

Distribution fitting and ANOVA test to analyze pavement sensing patterns for condition assessments

Distribution fitting and ANOVA test

Dada Zhang and Chun-Hsing Ho

*Durham School of Architectural Engineering and Construction,
College of Engineering, University of Nebraska-Lincoln, Lincoln, Nebraska, USA*

Received 13 November 2023
Revised 31 January 2024
Accepted 20 March 2024

Abstract

Purpose – The purpose of this paper is to investigate the vehicle-based sensor effect and pavement temperature on road condition assessment, as well as to compute a threshold value for the classification of pavement conditions.

Design/methodology/approach – Four sensors were placed on the vehicle's control arms and one inside the vehicle to collect vibration acceleration data for analysis. The Analysis of Variance (ANOVA) tests were performed to diagnose the effect of the vehicle-based sensors' placement in the field. To classify road conditions and identify pavement distress (point of interest), the probability distribution was applied based on the magnitude values of vibration data.

Findings – Results from ANOVA indicate that pavement sensing patterns from the sensors placed on the front control arms were statistically significant, and there is no difference between the sensors placed on the same side of the vehicle (e.g., left or right side). A reference threshold (i.e., 1.7 g) was computed from the distribution fitting method to classify road conditions and identify the road distress based on the magnitude values that combine all acceleration along three axes. In addition, the pavement temperature was found to be highly correlated with the sensing patterns, which is noteworthy for future projects.

Originality/value – The paper investigates the effect of pavement sensors' placement in assessing road conditions, emphasizing the implications for future road condition assessment projects. A threshold value for classifying road conditions was proposed and applied in class assignments (I-17 highway projects).

Keywords Road condition assessment, ANOVA, Pavement sensing patterns

Paper type Research paper

1. Introduction

Pavement distress assessment is crucial as it addresses issues caused by various factors, including traffic loading, materials and environmental conditions. This is vital for enhancing the driving experience and reducing traffic accidents. Poor road conditions and anomalies pose a risk of damage and can lead to serious traffic incidents. Consequently, timely pavement maintenance is essential for improving ride quality and transportation safety. The international roughness index (IRI) is typically used as a standard parameter in pavement evaluation, with higher IRI values indicating rougher road conditions (Du *et al.*, 2014; Arhin *et al.*, 2015). However, road condition assessments are challenging for road agencies and institutions due to heavy traffic, high labor and equipment costs, and weather conditions. Nowadays, scholars have employed multiple approaches to detect and assess pavement conditions, including mathematical methods, statistical analysis and machine learning techniques (Li *et al.*, 2019; Yan *et al.*, 2014; Zhao and Nagayama, 2017; Campillo, 2018; Ho *et al.*, 2020). These are based on data collected from mobile applications installed



on smartphones or from self-developed sensors to save costs (Wang *et al.*, 2020; Bhatt *et al.*, 2017). The equipment used for data collection is normally placed inside of the vehicle, and the pavement sensing data include vibration accelerations, time, speed and Global Positioning System (GPS) coordinates. A study by Douangphachanh and Oneyama (2013) demonstrated a linear relationship between vibration acceleration and road roughness. Furthermore, Ho *et al.* (2020) and Chen *et al.* (2019) have shown that acceleration data collected from both smartphone and vehicle-based sensors can be used to detect pavement distress (e.g., cracks) through mathematic algorithms. Therefore, monitoring road conditions using acceleration data has become a popular and common method among road agencies and institutions.

In this paper, we present the methods for assessing pavement conditions based on sensing patterns. The objectives of the paper are (1) to investigate the effect of vehicle-based sensor placement and pavement temperature in pavement distress detection and (2) to determine the threshold values for the identification of pavement conditions.

2. Literature review

Different approaches have been applied to assess road conditions based on acceleration data. For instance, Bridgelall used mathematical methodologies such as derivative and integral to analyze acceleration, velocity and displacement to evaluate pavement roughness. Yan *et al.* (2018) used vertical acceleration signals to determine the crack damage through fast Fourier transform analysis. Similarly, Ye *et al.* (2018) proposed a numerical model based on vertical acceleration that determines potential road conditions through acceleration extrema and frequency distribution. Vertical acceleration has been primarily analyzed when assessing pavement conditions. Harikrishnan and Gopi (2017) proposed a method to monitor the road surface based on vertical acceleration using the Gaussian model. In their study, a smartphone was placed on the vehicle's dashboard at the center, and they found that the threshold for identifying an abnormal event (i.e., pavement distress) varied by speed. Loprencipe *et al.* (2019) conducted a study comparing three approaches to assess pavement roughness: the IRI, road profile classification and vertical acceleration. Their results indicated that the vertical acceleration signal can be used to locate the pavement distress alongside the longitudinal profile.

The combination of statistical models (Bayesian, time series and Markov chain Monte Carlo) and machine learning techniques has garnered attention among scholars for assessing pavement conditions (Hong and Prozzi, 2006; Hunt and Bunker, 2003). Gao *et al.* (2021) presented a series of computational analysis based on the measurement of vertical acceleration using machine learning technology to classify the level of pavement distress. Additionally, an appropriate numerical model can help scholars in making prediction and identifying potential road deterioration. For example, Sandamal and Pasindu (2022) used the acceleration data to generate IRI values for specific pavement distress (e.g., bumps), where the data were collected from a smartphone-based application. The results showed that the proposed method was validated by comparison with real results from the conventional roughness measurement methods. Moreover, vertical acceleration data collected from vehicle response can be applied to predict road roughness through multivariate linear regression model (Wang *et al.*, 2020). Furthermore, Padarthy *et al.* (2020) developed a model to process lateral accelerations and speeds to determine potholes and identify the road anomalies. This model was validated in real-world conditions.

Various methodologies were applied in assessing pavement conditions, most of them used the measurement of acceleration response from single axis (i.e., vertical acceleration). However, when encountering pavement distress (e.g., cracks), the vehicle's response is affected by multiple directions. Therefore, the acceleration data of all three axes should be

considered, whether collected from smartphone application or self-developed sensors. According to the literature review, no study has established a reference threshold for identifying road deterioration using acceleration data. There is a clear need to systematically determine a threshold value to effectively facilitate road condition assessment. While previous studies have indicated that speed may influence the results of classifying pavement conditions, none have addressed how pavement temperature might affect the assessment of road conditions. Additionally, determining adequate threshold values for pavement condition assessments based on acceleration data remains a challenge. There is a need to systematically determine a range of threshold values to adequately support road condition assessment using vibration data and discuss the impact of sensor placement in the pavement condition surveys.

3. Materials and methods

The project collected pavement sensing patterns using a university-owned vehicle traveling along the I-10 corridors in Phoenix, Arizona, from March 2017 to February 2018. Data were collected monthly from two testing road sections, each approximately 3 miles in length. Four self-developed sensors (named M1, M2, M3 and M4) were placed on top of the control arms: M1 and M2 sensors on top of the front control arms of the vehicle, and M3 and M4 on top of the rear control arms of the vehicle. An additional sensor (M5) was placed on top of the cap inside of the vehicle. The vehicle-based sensor system consists of sensor boxes, an ADXL335 accelerometer, an Ada-fruit GPS and a TP-LINK 3G router. During the field test, pavement sensing patterns were captured by a sensor box, which included accelerometers and GPS, and transmitted to a laptop via the router. Concurrently, a GoPro camera was attached to the front of the vehicle to validate the occurrence of cracks, construction joints or road reflectors in the pavement signal data. Road temperature was recorded with an infrared thermometer before driving on the target road sections, and the vehicle's speed was maintained at 60 miles per hour (95 km/hour). The details about vehicle-based sensors and data collection were explained in the reference by [Ho *et al.* \(2020\)](#). In their paper, Ho *et al.* describe the entire development process of vehicle-based sensors and their configuration. The ArcGIS software was applied after the data collection to ensure that all pavement signals were displayed and matched the target road sections (i.e., I-10 corridors) on the GIS map as shown in [Figure 1](#). However, due to unexpected technical issues (a temporary sensor deficiency and malfunction sensing signals attributed to extreme heat in Phoenix), the data from July 2017 had an error ([Ho *et al.*, 2020](#)). Considering that Phoenix's high temperatures do not fluctuate substantially during the summer months (i.e., July to September), the July data were omitted from subsequent analysis.

During the field test, the acceleration data along three axes were captured from the pavement sensing patterns and wirelessly transferred to the computer server. The z -axis is directly used to detect anomalies such as cracks or bumps in the road condition assessment ([Ye *et al.*, 2018](#)), while the magnitude of the x -axis and y -axis represent vehicle maneuvers such as turning and acceleration-stop events ([Hsiao *et al.*, 2012](#)). This paper proposes a comprehensive approach that considers a resultant vector derived from all three axes to represent pavement conditions more accurately. The variable of "M" is named as total magnitude and is expressed as follows ([Zhao and Nagayama, 2017](#); [Ho *et al.*, 2020](#)):

$$M = \sqrt{x^2 + y^2 + z^2} \quad (1)$$

where x , y and z represent the vibration along three axes of x , y and z . The higher magnitude indicates more significant pavement deterioration and brings immediate attention to road agencies and engineers.

BEPAM

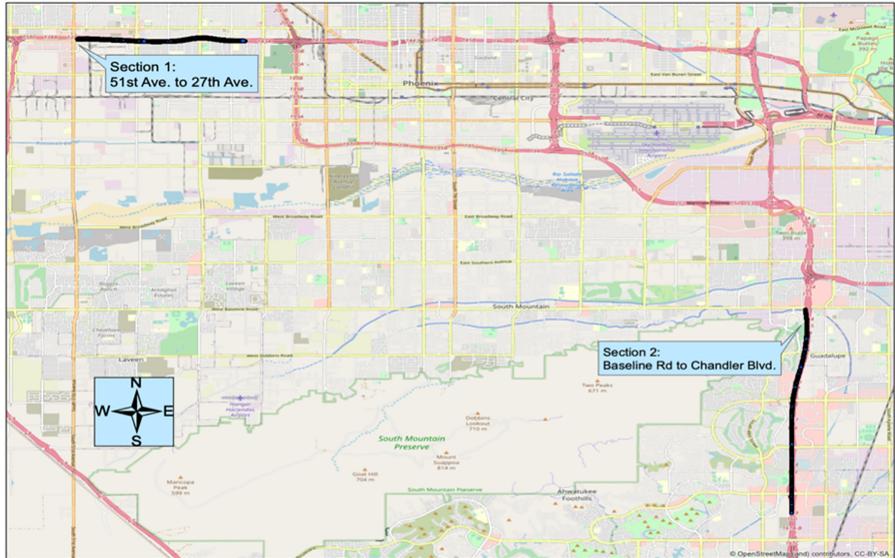


Figure 1.
Testing sections on the I-10 corridor in Phoenix, AZ

Source(s): Figure created by authors

As previously mentioned, sensors M1 to M4 were placed on the control arms, while M5 was placed on the dashboard of the vehicle. The time lag can be determined based on the driving speed and vehicle specifications, and it needs to be minimized to ensure that the vibration responses from all sensors reflect the exact location. To optimize the use of vehicle-based sensors, ANOVA tests were conducted to ascertain whether using a single sensor or all five sensors is appropriate. Before the statistical analysis, it is also essential to check the assumptions of the ANOVA tests, including normality, equal variance, and independence using visualization and statistical tests. **Figure 2** illustrates the magnitude values from all five sensors exhibit a right-skewed distribution by the boxplot visualization, and a data transformation is expected to normalize the distribution and meet the assumptions such as normality. In cases where the data display a positive skew, several transformation methods can be applied, including square root, logarithm and Box-Cox Power ([Olivier and Norberg](#),

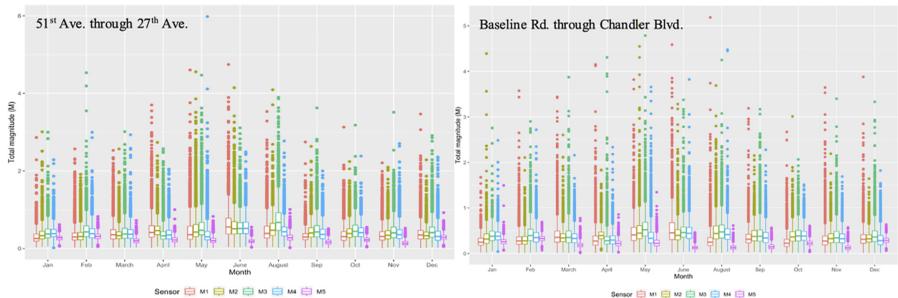


Figure 2.
Boxplots of magnitude values from two road sections

Source(s): Figure created by authors

2010). In this paper, all magnitude values were transformed to a logarithmic scale before conducting the ANOVA test as expressed as follows:

$$f(M) = \log(M) \quad (2)$$

where M is the total magnitude and $f(M)$ is a function to transform M values into a logarithmic scale. The transformation aims to meet the normality assumptions of ANOVA tests and reduces the skewness to ensure the results can be statistically significant.

3.1 Sensor placement

To diagnose the effect of vehicle-based sensor placement in the field, the ANOVA tests were conducted utilizing a cell means model. For instance, if the difference between either M1 and M3 (sensors placed on the left front and rear control arms, respectively) or M2 and M4 (sensors placed on the right front and rear control arms, respectively) is statistically insignificant through the tests, then it is sufficient to place selected sensors on either the left or right control arms at the front and rear of the vehicle. The cell means model is described by Kuehl (2000) and is structured as follows:

$$y_{ij} = \mu_i + e_{ij} \quad (3)$$

where $i = 1, 2, \dots, t$ and $j = 1, 2, \dots, r$. y_{ij} is vibration data of the j^{th} months from the i^{th} sensors, μ_i is the mean magnitude for all acceleration from the i^{th} sensor, e_{ij} is a random error, which should be independent and identically distributed to a normal distribution with zero mean and a constant variance. An expression for the hypothesis test is shown below:

$$H_0 : \mu_i = \mu_j, \text{ where } i \neq j \quad (4)$$

$$H_a : \text{not all the } \mu_i \text{ are equal} \quad (5)$$

$$\text{Test statistics : } F = \frac{MST}{MSE} \quad (6)$$

where MST is the mean square of treatment, and MSE is the mean square of error. μ_i, μ_j are defined the same as shown in Eq. 3.

The post-hoc procedures are indeed conducted after the ANOVA when the null hypothesis has been rejected to determine which specific groups have statistically significant differences. The Bonferroni correction is suggested for pairwise comparisons to control the family-wise error rate due to multiple comparisons. The small p -value (i.e., <0.05) would indicate that there is a significant difference between the groups. Conversely, if the p -value is large (i.e., p -value > 0.05), it implies that any one of the five sensors may be sufficiently reliable for exclusive use in pavement condition assessments.

3.2 Distribution fitting analysis

The paper notes that pavement distress is a point of interest (POI), with the concept of identifying POIs and classifying road conditions based on probability distributions. For instance, it fits all magnitude values of vibration data to various probability distributions, from which a specified percentile can estimate a critical threshold value for classifying pavement conditions. The distribution fitting approach involves selecting an appropriate probability distribution according to the magnitude values. The package `fitdistrplus` in R is utilized to fit multiple parametric distributions, with the best-fit models determined by comparing Akaike Information Criterion (AIC) scores and estimators of distribution parameters, as described by Gareth *et al.* (2013).

where $\ln(\hat{\theta})$ represents the maximum value of the log-likelihood function for the model, and p denotes the number of parameters. Moreover, various plots are generated from the fitdistplus package such as the histogram and theoretical densities plot, probability-probability (P-P) plot, quantile-quantile (Q-Q) plot and cumulative distribution function (CDF) plot, which assists in refining the fitting process and selecting the most appropriate model. After computing the 99th percentile from the fitted models, the corresponding magnitude values are set as the threshold for classifying pavement conditions. The top one percent of the magnitude values may suggest deteriorated pavement surfaces, referred to as POIs in the paper.

4. Results

4.1 Sensor placement in the field test

To satisfy the assumptions required for statistical tests such as ANOVA, the magnitude values were transformed into a logarithmic scale by equation (2) before performing ANOVA tests. In the context of the experimental setup, the confounding variable identified was pavement temperature, which correlates with both the sensors and the pavement sensing patterns. To mitigate the impact of this confounding variable, the study employed randomization tests to evaluate the mean differences. The cell means models (2) were constructed with pavement sensing data collected from the same road segment by all five sensors. In this scenario, the control variable is M5, which is the sensor placed inside the vehicle during the experiment. The results of the ANOVA (Table 1) indicate that the p -values are less than the significance level of 0.05 for both sections, and it suggests that the means of the five sensors differ on a logarithmic scale, as expressed in equations (3)-(5). Implying that each sensor plays a distinct role in collecting pavement sensing patterns. The results for checking assumptions are shown in Table 2, which indicates that the assumptions have not been violated. Skewness and kurtosis are used to determine if the normality assumption is

Table 1.
Results of ANOVA tests for two road sections

	df	Sum of squares	Mean square	F-value	p-value
<i>Analysis of variance table: Section 1 (51st Ave. through 27th Ave.)</i>					
Sensor	4	0.010	0.024	13.42	0.0004
Error	10	0.018	0.002		
<i>Analysis of variance table: Section 2 (Baseline Rd. through Chandler Blvd)</i>					
Sensor	4	0.048	0.012	5.74	0.011
Error	10	0.021	0.002		

Table 2.
Results of ANOVA assumptions

	p-value	Road section 1 Decision	p-value	Road section 2 Decision
Global stat	0.84	Assumptions acceptable	0.89	Assumptions acceptable
Skewness	0.57	Assumptions acceptable	0.88	Assumptions acceptable
Kurtosis	0.82	Assumptions acceptable	0.93	Assumptions acceptable
Link Function	1.00	Assumptions acceptable	1.00	Assumptions acceptable
Heteroscedasticity	0.30	Assumptions acceptable	0.26	Assumptions acceptable

violated, and heteroscedasticity indicates equal variance. The results of these assumptions suggest that all assumptions are acceptable.

Further analysis to identify which sensor(s) might account for the significant difference in pavement condition assessments is presented in Table 3 with the Bonferroni correction applied. The error rate for the Bonferroni correction is 5%, and the adjusted p -value is 0.005 (5%/10 comparisons). According to Table 3, sensor 5 (M5) placed inside the vehicle reveals a significant difference (adjust p -value <0.001) among the other four sensors (M1–M4) placed on the control arms of the vehicle. Consequently, the ANOVA results and Bonferroni correction strongly suggest that the fifth sensor (inside the vehicle) should not be used for pavement detection due to its significantly different vibration magnitudes.

It is also worth investigating whether the sensors could be placed on either side (left or right) or on the front and rear control arms of the vehicle (e.g., M1 and M2). The ANOVA and Bonferroni correction results as shown in Table 4 and Table 5 indicate that the means of magnitudes among the sensors placed on the front wheels differ in log scale. The result shows there is no significant difference between the sensors that are placed on the same side of the

Mixing comparison	Road section 1 p -value	Road section 2 p -value
M1–M2	0.073	0.005
M1–M3	0.010	0.009
M1–M4	0.910	0.600
M1–M5	<0.001	<0.001
M2–M3	0.722	0.097
M2–M4	0.980	0.099
M2–M5	<0.001	<0.001
M3–M4	0.403	0.078
M3–M5	<0.001	<0.001
M4–M5	<0.001	<0.001

Table 3.
Bonferroni correction for pairwise comparison among five sensors

	df	Sum square	Mean square	F -value	p -value
<i>Analysis of variance table: Section 1 (51st Ave. through 27th Ave.)</i>					
Sensor	3	0.014	0.005	5.31	0.016
Error	8	0.007	0.001		
<i>Analysis of variance table: Section 2 (Baseline Rd. through Chandler Blvd)</i>					
Sensor	3	0.015	0.005	5.44	0.025
Error	8	0.007	0.000		

Table 4.
Results of ANOVA tests for sensors placed on control arms (M1–M4)

Mixing comparison	Road section 1	Road section 2
M1–M2	0.003	0.005
M1–M3	0.033	0.054
M1–M4	0.145	0.065
M2–M3	0.209	0.977
M2–M4	0.901	0.993
M3–M4	0.401	0.999

Table 5.
Bonferroni correction for pairwise comparison among four sensors (M1 to M4)

vehicle, such as left and right (e.g., M1 and M3, M2 and M4). Additionally, sensor M5 significantly caused differences among all comparisons. Therefore, in combination with results from Table 2 to Table 5, the evidence supports the use of just two sensors placed on the front control arms of the vehicle for data collection, and M5 should be used individually for data analysis purposes. The findings suggest that utilizing two sensors, as opposed to multiple sensors as in previous work, would be sufficient for assessing pavement conditions in future projects.

4.2 Pavement condition classification

An example calculation using magnitude values from M1 (the sensor on the left front wheel) collected in March 2017 is detailed in Table 6. The fitted lognormal distribution emerged as the better model over the gamma distribution due to its lower AIC score. This is further supported by Figure 3, where the histogram, theoretical density plot, CDF, Q-Q plot and P-P plot all favor the lognormal distribution as a more suitable fit for M1’s data. Subsequently, the estimated parameters of mean and standard deviations from the lognormal distribution are determined (Table 6). With these parameters, a new distribution model was built. After computing the 99th percentile from the fitted model, the corresponding magnitude (critical) value is determined as a threshold value for road condition classification. This procedure

Table 6.
An example of pavement estimation results from fitted distribution models

Fitted “lognormal” distribution by maximum likelihood		
Log-likelihood	74.09	AIC: -144.19
Estimator	meanlog: -1.066	sdlog: 0.575
Fitted “Gamma” distribution by maximum likelihood		
Log-likelihood	17.79	AIC: -31.58
Estimator	shape: 3.177	rate: 7.817

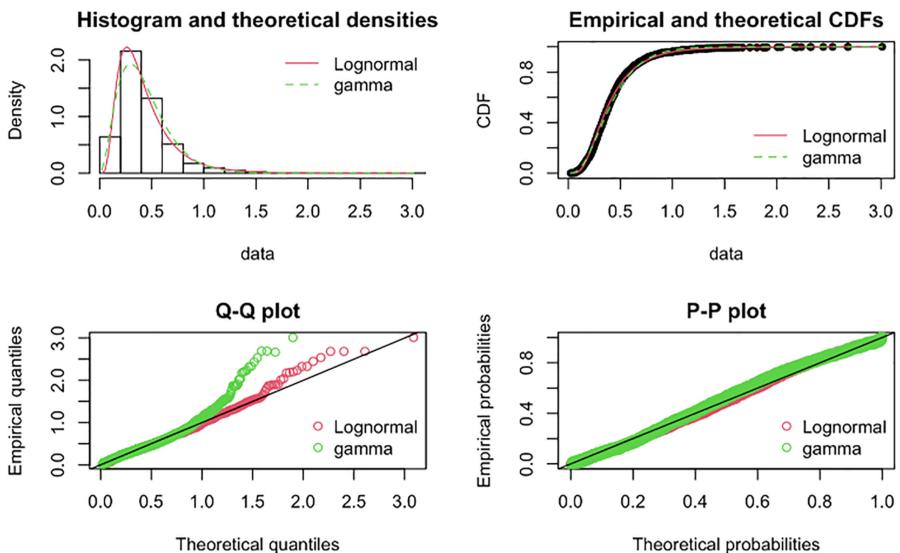
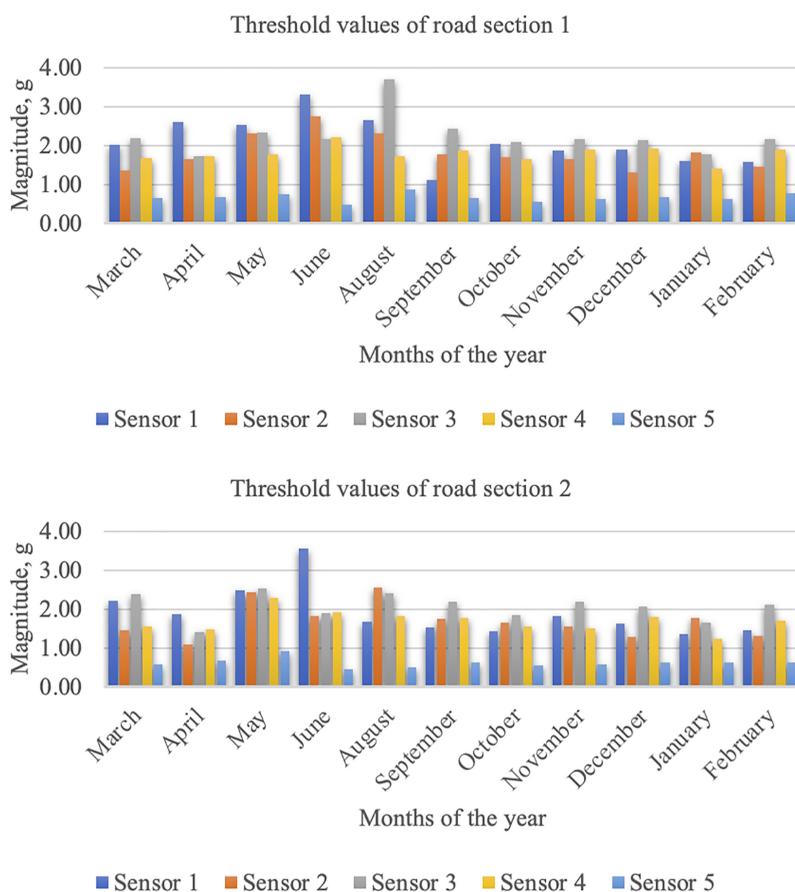


Figure 3.
Plot of fitted distribution model for M1 in March

Source(s): Figure created by authors



Source(s): Figure created by authors

Figure 4. Threshold values of road condition classification

was repeated and applied in the other sensors for all datasets collected from 11 months. Each month, all threshold values computed from the fitted models are shown in Figure 4. As shown in Figure 4, sensor 5 has the lowest threshold since it was placed inside the vehicle rather than on the top control arms in the field test. These threshold values fluctuate monthly, correlating with the pavement temperature changes, as illustrated in Figure 4. It is expected that higher magnitude values will be observed in road condition assessments during the summer compared to the winter. However, this does not imply that roads are rougher in summer than in winter. To further determine a threshold as a reference line to be used in the determination of POIs, all threshold values except for M5 were averaged over 11 months and are summarized in Table 7. It shows the average threshold value for 11 months obtained from sensors M1–M4. The selection of these values is due to the biased results caused by using thresholds from either the front or rear control arms (e.g., 1.8 and 1.6) in the GIS software. Therefore, a threshold of 1.7 was chosen to facilitate the determination of pavement conditions (i.e., POIs), which addresses the objective of determining a threshold value for identifying road conditions.

4.2.1 Temperature effect on pavement performance. The paper aims to investigate the impact of sensor placement and threshold determination for pavement condition assessments by also considering the impact of extreme heat events. A year-long data collection in Phoenix, AZ, was analyzed, and a threshold value of 1.7 g was used to determine POIs. To quantify the influence of pavement temperature on road conditions, two scattered plots were constructed and annotated with their R-squared value, as shown in Figure 5. It shows that the number of identified POIs, which signify pavement distress, varies with changes in pavement temperature. The data exhibit a strong correlation between pavement temperature and the frequency of POIs, with an 83% R-squared value for road section 1 and 70% for road section 2. This substantial correlation underscores temperature's significant effect on the quantity of POIs. It is also notable that the number of POIs increases as the pavement temperature rises. The pattern suggests that travelers on the I-10 corridor are more likely to encounter uncomfortable conditions during the summer months. Consequently, highway agencies are advised to intensify their monitoring efforts of pavement conditions throughout this period and ensure a comfortable and safe riding experience.

4.3 Comparison of POIs and IRI

To further corroborate the accuracy of the findings, georeferenced POIs obtained from the proposed method depicted in Figure 6 were cross-referenced with the IRI data obtained from the Arizona Department of Transportation (ADOT). The poor pavement conditions of IRI were related to IRI values >95 inches/mile. Since the IRI evaluation was performed in the winter, POIs from October through December were extracted and plotted on a GIS map for comparison. As seen in Figure 6, there is a direct accuracy comparison between the POIs and IRI data in the four segments, indicating that the adoption of 1.7 g as a threshold is justified. In previous work (Ho *et al.*, 2020), a threshold value of 2.5 g was used to identify pavement deterioration and the results were underestimated. It is also noted that some POIs appear on the map without corresponding IRI segments, this discrepancy occurs because IRI values are derived from averaging

Table 7. Threshold values for pavement condition classification obtained from different groups

Average threshold obtained from four sensors (M1 to M4)					
Road	M1	M2	M3	M4	Average
Section 1	1.80	2.08	1.59	1.75	1.80
Section 2	1.54	1.98	1.52	1.56	1.65
Average					1.73

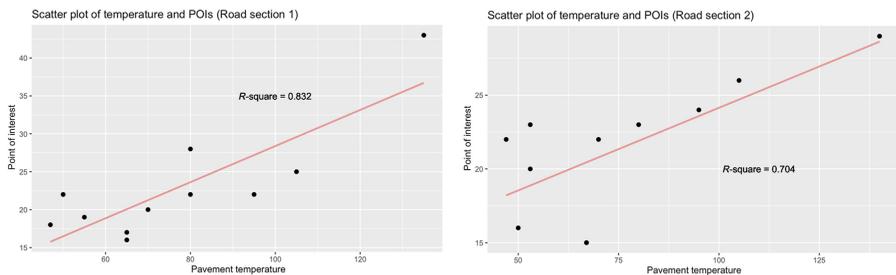
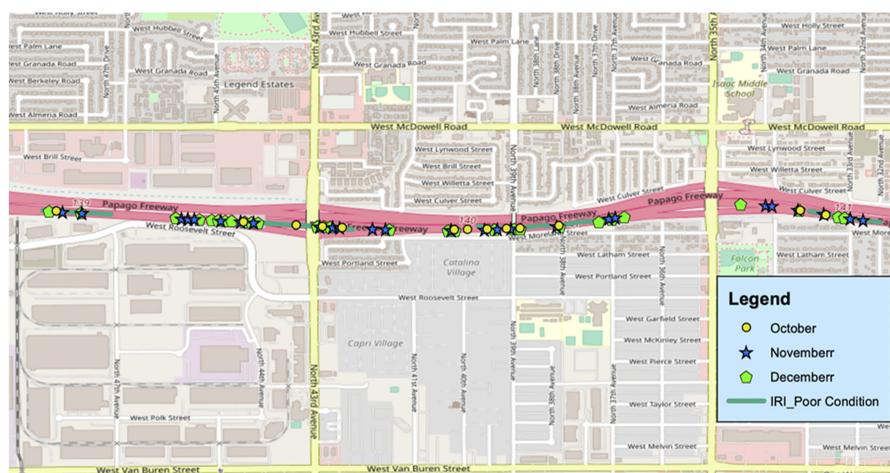


Figure 5. Pavement temperature and identified POIs in two road sections

Source(s): Figure created by authors



Source(s): Figure created by authors

Figure 6. Verification of POIs with IRI segments in GIS map

readings across a segment, which may not capture individual instances of the pavement distress. However, the sensing method studied in the paper employs sensors attached to the control arms, which are capable of detecting such localized distress (POIs), thereby providing a detailed survey of pavement conditions.

5. Discussion

The use of vibration acceleration data from smartphones and sensors to assess pavement conditions has been investigated by numerous scholars and agencies. However, the impact of sensor or smartphone location within a vehicle on the accuracy of pavement condition assessment has not been thoroughly discussed by all researchers. This paper used multiple sensors to collect acceleration data over a year-long period. It investigated the effect of sensor placements in predicting pavement conditions using statistical analysis (i.e., ANOVA test and Bonferroni correction). According to the results, sensor placement on the front control arms significantly influences data accuracy, while the precise side of the vehicle (left or right) where the sensors are mounted does not. If only one sensor is available for data collection, it is recommended to place the sensor inside the vehicle.

Furthermore, the study computed a constant threshold value of 1.7 g for pavement distress (POIs) identification, calculated using distribution fitting and percentile methods. The selected POIs were subsequently validated using a GIS map. The previous studies have yet to fully explore the effects of pavement temperature on pavement condition assessments, and the gap is addressed in the paper, potentially guiding future research in this area. For a new project, a threshold value of 1.7 g could serve as a reference for assessing road conditions.

The sensing technology introduced in the paper offers a cost-effective alternative for highway agencies, engineers and third-party entities needing pavement condition assessment but is constrained by budget limitations. These findings not only enhance the sensor selection and placement process for field testing but also highlight the appropriate timing for pavement condition monitoring, thereby attracting the attention of highway agencies interested in optimizing their maintenance strategies.

6. Conclusions

Through statistical analyses of pavement sensing patterns, the paper has the following conclusions:

- (1) ANOVA results indicate significant differences in the logarithmic mean magnitude values across all five sensors (M1–M5), with all assumptions being met. This means that each of the five sensors plays a distinct role in the assessment of road conditions. The Bonferroni correction was further performed which showed there were no significant differences between sensors positioned on the same side of the vehicle's top control arms (e.g., M1 and M3, or M2 and M4).
- (2) For future road condition assessments, the pair of sensors M1 and M2 is recommended. Utilizing this specific pair of sensors can help reduce the workload while still yielding accurate results. Therefore, the M1 and M2 sensor pair are highly recommended for use in future projects aimed at assessing pavement conditions.
- (3) The sensor placed on top of the dashboard inside the vehicle (M5) should be excluded from analyses that combine data from sensors on the vehicle's front or rear control arms to prevent biased results. Data from M5 should be analyzed separately.
- (4) Lognormal distributions fitting the magnitude values from each sensor have been validated, with the 99th percentile leading to the establishment of a 1.7 g threshold. This threshold is deemed reasonable for classifying poor pavement conditions when cross-referenced with IRI segments provided by ADOT in a GIS map. This threshold value would be a reference in a new project for pavement condition assessment.
- (5) Statistical analysis confirms a significant impact of pavement temperature on pavement condition prediction.

While the study gathered year-long pavement sensing data, it was limited to a single vehicle type used in the field test. Consequently, the computed threshold may not be universally applicable across different vehicle types. The determined threshold values should serve as a reference point for future research, such as employing machine learning techniques to classify and predict pavement distress based on acceleration data. This threshold could facilitate feature extraction in the training of machine learning models to identify cracks and bumps. Looking ahead, the authors anticipate the application of machine learning techniques to enhance predictions of pavement distress.

References

- Arhin, S.A., Williams, L.N., Ribbiso, A. and Anderson, M.F. (2015), "Predicting pavement condition index using international roughness index in a dense urban area", *Journal of Civil Engineering Research*, Vol. 5 No. 1, pp. 10-17.
- Bhatt, U., Mani, S., Xi, E. and Kolter, J.Z. (2017), "Intelligent pothole detection and road condition assessment", *Bloomberg Data for Good Exchange Conference*, Chicago, U.S.
- Campillo, J.R. (2018), *A Simplified Pavement Condition Assessment and its Integration to a Pavement Management System*, Doctoral dissertation, Arizona State University, Tempe, May.
- Chen, C., Seo, H.S., Zhao, Y., Chen, B., Kim, J.W., Choi, Y. and Bang, M. (2019), "Pavement damage detection system using big data analysis of multiple sensor", *Proceedings of International Conference on Smart Infrastructure and Construction 2019 (ICSIC) Driving data-informed decision-making*, U.S, doi: [10.1680/icsic.64669.559](https://doi.org/10.1680/icsic.64669.559).

-
- Douangphachanh, V. and Oneyama, H. (2013), "Using smartphones to estimate road pavement condition", *Proceedings of International Symposium for Next Generation Infrastructure*, Wollongong, Australia.
- Du, Y., Liu, C., Wu, D. and Jiang, S. (2014), "Measurement of international roughness index by using Z-axis accelerometers and GPS", *Mathematical Problems in Engineering*, Vol. 2014, p. 10, doi: [10.1155/2014/928980](https://doi.org/10.1155/2014/928980).
- Gao, J., Ho, C.H., Wiese, I., Zhang, D. and Gerosa, M. (2021), "The pavement condition assessment and prediction system based on machine learning technology", *2021 Annual Meeting of Transportation Research Board*, Washington, D.C., U.S.
- Gareth, J., Daniela, W., Trevor, H. and Robert, T. (2013), *An Introduction to Statistical Learning: With Applications in R*, Springer Publication, U.S.
- Harikrishnan, P.M. and Gopi, V.P. (2017), "Vehicle vibration signal processing for road surface monitoring", *IEEE Sensors Journal*, Vol. 17 No. 16, pp. 5192-5197, doi: [10.1109/jsen.2017.2719865](https://doi.org/10.1109/jsen.2017.2719865).
- Ho, C.H., Snyder, M. and Zhang, D. (2020), "Application of vehicle-based sensing technology in monitoring vibration response of pavement conditions", *Journal of Transportation Engineering, Part B: Pavements*, Vol. 146 No. 3, 04020053, doi: [10.1061/jpeodx.0000205](https://doi.org/10.1061/jpeodx.0000205).
- Hong, F. and Prozzi, J.A. (2006), "Estimation of pavement performance deterioration using Bayesian approach", *Journal of Infrastructure Systems*, Vol. 12 No. 2, pp. 77-86, doi: [10.1061/\(asce\)1076-0342\(2006\)12:2\(77\)](https://doi.org/10.1061/(asce)1076-0342(2006)12:2(77)).
- Hunt, P. and Bunker, J. (2003), "Time series analysis of pavement roughness condition data for use in asset management", *the 3rd Asia-Pacific Conference on Systems Integrity and Maintenance*, Australia, Queensland University of Technology.
- Hsiao, W.C., Horng, M.F., Tsai, Y.J., Chen, T.Y. and Liao, B.Y. (2012), "A driving behavior detection based on a zigbee network for moving vehicles", *2012 Conference on Technologies and Applications of Artificial Intelligence*, Tainan, Taiwan, doi: [10.1109/taai.2012.65](https://doi.org/10.1109/taai.2012.65).
- Kuehl, R.O. (2000), *Design of Experiments: Statistical Principles of Research Design and Analysis*, 2nd ed., Cengage Learning, U.S., pp. 42-47.
- Li, C., Chen, S., Zhao, Y. and Chen, Y. (2019), "Road pavement identification based on acceleration signals of off-road vehicles using the batch normalized recurrent neural network", *IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, pp. 172-177, doi: [10.1109/icaica.2019.8873458](https://doi.org/10.1109/icaica.2019.8873458).
- Loprencipe, G., Zoccali, P. and Cantisani, G. (2019), "Effects of vehicular speed on the assessment of pavement road roughness", *Applied Sciences*, Vol. 9 No. 9, p. 1783, doi: [10.3390/app9091783](https://doi.org/10.3390/app9091783).
- Olivier, J. and Norberg, M.M. (2010), "Positively skewed data: revisiting the box-cox power transformation", *International Journal of Psychological Research*, Vol. 3 No. 1, pp. 68-95.
- Padarthy, M., Sami, M. and Heyns, E. (2020), "Investigation on identifying road anomalies using in-vehicle sensors for cooperative applications and road asset management", *Transportation Research Record*, Vol. 2674 No. 7, pp. 585-595, doi: [10.1177/0361198120923989](https://doi.org/10.1177/0361198120923989).
- Sandamal, R.M.K. and Pasindu, H.R. (2022), "Applicability of smartphone-based roughness data for rural road pavement condition evaluation", *International Journal of Pavement Engineering*, Vol. 23 No. 3, pp. 663-672, doi: [10.1080/10298436.2020.1765243](https://doi.org/10.1080/10298436.2020.1765243).
- Wang, G., Burrow, M. and Ghataora, G. (2020), "Study of the factors affecting road roughness measurement using smartphones", *Journal of Infrastructure Systems*, Vol. 26 No. 3, doi: [10.1061/\(asce\)is.1943-555x.0000558](https://doi.org/10.1061/(asce)is.1943-555x.0000558).
- Yan, G., Wang, L., Ye, Z. and Wang, W. (2014), "Effects of crack damage on acceleration response of asphalt pavement via numerical analysis", *Mathematical Problems in Engineering*, Vol. 32 No. 7, 04020163, doi: [10.1061/\(asce\)mt.1943-5533.0003221](https://doi.org/10.1061/(asce)mt.1943-5533.0003221).

BEPAM

- Yan, G.N., Ye, Z.J., Guo, M. and Wang, L.B. (2018), "Study on integral processing of pavement vibration acceleration", in *Advances in Materials and Pavement Prediction*, 1st ed., CRC Press, pp. 83-87.
- Ye, Z., Lu, Y. and Wang, L. (2018), "Investigating the pavement vibration response for roadway service condition evaluation", *Advances in Civil Engineering*, Vol. 2018, p. 14, doi: [10.1155/2018/2714657](https://doi.org/10.1155/2018/2714657).
- Zhao, B. and Nagayama, T. (2017), "IRI estimation by the frequency domain analysis of vehicle dynamic responses", *Procedia Engineering*, Vol. 188, pp. 9-16, doi: [10.1016/j.proeng.2017.04.451](https://doi.org/10.1016/j.proeng.2017.04.451).
-

Corresponding author

Dada Zhang can be contacted at: dzhang19@huskers.unl.edu

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgroupublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com