

Calibration of Smartphone Sensors to Evaluate the Ride Quality of Paved and Unpaved Roads

Xinyi Yang^a, Liuqing Hu^a, Hafiz Usman Ahmed^a, Raj Bridgelall^b and Ying Huang^{a*}

^aDepartment of Civil and Environmental Engineering, North Dakota State University, Fargo, USA; ^bDepartment of Transportation, Logistics and Finance, North Dakota State University, Fargo, USA

ABSTRACT

Transportation agencies report that millions of crashes are caused by poor road conditions every year, which makes the localization of roadway anomalies extremely important. Common methods of road condition evaluation require special types of equipment that are usually expensive and time-consuming. Therefore, the use of smartphones has become a potential alternative. However, differences in the sensitivity of their inertial sensors and their sample rate can result in measurement inconsistencies. This study validated those inconsistencies by using three different types of smartphones to collect data from the traversal of both a paved and an unpaved road. Three calibration methods were used including the reference-mean, reference-maximum, and reference-road-type methods. Statistical testing under identical conditions of device mounting using the same vehicle revealed that the roughness indices derived from each device and road type are normally distributed with unequal means. Consequently, applying a calibration coefficient to equalize the means of the distributions of roughness indices produced from any device using the reference mean method resulted in consistent measurements for both road types.

ARTICLE HISTORY

KEYWORDS

Road roughness; Road Impact Factor; noise reduction; calibration methods; margin of error

1. Introduction

There are more than four million miles of the highway within the United States (NTSB 2018), all of which carry 58.8% of the total travel through roadways. There is risk associated with road travel due to severe road roughness from excessive loading, chemical reactivity, traffic abrasion, aging, and others (Agbovi 2012, Cebon 1998). Rougher roadways may lead to vehicle delays, potential accidents, and injuries resulting in death. Among twenty two different influencing factors such as congestion, traffic density, super-elevation, insufficient sight, water, icy conditions, narrow shoulder or road, and ramp speed, the poor road surface condition (ride quality) was ranked as the major contributor for car crashes (Zaloshnja and Miller 2009). Therefore, it is essential for state agencies to assess the road quality and fix road anomalies.

Agencies use different devices to measure the road roughness, such as the level pavement profiler, profilograph, high-speed inertial profilers, lightweight inertial profilers, and vehicles equipped with smartphones sensors (Smith and Ram 2016). An associated pavement management system (PMS) (Golabi and Way 1982, Chan et al. 2010) stores a roughness index such as the international roughness index (IRI) (Paterson 1986, Park et al. 2007, Pawar et al. 2018, Múčka 2017) and the profilograph index (PrI) (Múčka 2015, Leverett 2008). These roughness indices provide important indicators of road surface performance in terms of their ride-quality (Brown et al. 2010). However, deriving those indices requires particular equipment driving at a fixed speed and trained staff, resulting in high cost and long labor time.

Therefore, using smartphones onboard of conventional vehicles has become a potential alternative. In recent years, technologies in smartphones are developing very rapidly. Sensors are commonly embedded inside smartphones, such as accelerometer, gyroscope, magnetometer, microphone, and camera. Many sensing-based applications based on embedded sensors inside smartphones become available not only to improve the conveniences of lives but also to provide various

possibilities for research (Liu 2013). Assessing road condition using smartphones is one among these researches. The use of smartphones can detect or localize the road surface anomalies such as potholes, bumps, braking, and honking (Ghose et al. 2012, Lekshmiathy et al. 2020, Mednis et al. 2011, Mukherjee et al. 2016, Seraj et al. 2015). For example, the Pothole Patrol system was developed to monitor road surface anomalies with accelerometers and GPS sensors embedded in the smartphone, it tackled the challenge of localizing the potholes and differentiates potholes from other road anomalies (Eriksson et al. 2008). On the other hand, smartphones can also be used to evaluate the road roughness (Aleadelat et al. 2018, Alessandroni et al. 2015, Buttlar and Islam 2014, Douangphachanh and Oneyama 2016, Janani et al. 2020, Kumar et al. 2017). For example, the NeriCell system focused on monitoring road roughness and traffic conditions using smartphones with data from the accelerometer, microphone, GSM Radio, and GPS sensors (Mohan et al. 2008).

More recently, to assess the road using smartphones more effectively, a Road Impact Factor (RIF) was developed to correlate the inertial and geospatial position data collected from smartphones to the common road roughness factor, IRI, under identical measurement conditions, (Bridgelall 2013, 2017). When using smartphones for roughness measurements, the detection and isolation of the sensor fault is extremely important to avoid the unacceptable deviations between the real and predicted sensor output values (Bhushan and Rengaswamy 2020, Huang et al. 2017a, Huang et al. 2017b, I-Iarte et al. 2005, Kronic et al. 2007). In between smartphones from different models or brands, the sensing characteristics of embedded inside smartphones, such as resolution, sampling frequency, and accuracy, vary significantly. Therefore, even for the same road segments, the roughness detected from the use of different uncalibrated smartphone models will result in measurement discrepancies and inconsistencies (Darawade et al. 2016). Calibrations between smartphone models are necessary before the wide-spread use of smartphones in road roughness measurements.

Statistically, calibration can be performed through either the

univariate or the multivariate methods. The commonly used univariate calibration methods include the classical, inverse, Bayesian, and non-parametric approaches together with the approaches via tolerance regions. For multivariate calibration, it considers multiple linear regression, the ridge regression, partial least squares regression, and other Bayesian and profile likelihood approaches (Osborne 1991). For whichever calibration methods used, the key for a model calibration is to adjust the parameters to obtain appropriate calibration constants either univariate or multivariate based on natural state and historical data to improve the model prediction in between measurements from different devices (O'Sullivan and O'Sullivan 2016). There are a variety of methods to calculate the calibration constants depending on the study fields and objectives. For an effective calibration, the calibration constants can also be compared for its correlate and transferability of measures from various equipment with the standards and guidelines if available (Sayers 1986, Ghasemi et al. 2014). The calibrations between various smartphone models for road roughness measurement are needed but not yet developed.

To provide guidance on calibration in between different smartphones with various embedded sensors for road roughness measurements and address the inconsistency from different smartphones, this study investigates the discrepancies and evaluates the effects of using three different univariate calibration methods, including formulating calibration constants from reference smartphones using the average mean road roughness, the maximum road roughness, and roughness from different road surface types. Field experiments were performed to validate the effectiveness of the three calibration methods using three different smartphones with two different operation considered including the iOS® for iPhones and the Android system for Google Pixel, which have three different sensing specifications for accelerometers in road roughness measurements. The statistical testing demonstrated that the accelerometers in the three different smartphones yield very different sensitivities for roughness measurement. With calibrations using either of the calibration methods, the smartphones can be an effective tool that produces consistent results in the analysis of ride-quality indices for both paved and unpaved roads.

The organization of the remainder of this paper is as follows: Section II introduces the road impact factor RIF-index and the applied three approaches to calibration; Section III describes the experimental design, the method of data collection, and signal processing; Section IV evaluates each calibration method to determine the best approach; Section V concludes the work and suggests future applications.

II. Road impact factor and calibration methods

This study quantified the ride quality of either paved or unpaved roads by applying a Road Impact Factor (RIF) transform to the inertial and geospatial position data collected and the RIF-transform produces RIF-indices, which is linearly proportional to the common road roughness factor, IRI, under identical measurement conditions (Bridgelall 2013, 2017). With a smartphone mounted onboard of a regular vehicle, the Road Impact factor (RIF) transform ($R_{\Delta L[w]}$) for a certain

distance window (w) along the traversal path and window size (ΔL) integrates a product of the n th vertical acceleration signal $a_z[n]$ and longitudinal velocity $v[n]$ measured from the smartphone as follows (Bridgelall 2013):

$$R_{\Delta L[w]} = \sqrt{\frac{1}{\Delta L[w]} \sum_{n=0}^{N_w-1} (g_z[n]v[n])^2 \Delta t[n]} \quad (1)$$

$$\text{where } g_z[n] = a_z[n]/g \quad (2)$$

and g is the earth-gravity factor equal to $9.81 \text{ m}\cdot\text{s}^{-2}$.

Within each window (ΔL), n is the sampling instant. The index of the first and last samples of each window is 0 and N_w-1 , respectively, and N_w is the number of samples in window w . Since the smartphones do not sample the signals at a uniform rate, the sampling time, ΔT , varies with each sample. The RIF-index is interpreted as the average g -force magnitude experienced per unit of distance L traveled. The segment RIF-index, $R_{\Delta L}$, is the mean of the RIF-indices across all windows of the entire section, which can be given as:

$$R_{\Delta L} = \frac{1}{N_w} \sum_{w=1}^{N_w} R_{\Delta L[w]} \quad (3)$$

Since different smartphones have various sensitivity and sampling frequency on the vertical acceleration signal $a_z[n]$ or $g_z[n]$ and longitudinal velocity $v[n]$, the measured RIF-index $R_{\Delta L[w]}$ and the resulted mean RIF-indices ($R_{\Delta L}$) will vary between smartphones with different operation systems. These variances are the source of the inconsistency of measured road roughness RIF-index induced in between different smartphones.

To calibrate different smartphones for a consistent RIF-index, this study evaluated three calibration methodologies, including 1) the reference mean method which calibrates smartphone signals toward the mean RIF, 2) the reference max method which calibrates the smartphone signals toward the maximum RIF, and 3) the reference road type method which uses the calibration coefficient determined for the paved road as the calibration coefficient for unpaved roads. The details of each calibration method are as follows:

A. Reference Mean

This method calculates the calibration constant from the mean of the segment RIF-indices (C_{mean}) across all traversals of a reference phone as:

$$C_{\text{mean}} = \frac{\frac{\sum_0^N RIF_{\text{mean},i,n}}{N}}{\frac{\sum_0^N RIF_{\text{mean},ref,n}}{N}} \quad (4)$$

in which $RIF_{\text{mean},i,n}$ is the mean RIF from the i^{th} smartphone from the n^{th} traversal and $RIF_{\text{mean},ref,n}$ is the mean RIF from the reference smartphone from the n^{th} traversal. N is the number of traversals of field calibration test and $0 < n < N$. Subsequently, scaling the segment RIF-indices from the other phones by the ratio of its mean segment RIF-index and the calibration constant, C_{mean} , to perform the calibration as:

$$RIF_{\text{mean},i,\text{cali}} = \frac{RIF_{\text{mean},i,\text{uncali}}}{C_{\text{mean}}} \quad (5)$$

where, $RIF_{\text{mean},i,\text{cali}}$ is the calibrated RIF from i^{th} smartphone

using the Ref-mean method and $RIF_{mean,i,uncali}$ is the uncalibrated RIF from i^{th} smartphone.

B. Reference Max

This method calculates the calibration constant from the maximum segment RIF (C_{\max}) across all traversals of a reference phone as:

$$C_{\max} = \frac{\frac{\sum_0^N RIF_{max,i,n}}{N}}{\frac{\sum_0^N RIF_{max,ref,n}}{N}} \quad (6)$$

in which, $RIF_{max,i,n}$ is the maximum RIF from the i^{th} smartphone from the n^{th} traversal and $RIF_{max,ref,n}$ is the maximum RIF from the reference smartphone from the n^{th} traversal. Subsequently, scaling the segment RIF-indices from the other phones by the ratio of its maximum segment RIF-index and the calibration constant, C_{\max} , to achieve the calibration as:

$$RIF_{mean,i,cali} = \frac{RIF_{mean,i,uncali}}{C_{\max}} \quad (7)$$

C. Reference Road Type

Since the paved roads are generally much smoother than unpaved roads and the larger vibrations from unpaved roads will likely emphasize any non-linearities in the sensing system, the use of a calibration constant from paved roads may produce a normalizing effect across all road types. Hence, the third calibration method used the calibration coefficient from the paved road as:

$$C_{paved} = \frac{\frac{\sum_0^N RIF_{mean,unpaved,i,n}}{N}}{\frac{\sum_0^N RIF_{mean,paved,ref,n}}{N}} \quad (8)$$

in which, $RIF_{mean,unpaved,i,n}$ is the mean RIF from the i^{th} smartphone from the unpaved road from the n^{th} traversal, and $RIF_{mean,paved,ref,n}$ is the mean RIF from the reference smartphone from the n^{th} traversal measured from the paved road. Then the calibration of smartphones for road roughness on unpaved roads can be performed as:

$$RIF_{mean,i,cali} = \frac{RIF_{mean,i,uncali}}{C_{paved}} \quad (9)$$

From Equation (4, 6, 8), it can be seen that for any of the three calibration methods, a reference phone which was pre-calibrated with known road roughness indexes such as IRI or other indexes is required.

To investigate the performances of different calibration methods, both laboratory tests and field testing are good approaches. Since the laboratory tests are set up in a controlled environment, various practical factors may be ignored, such as the noise generated by different road type, mounting of the smartphones, and driving speed variances, etc. Performing field tests would increase the confidence of the practical application of the investigated calibration methods. Therefore, in this study, field testing was conducted to evaluate the performances of various calibration methods.

III. EXPERIMENT METHODS

A. Experimental Design

Most smartphones currently available in market embed micro-electromechanical systems inertial motion unit (MEMS IMU) to measure accelerations. The embedded MEMS IMU in different smartphone models is expected to have different responses since they use different hardware, antenna size, and installation locations (Martin et al. 2016). More specifically, differences in smartphone responses to acceleration can be introduced by variations in their mechanical construction, materials used, internal construction, mount, and the internal location of the sensor. Such differences can result in variations in the mechanical damping effects. Differences in the maximum sampling rate and the distribution of the sample rate also plays a role. A sampling frequency of 100 Hz or more could also improve the quality of the recorded signal. The resolution of the embedded accelerometers in smartphones is another crucial factor for their applicability in vehicles to measure ride quality. Accelerometers with resolution of 10 mm/s² introduces inherent noise in the range of 100 mm/s² and that of 1 mm/s² yields half the noise (Feldbusch et al 2017).

In this study, three different smartphones were selected to perform the field experiments. They were the iPhone® 8 (i8), iPhone 10® (iX), and Google Pixel (GP). The selection includes two most popular operational systems in smart phones, the iOS® for iPhones and the Android system for Google Pixel. Within the same iOS® operational system, two different generations of smart phones, iPhone® 8 and 10 were selected to see the impact of different versions of the same operational system. Both iPhone 8 and iPhone X implemented the customized Bosch's 6-Axis IMU accelerometer (Bosch Sensortec, Gerlingen, Germany) which are made specifically for Apple. The Google Pixel installed the Bosch's BMI160 IMU (Bosch Sensortec, Gerlingen, Germany). The iPhone customized 6-Axis IMU has a thin profile, resulting in a different packaging structure compared with the BMI160 IMU used in the Google Pixel phones (Odenwald and Bailey 2019).

Table 1 shows the estimated sensing characteristics of sampling frequency, accuracy, and resolution from the accelerometers embedded on the three selected smart phones from the app data. Table 1 indicates that there is a significant difference on the acceleration sensing characteristics between the apps on the iPhones and the GP. Among the three phones, the i8 was selected as the reference phone since it had a high resolution and it was also previously used to measure RIF for paved roads which was calibrated with the IRI using a traditional inertial road profiler (Bridgelall et al. 2016). The i8 measured the $RIF_{max,ref,n}$ in Equations (4, 6, 8).

Table 1 Sensing specification of the embedded accelerometers in the selected smart phones (estimated from app data)

| Smartphone Model | Sampling Frequency (Hz) | Accuracy | Resolution (mm/s ²) |
|------------------|-------------------------|----------|---------------------------------|
| i8 | 125 | 96% | ±1 |
| iX | 125 | 96% | ±1 |
| GP | 400 | 95% | ±10 |

A regular passenger vehicle, 2015 Volkswagen Jetta, was selected to perform the field testing to collect the data, as illustrated in Figure 1(a). The vehicle was selected not only based on availability and convenience but also to show that smartphones and their calibration can be used on commonly

available passenger vehicles instead of specially instrumented trucks or pickups. Although the smartphones can be mounted on various locations on the vehicle for acceleration measurements such as on the dashboard, windshields, or the floor of the vehicle, to reduce the potential measurement noise introduced by the distance between the car body and smartphones, the front seat floor was selected to mount the smartphones. The smartphones were mounted flat on the floor of the front passenger seat of the vehicle using industrial clear tape as seen in Figure 1(b). The tape ensured that the smartphones coherently sampled accelerations in the vertical direction while providing both good GPS reception and stable vibration coupling for acceleration measurements. Each smartphone accelerometer axes were oriented the same direction with respect to the axes of the vehicle. As seen in Figure 1(b), the longitudinal direction of all the smartphones were aligned with the longitudinal direction of the car to ensure that the z-axis is in the vertical direction for vertical acceleration measurements.

It is also worth noting that there were four smartphones as shown in Figure 1(b). In addition to the i8, iX, and GP, there was one HTC smartphone with the Android system, mounted in the same location. However, due to a failure of the GPS receiver in HTC smartphone, the data was excluded from the analysis.

The experiments used a data collection app called PAVVET (Bridgelall 2013) for the iOS® devices and another called RIVET (Lu et al. 2019) for the Android operating the GP. Fig. 1(c) shows the user's interface for PAVVET and RIVET. To ensure that the acceleration sensors inside all the three smartphones in this experiment performed in the expected manner, the data collected from the first run in the experiment for each road type from the iPhone 10® (iX), and Google Pixel (GP) was used to compare the calibrated iPhone® 8 (i8) for validation, before the testing of other runs for statistical analysis.



Figure 1. (a) Smartphones were mounted on the front seat floor of the vehicle, (b) Vehicle for road test, and (c) PAVVET and RIVET' interface.

The data was collected by traversing the same segment for 35 times each of both the paved road as shown in Figure 2 (a, b) and unpaved road as shown in Figs. 2(c, d). As shown in Fig. 2, the two road segments, selected in Fargo, North Dakota, was a 400-meter section of the paved road on 9th St NE, and a 300-meter section of the unpaved road on 57th St N. Since there is a railroad grade crossing in each road, the speed limit for the paved road is 25mph and that for the unpaved road is 15mph. To ensure the driving safety and to easily maintained the same speed for all 35 runs on the unpaved and paved roads with $N=35$ in Equations (4, 6, 8), the driving speed of the testing vehicle was set to approximately 11.5 m/s for all the test runs. To ensure that the car suspension would not increase the variance of the results measured among smartphones, the test setup used identical test conditions for every run on each pavement segment, including the use of the same vehicle, the same

mounting and orientation of all the smartphones, and the same traversal speed.

A previously reported technique (Bridgelall 2013) was used to align the spatial positions of each signal by using a known ground truth, which is a rail-grade crossing bump on each road type. Figs. 3(a, b) shows a satellite view of the two rail-grade crossings. As shown in Figs. 2 and Figs. 3, the paved road segment has a road width of 8m, and that for the unpaved road is 6m. The railroad grade crossing for the paved road as the maximum bump has a length of 3 m and a width of 10 m, and that for the unpaved road is 2.5 m in length and 6.7 m in width. The paved road segment for testing has a length of 580 m, while the unpaved road is 620 m. For the paved road, the railroad crossing is 420 m away from the beginning point, and that for the unpaved road is 280 m. With the known locations of the railroad crossing and using the signal peak produced from traversing the railroad as a reference for GPS locating, the distances can be interpolated outwards, and then all traversals can be truncated to approximately the same length to produce aligned signals for further processing.



Figure 2. (a) Satellite view and (b) street view of the paved road segment; (c) Satellite view and (d) Street view of the unpaved road segment. Source: Google Maps (2019).

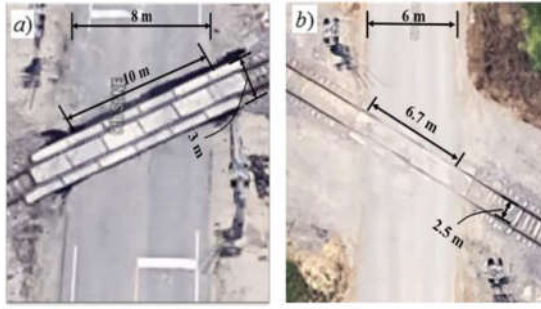


Figure 3. Satellite view of the rail-grade crossings of (a) the paved road segment and (b) the unpaved road segment. Source: Google Maps (2019).

B. Data Collection

The PAVVET and RIVET apps collected the data uploaded wirelessly to a remote database for post-analysis. Table 2 show a portion of the data collected using the PAVVET app with the iPhone® device as an example of data. The column ‘Time’ stored the time of the data sample in milliseconds (ms) offset from a date-time stamp encoded in the filename. ‘Lat’ and ‘Lon’ stored the GPS coordinates of latitude and longitudinal in decimal degrees, ‘v’ stored the ground speed of the vehicle. Pitch, Roll, and Yaw stored the sensor orientation angles in degrees. The variables g_x , g_y , g_z stored the g-forces sensed in the lateral, longitudinal, and vertical directions in units of $m \cdot s^{-2}$. From Equations (1) and (2), it can be seen that the calculations of the ride quality indices, RIF, used only the time (t), the g-force (g_z), and the vehicle speed (v). The PAVVET app integrated Equation (2) and calculated the g-force at each time instant, which can be easily used to estimate RIF using Equation (1) for post data analysis.

Table 3 shows part of the data collected using the RIVET app on the Android phone operating system (Google Pixel in this study). The column ‘Time’ stored the epoch time of the sampling instant in milliseconds. The columns labeled ‘Lat’ and ‘Lon’ stored the GPS coordinates of latitude and longitudinal in decimal degrees. The ‘v’ column stored the ground speed of the vehicle speed in units of $m \cdot s^{-1}$. The variables a_x , a_y , a_z stored the accelerometer lateral, longitudinal, and vertical accelerations in units of $m \cdot s^{-2}$. Yaw, Pitch, and Roll stored the azimuth, pitch, and roll angles, respectively, recorded from the gyroscope in decimal degrees. Rx, Ry, Rz stored the rotation rates around the x-axis, y-axis, and z-axis, respectively. Mx, My, Mz stored the magnetic field strength along the x-, y-, z-axis, respectively, in units of micro-Tesla (μT). The RIVET app directly collected the accelerations of the smartphones without using Equation (2). Thus, in the post data analysis, Equation (2) needs to be applied to compute the g-force (g_z) before estimating RIF using Equation (1).

C. Ride Quality Index

As shown in Table 1, the sampling frequency of each phone was different. Attempts had been made to adjust the sampling frequency of the sensors on each smartphone to be identical. However, the sampling frequency was integrated into each device and users have no access to change it. Thus, the default

maximum sampling frequency of the iOS® and Android systems of 125 Hz and 400 Hz was used on the iPhones and the Google phone, respectively. To ensure the data from each phone is compatible with different sampling frequency, a low-pass filter, the Wavelet Filter (Lane 2005), was used to normalize the effect of sampling at different rates by removing energy above the same frequency (Bridgelall and Tolliver 2020). The Wavelet Filter also can reduce the noise before feature extraction. Figs. 4 (a, c, e). Figs.4 (b, d, f) show the filtered g_z signal and the extracted RIF-indices for a feature extraction window size of 15 meters, for the first traversal of the paved and unpaved roads from the data of the i8, iX, and GP smartphones, respectively.

Each RIF spike computed from the collected data in Figure 4 (a~f) indicates a rougher road surface feature compare to the average road conditions on both the paved and unpaved roads. As in Figure 4 (a, c, e), from the first run on the paved road, the biggest RIF spike had a RIF value of 0.963 for i8, 1.063 for iX, and 0.955 for GP, demonstrating the RIF responses across the rail-grade crossing on the paved road for the three smartphones. Figure 4 (a, c, e) also indicated that for a regular paved road segment with smooth road surface, the measured RIF is smooth as well with a RIF value of approximately 0.10 for iPhones and 0.20 for GP. For small local road surface cracks or variances, there were small spikes with RIF values between 0.20 to 0.40 for all three smartphones. Compared to the regular smooth paved surface and small surface variances, the railroad crossing introduced significant roughness changes to the acceleration measurements. The start and end of the railroad grade crossing can be identified from Figure 4 (a, c, e) using all three smartphones.

Different from the smooth RIF measured from the paved road, due to the rough surface condition of unpaved roads, the measured RIF for the unpaved road had a much larger noise level, resulting in large variances in RIF values as shown in Figure 4 (b, d, f). Small local road surface variances for unpaved roads are not so obvious in RIF changes compared to paved road. The roughness changes of the unpaved road can be tracked by the RIF change trends. Figure 4 (b, d, f) show that the tested unpaved road segment had a regular RIF value changing from 0.25 to 0.50 for iPhones and from 0.40 to 0.60 for GP. As also shown in Figure 4 (b, d, f), from the first run on the unpaved road, the biggest RIF spike had a RIF value of 0.949 for i8, 0.909 for iX, and 0.862 for GP, demonstrating the RIF responses across the rail-grade crossing on the unpaved road for the three smartphones. The start and end of the railroad crossing position is also indicated in the data from all three smartphones.

Comparing Figure 4 (a, c, e), it can be seen that the three smartphones had small variations for RIF measurements from paved roads and the comparison between Figure 4 (b, d, f) shows that the smartphones had relatively larger variations for RIF measurements for unpaved roads. The differences in segment RIF-indices for the same road from three smartphones highlight the differences of the sensing characteristics of the accelerometers embedded in each smartphone. This result points to the need for calibration when using different smartphones to measure the ride quality of a roadway segment even under identical test conditions of the same vehicle, mounting, orientation, and traversal speed. Using different vehicles with the same smartphone and mount will also produce similar differences in the segment RIF-indices because of differences in

suspension systems and vehicle loading conditions which increase the variance of the results measured among smartphones.

Based on the RIF-index measured from each run (35 runs in total), the mean and peak RIF of each run of the road tests can be computed. The calibration can then be performed on iX and GP smartphones using the i8 phone as a reference since it was previously calibrated with commercially available road profiler. Three calibration methods are used to calibrate the measured data from the paved and unpaved roads including the reference mean (R-Mean), max (R-Max), and road type (R-Road) methods as mentioned in Section II. The effectiveness of

the three different calibration approaches is then evaluated by conducting statistical tests for equal means. The unpaired two-tailed t-test determines if the means of two independent sample sets are significantly different. The test assumes that the two populations follow a normal distribution. Evaluating the skewness and kurtosis of a distribution is a simple test to gauge the amount of departure from normality. The two-tailed F-test for equal variances determines the type of t-test to use. The t-test rejects the null hypothesis that the means are equal if the p value is less than the traditional threshold of 0.05. The standard t-test assumes that the variances of the two population are equal. A variant of the t test that does not assume equal variances is also available.

Table 2. PAVVET data format

| Time (ms) | Lat (°) | Lon (°) | v (m·s ⁻¹) | g _x (g-force) | g _y (g-force) | g _z (g-force) | Pitch (°) | Roll (°) | Yaw (°) | RotX (°) | RotY (°) | RotZ (°) |
|-----------|---------|---------|------------------------|--------------------------|--------------------------|--------------------------|-----------|----------|---------|----------|----------|----------|
| 34986.615 | 46.899 | -96.883 | 11.569 | 0.045 | -0.058 | -0.989 | 3.001 | 1.950 | 6.180 | -0.712 | -1.237 | -0.250 |
| 34994.377 | 46.899 | -96.883 | 11.569 | 0.049 | -0.061 | -0.969 | 3.003 | 1.944 | 6.181 | -0.712 | -1.237 | -0.250 |
| 35002.688 | 46.899 | -96.883 | 11.569 | 0.046 | -0.059 | -0.943 | 3.003 | 1.934 | 6.180 | -1.080 | -0.933 | -0.374 |
| 35010.222 | 46.899 | -96.883 | 11.569 | 0.060 | -0.065 | -0.936 | 3.003 | 1.934 | 6.180 | -1.080 | -0.933 | -0.374 |
| 35018.392 | 46.899 | -96.883 | 11.569 | 0.051 | -0.058 | -0.943 | 3.002 | 1.927 | 6.179 | -1.141 | -0.811 | -0.496 |

Table 3. RIVET data format

| Time (ms) | Lat (°) | Lon (°) | v (m·s ⁻¹) | a _x (m·s ⁻²) | a _y (m·s ⁻²) | a _z (m·s ⁻²) | Yaw (°) | Pitch (°) | Roll (°) | R _x (r·s ⁻¹) | R _y (r·s ⁻¹) | R _z (r·s ⁻¹) | M _x (μT) | M _y (μT) | M _z (μT) |
|-----------|---------|---------|------------------------|-------------------------------------|-------------------------------------|-------------------------------------|---------|-----------|----------|-------------------------------------|-------------------------------------|-------------------------------------|---------------------|---------------------|---------------------|
| 34985.49 | 46.90 | -96.88 | 11.58 | -0.42 | 0.32 | 10.16 | 56.76 | 1.56 | 2.26 | -0.014 | -0.018 | -0.002 | -64.65 | 42.45 | 21.00 |
| 34988.05 | 46.90 | -96.88 | 11.58 | -0.58 | 0.39 | 9.94 | 56.76 | 1.56 | 2.26 | -0.015 | -0.007 | -0.001 | -64.65 | 42.45 | 21.00 |
| 34990.48 | 46.90 | -96.88 | 11.58 | -0.58 | 0.39 | 9.94 | 56.76 | 1.56 | 2.26 | -0.015 | -0.007 | -0.001 | -64.65 | 42.45 | 21.00 |
| 34993.16 | 46.90 | -96.88 | 11.58 | -0.58 | 0.39 | 9.94 | 56.66 | 2.23 | 3.34 | -0.015 | -0.007 | -0.001 | -64.35 | 42.45 | 21.45 |
| 34995.67 | 46.90 | -96.88 | 11.58 | -0.46 | 0.22 | 10.01 | 56.38 | 1.72 | 4.11 | -0.012 | -0.004 | -0.001 | -64.35 | 42.45 | 21.45 |

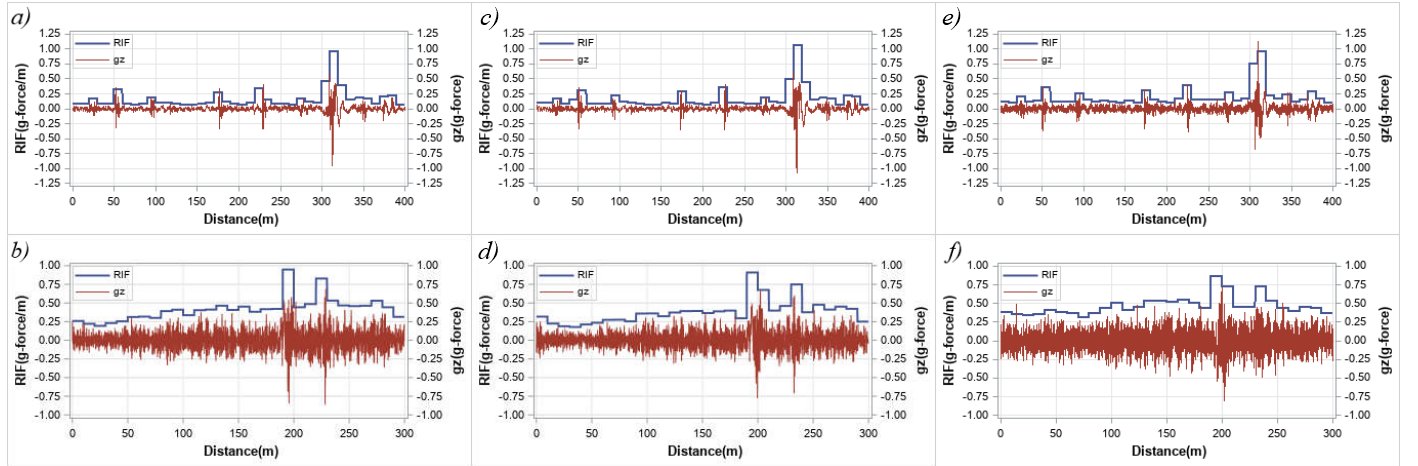


Figure 4. Collected g-force data and windowed calculated RIF-indices using i8 for (a) the paved road segment and for (b) the unpaved road segment; using iX for (c) the paved road segment and (d) the unpaved road segment; using GP for (e) the paved road segment and (f) the unpaved road segment.

D. Mean and Max RIF-Index

Fig.5 (a, b) show the measured mean RIF-indices from the three smartphones for all the 35 runs for paved and unpaved road segments, respectively. Fig.5 (c, d) illustrates the measured maximum RIF-indices from the three smartphones for all 35 runs on the paved and unpaved road segments, respectively. When comparing Fig.5 (a) and (b), and Fig.5 (c) and (d), it is observed that the unpaved road had a higher variance in RIF from all three smartphones. Table 4 and Fig.6 show the average mean and maximum (railroad crossing location) RIF calculated from all the 35 runs of the road tests on the paved and unpaved

roads for the three smartphones, respectively. As shown by the standard deviation in between three smartphones, the measured average mean and maximum RIF of 35 runs for the iPhones and Android phones vary from each other significantly since they have different sensitivities for g-force measurements. In general, the i8 and iX show smaller differences since they shared the same type of accelerometer and resulted in similar measurement accuracy.

Table 5 shows the calibration constants from the three different calibration methods. Since the i8 was previously calibrated using traditional RIF-indices, it was used as the reference smartphone for calibration by both Ref-mean and Ref-

max methods. The calibration constant for i8 was 1.0 as expected. For iX, the calibration constants from both Ref-mean and Ref-max methods were slightly above 1.0 for paved road segments and below 1.0 for unpaved road segments. For GP, the calibration constant varied significantly between paved and unpaved road surface types, in addition to large variances between Ref-mean and Ref-max for GP.

Table 4. Original average mean and peak RIF of all the 35 runs of road tests for paved and unpaved roads

| Road type | | i8 | iX | GP | Standard Deviation between Phones |
|--------------|------|-------|-------|-------|-----------------------------------|
| Paved Road | Mean | 0.202 | 0.207 | 0.190 | 0.0087 |
| | Max | 0.935 | 0.996 | 0.826 | 0.0861 |
| Unpaved Road | Mean | 0.527 | 0.481 | 0.464 | 0.0326 |
| | Max | 1.109 | 1.030 | 0.905 | 0.1029 |

Table 5. Calibration constants from different calibration methods

| Road Type | Phone Type | Ref-Mean (C _{Mean}) | Ref-Max (C _{Max}) | Ref-Paved (C _{Paved}) |
|-----------|------------|-------------------------------|-----------------------------|---------------------------------|
| Paved | i8 | 1.000 | 1.000 | - |
| | iX | 1.026 | 1.065 | - |
| | GP | 0.940 | 0.883 | - |
| Unpaved | i8 | 1.000 | 1.000 | 1.000 |
| | iX | 0.913 | 0.929 | 1.026 |
| | GP | 0.881 | 0.816 | 0.940 |

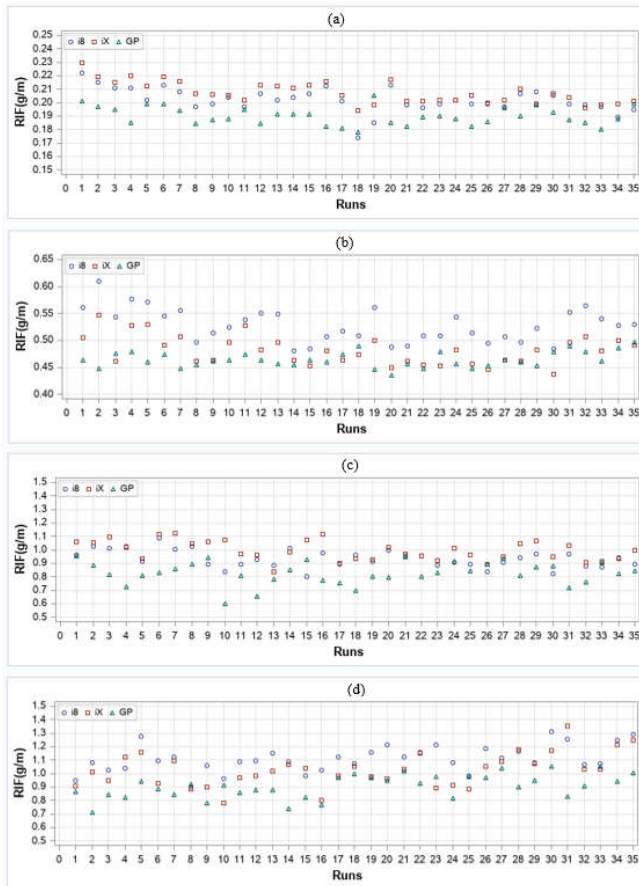


Figure 5. Measured mean RIF-indices for all 35 runs using the three smartphones (a) for the paved road segment, (b) for the unpaved road segment, and measured Maximum RIF -indices for all 35 runs using the three smartphones (c) for the paved road segment; and (d) for the unpaved road segment.

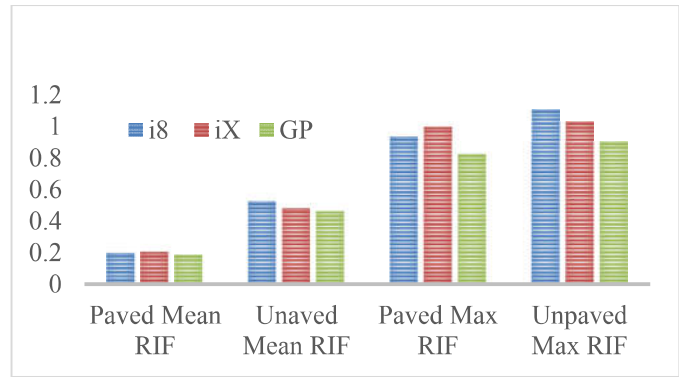


Figure 6. Comparison of original average mean and peak RIF from three smartphones

E. Calibration Effectiveness

Based on the average mean and max RIF in Table 5, calibrations following the three calibration methods were performed. Table 6 shows the calibrated average mean RIF and their standard deviation (SD) after calibration using the three calibration methods calculated from 35 runs of the road tests on the paved and unpaved roads for the three smartphones, respectively. Based on Table 6, for paved road, R-mean and R-max methods yield a standard deviation of around 0.008 for 35 runs, and the standard deviation for 35 runs varies from 0.015 to 0.03 in between the three calibration methods for unpaved road.

TABLE 6. Average mean RIF of the 35 runs of road tests for paved and unpaved roads after calibration

| Road Type | Calibration Methods | | i8 | iX | GP |
|-----------|---------------------|---------|--------|--------|--------|
| Paved | R-Mean | Average | 0.202 | 0.202 | 0.202 |
| | | SD | 0.0090 | 0.0080 | 0.0071 |
| | R-Max | Average | 0.202 | 0.195 | 0.215 |
| | | SD | 0.0090 | 0.0077 | 0.0070 |
| Unpaved | R-Mean | Average | 0.527 | 0.527 | 0.527 |
| | | SD | 0.0309 | 0.0294 | 0.0162 |
| | R-Max | Average | 0.527 | 0.518 | 0.569 |
| | | SD | 0.0308 | 0.0290 | 0.0175 |
| | R-Road | Average | 0.527 | 0.469 | 0.494 |
| | | SD | 0.0309 | 0.0262 | 0.0152 |

Table 7 analyzed the standard deviation of the average mean RIF of 35 runs in between three smartphones after calibration using the three different calibration methods. For both the paved and unpaved roads, the R-mean method has a standard deviation of 0 in between three phones which is expected since the average mean was used to calibrate each run. For the R-max method, the standard deviation in between three smartphones are 0.0101 for paved road and 0.0272 for unpaved road. Compared to the original data before calibration in Table 5, the R-max calibration method did not effectively reduce the standard deviation of mean RIF in between phones. Table 7 also shows that the reference road type method can effectively reduce the variation in between phones for unpaved roads. To further evaluate which calibration method would perform best, T-tests are needed.

Table 7. Standard deviation of average mean RIF of 35 runs in between three smartphones for paved and unpaved roads after calibration

| Road type | Standard Deviation between Phones | |
|--------------|-----------------------------------|--------|
| Paved Road | R-Mean | 0.0000 |
| | R-Max | 0.0101 |
| | R-Mean | 0.0000 |
| | R-Max | 0.0272 |
| Unpaved Road | R-Road | 0.0293 |

F. Normality Tests

A skewness of zero means that distribution is symmetrical, and an absolute value of less than 2.1 indicates that there is no significant departure from normality (West 1995). Similarly, an absolute value of kurtosis less than 7.1 indicates no significant departure from normality (West 1995). Table 8 summarizes the skewness and kurtosis for the distribution of segment RIF-indices from each smartphone for the paved and unpaved roads based on the original data. Table 8 suggests that the distribution of segment RIF-indices measured by each smartphone does not deviate significantly from the normal distribution.

Table 8. Kurtosis and skewness of paved and unpaved road

| Road type | | i8 | iX | GP |
|--------------|----------|--------|--------|--------|
| Paved Road | Skewness | -0.647 | 0.641 | 0.330 |
| | Kurtosis | 2.015 | -0.062 | -0.654 |
| Unpaved Road | Skewness | 0.444 | 0.576 | 0.414 |
| | Kurtosis | -0.184 | -0.273 | -0.371 |

G. F-test for Equal Variances

The upper and lower bounds of the F-statistic within a 95% confidence interval is 1.981 and 0.505 based on the degrees-of-freedom from the 35 samples of each smartphone. Table 9 and Table 10 summarize the F-statistic and their p-values for the paved and unpaved roads, respectively. The results suggest that before and after calibration, the variances of segment RIF-indices from the paved road are equal. However, for data from the unpaved road, only the variances of segment RIF-indices measured with the i8 and iX before and after calibration are equal. This result informs the appropriate type of t-test to use.

Table 9. F-Value and P-Value (paved road)

| | | i8+iX | i8+GP | iX+GP |
|----------|----------------------|-------|-------|-------|
| Original | F-Statistic | 1.206 | 1.856 | 1.541 |
| | P-Value (Two-Tailed) | 0.589 | 0.076 | 0.213 |
| | F-Statistic | 1.269 | 1.638 | 1.291 |
| | P-Value (Two-Tailed) | 0.491 | 0.155 | 0.461 |
| R-Mean | F-Statistic | 1.368 | 1.446 | 1.057 |
| | P-Value (Two-Tailed) | 0.366 | 0.287 | 0.872 |

Table 10. F-Value and P Value (unpaved road)

| | | i8+iX | i8+GP | iX+GP |
|----------|----------------------|-------|---------|--------|
| Original | F-Statistic | 1.313 | 4.816 | 3.559 |
| | P-Value (Two-Tailed) | 0.431 | 2.0E-05 | 0.0004 |
| | F-Statistic | 1.094 | 3.312 | 3.312 |
| | P-Value (Two-Tailed) | 0.795 | 0.0003 | 0.001 |
| R-Mean | F-Statistic | 1.132 | 3.112 | 2.748 |
| | P-Value (Two-Tailed) | 0.719 | 0.001 | 0.004 |
| R-Max | F-Statistic | 1.383 | 4.126 | 2.984 |
| | P-Value (Two-Tailed) | 0.349 | 7.9E-05 | 0.002 |

H. T-test for Equal Means

The null hypothesis of the t-test is that the mean of the two variables is equal, and the null hypothesis can be rejected only if the probability value (P-Value) is greater than 0.05. Table 11 and Table 12 summarize the results of the t-tests for the paved and unpaved roads, respectively. For the paved road, Table 11 shows that the P-value of all three smartphone from original and reference maximum calibration method is smaller than 0.05, indicating an unequal mean among the three smartphones. However, the P-value of all three smartphones equals to one for the reference mean calibration method, which is larger than 0.05. Thus, for paved roads, the reference mean calibration method would result in equal means for all three smartphones. For unpaved roads, Table 11 shows that both original and reference road type calibration method have P-value smaller than 0.05 and yield unequal means among the three smartphones. The reference maximum calibration method has a P-value larger than 0.05 in between i8 and iX, but smaller than 0.05 in between iPhone and GP, indicating that the two iPhone share the same mean but the iPhone and the Google phone have different means after using the reference maximum calibration method. On the other hand, the reference mean calibration method also resulted in a P-value larger than 0.05 among the three smartphones, thus, equal means among the three smartphones for unpaved road.

Table 11. T-Statistic and P-Value (paved road)

| | | i8+iX | i8+GP | iX+GP |
|----------|---------------------|---------|-----------|----------|
| Original | T-Statistic | -2.547 | 6.470 | 9.807 |
| | T-Critical | 1.995 | 1.999 | 1.995 |
| | P-Value (Two-Tails) | 0.013 | 1.759E-08 | 1.21E-14 |
| | T-Statistic | 2.9E-08 | 3.29E-08 | 2.35E-09 |
| R-Mean | T-Critical | 1.995 | 1.995 | 1.995 |
| | P-Value (Two-Tails) | 1 | 1 | 1 |
| R-Max | T-Statistic | 3.709 | -6.552 | -11.236 |
| | T-Critical | 1.995 | 1.995 | 1.995 |
| | P-Value (Two-Tails) | 0.001 | 8.99E-09 | 3.76E-17 |

Table 12. T-Statistic and P-Value (unpaved road)

| | | i8+iX | i8+GP | iX+GP |
|----------|---------------------|----------|----------|----------|
| Original | T-Statistic | 6.668 | 10.997 | 3.307 |
| | T-Critical | 1.995 | 2.011 | 2.007 |
| | P-Value (Two-Tails) | 5.59E-09 | 1.03E-14 | 0.002 |
| | T-Statistic | 3.6E-08 | 3.6E-08 | -4.4E-08 |
| R-Mean | T-Critical | 1.995 | 2.008 | 2.006 |
| | P-Value (Two-Tails) | 1 | 1 | 1 |
| R-Max | T-Statistic | 1.271 | -6.961 | -8.881 |
| | T-Critical | 1.995 | 2.005 | 2.003 |
| | P-Value (Two-Tails) | 0.208 | 4.8E-09 | 2.8E-12 |
| | T-Statistic | 8.524 | 5.715 | -4.906 |
| R-Road | T-Critical | 1.995 | 2.009 | 2.005 |
| | P-Value (Two-Tails) | 2.4E-12 | 6.0E-07 | 8.9E-06 |

IV. CONCLUSIONS AND FUTURE WORK

This study evaluated three calibration methods for calibrating the road roughness measurements using different smartphones for both paved and unpaved roads. The findings from this study can be summarized as below:

- 1) Road tests indicated that the measured RIF-index varies significantly among smartphones, especially among different phone models that use different operating systems.
- 2) Spikes in RIF reflect roadway anomalies.
- 3) The data analysis indicates that the segment RIF-indices are normally distributed, regardless of the type of phone used, and the type of road evaluated.
- 4) Among the three calibration methods (reference mean, max, and road type), the reference mean method performed the best for calibration. The reference road type method can also reduce inconsistency between phones for the unpaved road.
- 5) Statistical two-tailed t-tests suggest that using the mean of the ride-quality indices from a reference smartphone as the calibration factor for all smartphones results in equal means.

The normality of the distribution of ride-quality indices using the RIF-transform suggests that after calibrating a smartphone sensor, averaging the segment RIF-indices from multiple traversals will produce ever-increasing accuracy and precision of ride-quality measurements, for any type of road. Doing the same type of calibration with both a reference smartphone and a reference vehicle will make the approach practical for wide-spread use. Although the calibration methods investigated in this paper could be time-intensive, a single type of smartphone only need to do the calibration once on a short segment of any type of road type. The calibration coefficient can then be applied to all the measurements later using the same smartphone and vehicle. In future work, the authors will investigate the impacts of different factors during calibration to minimize the runs for statistical analysis and to optimize factors for less calibration time.

In addition, more studies are needed to incorporate more road, smartphone, and vehicle types to validate the performance of an automated means to conduct the suggested method of calibration in addition to attempts for correlating the measured roughness with the road distress. More importantly, future implementation in connected vehicles with a standardized method to achieve the calibration is an attractive way to perform calibrations of smartphones because the accelerometers are built into the vehicle. Potentially, if connected vehicles become prevalent, their use with same sensing characteristics as the smartphones on routinely driven routes can be used for calibration.

ACKNOWLEDGMENT

This work was supported by the U.S. Department of Transportation under the agreement of No. 69A3551747108 through MPC project No. 547.

References

Agbovi, H. K. 2012. Effects of low temperatures, repetitive stresses and chemical aging on thermal and fatigue cracking in asphalt cement pavements on Highway 417. Master Thesis, Queen's University at Kingston.

Aleadelat, W., Ksaibati, K., Wright, C. H., and Saha, P., 2018. Evaluation of pavement roughness using an android-based smartphone. *Journal of Transportation Engineering*, Part B: Pavements, 144(3), 04018033.

Alessandroni, G., Carini, A., Lattanzi, E., and Bogliolo, A., 2015. Sensing road roughness via mobile devices: A study on speed influence. 2015 9th International Symposium on

Image and Signal Processing and Analysis (ISPA). *IEEE*, 270-275.

Bhushan, M., and Rengaswamy, R., 2000. Design of sensor location based on various fault diagnostic observability and reliability criteria. *Computers & Chemical Engineering*, 24(2-7), 735-741.

Bridgelall, R., 2013. Connected vehicle approach for pavement roughness evaluation. *Journal of Infrastructure Systems*, 20 (1), 04013001.

Bridgelall, R., Huang, Y., Zhang, Z., and Deng, F., 2016. Precision enhancement of pavement roughness localization with connected vehicles. *Measurement Science and Technology*, 27 (2), 025012.

Bridgelall, R., Rahman, M. T., Tolliver, D., and Daleiden, J. F., 2017. Wavelength Sensitivity of a Connected Vehicle Method of Ride Quality Characterizations. *International Journal of Pavement Engineering*, 1-7.

Bridgelall, R., and Tolliver, D., 2020. Accuracy Enhancement of Anomaly Localization with Participatory Sensing Vehicles. *Sensors*, 20(2), p. 409.

Brown, D., Liu, W., and Henning, T. F., 2010. Identifying pavement deterioration by enhancing the definition of road roughness. *Journal of Transportation Research Board*, 430, p. 66.

Buttlar, W. G. and Islam, M., 2014. Integration of smart-phone-based pavement roughness data collection tool with asset management system. NEXTRANS Center (US). Technical Report for Project No. 098IY04.

Cebon, D., 1998. Theoretical road damage due to dynamic tyre forces of heavy vehicles Part 2: simulated damage caused by a tandem-axle vehicle. Proceedings of the Institution of Mechanical Engineers. Part C: *Journal of Mechanical Engineering Science*, 202 (2), 109-117.

Chan, C. Y., Huang, B., Yan, X., and Richards, S., 2010. Investigating effects of asphalt pavement conditions on traffic accidents in Tennessee based on the pavement management system (PMS). *Journal of advanced transportation*, 44(3), 150-161.

Darawade, K., Karmare, P., Kothmire, S., and Panchal, N., 2016. Estimation of road surface roughness condition from android smartphone sensors. *Int. J. Recent Trends in Eng. & Research*, 2(3), 339-346.

Douangphachanh, V., and Oneyama, H., 2016. A study on the use of smartphones for road roughness condition estimation. *Journal of the Eastern Asia Society for Transportation Studies*, 10, 1551-1564.

Eriksson, J., Girod, L., Hull, B., Newton, R., Madden, S., and Balakrishnan, H., 2008. The pothole patrol: using a mobile sensor network for road surface monitoring. *Proceedings of the 6th international conference on Mobile systems, applications, and services*, 29-39.

Ghose, A., Biswas, P., Bhaumik, C., Sharma, M., Pal, A., and Jha, A., 2012. Road condition monitoring and alert application: Using in-vehicle smartphone as internet-connected sensor. *2012 IEEE international conference on pervasive computing and communications workshops*. IEEE, 489-491.

Ghasemi, A. R., Taheri-Behrooz, F., & Shokrieh, M. M., 2014. Measuring residual stresses in composite materials using the simulated hole-drilling method. *Residual Stresses in Composite Materials*, 76-120.

Feldbusch, A., Sadegh-Azar, H., and Agne, P., 2017. Vibration

- analysis using mobile devices (smartphones or tablets). *Procedia Engineering*, 199, p.2790–2795.
- Golabi, K., Kulkarni, R. B., and Way, G. B., 1982. A statewide pavement management system. *Interfaces*, 12(6), 5-21.
- Huang, H.-B., Yi, T.-H., and Li, H.-N., 2017a. Bayesian combination of weighted principal-component analysis for diagnosing sensor faults in structural monitoring systems. *Journal of Engineering Mechanics*, 143(9), p. 04017088.
- Huang, H.-B., Yi, T.-H., and Li, H.-N., 2017b. Sensor fault diagnosis for structural health monitoring based on statistical hypothesis test and missing variable approach. *Journal of Aerospace Engineering*, 30(2), p. B4015003.
- I-larte, S., Rahman, A., and Razeed, K., 2005. Fault tolerance in sensor networks using self-diagnosing sensor nodes.
- Janani, L., Sunitha, V., and Mathew, S., 2020. Influence of surface distresses on smartphone-based pavement roughness evaluation. *International Journal of Pavement Engineering*, p.1-14.
- Krunic, V., Trumpler, E., and Han, R., 2007. NodeMD: Diagnosing node-level faults in remote wireless sensor systems. Proceedings of the 5th international conference on Mobile systems, applications and services, pp. 43-56.
- Kumar, R., Mukherjee, A., and Singh, V., 2017. Community sensor network for monitoring road roughness using smartphones. *Journal of Computing in Civil Engineering*, 31(3), 04016059.
- Lane, M., 2005. An introduction to wavelet analysis with SAS. *SAS Conference Proceedings NESUG 2005*. p.1-10.
- Lekshmi, J., Samuel, N. M., and Velayudhan, S., 2020. Vibration vs. vision: best approach for automated pavement distress detection. *International Journal of Pavement Research and Technology*, p. 1-9.
- Leverett, B. C., 2008. Effects of incentive/disincentive program on pavement performance. Ph. D. Dissertation, Michigan State University.
- Liu, M., 2013. A study of mobile sensing using smartphones. *International Journal of Distributed Sensor Networks*, 9(3), 272916.
- Lu, P., Bridgelall, R., Tolliver, D., Chia, L., and Bhardwaj, B., 2019. Intelligent Transportation Systems Approach to Railroad Infrastructure Performance Evaluation: Track Surface Abnormality Identification with Smartphone-Based App. U. S. Department of Transportation, Technical Report, No. MPC-19-384.
- Martin, H., Groves, P., and Newman, M., 2016. The Limits of In-Run Calibration of MEMS Inertial Sensors and Sensor Arrays. *Journal of The Institute of Navigation*, 63(2), p. 127-143.
- Mednis, A., Strazdins, G., Zviedris, R., Kanonirs, G., and Selavo, L., 2011. Real time pothole detection using android smartphones with accelerometers. *2011 International conference on distributed computing in sensor systems and workshops (DCOSS)*. IEEE, p. 1-6.
- Mohan, P., Padmanabhan, V. N., and Ramjee, R., 2008. Nericell: rich monitoring of road and traffic conditions using mobile smartphones. *Proceedings of the 6th ACM conference on Embedded network sensor systems*, p. 323-336.
- Múčka, P., 2015. Sensitivity of road unevenness indicators to distresses of composite pavements. *International Journal of Pavement Research and Technology*, 8 (2), 72.
- Múčka, P., 2017. Road roughness limit values based on measured vehicle vibration. *Journal of Infrastructure Systems*, 23 (2), 04016029.
- Mukherjee, A., and Majhi, S., 2016. Characterisation of road bumps using smartphones. *European Transport Research Review*, 8(2), p. 13.
- Odenwald, S. F., and Bailey, C. M., 2019. Gravimetric Detection of Earth's Rotation Using Crowdsourced Smartphone Observations. *IEEE Access*, 7, 148131-148141.
- Osborne, C., 1991. Statistical calibration: a review. *International Statistical Review/Revue Internationale de Statistique*. 309-336.
- O'Sullivan, M. J., & O'Sullivan, J. P., 2016. Reservoir modeling and simulation for geothermal resource characterization and evaluation. *Geothermal power generation*, 165-199.
- Park, K., Thomas, N. E., and Wayne Lee, K., 2007. Applicability of the international roughness index as a predictor of asphalt pavement condition. *Journal of Transportation Engineering*, 133(12), 706-709.
- Paterson, W., 1986. International roughness index: Relationship to other measures of roughness and riding quality. *Transportation Research Record*, (1084).
- Pawar, P. R., Mathew, A. T., and Saraf, M., 2018. IRI (International Roughness Index): An Indicator Of Vehicle Response. *Materials Today: Proceedings*, 5(5), 11738-11750.
- Sayers, M. W., 1986. Guidelines for conducting and calibrating road roughness measurements. University of Michigan, Ann Arbor, Transportation Research Institute.
- Seraj, F., van der Zwaag, B. J., Dilo, A., Luarasi, T., and Havinga, P., 2015. RoADS: A road pavement monitoring system for anomaly detection using smart phones. *Big data analytics in the social and ubiquitous context*, Springer, 128-146.
- Smith, K. and Ram, P., 2016. Measuring and Specifying Pavement Smoothness. U. S. Federal Highway Administration, Technical Report No. FHWA-HIF-16-032.
- National Transportation Safety Board (NTSB). 2018. Road Statistics in 2018. *National Transportation Safety Board*.
- West, S. G., Finch, J. F., and Curran, P. J., 1995. Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (p. 56–75). Sage Publications, Inc.
- Zaloshnja, E. and Miller, T. R., 2009. Cost of crashes related to road conditions, United States, 2006. In: *Annals of Advances in Automotive Medicine/Annual Scientific Conference. Association for the Advancement of Automotive Medicine*, 53, p. 141.