness of segmentation results.

C. Deep Learning-Based Segmentation Methods

Recent advancements in deep learning have greatly improved road segmentation based on images, leading to various innovative methodologies. Key architectures include Fully Convolutional Networks (FCN), U-Net, and DeepLab. FCNs introduced end-to-end training for pixel-wise predictions but often struggle with fine details, particularly for thin objects like roads [55]. U-Net features a symmetric encoder-decoder structure that captures high-resolution features through skip connections [56]; however, it can produce noisy segmentations in complex environments. DeepLab utilizes convolutions to capture multi-scale context, achieving high accuracy while increasing computational complexity [57]. In addition, Cheng et al. proposed a method that fuses geometric and appearance information to enhance segmentation accuracy [58], while Kong et al. developed a technique for general road detection from a single image, improving the effectiveness of image-based detection methods [59].

Despite these advancements, methods based on deep learning face limitations. A prominent issue is class imbalance, where road pixels vastly outnumber background pixels, leading to poor precision and recall for narrow roads. Many models also tend to overfit, particularly when trained on small datasets, posing challenges in generalization. In addition, the quality of images can significantly impact segmentation accuracy, making it difficult to rely solely on image-based methods. For governments needing to make long-term maintenance decisions, collecting images of all roads results in a massive volume of data, complicating the analysis.

In contrast, the proposed Gaussian Mixture Segmentation (GMS) model offers distinct advantages that enhance road segmentation accuracy and management. The GMS model does not depend on image information; instead, it utilizes the inherent distribution characteristics of road deterioration data, resulting in higher precision less affected by image quality. This approach requires significantly less computation time and eliminates the need to collect vast amounts of image data. Furthermore, the model's adaptability to different environmental conditions, combined with its emphasis on the inherent distribution characteristics of road data, positions it as a powerful tool for managing large-scale road networks.

III. MATERIALS AND METHODS

A. Study Area and Data

This study assesses road deterioration across the South West region of Western Australia, selected for its diverse range of land use and road network characteristics, as illustrated in Fig. 1. Located approximately 300 km south of Perth, the capital of Western Australia, the South West region encompasses nearly all types of terrain found across Australia. It is characterized by diverse topography, fertile farmland, and favorable climatic conditions, making it one of the most densely populated areas in the state. In addition, the region is rich in mineral resources, including iron ore, gold, nickel, and lead-zinc, which necessitates numerous highways for transporting agricultural products and minerals. This diversity in terrain and road network characteristics makes the South West a highly representative area for our study. The effectiveness of our proposed GMS model in this environment suggests its applicability in similar terrains and road network contexts throughout the country, highlighting the substantial research value of the road network in this region.

This study primarily collected road deterioration data using the Traffic Speed Deflectometer (TSD) system in West Australia. Table I presents the road deterioration data collected in the South West Region using the TSD system. A total of 139,894 observations were collected at a resolution of 20 meters, covering 23 highways totaling 2,797.88 km. The South Western Hwy (H009) has the longest length of 520.30 km. Regarding the selected variables for road deterioration, this study focused on curvature and deflection, and Table I provides detailed statistics on the curvature and deflection of all highways. Curvature represents specific shape changes in the road surface under asphalt fatigue. Deflection indicates the maximum surface deflection of the road under standard load, reflecting its strength. It can be observed that the road condition of the M036 highway is the poorest, with an average curvature of $188.9 \mu\text{m}$ and an average deflection of $617.6 \mu\text{m}$.

B. Methods

This study develops a Gaussian mixture segmentation model for managing the deterioration of large-scale road networks. The proposed method's flowchart is illustrated in Fig. 2, consisting of four main steps: data preprocessing, explanatory variable selection for road segmentation, Gaussian mixture segmentation model, and model comparison and validation.

1) *Data Processing*: Data processing organizes and prepares road deterioration data collected through the Traffic Speed Deflectometer (TSD) in our study. Our structured approach converts raw monitoring data into a coherent dataset suitable for subsequent spatial line segmentation analysis. This comprehensive process includes four steps.

The first step involves processing the road deterioration data, which are monitored at a 20-meter interval using an updated Traffic Speed Deflectometer (TSD) system across the road network, without any missing data. The TSD is a heavy vehicle-based road deterioration monitoring system consisting of Doppler lasers, 3D sensors, GPS, cameras, and temperature probes, providing a comprehensive dataset on road conditions.

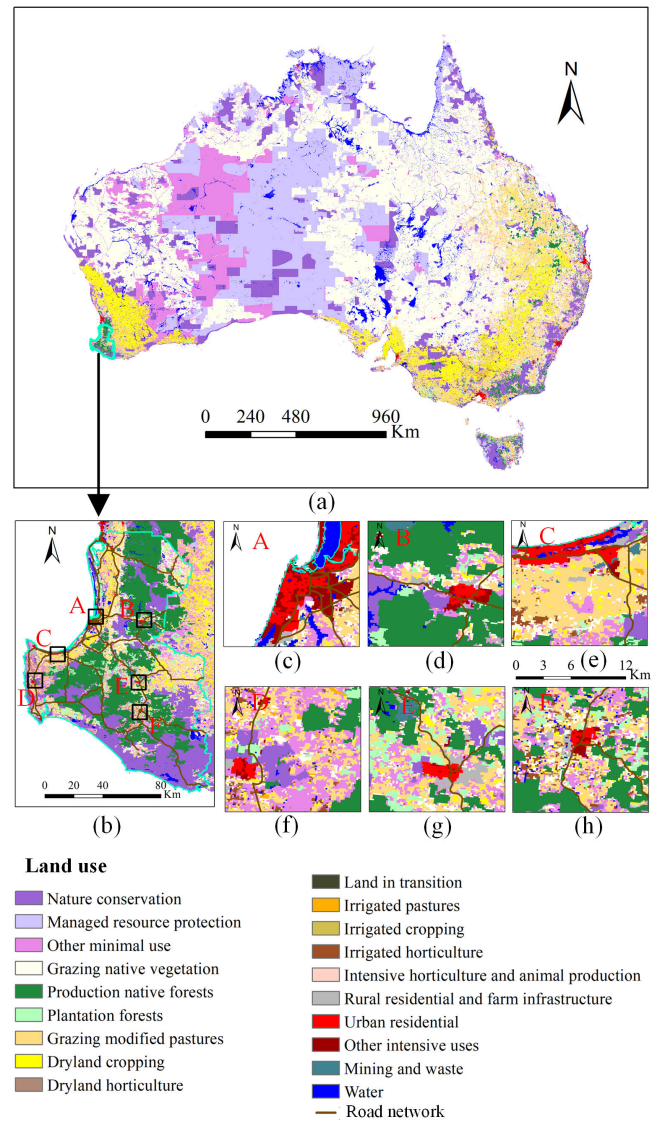


Fig. 1. Location and land use distribution in (a) Australia, (b) South West, (c)-(h) the inner part of South West.

The second step focuses on computing the road lengths for each observation. Each observation represents a 20-meter segment of road conditions examined in the study. We establish a consistent basis for comprehensive analysis by determining the total lengths of the roads, such as calculating the M002 road to be 291.72 kilometers based on its 14,585 sampling points. This calculation allows our segmentation model to assess road deterioration across the entire segment accurately.

Furthermore, the spatial sorting of observation locations was carefully executed, aligning the data points along the roadway to reflect their actual physical positions. This alignment ensures the geographical accuracy of our data, facilitating precise spatial analysis and road deterioration assessment.

Lastly, our data processing included the identification and separate treatment of geographically distinct roads sharing identical identifiers, such as road numbers and names. It's imperative to distinguish between these roads to prevent any analytical ambiguities that might arise from conflating unconnected segments.

TABLE I
A SUMMARY OF OBSERVATIONS OF ROAD DETERIORATION RECORDS

Road No.	Road Name	Length (km)	Curvature (μm)			Deflection (μm)		
			min	max	mean	min	max	mean
H002	Melville Mandurah Hwy	291.72	13.0	298.0	100.0	57.0	763.0	277.1
H009	South Western Hwy	520.30	13.0	970.0	140.2	80.0	2089.0	434.6
H043	Bussell Hwy	164.82	27.0	319.0	130.9	144.0	876.0	378.9
H045	Coalfields Hwy	31.38	29.0	394.0	111.2	143.0	1150.0	369.5
H061	Australind Roelands Link	6.76	37.0	208.0	128.4	150.0	548.0	387.1
M003	Bannister Marradong	22.00	18.0	458.0	131.6	120.0	1404.0	419.4
M006	Bridgetown Boyup Brook	30.54	9.0	429.0	117.4	107.0	1415.0	374.8
M008	Vasse	151.16	12.0	554.0	122.6	104.0	1721.0	390.2
M013	Donnybrook Kojonup	157.22	14.0	448.0	117.8	108.0	1190.0	369.5
M023	Pinjarra	32.96	63.0	185.0	122.9	206.0	813.0	414.5
M024	Muir	146.40	24.0	400.0	141.4	117.0	1832.0	428.7
M027	Brockman	71.70	15.0	315.0	127.7	54.0	1462.0	372.1
M036	Pemberton Northcliffe	27.90	37.0	448.0	188.9	143.0	1591.0	617.7
M037	Collie Lake King	354.28	15.0	466.0	131.0	74.0	1532.0	375.1
M042	Goodwood	37.78	23.0	639.0	110.1	151.0	2119.0	353.4
M043	Caves	111.22	15.0	485.0	131.6	109.0	1252.0	389.4
M046	Collie Mumballup	21.32	36.0	358.0	100.1	121.0	733.0	330.9
M052	Boyanup Picton	16.46	11.0	192.0	117.8	153.0	836.0	435.1
M053	Pinjarra Williams	124.98	18.0	453.0	118.7	68.0	1207.0	349.3
M061	Boyanup Capel	17.32	34.0	194.0	118.1	152.0	876.0	412.2
M066	Collie Williams	63.58	32.0	534.0	142.4	137.0	1278.0	377.8
M080	Stewart Rd	27.60	29.0	227.0	138.9	141.0	581.0	319.1
M081	Sues Rd	58.90	47.0	207.0	125.2	135.0	606.0	327.3

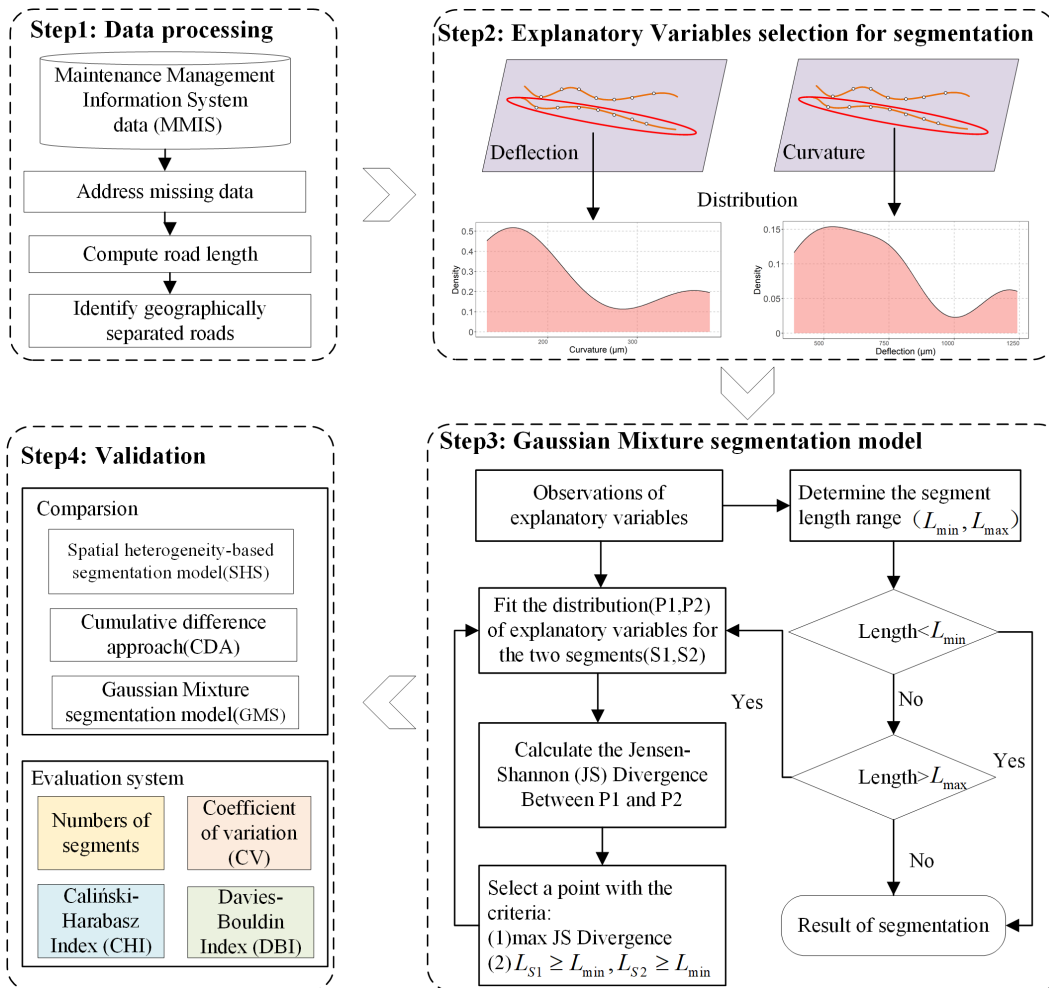


Fig. 2. Flowchart of Gaussian mixture segmentation Model for managing deterioration of large-scale road networks.

2) *Explanatory Variables Collection for Segmentation*: This study considers variables for determining the spatial line of road segments, explicitly aligning with the practical requirements of road construction projects and network

management. These requirements include the need for accurate performance assessment, informed maintenance decision-making, and ensuring the safety and longevity of the road network [60]. To meet these needs, variables such

as curvature and deflection were selected, which are critical indicators of road performance. These variables allow us to quantify the degradation of road segments over time, facilitating timely maintenance actions.

3) *Gaussian Mixture Segmentation Model*: The Gaussian mixture segmentation model is developed because explanatory variables of the same road segment should present similar distributions. It mainly consists of two steps.

In the first step, the spatial line segment data is compared to a predetermined threshold for road segment length. If the length of the segment falls below the minimum requirement, it is considered a valid segment. Conversely, if the length exceeds the maximum requirement, the segment undergoes further division using the approach described in the second step. Appropriate minimum and maximum segment lengths significantly influence effective road deterioration management. Shorter segments facilitate more localized analysis and targeted maintenance strategies, while longer segments offer a broader perspective on the overall network. In this study, we set the minimum segment length to 1 km and the maximum to 5 km based on our understanding of the road network characteristics and discussions with infrastructure management experts. This range enables us to capture the spatial variability in road deterioration while maintaining a reasonable number of segments for practical implementation of deterioration mitigation measures.

In the second step, the road is segmented into two segments by identifying a point of change that maximizes the difference in probability distributions between the two segments. This process aims to meet specific requirements: the length of the segmented road should be within a predefined threshold, and the two segments should exhibit the largest possible difference in the probability distribution of explanatory variables. To measure the dissimilarity between the probability distributions within each road segment, a Gaussian Mixture Segmentation (GMS) model is employed. This model applies a Gaussian mixture distribution to capture the variability in the explanatory variables. The implementation of the Gaussian mixture distribution involves the following steps: First, the explanatory variables are represented mathematically as a weighted sum of multiple Gaussian components. This is formulated as:

$$p(x) = \sum_{k=1}^K \pi_k N(X|\mu_k, \Sigma_k) \quad (1)$$

where $p(x)$ is the probability density function of the explanatory variables, π_k represents the mixing weight of the k -th Gaussian component, and μ_k is the mean of explanatory variables, Σ_k is the covariance matrix, and K is the number of components. Second, the Expectation-Maximization (EM) algorithm was used to estimate the parameters of the Gaussian mixture distribution, including the mixing weights, means, and covariances. In the E-step, the posterior responsibilities of each data point belonging to each Gaussian component was computed:

$$r_{ik} = \frac{\pi_k N(X_i|\mu_k, \Sigma_k)}{\sum_{j=1}^K \pi_j N(X_i|\mu_j, \Sigma_j)} \quad (2)$$

where r_{ik} is the responsibility of the k -th component for the i -th data.

In the M-step: the parameters based on these responsibilities were updated:

$$\pi_k = \frac{1}{n} \sum_{i=1}^n r_{ik} \quad (3)$$

$$\mu_k = \frac{\sum_{i=1}^n r_{ik} X_i}{\sum_{i=1}^n r_{ik}} \quad (4)$$

$$\Sigma_k = \frac{\sum_{i=1}^n r_{ik} (X_i - \mu_k)(X_i - \mu_k)^T}{\frac{1}{n} \sum_{i=1}^n r_{ik}} \quad (5)$$

After updating the parameters, convergence check is performed to determine if the algorithm has stabilized. This involves evaluating the change in log-likelihood between iterations and ensuring that it falls below a predefined threshold ϵ .

$$|\log L(\theta) - \log L(\theta_{previous})| < \epsilon \quad (6)$$

To quantify the dissimilarity between the explanatory variable distributions of the two road segments, this study utilizes the JS divergence as a metric [61]. The advantage of using the JS divergence is that it provides a reliable measure to quantify the dissimilarity or similarity between two probability distributions, allowing for a robust evaluation of the differences in explanatory variable distributions and facilitating accurate segmentation and analysis in various applications [62]. JS divergence is calculated as follows:

$$\begin{aligned} JS(P\|Q) = \frac{1}{2} & \left(\sum p_i(x) \log \left(\frac{2p_i(x)}{p_i(x) + p_j(x)} \right) \right. \\ & \left. + \sum p_j(x) \log \left(\frac{2p_j(x)}{p_i(x) + p_j(x)} \right) \right) \quad (7) \end{aligned}$$

where p_i denotes the probability distribution of the explanatory variables in road segment i , whereas p_j represents the probability distribution of the explanatory variables in road segment j .

The dissimilarity between the explanatory variable distributions of the two road segments can be effectively evaluated by utilizing the JS divergence, providing insight into the segmentation quality and the degree of dissimilarity between the identified segments. If the segmented data exceeds the specified maximum segment length, the segmentation process will be iterated until all segment lengths conform to the predetermined length threshold.

4) *Validation for Segmentation*: This study evaluates the segmented road segments based on a set of evaluation metrics, which include the number of segments, coefficient of variation, Caliński-Harabasz Index (CHI), and Davies-Bouldin Index (DBI). These metrics are universal and are widely applied in various segmentation and clustering methods [63], [64], [65]. These metrics are utilized to compare the effectiveness of the segmentation methods employed in this study with four commonly used road segmentation methods: the constant length (CL) method, the minimization coefficient of variation method, the cumulative difference approach (CDA), and the spatial heterogeneity-based segmentation model (SHS).

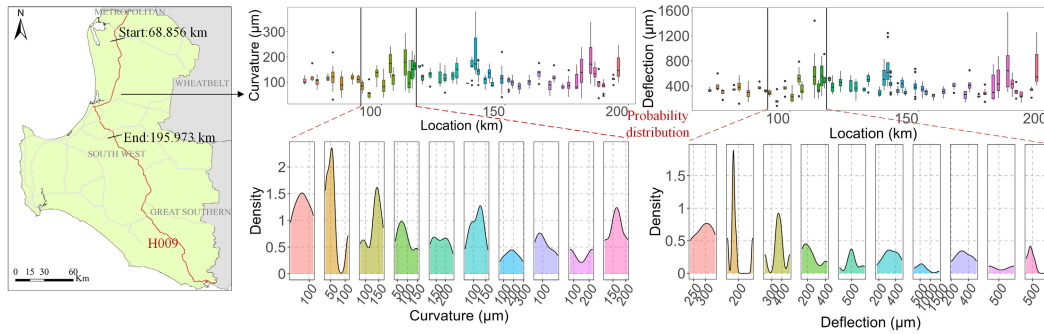


Fig. 4. Statistic distribution of Deflection and Curvature of a selected road in South West Region.

deflection are primarily concentrated in the western and southern regions. Similarly, regions with higher curvature also show a concentration in these areas. Factors such as the abundance of agricultural and mineral resources contribute to this pattern, as heavy vehicles transporting these resources can accelerate the deterioration of road infrastructure. The segmentation and spatial analysis of road deterioration variables provide valuable information for developing targeted maintenance strategies and optimizing resource allocation to manage the road network effectively.

The study focused on Highway H009, which is the longest road running from the southern to the northern regions and serving as a major freight route. Specifically, the study examined the complex traffic area from 68.856 km to 195.973 km along this road. In Fig. 4, the curvature and deflection of each segment in this middle section were plotted, showing variations among segments. The GMS segmentation method revealed significant heterogeneity between segments. In addition, the study observed distinct Gaussian mixture distributions of curvature and deflection by focusing on several segments, indicating that the proposed method effectively identifies the distribution characteristics of explanatory variables around road segments while ensuring clear differentiation between segments.

To further analyze the robustness of the GMS model, an in-depth analysis is conducted of a specific road segment, H002, comparing the segmentation results obtained from GMS with those from four other methods: CDA, SHS, CL, and MCV. Our evaluation focused on the model's performance concerning outliers and abrupt changes, specifically examining the deflection and curvature. As illustrated in Fig. 5, our findings demonstrate that the GMS method effectively distinguishes between high and low values, proving its capability to accurately segment road conditions compared to the other methods. The GMS model exhibits superior resilience to the adverse effects of outliers and unexpected data variations, unlike SHS, CDA, and MCV, which show a propensity for erroneous segmentation under these conditions. These methods often mis-segment by dividing a continuous segment into two based on isolated extreme values, undermining the integrity of the analysis. In contrast, the enhanced robustness of the GMS model confirms its effectiveness and reliability in identifying data patterns, even in the presence of significant perturbations, making it a strong choice for complex data segmentation tasks.

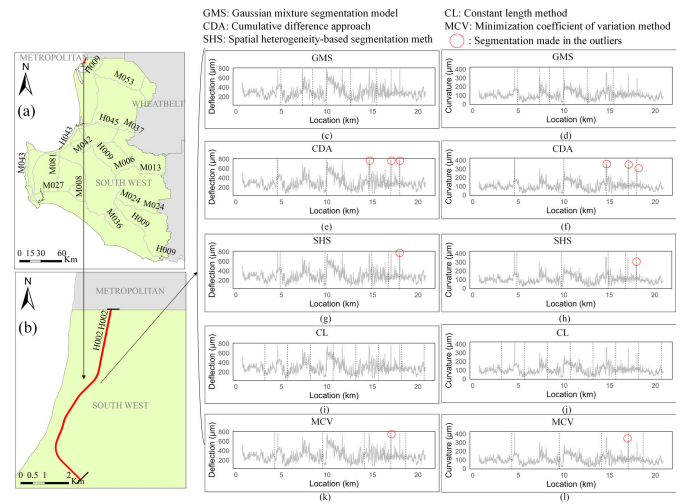


Fig. 5. Comparison of five homogeneous segmentation methods: GMS, CDA, SHS, CL and MCV ((a) and (b): Spatial distribution of selected roads, (c)-(l) deflection and curvature segmented by GMS, CDA, SHS, CL and MCV methods).

This capability is crucial for recognizing deteriorating road segments as single continuous issues, which is essential for effective maintenance strategies.

Lastly, the study also analyzed high-risk situations for each road segment. Segments with curvature greater than 300 micrometers and deflection greater than 700 micrometers are generally classified as high-risk segments. Fig. 6 illustrates the proportion of high-risk segments on each highway, revealing that the M036 segment has the highest proportion of high-risk segments. The location of this road near mining and agricultural areas contributes to its classification. In addition, the limited availability of alternative routes leads to a higher volume of freight transportation on this road, resulting in increased levels of damage.

B. Model Validation

In order to assess the effectiveness of the proposed GMS model, this study utilized an evaluation system to compare it with four commonly used methods: CDA, SHS, CL, and MCV. The assessment methods employed in this study include the number of segments, the CV, the Caliński-Harabasz Index, and the Davies-Bouldin Index.

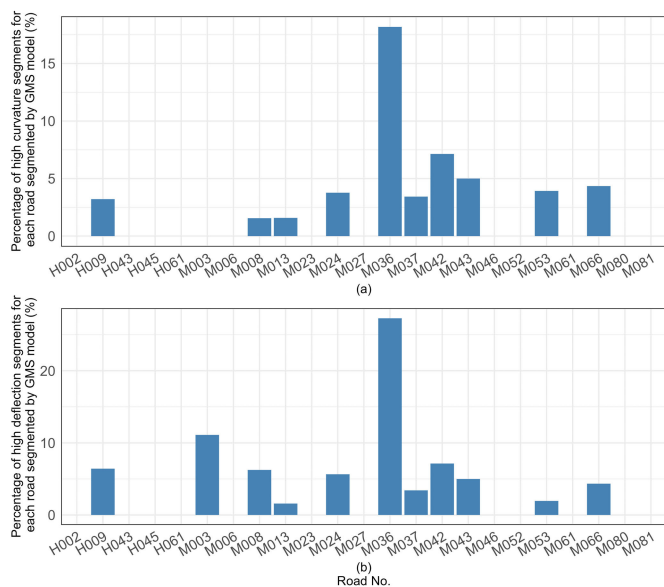


Fig. 6. Comparison of high deterioration risk roads. (a) Percentage of High-curvature-risk segments for each road segmented by GMS model. (b) Percentage of High-deflection-risk segments for each road segmented by GMS model.

Firstly, the number of segments after segmentation was considered. Table II compares the number of segments among the five methods. It can be observed that the GMS method falls between other four methods in terms of the number of segments. Generally, having too many segments can result in excessively short segment lengths, which may hinder the analysis of specific traffic issues. Conversely, if the number of segments is too small, it may fail to ensure homogeneity within each segment. Therefore, the GMS method helps maintain a reasonable number of segments after segmentation.

The second metric under consideration is the CV, which evaluates the homogeneity or consistency within each segment. The percentage of segments with a CV less than 0.25 is commonly employed as a benchmark for the practical assessment of road segmentation methods. As depicted in Table III, a comparative analysis of this percentage across five segmentation methods is provided. The proposed GMS method demonstrates superior performance in terms of this metric. Specifically, the average percentage of segments with a CV lower than 0.25 in the GMS model is 46.2% higher than that of the CDA method and 23.5% higher than that of the SHS method. In addition, when compared to other methods, the GMS model shows a 98.8% improvement over the CL method and a 48.7% improvement over the MCV method. These results highlight the effectiveness of the GMS method in achieving more accurate segmentation. These findings show that the GMS method produces segments with more homogeneous data points, as evidenced by the lower CV values compared to the other methods. Consequently, the GMS approach effectively generates segments with internally consistent patterns or properties, which is a critical aspect of segmentation analysis.

Table III also presents a comparative evaluation of the Davies-Bouldin Index and the Caliński-Harabasz Index across

TABLE II
COMPARISON OF THE NUMBER OF SEGMENTS AMONG THE FIVE SEGMENTATION METHODS

Road No.	Road Name	CL	MCV	CDA	SHS	GMS
H002	Melville Mandurah Hwy	116	93	114	97	104
H009	South Western Hwy	208	220	221	189	218
H043	Bussell Hwy	65	64	63	57	60
H045	Coalfields Hwy	12	11	13	10	12
H061	Australind Roelands Link	2	3	4	3	3
M003	Bannister Marradong	8	9	10	8	9
M006	Bridgetown Boyup Brook	12	10	11	11	11
M008	Vasse	60	59	59	52	57
M013	Donnybrook Kojonup	62	59	59	52	57
M023	Pinjarra	13	12	9	9	9
M024	Muir	58	61	52	52	52
M027	Brockman	28	28	29	25	28
M036	Pemberton Northcliffe	11	11	13	13	13
M037	Collie Lake King	141	145	146	130	144
M042	Goodwood	15	15	14	13	13
M043	Caves Collie	44	31	42	40	40
M046	Mumballup Boyanup	8	9	8	7	7
M052	Picton Pinjarra	6	7	7	7	7
M053	Williams	50	49	48	44	46
M061	Boyanup Capel	6	7	8	6	6
M066	Collie Williams	25	22	23	20	23
M080	Stewart Rd	11	11	11	10	10
M081	Sues Rd	23	24	24	20	23

CL: Constant length method
 MCV: Minimization coefficient of variation method
 CDA: Cumulative difference approach
 SHS: Spatial heterogeneity-based segmentation method
 GMS: Gaussian mixture segmentation model

five distinct road segmentation methods. The results indicate that the road segments delineated by the proposed GMS method are characterized by a higher Caliński-Harabasz Index and a lower Davies-Bouldin Index, relative to those obtained using the other methods. Specifically, the GMS model exhibits a remarkable reduction in the average Davies-Bouldin Index by 59.8% compared to the CDA method and by 25.2% in comparison to the SHS method. In addition, the GMS model shows a 64.4% reduction compared to the CL method and a 41.0% reduction compared to the MCV method. Furthermore, the average Caliński-Harabasz Index for the GMS model shows a substantial improvement, increasing by 98.5% over the CDA method and by 32.0% when contrasted with the SHS method. The GMS model also demonstrates a 131.1% improvement over the CL method and a 23.0% improvement over the MCV method. These results underscore the effectiveness of the GMS method in achieving superior segmentation quality.

These metrics underscore the efficacy of the GMS method in achieving a segmentation that ensures distinct separation between segments and maintains high internal similarity within each segment. The enhanced separation and

TABLE III
A SUMMARY OF COMPARISON OF THREE METRICS AMONG THE FIVE SEGMENTATION METHODS

Road No.	Road Name	Percentage of CV lower than 0.25				The Davies-Bouldin Index				The Caliński-Harabasz Index						
		CL	MCV	CDA	SHS	GMS	CL	MCV	CDA	SHS	GMS	CL	MCV	CDA	SHS	GMS
H002	Melville Mandurah Hwy	43.1%	49.5%	46.4%	52.6%	69.2%	1770.929	835.688	1665.734	1200.103	518.951	500.233	1010.051	401.114	741.124	1025.194
H009	South Western Hwy	21.2%	23.2%	22.2%	23.8%	36.2%	1673.563	1470.816	1506.206	1434.532	1410.129	149.527	270.348	191.187	275.795	281.369
H043	Bussell Hwy	38.5%	40.6%	42.9%	43.8%	61.7%	892.630	338.862	803.058	273.528	238.219	60.6732	94.0012	77.794	84.314	126.512
H045	Coalfields Hwy	8.3%	18.2%	15.4%	20.0%	25.0%	125.278	99.718	115.887	77.761	74.318	9.223	29.377	21.231	28.463	43.831
H061	Australind Roelands Link	50.0%	66.7%	50.0%	66.7%	66.7%	10.401	6.374	10.401	6.374	6.374	6.899	35.660	20.551	35.660	35.660
M003	Bannister Marradong	0.0%	11.1%	10.0%	12.5%	22.2%	152.593	30.544	64.336	23.754	22.436	11.8475	30.552	16.450	25.604	33.688
M006	Bridgetown Boyup Brook	0.0%	0.0%	9.1%	9.1%	18.2%	290.254	178.479	286.207	163.071	50.883	12.162	24.309	28.300	42.365	42.815
M008	Vasse	16.7%	27.1%	18.6%	30.7%	31.6%	664.995	421.586	503.015	330.045	275.782	27.078	50.856	52.194	60.627	96.841
M013	Donnybrook Kojonup	14.5%	18.6%	16.9%	21.0%	21.1%	845.867	506.611	750.704	484.440	480.435	21.967	45.305	33.908	45.350	59.527
M023	Pinjarra	61.5%	66.7%	77.8%	77.8%	88.9%	223.556	33.969	216.097	43.322	26.454	7.857	14.407	8.181	12.873	28.102
M024	Muir	8.6%	21.3%	13.5%	25.0%	48.1%	1343.575	444.214	1104.443	310.770	184.378	31.089	42.487	45.835	55.782	62.705
M027	Brockman	7.1%	21.4%	24.1%	28.0%	32.1%	410.809	328.836	343.467	190.350	172.930	16.705	35.546	40.377	41.707	62.674
M036	Pemberton Northcliffe	9.1%	18.2%	23.1%	30.8%	53.8%	190.281	94.856	180.079	46.546	22.551	9.922	15.629	17.908	20.857	24.339
M037	Collie Lake King	14.9%	17.2%	16.4%	19.2%	21.5%	2113.554	1306.667	1944.437	1150.753	966.919	37.903	59.730	51.126	64.046	89.588
M042	Goodwood	20.0%	20.0%	28.5%	30.7%	38.4%	195.741	175.710	161.497	153.389	73.517	19.996	22.766	32.395	37.871	40.062
M043	Caves	15.9%	29.0%	19.0%	32.5%	35.0%	411.681	303.461	375.874	226.805	200.045	60.117	79.429	53.406	86.855	91.745
M046	Collie Mumballup	0.0%	11.1%	12.5%	14.3%	14.3%	195.987	144.903	88.982	75.417	70.290	20.417	26.655	21.937	41.629	61.514
M052	Boyanup Picton	0.0%	0.0%	14.2%	14.2%	28.6%	82.224	73.417	75.017	46.723	16.315	18.387	26.023	26.923	42.215	45.859
M053	Pinjarra Williams	12.0%	22.4%	15.5%	25.0%	30.4%	463.454	326.721	383.580	233.319	180.965	51.012	71.282	57.660	72.899	94.101
M061	Boyanup Capel	33.3%	42.9%	37.5%	50.0%	50.0%	286.197	123.715	275.097	36.593	18.021	17.376	42.343	25.996	63.607	64.197
M066	Collie Williams	4.0%	4.5%	4.3%	5.0%	13.0%	1907.155	1275.671	1810.552	186.627	114.623	25.847	77.371	59.788	93.912	108.767
M080	Stewart Rd	45.5%	45.5%	72.7%	70.0%	80.0%	94.903	86.903	85.042	83.798	52.234	56.951	101.138	63.678	71.758	166.716
M081	Sues Rd	65.2%	79.2%	75.0%	85.0%	87.0%	355.356	272.147	285.968	229.319	64.464	19.124	32.967	40.007	42.527	69.317

CL: Constant length method

MCV: Minimization coefficient of variation method

CV: Coefficient of variation

CDA: Cumulative difference approach

SHS: Spatial heterogeneity-based segmentation method

GMS: Gaussian mixture segmentation model

homogeneity are advantageous for conducting meaningful analysis and interpretation of the data within the segmented road network. The refined segmentation facilitates a comprehensive understanding of road segments' diverse characteristics and behaviors.

The analysis of the segmented road network proves instrumental for in-depth traffic studies, including examining traffic patterns, identifying congestion hotspots, and scrutinizing travel behavior within specific segments. The actionable intelligence derived from this analysis can be leveraged to optimize traffic flow, pinpoint bottlenecks, and implement efficacious traffic management strategies, thereby improving transportation systems.

V. DISCUSSION

In this study, an innovative approach to road segmentation was introduced, designed to tackle the challenges of managing deterioration within extensive road networks, focusing on spatial line segmentation. A significant advantage of the method lies in its incorporation of probability distributions for variables related to road deterioration. The inherent variability within road segments was captured by accounting for variances among these variables and modeling their probability distributions. The GMS model enabled the definition of more meaningful segment boundaries, enhancing the accuracy of the segmentation process. Notable improvements were observed, evidenced by an increase in the average percentage of segments with a CV lower than 0.25, which was 23.5% to 98.8% higher compared to other methods. In addition, a reduction in the average Davies-Bouldin Index was observed, ranging from 25.2% to 64.4% lower than other models. Improvements in the average Caliński-Harabasz Index showed an increase ranging from 23.0% to 131.1% over other methods. In addition, the computation time of the GMS model under varying total lengths of roads has been discussed. As indicated in Table IV, even when processing road segments totaling up to 500 km

TABLE IV
COMPUTATION TIME FOR DIFFERENT TOTAL LENGTH OF ROADS

Total Length of Road (km)	Computation Time (s)
10	0.03
50	0.11
100	0.29
200	0.73
300	1.78
400	2.70
500	4.54

in length, our algorithm completes the computations in under 5 seconds. This result demonstrates the method's efficiency and scalability, confirming that it can effectively handle larger datasets without significant increases in computation time. These results underscore the effectiveness and efficiency of the GMS method in achieving more accurate segmentation and its robustness in diverse conditions. The GMS model also demonstrated resilience against sudden changes in observations, particularly the presence of outliers, contributing to its robustness.

A key aspect of the study involved applying probability distributions to accurately reflect the underlying characteristics of road variables. Appropriate distributions were selected to model the complex interplay between road infrastructure, traffic dynamics, and environmental conditions, thereby strengthening the robustness of the segmentation process. The use of Gaussian mixture distributions to represent variability in road deterioration data demonstrated the advantages of employing Gaussian models, effectively approximating a wide range of distributions in mixture form. This approach addressed the challenges presented by heterogeneous data. Furthermore, the method's flexibility allows for the incorporation of different probability distribution models based on specific variable characteristics. Utilizing the Expectation-Maximization (EM) algorithm to estimate the parameters of

the Gaussian mixture distribution enhanced the robustness and efficiency of the segmentation model in analyzing road deterioration.

In addition to deflection and curvature, there are numerous factors that influence road deterioration, such as Average Annual Daily Traffic (AADT) and weather conditions. When multiple indicators are available, the entropy weight method can assign weights to these variables. This method involves calculating the entropy value for each indicator to determine its degree of disorder or uncertainty. The entropy weight is then derived by normalizing these values, allowing us to quantify the relative importance of each factor [67]. Integrating these weighted indicators with our GMS model enables a more precise segmentation of road deterioration. This approach enhances the accuracy of our analysis and provides a more comprehensive understanding of the factors contributing to road degradation, ultimately leading to more effective maintenance and management strategies.

The practical implications of our findings are substantial, as our research enhances road maintenance planning by precisely identifying segments with high deterioration rates. The GMS model allows government agencies to effectively prioritize maintenance efforts by pinpointing stretches of road that deteriorate more rapidly due to factors such as traffic volume and environmental conditions. This strategic approach optimizes resource allocation, prolongs infrastructure lifespan, and improves safety and driving conditions. By integrating the GMS model into existing Pavement Management Systems (PMS) and utilizing real-time deterioration data from sensors, agencies can monitor high deterioration segments effectively, which enables better real-time assessment of road conditions, allowing for proactive allocation of maintenance resources to minimize traffic disruptions and boost overall network efficiency. The model's adaptability to different environmental conditions further underscores its potential applicability across varied geographical regions, supporting a transition toward sustainable and resilient transport infrastructure while aiding planners in predicting potential problem areas before they escalate.

While the case study focused on Western Australia's South West region, the Gaussian Mixture Segmentation (GMS) model demonstrates high adaptability for application in other regions and network characteristics. Future research could refine and optimize the GMS model by integrating additional variables into the segmentation process and exploring its applicability to other infrastructure domains. Our case study in the South West region illustrates a diverse range of land use and road conditions, demonstrating that the GMS model performs effectively across various road environments. While this study primarily focuses on road infrastructure deterioration, it is recognized that the theoretical framework of the model can extend to other types of infrastructure, such as bridges and railways, where deterioration patterns may exhibit similar probabilistic characteristics. The potential for developing decision-support systems that utilize data from GMS segmentation is emphasized, which would assist government agencies in making informed long-term maintenance decisions and ensuring effective resource allocation for segments identified

as high deterioration areas. Future research can enhance the applicability and impact of the GMS model across multiple infrastructure contexts by exploring these avenues.

VI. CONCLUSION

In this study, a novel Gaussian mixture segmentation (GMS) model was proposed, utilizing the probabilistic distribution of road deterioration data to capture the inherent variability within road segments. The GMS model provides a management approach for road deterioration in large-scale road networks. A thorough evaluation of segmentation results was achieved by integrating evaluation metrics that consider the features of data distribution. This study emphasizes the significance of further research to improve the robustness and adaptability of the GMS model across various transportation contexts.

However, limitations exist, as the model could benefit from considering additional variables such as traffic characteristics and environmental factors, as well as a broader range of infrastructure types. Future work should focus on integrating the GMS model with traffic simulation and predictive maintenance strategies, as well as exploring interdisciplinary collaborations. Expanding its applications to include environmental assessments and sustainable transportation planning will enable the GMS model to contribute to the development of transportation systems that are resilient, sustainable, and efficient, thereby reducing unexpected failures and enhancing maintenance strategies.

REFERENCES

- [1] L. Wang, X. Xue, Z. Zhao, and Z. Wang, "The impacts of transportation infrastructure on sustainable development: Emerging trends and challenges," *Int. J. Environ. Res. Public Health*, vol. 15, no. 6, p. 1172, Jun. 2018.
- [2] Y. Song, D. Thatcher, Q. Li, T. McHugh, and P. Wu, "Developing sustainable road infrastructure performance indicators using a model-driven fuzzy spatial multi-criteria decision making method," *Renew. Sustain. Energy Rev.*, vol. 138, Mar. 2021, Art. no. 110538.
- [3] S. Pan, H. Yan, J. He, and Z. He, "Vulnerability and resilience of transportation systems: A recent literature review," *Phys. A, Stat. Mech. Appl.*, vol. 581, Nov. 2021, Art. no. 126235.
- [4] J.-P. Rodrigue, *The Geography of Transport Systems*. Evanston, IL, USA: Routledge, 2020.
- [5] M. Kohan and J. M. Ale, "Discovering traffic congestion through traffic flow patterns generated by moving object trajectories," *Comput., Environ. Urban Syst.*, vol. 80, Mar. 2020, Art. no. 101426.
- [6] A. P. Afghari, M. M. Haque, and S. Washington, "Applying a joint model of crash count and crash severity to identify road segments with high risk of fatal and serious injury crashes," *Accident Anal. Prevention*, vol. 144, Sep. 2020, Art. no. 105615.
- [7] V. Mavrin, K. Magdin, V. Shepelev, and I. Danilov, "Reduction of environmental impact from road transport using analysis and simulation methods," *Transp. Res. Proce.*, vol. 50, pp. 451–457, Jan. 2020.
- [8] S. A. Blaauw, J. W. Maina, G. A. J. Mturi, and A. T. Visser, "Flexible pavement performance and life cycle assessment incorporating climate change impacts," *Transp. Res. D, Transp. Environ.*, vol. 104, Mar. 2022, Art. no. 103203.
- [9] P. Zhang, W. Yi, Y. Song, G. Thomson, P. Wu, and N. Aghamohammadi, "Geospatial learning for large-scale transport infrastructure depth prediction," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 132, Aug. 2024, Art. no. 103986.
- [10] P. Paplauskas, A. Vaitkus, and R. Simanavičienė, "Road pavement condition index deterioration model for network-level analysis of national road network based on pavement condition scanning data," *Baltic J. Road Bridge Eng.*, vol. 18, no. 3, pp. 70–101, Sep. 2023.

- [11] A. Setyawan, J. Nainggolan, and A. Budiarto, "Predicting the remaining service life of road using pavement condition index," *Proc. Eng.*, vol. 125, pp. 417–423, Jan. 2015.
- [12] Y. Song et al., "Segment-based spatial analysis for assessing road infrastructure performance using monitoring observations and remote sensing data," *Remote Sens.*, vol. 10, no. 11, p. 1696, Oct. 2018.
- [13] A. I. Rifai, S. Handayani, and R. A. Rasyid, "Data mining applied for national road maintenance decision support system," in *Proc. MATEC Web Conf.*, vol. 195, Jan. 2018, p. 4007.
- [14] S. Amin, U. Tamima, and L. E. Amador-Jiménez, "Optimal pavement management: Resilient roads in support of emergency response of cyclone affected coastal areas," *Transp. Res. A, Policy Pract.*, vol. 119, pp. 45–61, Jan. 2019.
- [15] J. Li, S. Zhang, and X. Wang, "Data-driven multi-objective optimization of road maintenance using XGBoost and NSGA-II," *Autom. Construction*, vol. 168, Dec. 2024, Art. no. 105750.
- [16] M. Romano, M. S. Siegel, and H. Y. Chan, "Creating a predictive model for pavement deterioration using geographic weighted regression," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2672, no. 40, pp. 166–175, Dec. 2018.
- [17] B. Wang, W. Zhang, and Y. Li, "A multi-task deep learning approach for lane-level pavement performance prediction with segment-level data," 2024, *arXiv:2408.01967*.
- [18] S. Hu et al., "Urban function classification at road segment level using taxi trajectory data: A graph convolutional neural network approach," *Comput., Environ. Urban Syst.*, vol. 87, May 2021, Art. no. 101619.
- [19] F. Thomas, "Statistical approach to road segmentation," *J. Transp. Eng.*, vol. 129, no. 3, pp. 300–308, May 2003.
- [20] S. Cafiso and A. Di Graziano, "Definition of homogenous sections in road pavement measurements," *Proc.-Social Behav. Sci.*, vol. 53, pp. 1069–1079, Oct. 2012.
- [21] H. Cui, J. Dong, M. Zhu, X. Li, and Q. Wang, "Identifying accident black spots based on the accident spacing distribution," *J. Traffic Transp. Eng.*, vol. 9, no. 6, pp. 1017–1026, Dec. 2022.
- [22] M. Ghadi and Á. Török, "A comparative analysis of black spot identification methods and road accident segmentation methods," *Accident Anal. Prevention*, vol. 128, pp. 1–7, Jul. 2019.
- [23] J. Lu, A. Gan, K. Haleem, and W. Wu, "Clustering-based roadway segment division for the identification of high-crash locations," *J. Transp. Saf. Secur.*, vol. 5, no. 3, pp. 224–239, Sep. 2013.
- [24] M. Bonera, B. Barabino, G. Yannis, and G. Maternini, "Network-wide road crash risk screening: A new framework," *Accident Anal. Prevention*, vol. 199, May 2024, Art. no. 107502.
- [25] B. Barabino, M. Bonera, G. Maternini, F. Porcu, and R. Ventura, "Refining a crash risk framework for urban bus safety assessment: Evidence from Sardinia (Italy)," *Rel. Eng. Syst. Saf.*, vol. 245, May 2024, Art. no. 110003.
- [26] Q. Cai, M. Abdel-Aty, J. Lee, L. Wang, and X. Wang, "Developing a grouped random parameters multivariate spatial model to explore zonal effects for segment and intersection crash modeling," *Anal. Methods Accident Res.*, vol. 19, pp. 1–15, Sep. 2018.
- [27] G. Fountas, P. C. Anastopoulos, and F. L. Mannering, "Analysis of vehicle accident-injury severities: A comparison of segment-versus accident-based latent class ordered probit models with class-probability functions," *Anal. Methods Accident Res.*, vol. 18, pp. 15–32, Jun. 2018.
- [28] A. Sadeghi, E. Ayati, and M. Pirayesh Neghab, "Identification and prioritization of hazardous road locations by segmentation and data envelopment analysis approach," *PROMET-Traffic Transp.*, vol. 25, no. 2, pp. 127–136, Jan. 1970.
- [29] H. B. Tahir, S. Washington, S. Yasmin, M. King, and M. M. Haque, "Influence of segmentation approaches on the before-after evaluation of engineering treatments: A hypothetical treatment approach," *Accident Anal. Prevention*, vol. 176, Oct. 2022, Art. no. 106795.
- [30] S. R. Nair and B. K. Bhavathrathan, "Hybrid segmentation approach to identify crash susceptible locations in large road networks," *Saf. Sci.*, vol. 145, Jan. 2022, Art. no. 105515.
- [31] L. Janani, V. Sunitha, and M. Samson, "Dynamic segmentation of smartphone sensor-based pavement functional condition," *Can. J. Civil Eng.*, vol. 50, no. 7, pp. 571–583, Jul. 2023.
- [32] Z. Li, D. P. Filev, I. Kolmanovsky, E. Atkins, and J. Lu, "A new clustering algorithm for processing GPS-based road anomaly reports with a Mahalanobis distance," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 7, pp. 1980–1988, Jul. 2017.
- [33] F. Thomas, "Automated road segmentation using a Bayesian algorithm," *J. Transp. Eng.*, vol. 131, no. 8, pp. 591–598, Aug. 2005.
- [34] N. Badshah, K. Chen, H. Ali, and G. Murtaza, "Coefficient of variation based image selective segmentation model using active contours," *East Asian J. Appl. Math.*, vol. 2, no. 2, pp. 150–169, May 2012.
- [35] G. Murtaza, H. Ali, and N. Badshah, "A robust local model for segmentation based on coefficient of variation," *J. Inf. Commun. Technol.*, vol. 5, no. 1, pp. 30–39, 2011.
- [36] Y. Gao, L. Huang, J. Feng, and X. Wang, "Semantic trajectory segmentation based on change-point detection and ontology," *Int. J. Geographical Inf. Sci.*, vol. 34, no. 12, pp. 2361–2394, Dec. 2020.
- [37] W. Liu, J. Qiang Li, W. Yu, and G. Yang, "Change-point detection approaches for pavement dynamic segmentation," *J. Transp. Eng., B, Pavements*, vol. 147, no. 2, Jun. 2021, Art. no. 06021001.
- [38] S. Biswas and K. K. Kuna, "Optimal change-point analysis of pavement condition data for identification of homogeneous sections," *J. Transp. Eng., B, Pavements*, vol. 149, no. 1, Mar. 2023, Art. no. 04022078.
- [39] R. Xu and D. C. Wunsch, "Survey of clustering algorithms," *IEEE Trans. Neural Netw.*, vol. 16, no. 3, pp. 645–678, Jun. 2005.
- [40] B. Depaire, G. Wets, and K. Vanhoof, "Traffic accident segmentation by means of latent class clustering," *Accident Anal. Prevention*, vol. 40, no. 4, pp. 1257–1266, Jul. 2008.
- [41] M. De Luca, R. Mauro, R. Lamberti, and G. Dell'Acqua, "Road safety management using Bayesian and cluster analysis," *Proc.-Social Behav. Sci.*, vol. 54, pp. 1260–1269, Oct. 2012.
- [42] S. Kumar and D. Toshniwal, "A data mining framework to analyze road accident data," *J. Big Data*, vol. 2, no. 1, pp. 1–18, Dec. 2015.
- [43] M. Ghadi, Á. Török, and K. Tánzos, "Integration of probability and clustering based approaches in the field of black spot identification," *Periodica Polytechnica Civil Eng.*, vol. 63, no. 1, pp. 46–52, Oct. 2018.
- [44] M. Q. Ghadi and Á. Török, "Comparison of different road segmentation methods," *PROMET-Traffic Transp.*, vol. 31, no. 2, pp. 163–172, Apr. 2019.
- [45] Z. Zhang, Z. Li, and Y. Song, "On ignoring the heterogeneity in spatial autocorrelation: Consequences and solutions," *Int. J. Geographical Inf. Sci.*, vol. 38, no. 12, pp. 2545–2571, Dec. 2024.
- [46] Y. Song, P. Wu, D. Gilmore, and Q. Li, "A spatial heterogeneity-based segmentation model for analyzing road deterioration network data in multi-scale infrastructure systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 11, pp. 7073–7083, Nov. 2021.
- [47] A. Boulmerka and M. S. Allili, "Foreground segmentation in videos combining general Gaussian mixture modeling and spatial information," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 6, pp. 1330–1345, Jun. 2018.
- [48] J. Liu and H. Zhang, "Image segmentation using a local GMM in a variational framework," *J. Math. Imag. Vis.*, vol. 46, no. 2, pp. 161–176, Jun. 2013.
- [49] F. Riaz et al., "Gaussian mixture model based probabilistic modeling of images for medical image segmentation," *IEEE Access*, vol. 8, pp. 16846–16856, 2020.
- [50] Y. Chen, M. Cai, X. Zhou, C. Ning, C. Cao, and J. Yang, "A robust spatial information-theoretic GMM algorithm for bias field estimation and brain MRI segmentation," *IEEE Access*, vol. 8, pp. 89617–89629, 2020.
- [51] B. Eckart, K. Kim, and J. Kautz, "HGMR: Hierarchical Gaussian mixtures for adaptive 3D registration," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 705–721.
- [52] J. Wang and J. Jiang, "An unsupervised deep learning framework via integrated optimization of representation learning and GMM-based modeling," in *Proc. Asian Conf. Comput. Vis.* Cham, Switzerland: Springer, Jan. 2020, pp. 249–265.
- [53] E. Tatulli and T. Hueber, "Feature extraction using multimodal convolutional neural networks for visual speech recognition," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2017, pp. 2971–2975.
- [54] J. Wang, J. Gao, Y. Yuan, and Q. Wang, "Crowd localization from Gaussian mixture scoped knowledge and scoped teacher," *IEEE Trans. Image Process.*, vol. 32, pp. 1802–1814, 2023.
- [55] O. Ozturk, B. Saritürk, and D. Z. Seker, "Comparison of fully convolutional networks (FCN) and U-Net for road segmentation from high resolution imageries," *Int. J. Environ. Geoinformatics*, vol. 7, no. 3, pp. 272–279, Dec. 2020.
- [56] N. Y. Q. Abderrahim, S. Abderrahim, and A. Rida, "Road segmentation using U-Net architecture," in *Proc. IEEE Int. Conf. Moroccan Geomatics (Morgeo)*, Casablanca, Morocco, May 2020, pp. 1–4.

- [57] M. N. Mahmud, M. K. Osman, A. P. Ismail, F. Ahmad, K. A. Ahmad, and A. Ibrahim, "Road image segmentation using unmanned aerial vehicle images and DeepLab V3+ semantic segmentation model," in *Proc. 11th IEEE Int. Conf. Control Syst., Comput. Eng. (ICCSCE)*, Aug. 2021, pp. 176–181.
- [58] G. Cheng, Y. Qian, and J. H. Elder, "Fusing geometry and appearance for road segmentation," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2017, pp. 166–173.
- [59] H. Kong, J.-Y. Audibert, and J. Ponce, "General road detection from a single image," *IEEE Trans. Image Process.*, vol. 19, no. 8, pp. 2211–2220, Aug. 2010.
- [60] Q. Li, "State of practice: Pavement modelling in Main Roads Western Australia," in *Proc. ARRB Int. Conf.*, Brisbane, QLD, Australia, 2018, pp. 1–14.
- [61] M. Menéndez, J. Pardo, L. Pardo, and M. Pardo, "The Jensen–Shannon divergence," *J. Franklin Inst.*, vol. 334, no. 2, pp. 307–318, 1997.
- [62] S. Dhillon, S. Mallela, and R. Kumar, "A divisive information theoretic feature clustering algorithm for text classification," *J. Mach. Learn. Res.*, vol. 3, no. 7, pp. 1265–1287, Oct. 2003.
- [63] X. Wang and Y. Xu, "An improved index for clustering validation based on silhouette index and Calinski–Harabasz index," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 569, no. 5, Jul. 2019, Art. no. 052024.
- [64] Z. Jalilibal, A. Amiri, P. Castagliola, and M. B. C. Khoo, "Monitoring the coefficient of variation: A literature review," *Comput. Ind. Eng.*, vol. 161, Nov. 2021, Art. no. 107600.
- [65] I. F. Ashari, E. D. Nugroho, R. Baraku, I. N. Yanda, and R. Liwardana, "Analysis of elbow, silhouette, davies-bouldin, calinski-harabasz, and rand-index evaluation on K-means algorithm for classifying flood-affected areas in Jakarta," *J. Appl. Informat. Comput.*, vol. 7, no. 1, pp. 89–97, Jul. 2023.
- [66] M. Mughnyanti, S. Efendi, and M. Zarlis, "Analysis of determining centroid clustering X-means algorithm with davies-bouldin index evaluation," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 725, no. 1, Jan. 2020, Art. no. 012128.
- [67] Y. Zhu, D. Tian, and F. Yan, "Effectiveness of entropy weight method in decision-making," *Math. Problems Eng.*, vol. 2020, pp. 1–5, Mar. 2020.

Pengcheng Zhang received the B.S. and M.S. degrees from the School of Remote Sensing and Information Engineering, Wuhan University, China, in 2019 and 2022, respectively. He is currently pursuing the Ph.D. degree with The Hong Kong Polytechnic University.

His research interests include spatial analysis, construction management, and intelligent transportation systems.

Wen Yi received the Ph.D. degree in construction management from the Department of Building and Real Estate (BRE), The Hong Kong Polytechnic University, Hong Kong, in 2014.

She is currently an Assistant Professor with BRE. Prior to joining The Hong Kong Polytechnic University, she was a Senior Lecturer with Massey University, New Zealand. She had more than two years of postdoctoral experience in Hong Kong and Australia. Her research interests include optimization in construction management, construction safety and health, and construction engineering and management.

Yongze Song (Senior Member, IEEE) received the B.S. and M.S. degrees in surveying and mapping from China University of Geosciences, Beijing, China, in 2012 and 2015, respectively, and the Ph.D. degree from Curtin University, Perth, WA, Australia, in 2019.

He is currently a Senior Lecturer with Curtin University. His research interests include spatial statistics, geospatial methods, sustainable infrastructure, and sustainable development.

Dr. Song is an Associate Editor of the *International Journal of Applied Earth Observation and Geoinformation*, *GIScience and Remote Sensing*, and *All Earth*, and an Editor of *Geoscientific Model Development* and *Journal of Spatial Science*.

Peng Wu received the B.S. degree in project management from Tsinghua University, China, in 2006, the M.S. degree in construction management from Loughborough University, U.K., in 2007, and the Ph.D. degree in project management from the National University of Singapore, Singapore, in 2012.

He is currently a Professor with the Department of Construction Management and an Associate Director of the Australasian Joint Research Centre for Building Information Modeling, Curtin University. His research interests include sustainable construction, lean production and construction, production and operations management, and life cycle assessment.

Dr. Wu received the Discovery Early Career Research Award from the Australian Research Council in 2016 which is a prestigious award to support excellent basic and applied research by early career researchers.

Keith Hampson received the bachelor's degree (Hons.) in civil engineering, the M.B.A. degree from QUT, and the Ph.D. degree from Stanford University. He is currently the CEO of the Sustainable Built Environment National Research Centre (SBENrc), Australia. He is a fellow of Australia's Academy of Technology and Engineering (ATSE), an Honorary Fellow of Engineers Australia, and a fellow of the Australian Institute of Company Directors (AICD). He is recognized world-wide for his inclusive leadership and award-winning collaborative model translating research into practice in the construction industry, and serves as the Immediate Past President for the International Council for Research and Innovation in Building and Construction (CIB).

Ammar Shemery is currently an International Senior Manager with more than 20 years of experience in a variety of sectors, including construction, assets, facilities, engineering, and education, coupled with strong business transformation experience and stakeholder management across government, industry, and research clients. He is also the Chief Operating Officer of the Sustainable Built Environment National Research Centre, Curtin University, Australia.