CALIBRATING SMARTPHONES FOR MONITORING ROAD CONDITION ON PAVED AND UNPAVED ROADS

A Thesis
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Liuqing Hu

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Department:
Civil and Environmental Engineering

June 2018

Fargo, North Dakota

North Dakota State University
Graduate School

Title

CALIBRATING SMARTPHONES FOR MONITORING ROAD CONDITION ON PAVED AND UNPAVED ROADS

By

Liuqing Hu

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

SUPERVISORY COMMITTEE:

Dr. Ying Huang
Chair

Dr. Raj Bridgelall

Dr. Simone Ludwig

Approved:

July 6, 2018
Date

Dr. David Steward
Department Chair

ABSTRACT

Transportation agencies report the localization of roadway anomalies that could cause serious hazards to the traveling public. However, the high cost and limitations of present technical prevent scaling the road monitoring to all roadways. Especially the unpaved road, because of the complexity of unpaved road. Using smartphone application as road condition data collection tool offer an attractive alternative because of its potential to monitor all roadways in real time and its low cost. However, the sensor sensitivity and sampling frequency of different smartphones may vary significantly, which challenge the confidence of using smartphones for actual pavement condition assessment applications. This study tends to solve this challenge by calibrating different smartphones using two different calibrating methods including calibrating towards reference or average road roughness.

ACKNOWLEDGEMENTS

I want to express my gratitude to my committee members, Dr. Ying Huang, Dr. Raj Bridgelall, and Dr. Simone Ludwig, my family, and my friends. Without help from all of them, I would never finish this thesis.

Firstly, I really want to thank my advisor, Dr. Ying Huang, without her support I would never achieve this. She is so kind, encouraging, especially when I was in a really bad situation, Dr. Ying Huang helped me out without any doubt. Dr. Ying Huang is such a great advisor, following her instructing, I learned so much. When I am facing a big problem, and don't know what to do, I learned from her how to propose a big problem to some small problems and solve them one by one. I also learned how to think and do things logically, which are extremely helpful to my research.

I would also like to thank my co-advisor, Dr. Raj Bridgelall. During this research, he provided so much help and so many ideas for the research, thanks for the supporting throughout the process of my research.

Finally, I would like to thank my friend, Fodan Deng, who helped me during this project, and my parents for their continuous support and always encouraging me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

1. INTRODUCTION

1.1. Background

There are four modes of transportation including road, rail, air, and waterways. In 2012, 4,273,876 miles of travel were made by passenger vehicles, motorcycles, trucks, or buses; 580,501 miles by air carriers, 37,757 miles by rail, and 4156 miles by other methods. Road travel accounts for 86.93% of the entire passenger transportation (Transportation, 2017). For freight transportation, 60% the largest percentage of US freight transportation were carried by trucks, compared to about 18% by pipelines, 10% by rail, 8% by waterways, and 0.01% by airplane, and 3% for other modes of transportation (Statistics, 2017). According to the statistics, it can be seen that road plays a critical role in both passenger and freight transportation. Road can be categorized as paved or unpaved roads. To monitor the surface conditions of the road, various methods exist which still have certain limitations.

1.2. Paved Road Surfacing Methods

The paved road has different surfacing materials such as asphalt or concrete on it to sustain vehicle or foot traffic. Asphalt concrete as shown in Figure 1(a), also known as flexible pavement, is widely used since the 1920s. The bitumen binder of asphalt concrete allows it to sustain plastic deformation. Some asphalt surfaces are laid directly on the native subgrade but most of them are laid on base especially when the subgrade is very soft or expansive like clay or peat, the base can be gravel, cement, lime, polypropylene and polyester geosynthetics. Asphalt can be categorized as hot mix, warm mix, cold mix depending on the temperature when it is applied (Administration, Warm Mix Asphalt Technologies and Research, 2008).

Asphalt concrete road usually constructed for highways which are high-volume having daily traffic load more than 1200 vehicles per day (Gerbrandt, Makahoniuk, Borbely, &

Berthelot, 2000). Bituminous/asphalt surface (Base, n.d.) (Gransberg, 2005) is popular due to its relatively low noise and low cost comparing to other paving methods, and ease to repair, flexibility especially in the area where unstable terrain thaws and softens in the spring. The disadvantage is its less durability and tensile strength comparing to other paving methods. During the hot weather, it will become slick and soft and have hydrocarbon pollution problem to soil and groundwater or river.

Portland cement concrete as shown in Figure 1(b) is constructed by using a mix of sand, water, Portland cement, and coarse aggregate. In virtually nearly all concrete will be mixed by various admixtures to increase its workability, reduce the required water and harmful chemical reactions. In many cases, Portland cement substitute such as fly ash added to the concrete to reduce the cost and the physical properties of the concrete improved. There are three types of concrete surfaces, jointed plain, jointed reinforced and continuously reinforced. Each type of these is distinguished by jointing system which used to control crack development. Comparing to the asphalt road, the advantages of concrete road is they are stronger and more durable, they can be grooved to provide a durable skid-resistant surface. The disadvantage of concrete road is the higher initial cost and takes more time to construct. There are many methods used to maintain concrete road including dowel bar retrofits, diamond grinding, cross-stitching, joint and crack sealing (Pavement, 2004) (Administration, Concrete Pavement Rehabilitation - Guide for Diamond Grinding, 2017).

(a) (b)

Figure 1. (a) Asphalt surface road (Roads&Bridges, 2017) and (b) Concrete surface road (NBM&CW, 2016)

1.3. Unpaved Road

Unpaved roads do not have a pavement laid on it. It uses gravel or subgrade material as surface materials for the roads. Gravel road is built by placing and compacting large stones followed by placing and compacting small stones (WordWeb, n.d.). Another type of unpaved road is dirt road, which is a road without any pavement on it. The road surface is the native material of land surface. Which is also be called as subgrade material to highway engineers. Dirt road is usually don’t have graded camber to let rainwater drain off the road, or drainage ditches at the sides, or embankments. All these lacks lead to further waterlogging and erosion, and the road will be impassable even off-road vehicles after heavy rain. for this reason, some countries call the dirt road as dry-weather roads. Depending on the soils and geology where the road passes, dirt roads might have different characteristics, like sandy, rocky, stony or just bare earth surface. Dirt road is commonly used in rural areas, usually narrow and have low traffic frequency.

(a) (b)

Figure 2. (a) Gravel surface road (SkidSteer, 2015) and (b) unpaved road (Cookaa, 2003)

1.4. Road Surface Distress

If lack of properly maintenance, poor road condition will occur which will affect the ride quality and may cause traffic accident. Statistic from U.S. Department of Transportation has shown that highway quality is significantly affected by road condition. An important metric of road condition is the road surface distress, the road surface distress related to the roughness as well as structural integrity of road. Road surface distress indicates a decline in road surface condition, is "any indication of poor or unfavorable pavement performance or signs of impending failure; any unsatisfactory performance of a pavement short of failure" (Board).

1.4.1. Surface Distress Type

Surface distress can be mainly classified into three types: fracture, distortion, and disintegration. Cracking or spalling can be classified into fracture, which may be resulted from excessive loading, moisture damage, fatigue, thermal changes, slippage or contraction. Deformation like rutting, shoving and corrugation can be classified into distortion. The reason for distortion can be excessive loading, densification, creep, consolidation, frost action or swelling. Stripping, spalling or raveling can be classified into disintegration. The reason for disintegration can be loss of bonding, traffic abrasion, chemical reactivity, aggregate degradation, binder aging or poor consolidation (Base, n.d.).

1.4.2. Surface Distress Characterization

The road surface distress can be characterized by using some indicators, mainly can be divided into two groups: direct statistics and indirect statistics. The direct statistical indicators were based on processing of the vertical road elevation data, such as statistics in particular wavebands (Delanne, 2001), parameters relating to the power spectral density (PSD) (Standardization, 1995) (Andren, 2006), or various straightedge indexes (Song, 2006) (Mucka P. , 2012). The indirect statistical indicators were based on processing of the vibration response of a measuring device to the road surface. Such as international roughness index (IRI), ride number (RN), profilograph index (PrI), mean roughness index (MRI), Mays ride number (MRN), half-car roughness index (HRI), profile index (PI), average rectified slope, average rectified velocity, etc (Mucka P. , 2016). There are other road roughness indexes proposed recently, such as road impact factor (RIF) (Bridgelall, 2014), spectrum evenness index, truck ride index, longitudinal evenness index, dynamic load index, profile index for truck, full-car roughness index, pavement quality index, heavy articulated truck index, vehicle response index, weighted longitudinal profile, corrected unevenness index, novel roughness, ride quality index, heavy vehicle roughness band index, and health index (Mucka P. , 2016). In the pavement management systems and transportation engineering community, the most popular indexes are IRI and PrI, and this paper focuses on another index road impact factor (RIF) (Bridgelall, 2014) which is direct proportionality to the IRI.

The IRI was introduced in 1986, most commonly used as a measure of ride quality (Sayers M. G., 1986) (Sayers M. G., 1986). An annual reporting of the IRI was required to report to the Federal Highway Administration (FHWA). IRI was obtained from longitudinal road profiles. There are many ways to measure the road profile, most common way to do this is either use the instrument measuring the road profile directly or use certain instrument measure with correlation equations for different speeds to relate the actual measurements to IRI.

PrI is sometimes called profile index (PI), it is based on the results of the profilograph. In determining the PrI, there are three key steps: outline trace, position blanking band, compute profile index. Outline trace is for averaging out the spikes and minor deviations by drawing a new profile line through the mid-point of the spikes of the actual road trace. Position blanking band is for eliminating minor elevation deviations, usually use 5 mm width blanking band. The last step is computing profile index, the summation of the heights of the scallops which is the deviations or excursions from the reference line within a segment represents the PrI for that segment (Mucka P. , 2016).

RIF is a measure of ride quality, which is linearly proportional to IRI. To measure IRI, smartphones are used for data collection. Since almost all smartphones contain GPS module and accelerometer module, by mounting a smartphone inside a regular vehicle, the smartphone can detect and record the changes in accelerations and vehicle velocity which are the vehicle response to road roughness (R. Bridgelall, 2016). By measuring the vertical acceleration (G_z), the vehicle velocity (V_k), the RIF can be calculated as below (Bridgelall, 2014):

$$\mathrm{RIF} = \sqrt{\frac{\sigma_t}{\Delta \mathrm{L}} \sum_{\mathrm{k}} [\mathrm{G_z} * \mathrm{V_k}]^2} \tag{1}$$

in which, σ_t is the time interval and ΔL represents the average window size of calculating RIF which is the distance for averaging the road surface condition.

1.5. Road Surface Condition Monitoring

There are many different devices can be used to measure road roughness, such as rod and level pavement profiler, profilograph, high-speed inertial profilers, lightweight inertial profilers, smartphone equipped as data collection sensor vehicle, etc (FHWA, 2016).

1.5.1. Rod and Level Pavement Profiler

The rod and level pavement profiler need to operate manually to measure road roughness, as the Figure 3 shows. This method is easy to use and can get accurate measurement results. However, during operation, the lane needs to be closed, and it is not suitable for large-scale data collection, because of it's low efficiency and high-level labor requirement (FHWA, 2016).

Figure 3. Rod and level pavement profiler (Consultants, 2018)

1.5.2. Profilographs

Profilographs are systems that consist of a frame, a center profiling wheel, and a system to provide a datum as shown in Figure 4. Profilographs are low-speed, usually operated at 2 to 5 miles per hour. This system can collect continuous profile data but it is low speed which makes it is hard to do network-level data collection and it is also insensitive to certain wavelengths (FHWA, 2016).

Figure 4. Profilographs (Group, 2003) (Instruments, n.d.)

1.5.3. Inertial Profilers

The profiling equipment uses inertial reference systems is the most sophisticated. The system consisted of three parts: accelerometer for measuring the vehicle movement, non-contact sensors for measuring displacement between vehicle and road surface, device for measuring distance along roadway. There are two types of inertial profilers, high-speed inertial profilers and lightweight inertial profiler (FHWA, 2016).

High-speed inertial profilers are used by State Highway Agencies for road roughness measurement at network level, considered to be high accurate and efficient. The front and rear of the vehicle equipped measurement equipment, do measurements at posted speeds as shown in Figure 5. This system can do data collection at high speed, and have high repeatability and accuracy. Lane does not need to be closed when doing data collection (FHWA, 2016).

Figure 5. High-speed inertial profilers (AMES, n.d.) (NCAT, n.d.)

Lightweight inertial profiler uses the same technologies as high-speed systems, the difference is using smaller and lighter vehicle, which let them ideal for testing certain road such as new constructed concrete road which has not yet achieved enough strength for supporting regular traffic loading. Figure 6 shows the lightweight inertial profiler (FHWA, 2016).

Figure 6. Lightweight inertial profilers (Purplewave, 2018) (SSI, n.d.)

There are some limitations of inertial profilers, such as high-speed inertial profilers cannot accurately collect data at low speeds, lightweight inertial profilers usually require lane closure. And another limitation is while using the laser-based inertial profilers, especially footprint lasers (point lasers), there might an issue caused by pavement texture, such as artificially high roughness measurements can be caused by the texture of concrete surface which created through longitudinally grooving. Because of this, instead of using tire footprint laser, a line laser is more recommended when doing the measurement on the textured concrete surface.

1.5.4. Smartphone Based Road Condition Monitoring

Smartphone-based road condition monitoring system uses smartphone to represent the three parts of inertial profiler. Because GPS module and accelerometer module are built in smartphone, so the vehicle movement, displacement between vehicle and road surface, roadway distance can be measured by smartphone. This is a simplified style of inertial profiler system, more economy (Bridgelall, 2014). The monitored GPS locations, acceleration, and vehicle speeds will be used to calculate the RIF to indicate the roughness level of the road as shown in Equation (1).

1.6. Problem Statements

The traditional methods for assessing the road conditions using IRI or PrI indexes need special equipment driving at a fixed speed and trained staff, resulting in high cost. The use of RIF index can use the smartphone on a regular vehicle with different speeds for data collection, which is a cost-effective alternative to monitor road condition. However, in practical application, different smartphones may have different sensitivity, resulting in inconsistency. Thus, ways to calibrate different phones to achieve consistent assessment results are urgently needed. In addition, current smartphone-based RIF method is targeted for paved road, there is barely any method to monitor road condition for unpaved roads. Thus, it is needed to validate whether smartphone can be used to monitor unpaved roads and whether the smartphone calibration method is same or not for paved and unpaved roads.

1.7. Objectives and Arrangement of This Thesis

To meet the challenges identified above, the objective of this study is to develop and validate calibration methods for different smartphones to assess the surface conditions using RIF index of both paved and unpaved roads for consistent measurements. To achieve this objective, there are four chapters in addition to the introduction of this thesis as follow:

Chapter 2 introduces the data format and develops two different calibration methodologies for calibrating smartphones for RIF detection;

Chapter 3 sets up and validates the calibration methods for paved roads using three different smartphones through field testing, and compares the two different calibration methods for applications on paved roads;

Chapter 4 sets up and validates the calibration methods for unpaved roads using three different smartphones through field testing, and compares the two different calibration methods for applications on unpaved roads;

Chapter 5 Use different RIF window size to calculate and do all the test again to see if different window size affects the calibration effect.

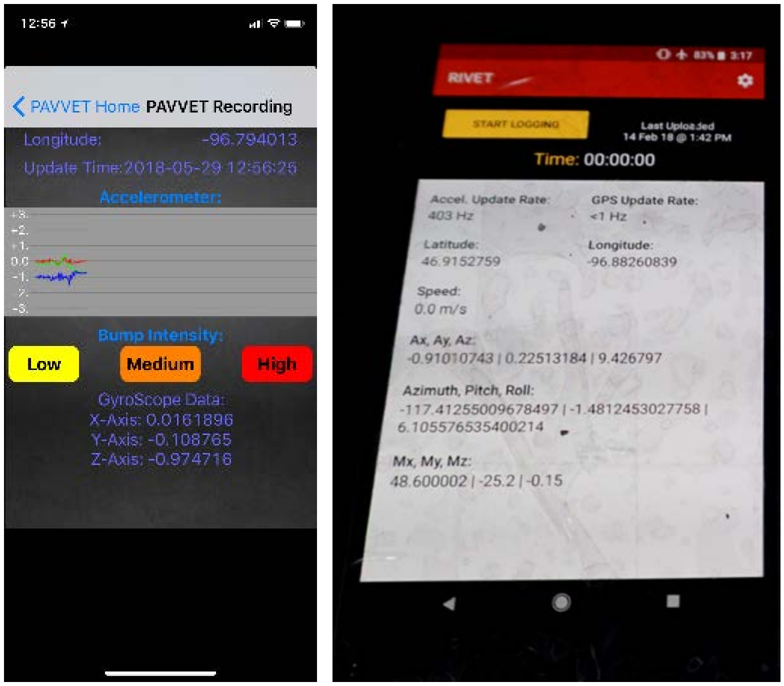
Chapter 6 concludes this study and introduces potential future work.

2. TECHNOLOGY AND CALIBRATION METHODS

Since this study uses smartphone for data collection, the RIF index is used for road condition assessment. As shown in Equation (1). σt represents time interval, which is determined by sampling rate of a particular smartphone. Different phones may have different sampling rate. ΔL represents window size, which is determined manually. Different phones use same window size. The velocity measured by different phones is consistent between most of the phones. However, the accelerations measured by different phones may vary significantly depending on the sensitivity of the accelerometers in each phone.

2.1. Data Format

There are two different apps in different smartphones for RIF data collection. The app used on an iPhone is named as PAVVET and the app used on an android or google phone is named as RIVET. After mounting the smartphones on a smooth surface in a regular vehicle, the user can run smartphone app application either PAVVET on iPhone or PIVET on Android or Google phone. The smartphone app will collect data of time, accelerometer z-value, vehicle speed, and GPS data, and transmitted them wirelessly to a secured web server for data collection and post-analysis. Time and vehicle speed can be used to calculate the vehicle movement and roadway distance and accelerometer z-value can be used to calculate the displacement between vehicle and road surface. We can also locate the specific location using the GPS data.



(a) (b)

Figure 7. (a) PAVVET interface and (b) RIVET interface

Table 1 shows an example of data collected from PAVVET on an iPhone and Table 2 shows an example of data collected from RIVET from an Android or Google phone. As shown in Table 1, for PAVVET on an iPhone, the time column represents the time in milliseconds, and Gx, Gy, Gz represent the g-forces sensed in the lateral, longitudinal and vertical directions and normalized to 9.81 m/s. Latitude and Longitude represent the GPS location, GSpeed represents the vehicle velocity in m/s, Pitch, Roll, and Yaw represent the sensor orientation angles in degrees. Intensity is the parameter input manually to mark the road roughness during the data collection process. RotationX, RotationY, and RotationZ represent the Gyroscope rotation around the x, y, z-axis. The major parameters which are needed to calculate RIF include Time, Gz, and GSpeed.

As shown in Table 2, for RIVET on an Android or google phone, DateTime represents the time in milliseconds, Lat and Lon represent the GPS location parameters, Speed represents the vehicle speed, Ax, Ay, Az represent the accelerometer x, y, z values in meters-per-second-squared, Azimuth is the gyroscope yaw angle in degrees, Pitch is the gyroscope pitch angle in degrees, and Roll is the gyroscope roll angle in degrees. Rx, Ry, Rz represent Gyroscope rotation rate around the x-axis y-axis and z-axis in radians-per-second. Mx, My, Mz represent the geomagnetic field strength along the x, y, z-axis in micro-Tesla. The major parameters needed to calculate RIF include DateTime, Az, and Speed.

Table 1. PAVVET data format

Time	Gz	Latitude	Longitude	GSpeed	Pitch	Roll	Yaw	Gx	Gy	Intensity	RotationX	RotationY	RotationZ
2.960086	-0.99847	46.90135	-96.8828	0.090558	3.19722	3.66581	175.072	0.056366	-0.04015	0	-1.21181	-0.07318	-0.37395
7.856011	-1.00343	46.90135	-96.8828	0.090558	3.19674	3.66556	175.071	0.051529	-0.04689	0	-1.21181	-0.07318	-0.37395
15.69998	-1.00783	46.90135	-96.8828	0.090558	3.19614	3.66636	175.072	0.053833	-0.04282	0	-1.4518	-0.32144	-0.25435
23.17905	-1.0041	46.90135	-96.8828	0.090558	3.19691	3.66812	175.072	0.054947	-0.04295	0	-1.14467	-0.62599	-0.19168
30.78008	-0.99074	46.90135	-96.8828	0.090558	3.19797	3.66609	175.073	0.060181	-0.04106	0	-1.28046	0.541183	-0.31279

Table 2. RIVET data format

DateTime	Lat	Lon	Speed	Ax	Ay	Az	Azimuth	Pitch	Roll	Rx	Ry	Rz	Mx	My	Mz
1.52E+12	46.89617	-96.8829	0	0.886157	0.450264	6.825806	-118.078	-2.98698	6.507976	-0.01551	-0.04347	0.036989	68.25	-36.45	-20.25
1.52E+12	46.89617	-96.8829	0	0.886157	0.450264	6.825806	-118.078	-2.98698	6.507976	-0.01551	-0.04347	0.036989	68.25	-36.45	-20.25
1.52E+12	46.89617	-96.8829	0	0.886157	0.450264	6.825806	-118.078	-2.98698	6.507976	-0.01551	-0.04347	0.036989	68.25	-36.45	-20.25
1.52E+12	46.89617	-96.8829	0	0.886157	0.450264	6.825806	-118.078	-2.98698	6.507976	-0.01551	-0.04347	0.036989	68.25	-36.45	-20.25
1.52E+12	46.89617	-96.8829	0	0.886157	0.450264	6.825806	-118.078	-2.98698	6.507976	-0.01551	-0.04347	0.036989	68.25	-36.45	-20.25

It is expected that the actual signal measured by the smartphones is a combination of low and high-frequency component. Low-frequency component will be relatively stationary while high-frequency components are the noise, usually shown as jumping discontinuities. To eliminate the effect from the white noise, before the calculation of RIF, the noise will be reduced using Wavelet Analysis in SAS (Lane, 2005). Wavelet Analysis can reduce both the low and high-frequency parts of signal noise simultaneously by breaking down and reconstruct the signal by certain rules. However, which filter type used here is not critical in this analysis, only the consistently using the same filter and parameter settings is important.

2.2. Methods for Calibration

After filtering the white noise, two methodologies to calibrate different smartphones are investigated, including 1) Method 1: calibrating toward the mean RIF and 2) Method 2: calibrating toward maximum RIF. The mean RIF is the average RIF value calculated from the entire distance of a calibrating road segment for all the different smartphones to be calibrated and the maximum RIF is the peak RIF value for the most severe bump on a calibrating road segment measured from all different phones to be calibrated. Figure 8 shows the examples of mean RIF and peak RIF for iPhone and Android or google phone for the same segment of paved road. The RIF calculation window size ΔL is set as 10 meters for obtaining Figure 8.

Figure 8 shows that the mean RIF for iPhone and Google phone are in different order due to different sensitivities for accelerometers in different phones. For the same anomaly on a paved road, an iPhone monitors the mean RIF around 0.215 and a Google phone monitors the mean RIF around 0.196. It is obvious that the two types of phones have significant different measurement sensitivity toward measuring RIF for road condition monitoring.

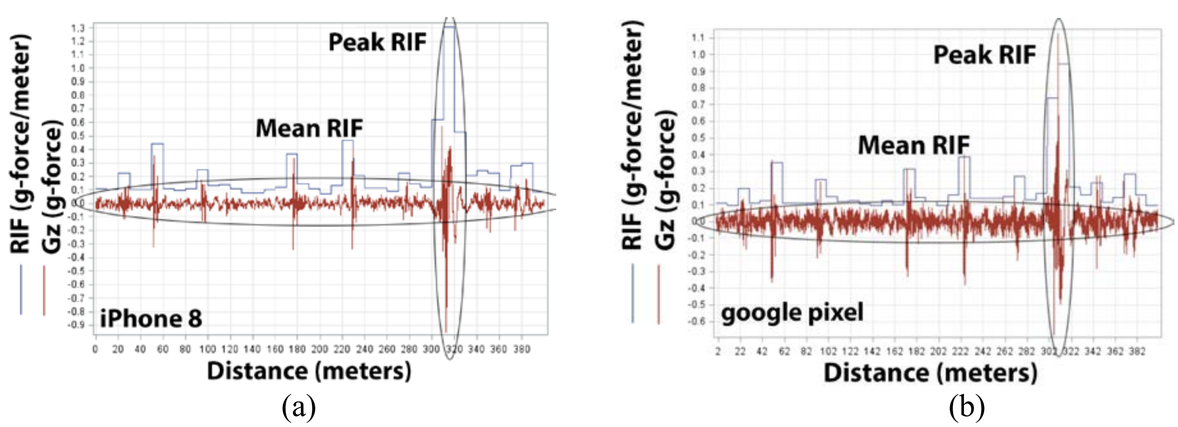


Figure 8. Comparison of Gz and RIF Indices for run 1

To calibrate toward the mean RIF, the first calibration method, the mean RIF value of all three phones from all travels will be used to calculate the reference mean RIF. Then the mean RIF values of all other phones will be scaled to the reference mean RIF for measuring the RIF for the same road segment. The mean RIF ratio in between the reference phone and all the others to be calibrated will be used to measure the RIF for all other road segments for road condition monitoring.

To calibrate toward the peak RIF, the second calibration method, the peak RIF value of a reference phone from all travels of a known calibrating road segment will be used to calculate the reference peak RIF. Then the peak RIF values of all other phones will be scaled to the reference peak RIF for measuring the RIF for the same road segment. The peak RIF ratio in between the reference phone and all the others to be calibrated will be used to measure the RIF for all other road segments for road condition monitoring.

To see which calibration method works better, the standard deviation will be calculated and compared. Analysis of variance (ANOVA) test will also be performed among three phones. Margin of error with 95% confidence (MOE95) will be calculated among multiple runs. Standard deviation is used to quantify the amount of variation among a set of data. Small standard deviation means the data tend to close to the mean value, and large standard deviation means the data are spread out (Bland, 1996). A smaller standard deviation for multiple runs among different phones indicates a better calibration method. ANOVA test is used to test the difference between the group means. The hypothesis of ANOVA test is that three phones have the same RIF means. After the ANOVA test, the statistic F value decides if the hypothesis should be rejected or not. If the F value is smaller than the critical F value, it means that the hypothesis fails to reject and the three phones have the same RIF means. If the F value is larger than the critical F value, it means that the hypothesis should be rejected and the three phones have different RIF means. The larger the F value is, the more the data are spread out. The MOE (margin-of-error) represents the reliability of the mean. The smaller the MOE, the more reliable the mean is.

2.3. Summary

Two calibration methodologies are introduced to calibrate the differences between different phones including iPhone, Android or Google phone or any other types of smartphones. The first method calibrates the phones toward the mean RIF and the second method calibrates the phones toward the maximum RIF. To prepare the data for calibration, all the data obtained from each phone will be filtered using Wavelet Analysis and using Equation (1) to calculate the RIF. The standard deviation, ANOVA test and MOE95 will be used to compare the two different calibration methods for different pavement conditions to determine a most appropriate calibration method for various road surface conditions.

3. FIELD TESTING ON PAVED ROADS

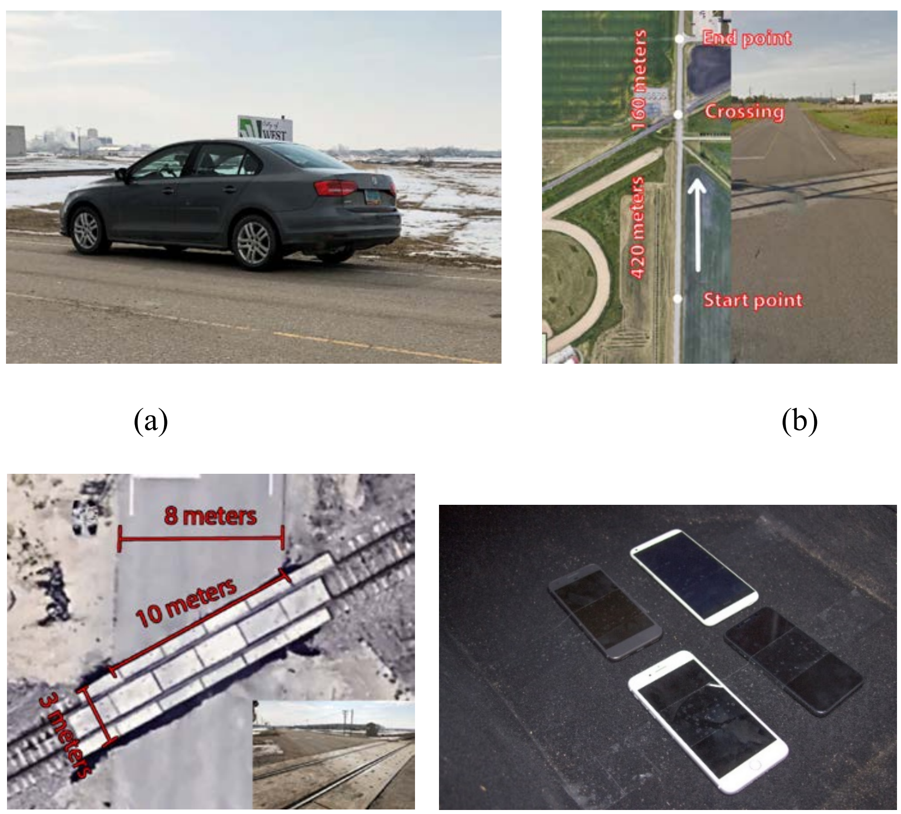
To test the two different calibration methods, field testing was performed on a paved road segment with a railroad grade crossing for 35 runs. The data was collected using three different smartphones on a regular passenger vehicle. More details are in the sections below.

3.1. Field Data Collection Setup

The regular passenger vehicle used for the field road test is a 2015 Volkswagen Jetta, as shown in Figure 9 (a). The field road test was performed on a segment of a paved road in Fargo, ND, USA, as shown in Figure 9 (b). This paved road segment has a road width of 8m. The railroad grade crossing as the maximum bump as shown in Figure 9 (c) has a length of 3 m and a width of 10 m. The total road segment for testing is 580m. From the beginning point of the road test to the crossing is 420 m, and from the ending point of the road test to the crossing is 160 m. Four smartphones were mounted on the front seat floor of the vehicle using tape horizontally, as shown in Figure 9 (d). The models of the smart phones were selected randomly based on the availability, which includes an iPhone 8, an iPhone X, a google pixel phone, and an HTC Android phone. The iPhones used the app PAVVET and the Google and Android phone used PIVET app to collect data. In practical application, all the phones needed to be calibrated no matter what model it belongs to for road condition monitoring. Different phones have different sampling frequency and sensor sensitivity. In this field test, the iPhones used iOS, which has a sampling frequency 90Hz, and the Google and Android phone has a sampling frequency of 390Hz.

Among these four phones, the model iPhone 8 was previously calibrated using the traditional road profiler, which was used as reference to check the reliability of the calibration (R. Bridgelall, 2016). In addition, the HTC Android phone failed to collect the correct GPS data during the data collection, so in the data analysis later, only three sets of data from the rest three phones were analyzed and the HTC Android was excluded in the analysis.

For a valid statistical analysis, 35 runs were performed on the same paved road segment as shown in Figure 9 (b). For each run, the vehicle started at the same start point and ended at the same end point, and each run used the same direction for road test as shown in Figure 9 (b). However, due to the fact that the activation of the app on each phone was operated manually, time delay was expected in between different phones and different runs. Thus, in the data analysis, based on the known start point and end point of the data collection, the data on each phone for each run was aligned into the same length 400 m for a valid comparison.



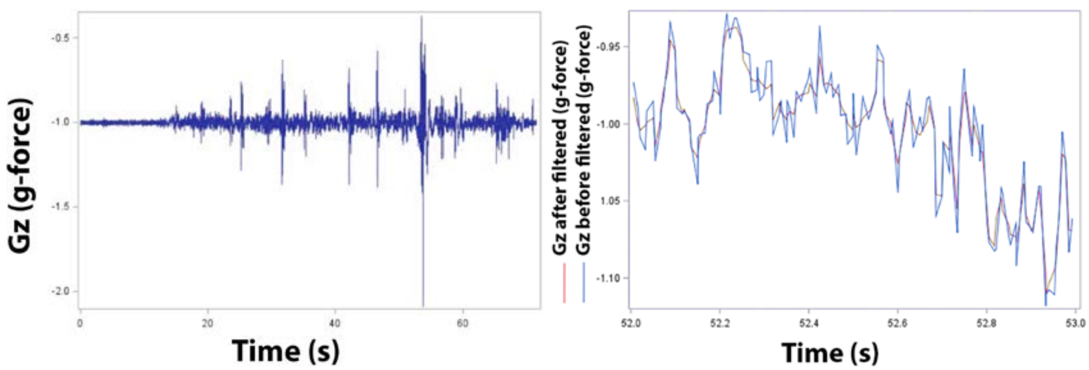
(a) (b)

(c) (d)

Figure 9. (a) Vehicle for road test (b) Paved test road (c) Railroad grade crossing (d) Four smartphones were mounted on the front seat floor

3.2. Original Data Without Calibration

Figure 10 (a) shows an example of the collected raw Gz signals from the 35 runs, in which the x-axis represents the driving time and the y-axis is the Gz values. It can be seen from Figure 10 (a) that the original Gz signal contains significant high-frequency noise. Thus, the Wavelet Filter was used to reduce the noise as shown in Figure 10 (b) to smooth the signal. In Figure 10 (b), the blue line represents the Gz before noise reduction and the red line represents the filtered Gz.



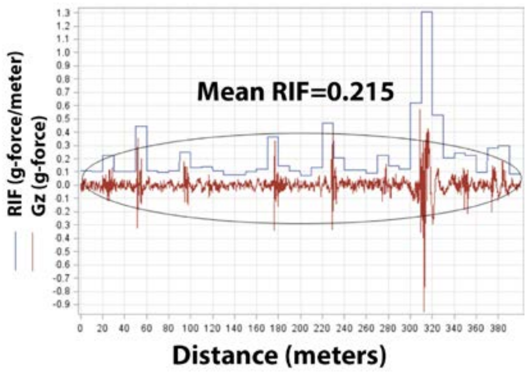
(a) (b)

Figure 10. (a) Raw Gz signal and (b) filtered Gz on the tested paved road segment

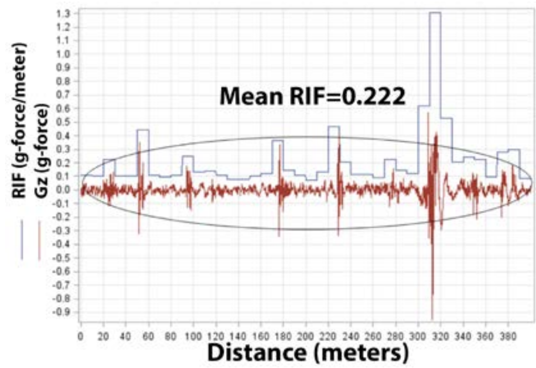
After filtering the noise, the RIF distributions were calculated for the 35 runs on the paved road segment from all the three correctly operated phones (iPhone 8, iPhone X, and Google Pixel) on the vehicle. Based on the known start point and end point of the data collection, the data on each phone for each run was aligned into the same length. Figures 11 (a~c) show the results of the first run from the different phones for the paved road segment. From Figure 11, it can be seen that all the three phones can identify the location and roughness of the railroad crossing on the paved road very clearly from the signal as the peak RIF.

The mean RIF for this paved road segment calculated from the iPhone 8, iPhone X and google pixel are 0.215, 0.222 and 0.196. The standard deviation among these three phones is

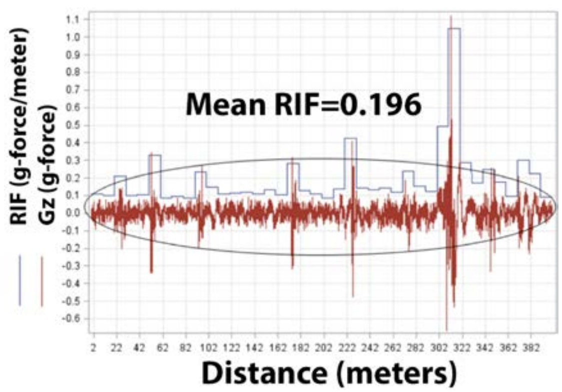
0.014. Indicates that different phones may have different sensor sensitivities for the embedded accelerometer in each phone.



(a)



(b)



(c)

Figure 11. Comparison of Gz and RIF Indices for Traversal 1 (a) iPhone 8, (b) iPhone x, and (c) google pixel

Table 3 shows the results for all the 35 runs of road testing on the paved road segment. From this table it can be seen that for this paved road, the average of 35 runs mean RIF before calibration for iPhone 8, iPhone x and google pixel are 0.195, 0.199, and 0.185, respectively. The standard deviation between the three phones for this paved road is 0.007.

Table 3. Mean RIF on a paved road with a railroad crossing obtained from iPhone 8, iPhone X, and google pixel phones

Run	iPhone 8 mean RIF (g-force/meter)	iPhone X mean RIF (g-force/meter)	google pixel mean RIF (g-force/meter)
1	0.215	0.222	0.196
2	0.209	0.213	0.192
3	0.203	0.208	0.189
4	0.204	0.209	0.178
5	0.194	0.201	0.194
6	0.208	0.211	0.194
7	0.199	0.210	0.186
8	0.196	0.198	0.178
9	0.193	0.198	0.187
10	0.193	0.198	0.180
11	0.190	0.194	0.191
12	0.199	0.201	0.175
13	0.194	0.201	0.186
14	0.196	0.202	0.184
15	0.196	0.206	0.190
16	0.203	0.209	0.176
17	0.193	0.196	0.175
18	0.173	0.184	0.172
19	0.178	0.189	0.199
20	0.206	0.206	0.178
21	0.189	0.190	0.181
22	0.190	0.194	0.182
23	0.193	0.193	0.183
24	0.195	0.192	0.183
25	0.194	0.198	0.181
26	0.192	0.191	0.185
27	0.187	0.194	0.194
28	0.199	0.202	0.184
29	0.200	0.198	0.191
30	0.195	0.196	0.186
31	0.192	0.195	0.181
32	0.190	0.187	0.179
33	0.186	0.189	0.180
34	0.182	0.190	0.180
35	0.189	0.192	0.191
mean	0.195	0.199	0.185

Table 4 shows the ANOVA test result of mean RIF for all 35 runs on the tested paved road segment from all the three phones. The F value is 29.196, the critical F value is 3.085. Since F value is significantly larger than critical F value, the hypothesis that three phones have equal mean RIF is not true. Figure 12 shows the mean RIF distribution of the three phones. Table 5 and Figure 13 show the MOE95 for the three phones. From this table it can be seen that for this paved road, the MOE95 of iPhone 8, iPhone X began to fall below 2% percentage after 23 and 18 runs, google pixel began to fall below 2% percentage after 19 runs. Thus, if multiple tests are performed, the iPhone 8 and iPhone X will get more consistent results after 23 and 18 runs and google pixel will get more consistent results after 19 runs, which are very consistent.

Table 4. ANOVA test of mean RIF for all 35 runs on the tested paved road segment from all the three phones

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F critical
Between Groups	0.004	2.000	0.002	29.196	0.000	3.085
Within Groups	0.006	102.000	0.000			
Total	0.010	104.000				

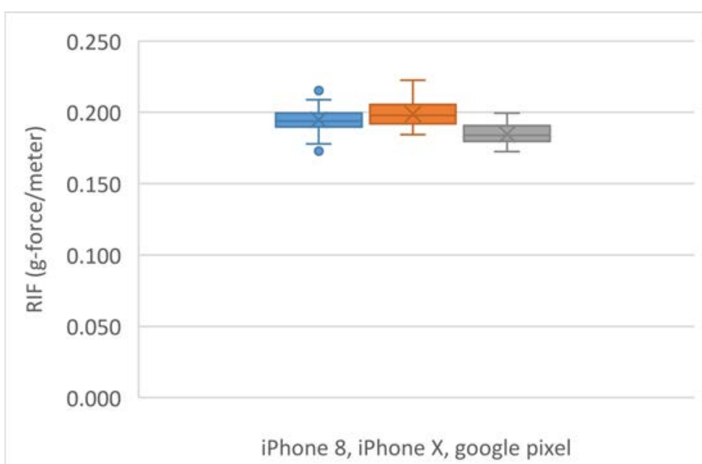


Figure 12. Mean RIF from paved roads from iPhone 8, iPhone X, and google pixel

Table 5. Mean RIF and MOE95 on a paved road with a railroad crossing obtained from iPhone 8

Run	iPhone8 MOE95	iPhone X MOE95	google pixel MOE95
2	2.9%	4.1%	2.0%
3	3.4%	3.9%	2.1%
4	2.7%	3.0%	4.0%
5	3.3%	3.3%	3.3%
6	2.8%	2.7%	2.8%
7	2.5%	2.3%	2.4%
8	2.4%	2.5%	2.6%
9	2.4%	2.5%	2.3%
10	2.4%	2.4%	2.2%
11	2.4%	2.5%	2.1%
12	2.2%	2.3%	2.2%
13	2.1%	2.1%	2.0%
14	2.0%	2.0%	1.9%
15	1.8%	1.9%	1.8%
16	1.7%	1.7%	1.8%
17	1.7%	1.7%	1.8%
18	2.2%	2.0%	1.9%
19	2.3%	2.0%	2.0%
20	2.2%	1.9%	1.9%
21	2.2%	1.9%	1.8%
22	2.1%	1.9%	1.7%
23	2.0%	1.8%	1.7%
24	1.9%	1.8%	1.6%
25	1.8%	1.7%	1.5%
26	1.8%	1.7%	1.5%
27	1.7%	1.7%	1.5%
28	1.7%	1.6%	1.4%
29	1.6%	1.5%	1.4%
30	1.6%	1.5%	1.3%
31	1.5%	1.5%	1.3%
32	1.5%	1.5%	1.3%
33	1.5%	1.5%	1.3%
34	1.5%	1.4%	1.2%
35	1.5%	1.4%	1.2%

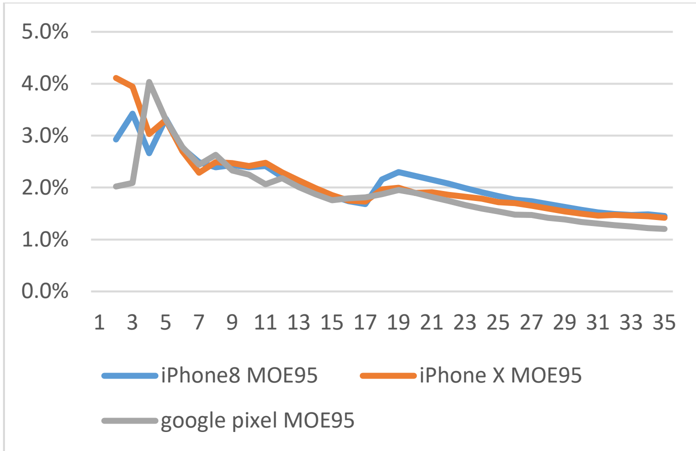


Figure 13. MOE95 of mean RIF from paved roads from iPhone 8, iPhone X, google pixel

The velocity should be the same among the three phones since they all fixed on the car collecting data simultaneously. Since the RIF is calculated based on Gz and vehicle velocity and, the deviation among three different phones may be attributed to the sensor sensitivity difference of Gz among different phones. The iPhone 8 and iPhone X measured the mean RIF very close while the Google pixel phone has much smaller results than the two iPhones. Since different models of iPhones have the same standard during manufacturing process, which makes the sensitivity between different iPhones are close to each other. The accelerometer iPhone 8 and iPhone X use is Bosch's 6-Axis IMU (Wire, 2018).

Although the accelerometer model Google pixel used was not reported, the manufacturing process and the sensors inside a google pixel is significantly different than the iPhones, which may result in the significant differences in the accelerometer sensor sensitivity. Thus, calibration is needed to achieve a consistent RIF measurement from different phones for paved roads.

3.3. Calibrating Phones Towards the Mean RIF for Paved Roads

To test the first calibration method for the paved roads, which is calibrating different phones towards the mean RIF, we used the mean RIF from the all 35 runs of the field test to perform the calibration and it was performed on all the 35 runs to do the validation. The mean RIF calculated from all 35 runs of the road test on this paved road segment from iPhone 8, iPhone X and google pixel are 0.195, 0.199 and 0.185. The average mean RIF of the paved road test from the three Phones is 0.193. The average mean RIF is used as a reference to calibrate all the three phones. To calibrate all the three phones, the mean RIF measured by each phone is divided by 0.193, resulting in a calibration coefficient of 1.011, 1.031, and 0.958 for the three phones on the paved road segment, respectively. Specifically, for the iPhone 8, the measured Gz value from all the 35 runs of the road test divide the calibration coefficient of 1.011 to get calibrated; for the iPhone X, the measured Gz value divide the calibration coefficient of 1.031; and for the Google pixel, the measured Gz value divide 0.958 to get calibrated. After calibrating the Gz based on the sensor sensitivity difference using calibration coefficients for each phone, the RIF for each phone and each run is recalculated as calibrated RIF.

Table 6 lists the calibrated mean RIF calculated from the three phones for all the 35 runs of the road test for the same paved road segment using the mean RIF calibration method.

Table 6. Towards mean RIF method calibrated mean RIF from the iPhone 8, iPhone X, and google pixel for the paved road

Run	iPhone 8 mean RIF (g-force/meter)	iPhone X mean RIF (g-force/meter)	google pixel mean RIF (g-force/meter)
1	0.213	0.216	0.204
2	0.207	0.207	0.200
3	0.200	0.201	0.197
4	0.202	0.203	0.186
5	0.192	0.195	0.203
6	0.206	0.205	0.202
7	0.197	0.203	0.194
8	0.194	0.192	0.186
9	0.191	0.192	0.195
10	0.191	0.192	0.188
11	0.188	0.188	0.200
12	0.197	0.195	0.183
13	0.192	0.195	0.194
14	0.194	0.196	0.192
15	0.194	0.199	0.198
16	0.201	0.202	0.184
17	0.191	0.190	0.183
18	0.171	0.179	0.180
19	0.176	0.183	0.208
20	0.204	0.200	0.185
21	0.187	0.184	0.189
22	0.188	0.188	0.190
23	0.191	0.187	0.191
24	0.193	0.186	0.191
25	0.192	0.192	0.189
26	0.190	0.185	0.193
27	0.185	0.188	0.203
28	0.197	0.195	0.192
29	0.198	0.192	0.199
30	0.193	0.190	0.195
31	0.190	0.189	0.189
32	0.188	0.181	0.187
33	0.184	0.183	0.188
34	0.180	0.184	0.188
35	0.187	0.186	0.200
mean	0.193	0.193	0.193

Table 7 shows the ANOVA test result of mean RIF for all 35 runs on the tested paved road segment from all the three phones. The F value is 0 and the critical F value is 3.085. Since the F value is smaller than the critical F value, the hypothesis that three phones have equal mean RIF is true. Thus, the calibration towards mean RIF for paved roads works well. Figure 14 shows the three phones calibrated mean RIF distribution.

Table 7. ANOVA test of towards mean RIF method calibrated mean RIF for all 35 runs on the tested paved road segment from all the three phones

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F critital
Between Groups	0.000	2.000	0.000	0.000	1.000	3.085
Within Groups	0.006	102.000	0.000			
Total	0.006	104.000				

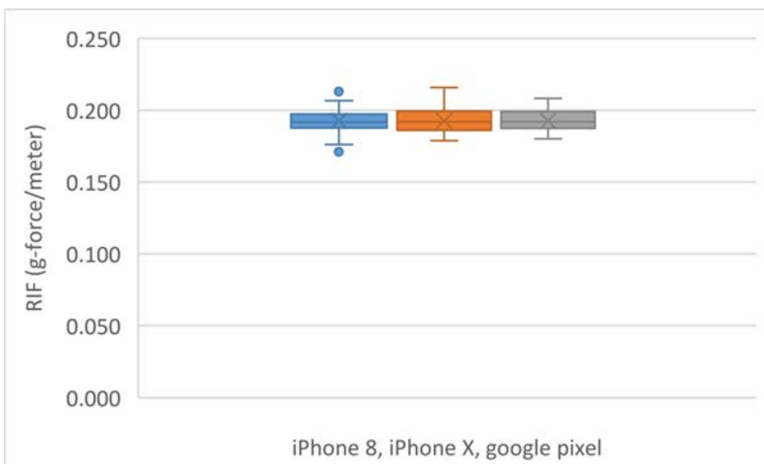


Figure 14. Towards mean RIF method calibrated mean RIF for paved roads from iPhone 8, iPhone X, and google pixel

3.4. Calibrating Phones Towards Maximum RIF for Paved Roads

The second calibration method for paved roads between different phones is the reference maximum RIF method. For the all 35 runs of the paved road test, the mean peak RIF measured from the iPhone 8, iPhone X and google pixel are 1.113, 1.177 and 0.955, Due to the fact that the iPhone model 8 was calibrated using traditional road profilers, the iPhone 8 mean peak RIF is used as a reference to calibrate all the three phones. Dividing the peak RIF from each phone using the iPhone 8 mean peak RIF, which is 1.113, the calibration coefficients for each phone are 1, 1.058 and 0.858 respectively. Thus, for iPhone 8, all the measured Gz values do not change; for the iPhone X, all the measured Gz values were divided by 1.058 for calibration; and for the Google pixel phone, all the measured Gz values were multiplied by 0.858 for calibration. After the calibration of Gz for each phone, the calibrated RIF for each phone and each run on the paved road segment were recalculated. Table 8 shows the calibrated mean RIF for all 35 runs on the tested paved road segment from all the three phones.

Table 8. Towards maximum RIF method calibrated mean RIF from the iPhone 8, iPhone X, and google pixel for the paved road

Run	iPhone 8 mean RIF (g-force/meter)	iPhone X mean RIF (g-force/meter)	google pixel mean RIF (g-force/meter)
1	0.215	0.210	0.228
2	0.209	0.202	0.224
3	0.203	0.196	0.220
4	0.204	0.198	0.207
5	0.194	0.190	0.226
6	0.208	0.199	0.226
7	0.199	0.198	0.217
8	0.196	0.187	0.207
9	0.193	0.188	0.218
10	0.193	0.187	0.210
11	0.190	0.183	0.223
12	0.199	0.190	0.204
13	0.194	0.190	0.216
14	0.196	0.191	0.214
15	0.196	0.194	0.221
16	0.203	0.197	0.205
17	0.193	0.185	0.204
18	0.173	0.174	0.201
19	0.178	0.179	0.232
20	0.206	0.195	0.207
21	0.189	0.179	0.211
22	0.190	0.183	0.212
23	0.193	0.183	0.213
24	0.195	0.182	0.213
25	0.194	0.187	0.211
26	0.192	0.180	0.215
27	0.187	0.184	0.226
28	0.199	0.191	0.215
29	0.200	0.187	0.222
30	0.195	0.185	0.217
31	0.192	0.184	0.211
32	0.190	0.177	0.209
33	0.186	0.179	0.209
34	0.182	0.179	0.210
35	0.189	0.182	0.223
mean	0.195	0.188	0.215

Table 9 shows the ANOVA test result of mean RIF for all 35 runs on the tested paved road segment from all the three phones. The F value is 106.011 and the critical F value is 3.085.

Since the F value is larger than the critical F value, and even larger than the F value before calibrated which is 29.196, the hypothesis that three phones have equal mean RIF is not true. Thus, the calibration towards maximum RIF for paved roads did not work well. Figure 15 show the calibrated mean RIF distribution for the three phones. Can be seen that after calibration, the mean RIF still spread out.

Table 9. ANOVA test of towards maximum RIF method calibrated mean RIF for all 35 runs on the tested paved road segment from all the three phones

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F critical
Between Groups	0.014	2.000	0.007	106.011	0.000	3.085
Within Groups	0.007	102.000	0.000			
Total	0.021	104.000				

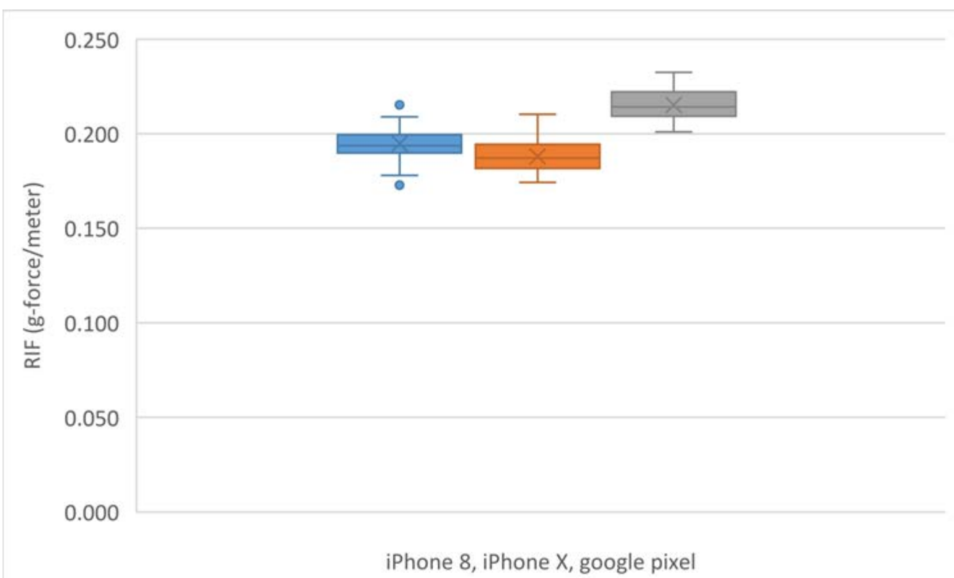
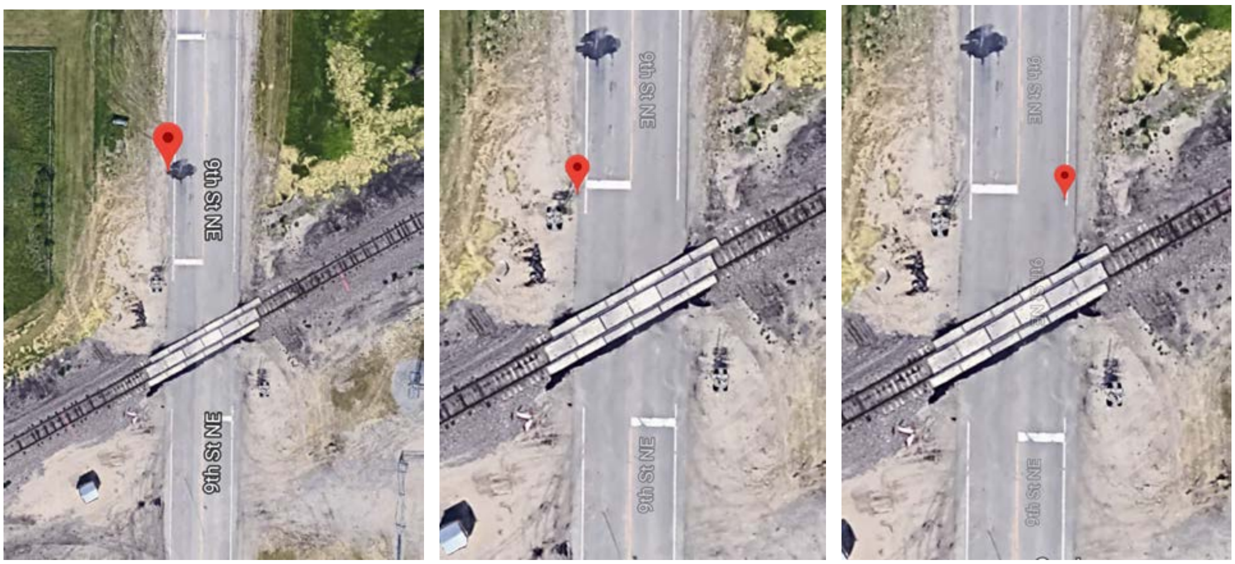


Figure 15. Towards maximum RIF method calibrated mean RIF from paved roads from iPhone 8, iPhone X, and google pixel

3.5. GPS Output To Locate The Reference Peak RIF

According to the measured peak RIF, the GPS output location, Figure 16 shows the located railroad crossing locations on google map. Although the location of the peak RIF was expected to be at the railroad crossing, time delay causes the measurement location deviation of the locations, which may result in measurement error. However, since the RIF is an average road roughness index for every ΔL, which is 10 m in this study, the deviation of the GPS locations is still within the acceptable range of the measurement of 10m resolution.



(a) (b) (c)

Figure 16. Paved road peak RIF location from the (a) iPhone 8, (b) iPhone X, and (c) google pixel phone

3.6. Comparison of The Two Different Calibration Methods

ANOVA test F value is used to compare different calibration methods. Before the calibration, the F value among the three phones on the paved road segment is 29.196. Calibrating toward the mean RIF, the first calibration method results in F value to be 0. Calibrating toward the peak RIF, the second calibration method results in F value to be 106.011. Thus, for paved roads, calibrating based on the mean RIF shows a smaller F value of 0 when compared to calibrating based on the peak RIF of 106.011, indicating a better way to calibrate different phones on paved roads.

3.7. Summary

In this chapter, the two proposed phone calibration methods, calibrating towards the mean RIF (method 1) and the peak RIF (method 2), were tested on a paved road segment with a railroad crossing as a known reference bump for the peak RIF. The original data from the three phones showed that different phones may deliver a measurement different in different orders (one order difference between the iPhone and Google phone) without a calibration. The calibration for different phones before practical applications is a must to get a consistent measurement from different phones for the same paved road segment. The analysis in this chapter indicated that for the paved roads, calibrating towards the mean RIF leads to a better result when compared to calibrating toward the peak RIF method.

4. FIELD TESTING ON UNPAVED ROADS

To compare whether the calibration method would make a difference for different road surface materials, field testing was also performed on an unpaved road segment with a railroad grade crossing for 35 runs. The data was collected using the same setup as for the paved road with three different smartphones on a regular passenger vehicle. More details are in the sections below.

4.1. Field Data Collection Setup

The same passenger vehicle used for the paved road testing was used for testing the unpaved road segment as shown in Figure 9 (a). The same four smartphones were used as in the paved road testing on the car front seat floor, as shown in Figure 9 (d), including an iPhone 8, an iPhone X, a google pixel phone, and an HTC Android phone. The same apps were used to collect data from the phones including PAVVET for iPhones and PIVET for the Google and Android phones. The HTC Android phone still failed to collect the correct GPS data and was excluded in the analysis.

The field road test was performed on a segment of an unpaved road in Fargo, ND, USA, as shown in Figure 17 (b). To be comparable to the paved road testing, the selected unpaved road segment in this field testing has a width of 6 m. This unpaved road segment also includes a railroad grade crossing as the maximum bump as shown in Figure 17 (c), which has a length of 2.5 m and a width of 6.7 m. The total length of the road testing is 620 m. From the beginning point of the road test to the crossing is 280 m, and from the ending point of the road test to the crossing is 340 m. 35 different runs were also performed on this same unpaved road segment as shown in Figure 17 (b). For each run, the vehicle also started at the same start point and ended at the same end point, and each run used the same direction for road test. The data on each phone for each run was also aligned into the same length 300 m for a valid comparison.

(a)

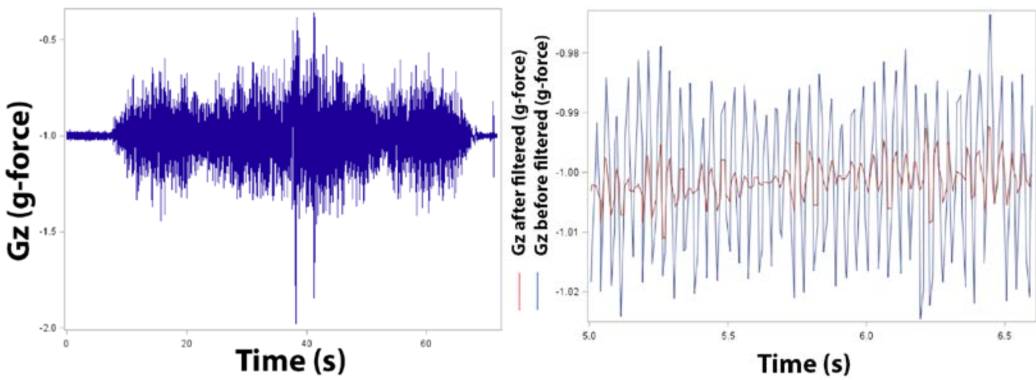
(b)

(c)

Figure 17. (a) Vehicle for unpaved road test, (b) google photo of the unpaved road segment, and (c) railroad crossing on the unpaved road segment

4.2. Original Data Without Calibration

Figure 18 (a) shows an example of the original Gz signals from the 35 runs. When compare to Figure 10 (a). Can be seen that the original Gz signal for the unpaved roads has a much bigger noise level, which contains significant high-frequency noise. Thus, the Wavelet Filter was also used to reduce the noise as shown in Figure 18 (b) to smooth the signal.



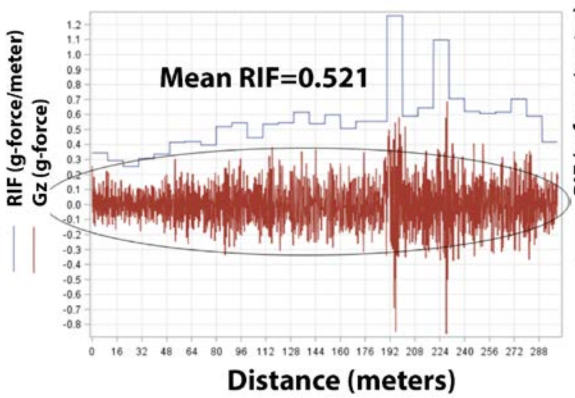
(a) (b)

Figure 18. (a) Raw Gz signal on unpaved road and (b) filtered Gz

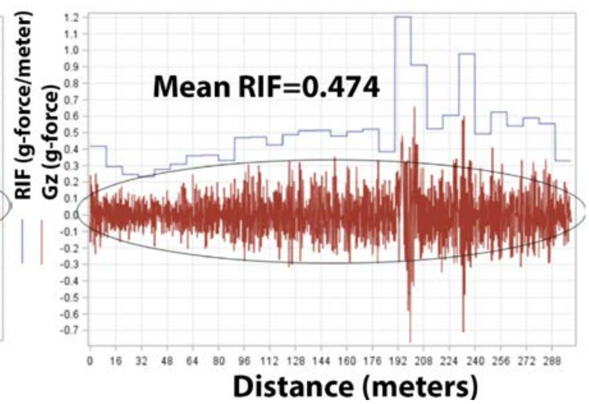
After filtering the noise, the RIF distributions were calculated for the 35 runs from all the three correctly operated phones (iPhone 8, iPhone X, and Google Pixel) on the vehicle. Based on the known start point and endpoint of the data collection, the data on each phone for each run was aligned into the same length. Figures 19 (a~c) show the results of the first run from the different phones on the unpaved road segment.

From Figure 19, Can be seen that all the three phones can identify the location and roughness of the railroad crossing from the signal as the peak RIF. For the unpaved road, the mean RIF calculated from the iPhone 8, iPhone X and google pixel are 0.521, 0.474 and 0.461. The standard deviation between the three phones for this unpaved road is 0.032. Comparing to the paved road standard deviation 0.007, the original mean RIF measured on an unpaved road indicated a larger derivation due to the poor and rough surface conditions when compared a RIF measurement on a paved road.

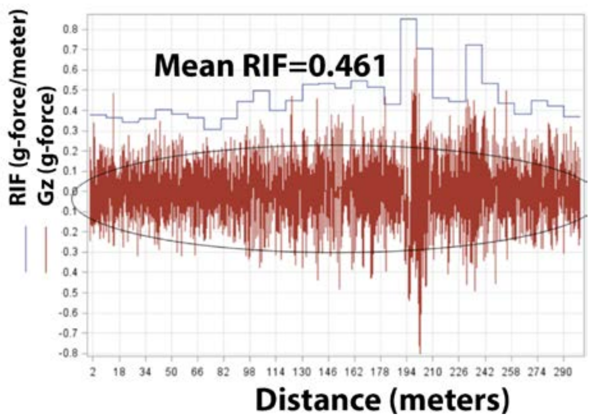
Figure 19 also validate that different phones have different sensor sensitivities for the embedded accelerometer in each phone. The standard deviation among these three phones is 0.045, indicating that different phones may have different sensor sensitivities for the embedded accelerometer in each phone. Table 10 shows the results for all the 35 runs of road testing for the unpaved road segment.



(a)



(b)



(c)

Figure 19. Measured original RIF for unpaved road segment using (a) iPhone 8, (b) iPhone X, and (c) google pixel phone

Table 10. Mean RIF on an unpaved road with a railroad crossing obtained from iPhone 8, iPhone X, and google pixel phones

Run	iPhone 8 mean RIF (g-force/meter)	iPhone X mean RIF (g-force/meter)	google pixel mean RIF (g-force/meter)
1	0.552	0.498	0.463
2	0.604	0.541	0.442
3	0.537	0.457	0.467
4	0.563	0.519	0.474
5	0.571	0.526	0.461
6	0.539	0.481	0.466
7	0.548	0.497	0.447
8	0.490	0.455	0.455
9	0.515	0.458	0.456
10	0.517	0.485	0.466
11	0.534	0.517	0.468
12	0.541	0.475	0.460
13	0.538	0.492	0.449
14	0.473	0.454	0.447
15	0.479	0.453	0.464
16	0.507	0.473	0.451
17	0.521	0.458	0.474
18	0.499	0.465	0.487
19	0.553	0.486	0.444
20	0.489	0.434	0.436
21	0.482	0.448	0.455
22	0.498	0.449	0.449
23	0.507	0.438	0.478
24	0.534	0.469	0.455
25	0.508	0.450	0.448
26	0.495	0.442	0.451
27	0.500	0.455	0.463
28	0.488	0.456	0.458
29	0.512	0.477	0.449
30	0.483	0.438	0.474
31	0.541	0.495	0.483
32	0.555	0.500	0.475
33	0.534	0.473	0.462
34	0.514	0.493	0.481
35	0.528	0.489	0.491
mean	0.521	0.474	0.461

Table 11 shows the ANOVA test result of mean RIF for all 35 runs on the tested unpaved road segment from all the three phones. The F value is 58.427 and the critical F value is 3.085. Since the F value is larger than the critical F value, the hypothesis that three phones have equal mean RIF is not true. Figure 20 shows the mean RIF distribution for the three phones for the unpaved roads. Table 12 and Figure 21 show the MOE95 for iPhone 8, iPhone X, google pixel. From this table can be seen that for this unpaved road, the MOE95 of iPhone 8 began to fall below 2% percentage after 34 runs, iPhone X began to fall below 2% percentage after 33 runs, google pixel began to fall below 2% percentage after 6 runs. Thus, for iPhone 8 and iPhone X, to get a consistent result, 34 and 33 runs are needed and for google pixel, 6 runs are needed.

Table 11. ANOVA test of mean RIF for all 35 runs on the tested unpaved road segment from all the three phones

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F critical
Between Groups	0.070	2.000	0.035	58.427	0.000	3.085
Within Groups	0.061	102.000	0.001			
Total	0.131	104.000				

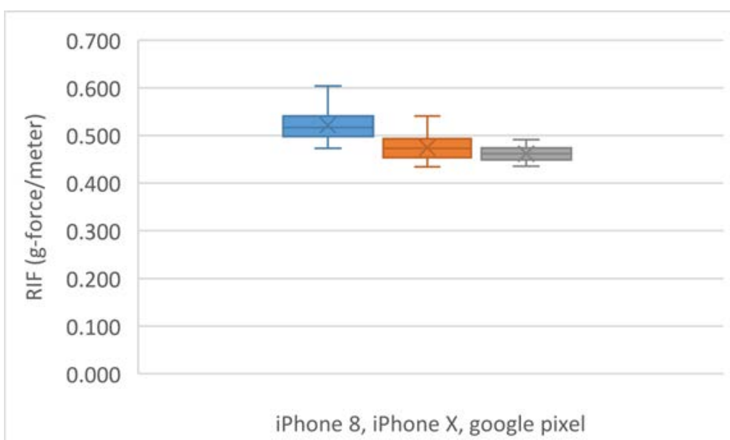


Figure 20. Mean RIF from unpaved roads from iPhone 8, iPhone X, and google pixel

Table 12. Mean RIF and MOE95 on unpaved road with a railroad crossing obtained from iPhone 8, iPhone X, google pixel

Run	iPhone8 MOE95	iPhone X MOE95	google pixel MOE95
2	8.74%	8.03%	4.56%
3	7.05%	9.55%	3.35%
4	4.99%	6.96%	2.91%
5	3.88%	5.63%	2.26%
6	3.54%	4.97%	1.86%
7	3.08%	4.22%	1.84%
8	4.10%	4.38%	1.62%
9	3.91%	4.26%	1.44%
10	3.67%	3.83%	1.32%
11	3.34%	3.57%	1.23%
12	3.05%	3.33%	1.13%
13	2.81%	3.06%	1.10%
14	3.19%	3.05%	1.09%
15	3.32%	3.02%	1.03%
16	3.17%	2.86%	0.98%
17	2.99%	2.77%	1.00%
18	2.91%	2.66%	1.15%
19	2.79%	2.51%	1.16%
20	2.76%	2.60%	1.22%
21	2.76%	2.56%	1.16%
22	2.69%	2.51%	1.12%
23	2.59%	2.51%	1.13%
24	2.48%	2.41%	1.09%
25	2.39%	2.36%	1.06%
26	2.34%	2.33%	1.03%
27	2.28%	2.27%	0.99%
28	2.25%	2.20%	0.95%
29	2.17%	2.12%	0.93%
30	2.16%	2.11%	0.93%
31	2.10%	2.06%	0.95%
32	2.07%	2.02%	0.95%
33	2.01%	1.96%	0.92%
34	1.96%	1.92%	0.93%
35	1.90%	1.87%	0.97%

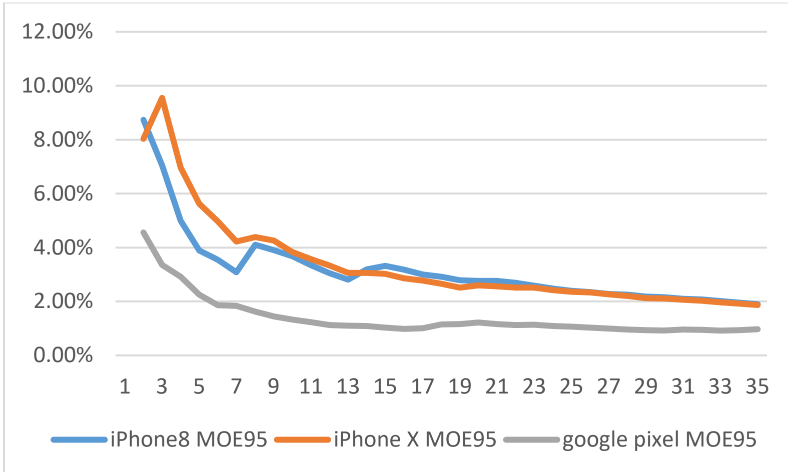


Figure 21. MOE95 of mean RIF from unpaved roads from iPhone 8, iPhone X, google pixel

4.3. Calibrating Phones Towards the Mean RIF for Unpaved Roads

To test the first calibration method for the unpaved roads based on the mean RIF, the mean RIF from all 35 runs of the field test was calculated and calibration was performed on all the 35 runs to do the validation. The mean RIF calculated from the unpaved road test from iPhone 8, iPhone x and google pixel are 0.521, 0.474 and 0.461. The average mean RIF of the unpaved road test is 0.486. When compared to paved road, it can be seen that the paved road has an average mean RIF of 0.193 for the three phones. The mean RIF of the unpaved road is around two times bigger than the mean RIF of paved road, indicating a significantly rougher road surface condition.

To calibrate all the three phones, the mean RIF measured by each phone from all 35 runs on the unpaved road is divided by the average mean RIF from the three phones, which is 0.486, resulting in a calibration coefficient of 1.074, 0.976, and 0.950 for the three phones, respectively. Specifically, for the iPhone 8, the measured Gz value from all the 35 runs of the road test divide the calibration coefficient of 1.074 to get calibrated; for the iPhone X, the measured Gz value divide the calibration coefficient of 0.976; and for the Google pixel, the measured Gz value divide 0.950 to get calibrated. After calibrating the Gz based on the sensor sensitivity difference using calibration coefficients for each phone, the RIF for each phone and each run is recalculated as calibrated RIF. When compared to paved roads, the three phones have calibration coefficients of 1.011, 1.031 and 0.958, which is consistent with the unpaved roads. Table 13 lists the calibrated mean RIF calculated from the three phones for all the 35 runs of the road test for the same unpaved road segment using the mean RIF calibration method.

Table 13. Towards mean RIF method calibrated mean RIF from the iPhone 8, iPhone X, and google pixel for the unpaved road

Run	iPhone 8 mean RIF (g-force/meter)	iPhone x mean RIF (g-force/meter)	google pixel mean RIF (g-force/meter)
1	0.515	0.510	0.487
2	0.563	0.554	0.465
3	0.500	0.468	0.492
4	0.525	0.531	0.499
5	0.532	0.539	0.486
6	0.502	0.492	0.490
7	0.510	0.509	0.470
8	0.457	0.466	0.478
9	0.480	0.469	0.480
10	0.481	0.497	0.491
11	0.498	0.530	0.492
12	0.504	0.487	0.484
13	0.502	0.503	0.473
14	0.441	0.465	0.470
15	0.446	0.464	0.488
16	0.472	0.484	0.475
17	0.485	0.470	0.499
18	0.465	0.476	0.513
19	0.515	0.497	0.467
20	0.455	0.445	0.458
21	0.449	0.459	0.479
22	0.464	0.460	0.473
23	0.472	0.449	0.503
24	0.497	0.480	0.479
25	0.473	0.461	0.471
26	0.461	0.453	0.475
27	0.466	0.466	0.487
28	0.454	0.467	0.482
29	0.477	0.488	0.473
30	0.449	0.449	0.499
31	0.504	0.507	0.508
32	0.517	0.512	0.500
33	0.497	0.485	0.487
34	0.479	0.505	0.507
35	0.492	0.501	0.517
mean	0.486	0.486	0.486

Table 14 shows the ANOVA test result of mean RIF for all 35 runs on the tested unpaved road segment from all the three phones. The F value is 0 and the critical F value is 3.085. Since the F value is smaller than the critical F value, the hypothesis that three phones have equal mean RIF is true. Thus, the calibration towards mean RIF for paved roads worked well. Figure 22 shows the calibrated mean RIF distribution of the three phones.

Table 14. ANOVA test of towards mean RIF method calibrated mean RIF for all 35 runs on the tested unpaved road segment from all the three phones

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F critical
Between Groups	0.000	2.000	0.000	0.000	1.000	3.085
Within Groups	0.059	102.000	0.001			
Total	0.059	104.000				

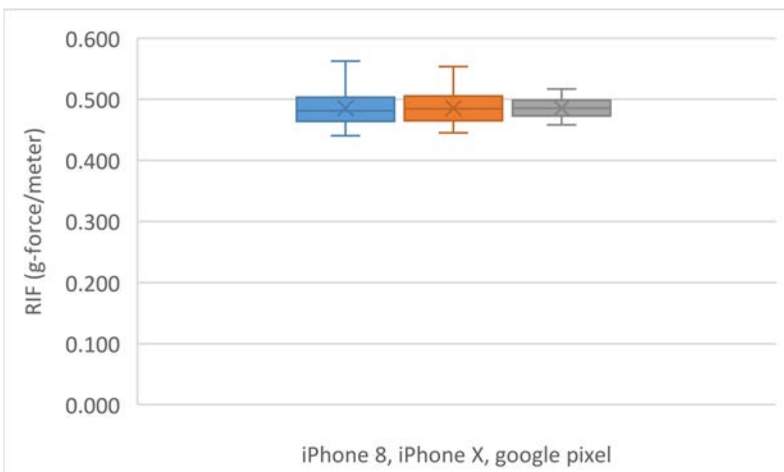


Figure 22. Towards mean RIF method calibrated mean RIF from unpaved roads from iPhone 8, iPhone X, and google pixel

4.4. Calibrating Phones Towards Maximum RIF for Unpaved Roads

The second calibration method for the unpaved roads between different phones is the reference maximum RIF method. For the road test, the mean peak RIF measured from the iPhone 8, iPhone X and google pixel are 1.314, 1.248 and 1.011. The peak RIF from iPhone 8 is used as a reference to calibrate all the three phones. The average peak RIF from the three phones is

1.191. When compared with the paved roads, the three iPhones have an average peak RIF of 1.082. The peak RIF for a railroad crossing on an unpaved road is around 91% of that on a paved road from all three different phones, which are very consistent for either for a paved road or an unpaved road since it is the same type of road bump.

Dividing the peak RIF from each phone using the peak RIF measured from iPhone 8, which is 1.314, the calibration coefficients for each phone are 1, 0.950, and 0.769, respectively. Thus, for iPhone 8, all the measured Gz values were divided by 1 to get calibrated; for the iPhone X, all the measured Gz values were divided by 0.950 for calibration; and for the Google pixel phone, all the measured Gz values were divided by 0.769 for calibration. After the calibration of Gz for each phone, the calibrated RIF for each phone and each run were recalculated. Table 15 shows the calibrated mean RIF for all 35 runs from all the three phones for the unpaved road.

Table 15. Towards maximum RIF method calibrated mean RIF from the iPhone 8, iPhone X, and google pixel for the unpaved road

Run	iPhone 8 mean RIF (g-force/meter)	iPhone x mean RIF (g-force/meter)	google pixel mean RIF (g-force/meter)
1	0.552	0.524	0.602
2	0.604	0.569	0.575
3	0.537	0.481	0.607
4	0.563	0.546	0.616
5	0.571	0.554	0.600
6	0.539	0.506	0.605
7	0.548	0.523	0.581
8	0.490	0.479	0.591
9	0.515	0.482	0.593
10	0.517	0.510	0.606
11	0.534	0.545	0.608
12	0.541	0.500	0.598
13	0.538	0.517	0.584
14	0.473	0.478	0.581
15	0.479	0.477	0.603
16	0.507	0.498	0.587
17	0.521	0.482	0.617
18	0.499	0.489	0.634
19	0.553	0.511	0.577
20	0.489	0.457	0.566
21	0.482	0.472	0.592
22	0.498	0.473	0.584
23	0.507	0.461	0.621
24	0.534	0.493	0.591
25	0.508	0.473	0.582
26	0.495	0.466	0.586
27	0.500	0.479	0.601
28	0.488	0.480	0.595
29	0.512	0.502	0.584
30	0.483	0.461	0.616
31	0.541	0.521	0.627
32	0.555	0.526	0.618
33	0.534	0.498	0.601
34	0.514	0.519	0.626
35	0.528	0.515	0.639
mean	0.521	0.499	0.600

Table 16 shows the ANOVA test result of mean RIF for all 35 runs on the tested paved road segment from all the three phones. The F value is 147.079 and the critical F value is 3.085.

Since the F value is larger than the critical F value and even larger than the F value before calibration which is 58.427, the hypothesis that three phones have equal mean RIF is not true. Thus, the calibration towards maximum RIF for paved roads die not work. Figure 23 show the calibrated mean RIF distribution for the three phones, and we can see after calibration they are still spread out.

Table 16. ANOVA test of towards maximum RIF method calibrated mean RIF for all 35 runs on the tested unpaved road segment from all the three phones

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F critical
Between Groups	0.196	2.000	0.098	147.079	0.000	3.085
Within Groups	0.068	102.000	0.001			
Total	0.264	104.000				

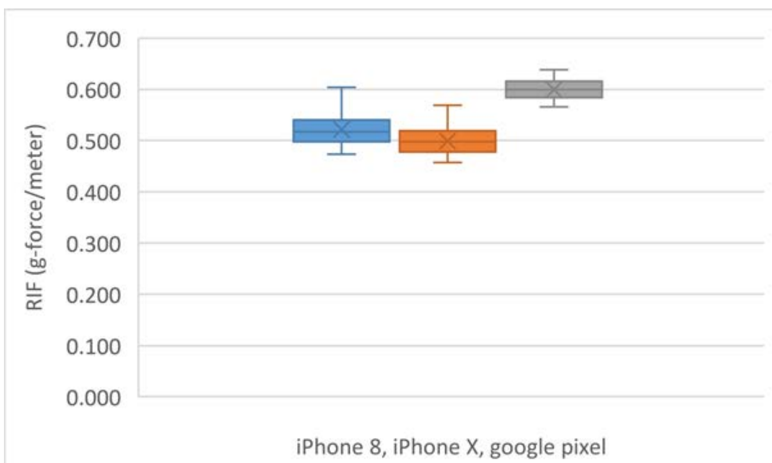


Figure 23. Towards maximum RIF method calibrated mean RIF from paved roads from iPhone 8, iPhone x, and google pixel

4.5. Calibrating Phones for Unpaved Road Using the Calibration Coefficient from Paved Road

The third calibration method for the unpaved roads between different phones is using the calibration coefficient from paved road, which is based on mean RIF calibration of paved road, which is 1.011, 1.031, and 0.958 for the three phones respectively. Specifically, for the iPhone 8, the measured Gz value from all the 35 runs of the road test divide the calibration coefficient of 1.011 to get calibrated. For the iPhone X, the measured Gz value divide the calibration coefficient of 1.031; and for the Google pixel, the measured Gz value divide 0.958 to get calibrated. After calibrating the Gz based on the sensor sensitivity difference using calibration coefficients for each phone, the RIF for each phone and each run is recalculated as calibrated RIF. Table 17 lists the calibrated mean RIF calculated from the three phones for all the 35 runs of the road test for the same unpaved road segment using this calibration method.

Table 17. Calibration coefficient from paved road method calibrated mean RIF from the iPhone 8, iPhone X, and google pixel for the unpaved road

Run	iPhone 8 mean RIF (g-force/meter)	iPhone X mean RIF (g-force/meter)	google pixel mean RIF (g-force/meter)
1	0.547	0.483	0.483
2	0.598	0.524	0.461
3	0.531	0.443	0.488
4	0.557	0.503	0.494
5	0.565	0.510	0.482
6	0.534	0.466	0.486
7	0.542	0.482	0.466
8	0.485	0.441	0.474
9	0.510	0.444	0.476
10	0.511	0.470	0.486
11	0.528	0.502	0.488
12	0.535	0.461	0.480
13	0.533	0.477	0.469
14	0.468	0.440	0.466
15	0.474	0.439	0.484
16	0.501	0.459	0.471
17	0.515	0.445	0.495
18	0.493	0.451	0.509
19	0.547	0.471	0.463
20	0.483	0.421	0.455
21	0.477	0.434	0.475
22	0.493	0.436	0.469
23	0.502	0.425	0.499
24	0.528	0.454	0.475
25	0.502	0.436	0.467
26	0.490	0.429	0.471
27	0.495	0.441	0.483
28	0.483	0.443	0.478
29	0.506	0.462	0.469
30	0.477	0.425	0.495
31	0.536	0.480	0.504
32	0.549	0.484	0.496
33	0.528	0.459	0.483
34	0.509	0.478	0.502
35	0.522	0.474	0.513
mean	0.516	0.460	0.482

Table 18 shows the ANOVA test result of mean RIF for all 35 runs on the tested unpaved road segment from all the three phones. The F value is 47.936 and the critical F value is 3.085. Since F value is larger than critic F value, the hypothesis that three phones have equal mean RIF is not true. Comparing to the original data which F value 58.427, the calibration towards the maximum RIF which F value is 147.097, there is some improvement. Figure 24 shows the calibrated mean RIF distribution for the three phones.

Table 18. ANOVA test of calibration coefficient from paved road method calibrated mean RIF for all 35 runs on the tested unpaved road segment from all the three phones

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F critical
Between Groups	0.056	2.000	0.028	47.936	0.000	3.085
Within Groups	0.059	102.000	0.001			
Total	0.115	104.000				

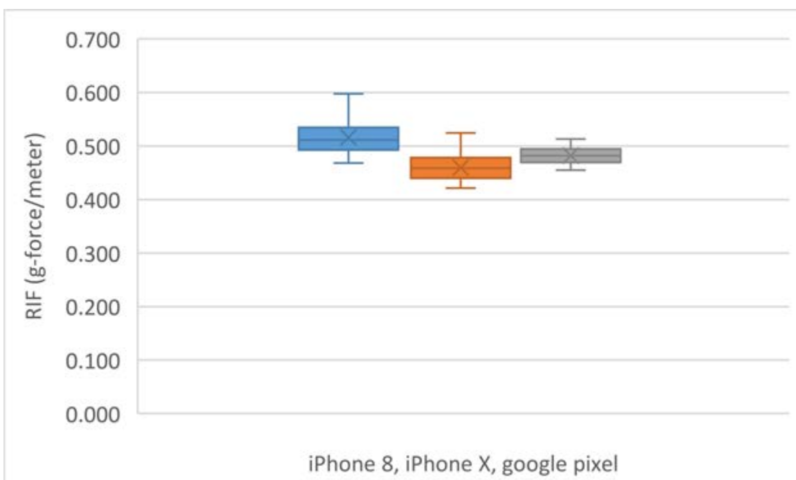


Figure 24. Calibration coefficient method calibrated mean RIF from unpaved roads from iPhone 8, iPhone X, and google pixel

4.6. GPS Output to Locate the Reference Peak RIF Location on Unpaved Road

According to the measured peak RIF, the GPS output location, Figure 25 shows the located railroad crossing locations of the unpaved road segment on google map. Although the location of the peak RIF was expected to be at the railroad crossing, time delay also causes the measurement location deviation of the locations, which may result in measurement error. However, since the RIF is an average road roughness index for every ΔL, which is 10 m in this study, the deviation of the GPS locations is still within the acceptable range of the measurement of 10m resolution.

(a) (b) (c)

Figure 25. Unpaved road peak RIF location from the (a) iPhone 8, (b) iPhone X, and (c) google pixel phone

4.7. Comparison of The Three Calibration Methods

Before the calibration, the F value among the three phones on the unpaved road segment is 58.427. Calibrating toward the mean RIF, the first calibration method, results in F value to be 0. Calibrating toward the peak RIF, the second calibration method, results in F value to be 147.079. Calibrating using the paved road coefficient, the third calibration method, results in F value to be 47.936.

Thus, for unpaved roads, calibrating based on the mean RIF shows a smallest F value of 0 when compared to calibrating based on the peak RIF of 147.079 and calibrating based on the paved road calibration coefficient, indicating a best way to calibrate different phones on unpaved roads.

4.8. Summary

In this chapter, the three proposed phone calibration methods, calibrating towards the mean RIF (method 1), the peak RIF (method 2) and use paved road calibration coefficient (method 3), were tested on an unpaved road segment with a railroad crossing as a known reference bump for the peak RIF. The original data from the three phones showed that different phones deliver a significant measurement difference without a calibration. The calibration for different phones before practical applications is a must to get a consistent measurement from different phones for the same unpaved road segment. The analysis in this chapter indicated that for the unpaved roads, calibrating towards the mean RIF leads to a smallest F value when do ANOVA test among three phones. The peak RIF will not differ much due to the same type of defect. Thus, for unpaved road, calibrating different phones towards mean RIF is recommended.

5. SENSITIVITY ANALYSIS ON DIFFERENT RIF WINDOW SIZE

For an accurate smart phone calibration, window size of RIF selection may be a critical factor to be considered. In this chapter, the effect of RIF window size is investigated different RIF window sizes including 1m, 5m, 10m, 15m, 20m, and 40m.

5.1. Window Size 1 Meter

When the window size is set to 1 meter, for paved road, before calibration, the average mean RIF of iPhone 8, iPhone X and google pixel for of the 35 runs are 0.169, 0.173 and 0.166, respectively. The standard deviation among them is 0.004, and the F value of the ANOVA test is 10.991. The MOE95 fall below 2% after 24 runs for iPhone 8, after 13 runs for iPhone X, and after 10 runs for google pixel. Thus, a valid road roughness road test for paved roads needs at least more than 24 runs for iPhone 8, 13 runs for iPhone X, and 10 runs for google pixel to get reliable results for each phone. Since the calibration method towards the mean RIF has been validated to be a better calibration method from the previous chapters, only this method is used to perform sensitivity analysis on different window sizes on calibration. For the tested paved road, after calibration towards the mean RIF, the average of calibrated mean RIF for iPhone 8, iPhone X and google pixel are 0.169, 0.169 and 0.169, respectively, the standard deviation among them is 0, and the F value of ANOVA test is 0.

For the tested unpaved road, when the window size is set to be 1m, before calibration, the average mean RIF of iPhone 8, iPhone X and google pixel for the 35 runs are 0.490, 0.444 and 0.441, respectively. The standard deviation among them is 0.027, and the F value of ANOVA test is 47.216. The MOE95 fall below 2% after 35 runs for iPhone 8, after 34 runs for iPhone X, and after 5 runs for google pixel. Thus, an effective road test for unpaved roads will need more than 35 runs for iPhone 8, 34 runs for iPhone X, 5 runs for google pixel to get reliable results for each phone. After calibration towards the mean RIF, the average of calibrated mean RIF for iPhone 8, iPhone X and google pixel are 0.459, 0.459 and 0.459, respectively, the standard deviation among them is 0, and the F value of ANOVA test is 0.

5.2. Window Size 5 Meters

When the window size is set to 5 meters, for paved road, before calibration, the average mean RIF of iPhone 8, iPhone X and google pixel for the 35 runs are 0.186, 0.190 and 0.178, respectively. The standard deviation among them is 0.006, and the F value of ANOVA test is 22.115. For iPhone 8, the MOE95 fall below 2% after 24 runs, for iPhone X, after 14 runs, and for google pixel, after 13 runs. Thus, with a window size of 5m, an effective road test for paved roads will need more than 24 runs for iPhone 8, 14 runs for iPhone X, 13 runs for google pixel. After calibration towards the mean RIF, the average of calibrated mean RIF for iPhone 8, iPhone X and google pixel are 0.184, 0.184 and 0.184, respectively, the standard deviation among them is 0, and the F value of ANOVA test is 0.

For unpaved road, when the window size is set to 5m, before calibration, the average mean RIF of iPhone 8, iPhone X and google pixel for the 35 runs are 0.515, 0.468 and 0.456, respectively. The standard deviation among them is 0.031, and the F value of ANOVA test is 56.661. For iPhone 8, the MOE95 fall below 2% after 34 runs, for iPhone X, after 33 runs, and for google pixel, after 6 runs. Thus, an unpaved road will need to test more than 34 runs for iPhone 8, 33 runs for iPhone X, 6 runs for google pixel for reliable results. After calibration towards the mean RIF, the average of calibrated mean RIF of the 35 runs for iPhone 8, iPhone X and google pixel are 0.480, 0.480 and 0.480, respectively, the standard deviation among them is 0, and the F value of ANOVA test is 0.

5.3. Window Size 10 Meters

When the window size is set to 10 meters, for paved road, before calibration, the average mean RIF of iPhone 8, iPhone X and google pixel for the 35 runs are 0.195, 0.199 and 0.185, respectively. The standard deviation among them is 0.007, and the F value of ANOVA test is 29.196. For iPhone 8, the MOE95 fall below 2% after 23 runs, for iPhone X, after 18 runs, and for google pixel, after 19 runs. Thus, more than 23 runs are needed for iPhone 8, 18 runs for iPhone X, 19 runs for google pixel for reliable results. After calibration towards the mean RIF, the average of calibrated mean RIF for iPhone 8, iPhone X and google pixel of the 35 runs are 0.193, 0.193 and 0.193, respectively, the standard deviation among them is 0, and the F value of ANOVA test is 0.

For unpaved road, with a window size of 15m, before calibration, the average mean RIF of iPhone 8, iPhone X and google pixel for the 35 runs are 0.521, 0.474 and 0.461, respectively. The standard deviation among them is 0.032, and the F value of ANOVA test is 58.427. For iPhone 8, the MOE95 fall below 2% after 34 runs, for iPhone X, after 33 runs, and for google pixel, after 6 runs. Thus, more than 34 runs are needed for iPhone 8, 33 runs for iPhone X, 6 runs for google pixel to get reliable results for each phone. After calibration towards the mean RIF, the average of calibrated mean RIF for iPhone 8, iPhone X and google pixel of the 35 runs are 0.486, 0.486 and 0.486, respectively, the standard deviation among them is 0, and the F value of ANOVA test is 0.

5.4. Window Size 15 Meters

When the window size is set to 15 meters, for paved road, before calibration, the average mean RIF of iPhone 8, iPhone X and google pixel for the 35 runs are 0.202, 0.207 and 0.190, respectively. The standard deviation among them is 0.009, and the F value of ANOVA test is

43.888. For iPhone 8, the MOE95 fall below 2% after 24 runs, for iPhone X, after 13 runs, and for google pixel, after 10 runs. Thus, more than 24 runs are needed for iPhone 8, 13 runs for iPhone X, 10 runs for google pixel for reliable results. After calibration towards the mean RIF, the average of calibrated mean RIF for iPhone 8, iPhone X and google pixel of the 35 runs are 0.200, 0.200 and 0.200, respectively, the standard deviation among them is 0, and the F value of ANOVA test is 0.

For unpaved road, with a window size of 15m, before calibration, the average mean RIF of iPhone 8, iPhone X and google pixel for the 35 runs are 0.527, 0.481 and 0.464, respectively. The standard deviation among them is 0.033, and the F value of ANOVA test is 59.684. For iPhone 8, the MOE95 fall below 2% after 34 runs, for iPhone X, after 33 runs, and for google pixel, after 8 runs. Thus, more than 34 runs are needed for iPhone 8, 33 runs for iPhone X, 8 runs for google pixel to get reliable results for each phone. After calibration towards the mean RIF, the average of calibrated mean RIF for iPhone 8, iPhone X and google pixel of the 35 runs are 0.491, 0.491 and 0.491, respectively, the standard deviation among them is 0, and the F value of ANOVA test is 0.

5.5. Window Size 20 Meters

When the window size is set to 20 meters, for paved road, before calibration, the average mean RIF of iPhone 8, iPhone X and google pixel of 35 runs are 0.210, 0.212 and 0.196, respectively. The standard deviation among them is 0.009, and the F value of ANOVA test is 32.168. For iPhone 8, the MOE95 fall below 2% after 26 runs, for iPhone X, after 25 runs, and for google pixel, after 21 runs. Thus, more than 26 runs are needed for iPhone 8, 25 runs for iPhone X, 21 runs for google pixel for reliable results. After calibration towards the mean RIF, the average of calibrated mean RIF for iPhone 8, iPhone X and google pixel are 0.206, 0.206 and 0.206, respectively, the standard deviation among them is 0, and the F value of ANOVA test is 0.

For unpaved road, before calibration, the average mean RIF of iPhone 8, iPhone X and google pixel of 35 runs are 0.530, 0.482 and 0.466, respectively. The standard deviation among them is 0.033, and the F value of ANOVA test is 63.591. For iPhone 8, the MOE95 fall below 2% after 34 runs, for iPhone X, after 32 runs, and for google pixel, after 6 runs. So, more than 34 runs are needed for iPhone 8, 32 runs for iPhone X, and 6 runs for google pixel to get reliable results for each phone. After calibration towards the mean RIF, the average of calibrated mean RIF for iPhone 8, iPhone X and google pixel are 0.493, 0.493 and 0.493, respectively, the standard deviation among them is 0, and the F value of ANOVA test is 0.

5.6. Window Size 40 Meters

When the window size is set to 40 meters, for paved road, before calibration, the average mean RIF of iPhone 8, iPhone X and google pixel of 35 runs are 0.229, 0.234 and 0.211, respectively. The standard deviation among them is 0.012, and the F value of ANOVA test is 60.951. For iPhone 8, the MOE95 fall below 2% after 23 runs, for iPhone X, after 21 runs, and for google pixel, after 11 runs. So, more than 23 runs are needed for iPhone 8, 21 runs for iPhone X, and 11 runs for google pixel for reliable results. After calibration towards the mean RIF, the average of calibrated mean RIF for iPhone 8, iPhone X and google pixel are 0.224, 0.224 and 0.224, respectively, the standard deviation among them is 0, and the F value of ANOVA test is 0.

For unpaved road, with a window size of 40m, before calibration, the average mean RIF of iPhone 8, iPhone X and google pixel of 35 runs are 0.529, 0.482 and 0.464, respectively. The standard deviation among them is 0.034, and the F value of ANOVA test is 61.279. For iPhone 8, the MOE95 fall below 2% after 35 runs, for iPhone X, after 32 runs, and for google pixel, after

8 runs. Thus, more than 35 runs are needed for iPhone 8, 32 runs for iPhone X, and 8 runs for google pixel for reliable results. After calibration towards the mean RIF, the average of calibrated mean RIF for iPhone 8, iPhone X and google pixel are 0.492, 0.492 and 0.492, respectively, the standard deviation among them is 0, and the F value of ANOVA test is 0.

5.7. Summary

Table 19 and 20 show the summary of the mean RIF, standard deviation and F value among three phones in different window size before calibration on paved road and unpaved road. From Table 19 and Table 20, can be seen that without calibration, different window size lead to different mean RIF. The smaller the window size is, the smaller standard deviation and F value among three phones for a more consistent result.

Table 19. Summary of before calibrated mean RIF, standard deviation, and F value among three phones in different window size on paved road

Window size	Three phone Mean RIF (g-force/meter)	STD	F value
1 meter	0.169	0.004	10.991
5 meters	0.184	0.006	22.115
10 meters	0.193	0.007	29.196
15 meters	0.200	0.009	43.888
20 meters	0.206	0.009	32.168
40 meters	0.224	0.012	60.951

Table 20. Summary of before calibrated mean RIF, standard deviation, and F value among three phones in different window size on unpaved road

Window size	Three phone Mean RIF (g-force/meter)	STD	F value
1 meter	0.459	0.027	47.216
5 meters	0.480	0.031	56.661
10 meters	0.486	0.032	58.427
15 meters	0.491	0.033	59.684
20 meters	0.493	0.033	63.591
40 meters	0.492	0.034	61.279

Table 21 and 22 show the summary of the mean RIF, standard deviation and F value among three phones in different window size after calibration towards the mean RIF on paved road and unpaved road. It can be seen that after calibration, the window size has little influence on the RIF measurements, and will not be a controlling parameter for an effective road roughness measurement.

Table 21. Summary of towards mean RIF method calibrated mean RIF, standard deviation, and F value among three phones in different window size on paved road

Window size	Three phone Mean RIF (g-force/meter)	STD	F value
1 meter	0.169	0	0
5 meters	0.184	0	0
10 meters	0.193	0	0
15 meters	0.200	0	0
20 meters	0.206	0	0
40 meters	0.224	0	0

Table 22. Summary of towards mean RIF method calibrated mean RIF, standard deviation, and F value among three phones in different window size on unpaved road

Window size	Three phone Mean RIF (g-force/meter)	STD	F value
1 meter	0.459	0	0
5 meters	0.480	0	0
10 meters	0.486	0	0
15 meters	0.491	0	0
20 meters	0.493	0	0
40 meters	0.492	0	0

Table 23 shows the summarizes of segment length, average vehicle speed, and average sample rate for each phone for paved road. Table 24 shows the summarizes of segment length, average vehicle speed, and average sample rate for each phone for unpaved road.

Table 23. Summarizes of segment length, average vehicle speed, and average sample rate for each phone for paved road

	segment length (m)	average speed (m/s)	sample rate (Hz)
iPhone 8	399.821	11.772	86.924
iPhone X	399.819	11.751	87.894
google pixel	399.965	11.753	386.488

Table 24. Summarizes of segment length, average vehicle speed, and average sample rate for each phone for unpaved road

	segment length (m)	average speed (m/s)	sample rate (Hz)
iPhone 8	299.823	11.548	79.088
iPhone X	299.810	11.543	80.364
google pixel	299.968	11.559	385.548

Table 25 shows before calibration the number of runs after which the MOE95 falls below 2% for each window size, and each phone for paved road. Table 26 shows after calibration the number of runs after which the MOE95 falls below 2% for each window size, and each phone for paved road. Table 27 shows before calibration the number of runs after which the MOE95 falls below 2% for each window size, and each phone for unpaved road. Table 28 shows after calibration the number of runs after which the MOE95 falls below 2% for each window size, and each phone for unpaved road. We can see calibration does not affect the MOE95 at all.

Table 25. Before calibration the number of runs after which the MOE95 falls below 2% for each window size, and each phone for paved road

window size	iPhone 8	iPhone X	google pixel
1m	24	13	10
5m	24	14	13
10m	23	18	19
15m	24	13	10
20m	26	25	21
40m	23	21	11

Table 26. After calibration the number of runs after which the MOE95 falls below 2% for each window size, and each phone for paved road

window size	iPhone 8	iPhone X	google pixel
1m	24	13	10
5m	24	14	13
10m	23	18	19
15m	24	13	10
20m	26	25	21
40m	23	21	11

Table 27. Before calibration the number of runs after which the MOE95 falls below 2% for each window size, and each phone for unpaved road

window size	iPhone 8	iPhone X	google pixel
1m	35	34	5
5m	34	33	6
10m	34	33	6
15m	34	33	8
20m	34	32	6
40m	35	32	8

Table 28. After calibration the number of runs after which the MOE95 falls below 2% for each window size, and each phone for unpaved road

window size	iPhone 8	iPhone X	google pixel
1m	35	34	5
5m	34	33	6
10m	34	33	6
15m	34	33	8
20m	34	32	6
40m	35	32	8

Table 29 shows the calibration coefficients use the towards mean RIF method for each phone and each window size for paved road. Table 30 shows the calibration coefficients use the towards mean RIF method for each phone and each window size for unpaved road.

Table 29. Calibration coefficients use the towards mean RIF method for each phone and each window size for paved road

window size	iPhone 8	iPhone X	google pixel
1m	1.000	1.023	0.977
5m	1.007	1.029	0.964
10m	1.011	1.031	0.958
15m	1.012	1.038	0.950
20m	1.018	1.032	0.950
40m	1.020	1.042	0.938

Table 30. Calibration coefficients use the towards mean RIF method for each phone and each window size for unpaved road

window size	iPhone 8	iPhone X	google pixel
1m	1.069	0.968	0.962
5m	1.073	0.976	0.951
10m	1.074	0.976	0.950
15m	1.074	0.980	0.946
20m	1.076	0.979	0.945
40m	1.076	0.980	0.944

6. CONCLUSIONS AND FUTURE WORK

6.1. Conclusions

Use smartphone as road condition monitor sensor can monitor road condition on either paved road or unpaved road. Comparing to the traditional road monitoring system which needs certain equipment and trained staff to do the monitoring, using smartphone as sensor and output RIF values to evaluate the road condition is low cost. The following conclusions can be drawn based on the investigations in this study:

1. This study showed that different phones have different accelerometer sensitivities resulting in inconsistency in measurements, thus, calibrating different phones is required before practical applications of the smartphone-based road condition monitoring.
2. There are two ways to calibrate different smartphone on paved road, including calibrating towards mean RIF or peak RIF as reference phone and one more way to calibrate different smartphone on unpaved road, which is using the paved road calibration coefficient to calibrate the unpaved road.
3. For both paved and unpaved road surfaces, calibration method based on the mean RIF yields the best results.
4. For the RIF calculation window size, without calibration, it will affect the consistency of different phones, which the smaller the window size, the more consistency different phones are. After calibration, the window size has little effects on the consistency among different phones.

Therefore, calibration for different phone is highly recommended for smartphone applications of road condition monitoring and for different road surface conditions.

6.2. Future Work

In the future, based on the calibration methods proposed in this study, more road test and more types of phones are required to validate the conclusions. In addition, the corrections on the GPS locations and time delay is also needed to improve the measurement accuracy of the smart phone-based road condition measurement methods.

REFERENCES

Administration, F. H. (2008, October 29). *Warm Mix Asphalt Technologies and Research*. Retrieved May 18, 2018, from https://www.fhwa.dot.gov/pavement/asphalt/wma.cfm

Administration, F. H. (2017, June 27). *Concrete Pavement Rehabilitation - Guide for Diamond Grinding.* Retrieved May 20, 2018, from https://www.fhwa.dot.gov/pavement/pub_details.cfm?id=168

AMES. (n.d.). *8300 Portable High Speed Profiler*. (AMES) Retrieved May 20, 2018, from https://amesengineering.com/products/8300-portable-high-speed-profiler/

Andren, P. (2006). Power Spectral Density Approximations of Longitudinal Road Profiles. *International Journal of Vehicle Design*, 2-14.

Base, T. L. (n.d.). *Sprayed Seal*. Retrieved May 18, 2018, from http://www.lgam.info/sprayed-seal

Bland, J. M. (1996). Measurement Error. *Bmj, 312*, 1654.

Board, H. R. (n.d.). *Special Report No. 113: Standard Nomenclature and Definitions for Pavement Components and Deficiencies.* Washington: National Academy of Sciences.

Bridgelall, R. (2014). Connected Vehicle Approach for Pavement Roughness Evaluation. *Journal of Infrastructure Systems*.

Consultants, A. (2018). *Auto Rod and Level*. (APR Consultants) Retrieved May 18, 2018, from http://www.aprconsultants.com/auto-rod-and-level-2/

Cookaa. (2003, February 19). *Dirt Road auf Kangaroo Island*. Retrieved May 12, 2018, from https://commons.wikimedia.org/wiki/File:Dirt_Road_SA.JPG

Delanne, Y. a. (2001). *Advantages and limits of.* Washington: Transportation Research Board.

FHWA. (2016, June). *Measuring and Specifying Pavement Smoothness.* Retrieved May 15, 2018, from https://www.fhwa.dot.gov/pavement/pubs/hif16032.pdf

Gerbrandt, R., Makahoniuk, T., Borbely, C. L., & Berthelot, C. (2000). Effect of Cold In-Place Recycling on The. *6th International Conference on Heavy Vehicle Weights.* Saskatoon.

Gransberg, D. D. (2005). Chip Seal Best Practices. *Transportation Research Board*, 13-20.

Group, I. (2003). *Profiler Systems*. (Inopave Group) Retrieved May 16, 2018, from http://www.innopave.com/en/pavement_profiler_profilograph.asp

Instruments, S. S. (n.d.). *CS8550 Bridge Profilograph Specifically designed for surface ride quality testing of bridge structures*. (Surface Systems & Instruments) Retrieved May 10, 2018, from https://www.smoothroad.com/cs8550-bridge-profilograph

Lane, M. (2005). An Introduction to Wavelet Analysis with SAS. *SAS Conference Proceedings NESUG 2005*, 1-10.

Mucka, P. (2012). Relationship Between International Roughness Index and Straightedge Index. *Journal of Transportation Engineering ASCE*, 1099-1112.

Mucka, P. (2016). Current Approaches to Quantify the Longitudinal Road Roughness. *International Journal of Pavement Engineering, 17*, 659-679.

NBM&CW. (2016, September). *Little Known Facts About Concrete Roads*. (NBM&CW) Retrieved May 15, 2018, from https://www.nbmcw.com/tech-articles/roads-and-pavements/35210-little-known-facts-about-concrete-roads.html

NCAT. (n.d.). *Automatic Road Analyzer (ARAN) Inertial Profiler*. (NCAT) Retrieved May 07, 2018, from http://www.eng.auburn.edu/research/centers/ncat/about/facilities/automatic-road-analyzer.html

Pavement, A. C. (2004, December 23). *Concrete Pavement Progress.* Retrieved May 18, 2018, from https://web.archive.org/web/20120417114823/http://www.pavement.com/CPP/2004/December04.pdf

Purplewave. (2018, March 20). *John Deere Gator TS UTV with Ames Engineering inertial profiler*. (purplewave) Retrieved May 03, 2018, from https://www.purplewave.com/auction/180320/item/EA9691

R. Bridgelall, Y. H. (2016). Precision Enhancement of Pavement Roughness Localization with Connected Vehicles. *Measurement Science and Technology, 27*, 25012.

Roads&Bridges. (2017, September 20). *ROADS: Research suggests asphalt roads cannot handle 21st century climate*. Retrieved May 13, 2018, from https://www.roadsbridges.com/roads-research-suggests-asphalt-roads-cannot-handle-21st-century-climate

Sayers, M. G. (1986). *Guidelines for the Conduct and Calibration of Road Roughness Measurements.* Washington DC: The World Bank.

Sayers, M. G. (1986). *The International Road Roughness Experiment: Establishing Correlation and a Calibration Standard for Measurements.* Washington DC: The World Bank.

SkidSteer. (2015, August 26). *One Gravel Road, One Machine and One Box Grader Attachment* . Retrieved May 12, 2018, from https://www.skidsteersolutions.com/blog/one-gravel-road-one-machine-and-one-box-grader-attachment/

Song, I. a. (2006). Airport Pavement Roughness Index Relationships Using the Federal Aviation Administration(Faa) Profiling System. *2006 airfield and highway pavement specialty conference.* Atlanta.

SSI. (n.d.). *CS8700 Lightweight Profiler Lightweight profiling system is made for quick access to fresh pavements*. (SSI) Retrieved May 05, 2018, from https://www.smoothroad.com/cs8700-lightweight-profiler

Standardization, I. O. (1995). *Mechanical Vibration-Road Surface Profiles-Reporting of Measured Data.*

Statistics, B. o. (2017, May 20). *Table 1 - Commercial Freight Activity in the United States by Mode of Transportation: 2002*. (United States Department of Transportation) Retrieved May 1, 2018, from https://www.bts.dot.gov/archive/publications/freight_in_america/table_01

Transportation, U. S. (2017, May 21). *Table 1-40: U.S. Passenger-Miles (Millions)*. (Bureau of Transportation Statistics) Retrieved may 1, 2018, from https://www.bts.gov/archive/publications/national_transportation_statistics/table_01_40

Wire, B. (2018, March 12). *Bosch's 6-Axis IMU in the Apple iPhone X Complete Teardown Report 2018 - ResearchAndMarkets.com*. (Business Wire) Retrieved June 23, 2018, from https://www.businesswire.com/news/home/20180312005463/en/Boschs-6-Axis-IMU-Apple-iPhone-Complete-Teardown

WordWeb. (n.d.). *Metalled Road*. Retrieved May 20, 2018, from https://www.wordwebonline.com/en/METALLEDROAD