Pre-Print Manuscript of Article:

Bridgelall, R., Hough, J., Tolliver D., "Characterizing Pavement Roughness at Non-Uniform Speeds Using Connected Vehicles," *International Journal of Pavement Engineering*, DOI: 10.1080/10298436.2017.1366768, pp. 1-7, August 28, 2017.

Characterizing pavement roughness at non-uniform speeds using connected vehicles

Methods of pavement roughness characterizations using connected vehicles are poised to scale beyond the frequency, span, and affordability of existing methods that require specially instrumented vehicles and skilled technicians. However, speed variability and differences in suspension behavior require segmentation of the connected vehicle data to achieve some level of desired precision and accuracy with relatively few measurements. This study evaluates the reliability of a Road Impact Factor (RIF) transform under stop-and-go conditions. A RIF-transform converts inertial signals from on-board accelerometers and speed sensors to roughness indices (RIF-indices), in real-time. The case studies collected data from 18 different buses during their normal operation in a small urban city. Within 30 measurements, the RIF-indices distributed normally with an average margin-of-error below 6%. This result indicates that a large number of measurements will provide a reliable estimate of the average roughness experienced. Statistical t-tests distinguished the relatively small differences in average roughness levels among the roadway segments evaluated. In conclusion, when averaging roughness measurements from the same type of vehicle moving at non-uniform speeds, the RIF-transform will provide ever-increasing precision and accuracy as the traversal volume increases.

Keywords: connected vehicle; inertial profiler; intelligent transportation systems; pavement management; probe vehicles; ride quality; roughness index; suspension systems

1 Introduction

State highway agencies measure pavement roughness to monitor the condition of the network and for other important purposes such as to assess the quality of construction and to forecast maintenance needs. The federal Highway Performance Monitoring System (HPMS) requires that states report the International Roughness Index (IRI) annually for the national highway system and on a 2-year maximum cycle for all other required sections (HPMS 2016). Apart from cycles in precipitation and temperature, heavy vehicles accelerate road deterioration (Bilodeau *et al*. 2015). Therefore, anomalies that become safety hazards are likely to appear between such long monitoring cycles.

The current practice is to measure the road elevation profile at a constant speed using specially instrumented vehicles called inertial profilers. A mathematical procedure subsequently calculates the IRI from the elevation profile samples. The calculations produce the mechanical response of a specific quarter-car traveling at precisely 80 km h^{-1} (Gillespie *et al.* 1986). The constant speed requirement makes ride quality measurements impractical for local roads and urban arterials where vehicles travel at non-uniform speeds. The National Cooperative Highway Research Program (NCHRP) recognized this shortcoming of the existing IRI procedure and commissioned a 5-year effort in 2013 "to identify/develop a means for measuring, characterizing, and reporting pavement roughness on low-speed and urban roads." This effort is part of the NCHRP 10-93 research project (Karamihas 2017).

Meanwhile, researchers have been developing methods to estimate the IRI from inertial, speed, and position data collected from sensors aboard regular vehicles (Islam, et al. 2014) (Nomura and Shiraishi 2015). In lieu of connected vehicles, the accelerometers, gyroscopes, and global positioning system (GPS) receivers embedded in smartphones serve as proxy sensors for collecting and analyzing the required data (Cruz and Castro 2015). The connected vehicle (CV) methods are likely to scale more cost-effectively than existing methods if the industry ratifies standards to make such data available for widespread use.

The authors previously reported a CV method called the Road Impact Factor (RIF) transform that is capable of measuring roughness at any speed (Bridgelall 2014a). That research demonstrated that, given any regular vehicle, the RIF-transform is directly proportional to the IRI when applied under the same constraint of a uniform speed. Practitioners know that roughness characterizations at a precise speed mask inertial responses to wavelength excitations that are outside of the sensitivity range of the fixed IRI model (Papagiannakis 1997). Therefore, the IRI cannot represent roughness that vehicles actually experience when traveling at speeds that are different from 80 km h^{-1}. Conversely, the RIF-transform can characterize roughness experienced at any speed because it operates on the inertial responses measured directly from actual vehicles. In other words, the CV method extends the wavelength sensitivity of the RIF-transform beyond that of the IRI, as recently demonstrated (Bridgelall *et al*. 2017).

The RIF-transform produces a direct measure of roughness called the RIF-index by averaging the RIF-indices derived from many CV traversals across a specified roadway segment. Speed variability and differences in suspension behavior require a large number of measurements to achieve some desired precision and accuracy (Bridgelall 2015). Hence, best practices of the method segment the data stream by vehicle type, for example, mid-sized sedans, and then by speed bands, for example 40 km h^{-1} ± 5 km h^{-1}. Previous studies show that such data segmentation enhances the precision of measurements within a margin-of-error that diminishes below 1.5% within 50 traversals (Bridgelall *et al.* 2016a). Transportation agencies recognize that in addition to improving the precision of roughness measurements, CV-based methods will enable continuous monitoring that could detect cyclical anomalies (Dennis and Spulber 2016).

The goal of this research is to evaluate the reliability of the RIF-transform under conditions of speed, suspension, and wheel-path variability. Subsequently, the objectives of this research are to determine the degree to which repeated measurements of RIF-indices from different vehicles of the same type, traveling at non-uniform speeds:

1) produce an unbiased estimate of the average roughness that riders experience
2) could distinguish among the roughness levels of different roadway segments.

The next section reviews the RIF-transform and its speed-dependent properties. The sub-sections describe data collection from public transit buses during their normal operations, and the statistical tests selected to achieve the above objectives. A description of the data analysis from the case studies then follows. The results and discussion section reveal that RIF-indices measured from the buses under stop-and-go conditions provide a reliable estimate of the mean roughness experienced, within a diminishing margin-of-error. The findings also indicate that the power to distinguish differences in segment roughness increases as the measurement volume of RIF-indices increase.

2 Methods

Given the detailed derivations of the RIF-transform referenced in the previous work (Bridgelall 2014a), this section provides only a brief review.

2.1 The RIF-Transform

The RIF-transform produces a RIF-index in units of g-force per meter (g/m) to characterize the actual roughness experienced in the sensing vehicle at any speed. For individual vehicle traversals, the RIF-transform summarizes roughness such that

$$R_{\bar{v}}^{L} = \sqrt{\frac{1}{L}\sum_{n=0}^{N-1}\left|g_{z[n]}v_n\right|^2 \delta t} \tag{1}$$

where the RIF-index $R_{\bar{v}}^{L}$ is the average g-force magnitude experienced per unit of distance

L traveled (Bridgelall *et al.* 2016b). An on-board accelerometer produces the vertical acceleration $g_{z[n]}$ for signal sample n and a speed sensor produces the instantaneous speed v_n. Previous research established that the sample rate, which is the inverse of the average sample period δt, should be at least 64 Hertz (Bridgelall 2014b).

The RIF-transform is stable at non-uniform speeds because it integrates the speed. That is, the RIF-index is zero when the speed is zero and compresses as the square root with increasing speed. For additional insights, setting the speed and the segment length in equation (1) to constants yield

$$R_{\bar{v}}^{L} = \sqrt{\frac{v_n^2}{L} \sum_{n=0}^{N-1} \left| g_{z[n]} \right|^2 \delta t} = \sqrt{v} \sqrt{\frac{1}{T} \sum_{n=0}^{N-1} \left| g_{z[n]} \right|^2 \delta t} = \mathrm{RMS}(g_z) \sqrt{v} \tag{2}$$

Although not used in this manner, this manipulation provides the insight that, for any fixed speed v, the RIF-index, which is the output of the RIF-transform, is directly proportional to the root-mean-square (RMS) of the vertical acceleration, scaled by the square root of a fixed speed. This result explains why using the RMS alone to characterize roughness produces results that do not account for speed variability (Wermers 1962). More recently, researchers attempted to account for the speed variations by applying heuristic calibration factors or machine-learning methods (Dawkins *et al*. 2011) (Du *et al*. 2014) (Stribling 2016) (Alessandroni *et al*. 2017).

2.2 Data collection

The ride quality from road roughness varies with speed and with changes in the behavior of the vehicle’s overall suspension system. The latter could include suspension adaptations to loading conditions and to changes in tire pressure. Hence, practicing the CV method requires data stratification by vehicle class to improve the precision of the roughness measurements. Therefore, the case studies use only buses for the data collection. The data set includes vertical acceleration, speed, and GPS receiver samples from 18 different buses traveling four different routes each day. The data collection device was a smartphone app called PAVVET (Bridgelall and Tolliver 2016). The authors selected two of the bus routes to contain a common roadway segment. This allowed for testing the ability of the selected statistical method to detect the overlap. By subjective observations, the tested roadway segment of one route was relatively smooth whereas one was relatively rough. The roughness of the overlapping roadway segment was intermediate.

The data collection covered bus operation at different times of the day, including during weekdays and on weekends. Figure 1 shows the nature of the sensor data collected. Figure 1a plots the typical vertical acceleration (g_z) signal and the corresponding RIF-indices derived from it. The vertical acceleration is in units of g-force, which the sensor sampled at a rate of 128 Hz. The RIF-indices shown characterize roughness in g-force/meter for the previous 10 meters of a traversal. Hence, the L in equation (1) is approximately 10 with variances due to GPS errors. The significance of selecting 10-meters is that it is the 95% confidence interval of GPS errors (Hughes 2016), and it is within one bus length. The reported RIF-index for the overall road segment is the average of the RIF-indices calculated for the consecutive 10-meter segments.

[Figure 1 near here].

Incidentally, in addition to reporting an overall segment roughness, the plots express how localized RIF-indices could identify roadway anomalies such as potholes and large cracks. Figure 1b shows the bus yaw that indicates turns in degrees, and the bus speed in m s^{-1}. The high variability of the speed and the stop-and-go conditions are evident.

2.3 Tests for normality of distribution

To achieve the first objective listed earlier, the authors apply a goodness-of-fit test to detect any significant departures of the RIF-indices from a normal distribution. Such a test hinges on the theory that the mean of a normally distributed variable from a population sample is an unbiased estimate of the true population average (Agresti and Finlay 2009). Most of the samples that form a normal distribution, also known as a Gaussian distribution, cluster towards the mean. Hence, a normal distribution of RIF-indices will express a central tendency towards the average roughness experienced. Deviations from the mean value will be due to variations in speed, tire path, and suspension behaviors. The central limit theorem posits that averages computed from independent and sufficiently large numbers of random measurements, taken from independent distributions, converge in distribution to the Gaussian (Papoulis 1991). A connected vehicle environment satisfies the condition for a large number of measurements because it is feasible to collect the inertial and speed data from thousands of similar connected vehicles traversing a road segment each day.

Statisticians use the Student's t-distribution instead of the normal distribution to estimate the mean of a population when the sample size is relatively small, such as less than 30 (Agresti and Finlay 2009). Nevertheless, the t-distribution approaches the Gaussian when the measurement volume exceeds 30. Therefore, the case study presented in the next section applies the goodness-of-fit tests for both Gaussian and t-distributions.

The Gaussian model $D_g(\xi)$ estimates the normal distribution of a random variable ξ such that

$$D_g(\xi) = \frac{\gamma_g}{\sqrt{2\pi\sigma_g^2}} \exp\left[-\frac{1}{2}\left(\frac{\xi - \mu_g}{\sigma_g} \right)^2 \right] \tag{3}$$

where γ_g, μ_g, and σ_g are estimates of the amplitude, mean, and standard deviation parameters, respectively. Similarly, the model for estimating the Student's *t*-distribution $D_t(\xi)$ is

$$D_t(\xi) = \frac{\gamma_t}{\sigma_t} t_{DF}\left[\frac{\xi - \mu_t}{\sigma_t}\right] \tag{4}$$

where $t_{df}(\xi)$ is the normalized Student's *t*-distribution, which is a gamma function of ξ and degrees-of-freedom *DF*. The parameters γ_t, μ_t, and σ_t are estimates of the amplitude, mean, and standard deviation parameters, respectively.

The chi-squared goodness-of-fit test (Papoulis 1991) determines whether there is a significant difference between the expected frequencies and the observed frequencies of measured values. The null hypothesis H_0 is that the observed distribution of the RIF-indices derived from the inertial measurements is the same as the candidate distribution. Failure to reject the null hypothesis will result in accepting the alternative hypothesis H_1: N that there was no significant departure of the observed distribution from the candidate distribution. Such an outcome will indicate a central tendency of roughness indices towards the mean value experienced.

The chi-squared statistic (χ^2) is

$$\chi^2 = \sum_{k=1}^{n} \frac{(O_k - E_k)^2}{E_k}. \tag{5}$$

The random variables O_k are the histogram values observed in bin *k* and E_k are the expected values of the hypothesized distribution. The chi-squared test rejects the null hypothesis if the χ^2-statistic exceeds the critical χ^2 value derived from the theoretical chi-squared distribution, evaluated at the *DF* and at a significance percentage α. Statisticians typically set α=0.05, which represents a low probability of 5% that the test will reject the null hypothesis when in fact it is true. The alternative approach is to calculate the chi-squared probability values (p-values) that correspond to the observed χ^2-statistic and the *DF*. The tests will reject the null hypothesis when the p-values are less than the selected significance percentage.

The margin-of-error (MOE) is a measure of the amount of clustering towards the mean. It expresses the amount of variability in the measurement, thereby indicating the reliability of the mean. The MOE percentage for the distribution of a random variable within a (1-α)% confidence interval with significance α (Papoulis 1991) is

$$MOE_{1-\alpha} = \pm \frac{\sigma \times t_{1-\alpha/2,DF}}{\mu\sqrt{N}} \quad (6)$$

where $t_{1-\alpha/2,df}$ is the *t*-score for a normalized cumulative t-distribution with *DF* degrees of freedom, μ is the mean, and σ is the standard deviation. The case study will calculate the MOE_{95}, which is the margin-of-error within a 95% confidence interval.

2.4 Tests for measurement distinguishability

To achieve the second objective listed earlier, the authors apply t-tests to compare the statistics of RIF-indices from each unique pair of road segment. A t-test determines if two sets of measurements are significantly different from each other (Agresti and Finlay 2009). Hence, the null hypothesis H_0:M for this test is that the two sets of measurements came from the same road segment. Setting the level of significance $\alpha = 5\%$ establishes a relatively low probability of observing a type I error, which is rejecting the null hypothesis when in fact it is true. Hence, a probability value that is less than 0.05 will result in a rejection of the null hypothesis H_0:M.

The standard t-test assumes that the two data sets have equal variances. The Cochran t-test (Cochran and Cox 1957) removes the assumption of equal variances by modifying the t-statistic to be

$$t = \frac{\left(\frac{s_1^2}{\sum_{i=1}^{N_1} f_{1i}}\right)t_1 + \left(\frac{s_2^2}{\sum_{i=1}^{N_2} f_{2i}}\right)t_2}{\left(\frac{s_1^2}{\sum_{i=1}^{N_1} f_{1i}}\right) + \left(\frac{s_2^2}{\sum_{i=1}^{N_2} f_{2i}}\right)} \tag{7}$$

where t_1 and t_2 are the critical values of the t-distribution corresponding to the selected significance level α and the sample sizes *N1* and *N2*, respectively. The parameters f_1 and f_2 are the frequencies of the i^{th} observation in their respective sample sets. The parameters s_1 and s_2 are the standard deviations of the respective sample sets.

The Folded F-test (Steel and Torrie 1980) checks for the homogeneity of variances among the sample sets. The null hypothesis H_0:V is that the variances of two distributions are equal. The F-statistic is

$$F = \frac{\max(s_1^2, s_2^2)}{\min(s_1^2, s_2^2)} \tag{8}$$

The F-distribution evaluated at the significance level α and the two degrees-of-freedom DF1 and DF2 produces the p-value. Hence, all of the tests in the case studies will reject the null hypothesis H_0:V if the p-values are less than 5%.

3 Case studies

This case study measured roughness from the traversals of 18 different buses across road segments of four different routes. A tested road segment along two of the routes overlaps. Figure 2 shows the relative positions of the road segments along the bus routes. The inset shows images of the two dominant bus types used.

[Figure 2 near here].

The bus routes are located in the small city of Fargo, North Dakota, and the lengths of the road segments are between two and five kilometers. The labels EM, SG, UG, and WM denote the road segments from Essentia Hospital to the Mall (5.1 km), Sanford Hospital to the Ground Transportation Center (2.1 km), University Drive to the Ground Transportation Center (2.3 km), and the Walmart-Mall loop (3.9 km), respectively. The overlap in the SG and UG segments is more than 90%.

From subjective assessments during the data collection, the ordered ranking of relative roughness from high to low is EM, SG, UG, and WM. Table 1 summarizes the number of measurement-trips from each of the 18 buses. The project sponsors supported the cost and time to obtain at least 30 roughness measurements from each route, thus assuring a statistically significant sample. The data collection occurred on weekdays, weekends, and at different times of the day when there was no precipitation on the roads. [Table 1 near here].

The same buses traversed the EM and SG road segments because those are part of their overall bus route. Similarly, the same buses traversed the UG and WM road segments. Only bus number 3 and 10 traversed all of the road segments at some time during the data collection period.

4 Results and discussion

Figure 3 plots the distribution of the RIF-indices for each road segments. The relative ranking of the mean RIF-indices agree with the subjective ranking. However, the distribution variances overlap significantly for each route. Therefore, only a statistical test can determine if the measurements are in fact from different road segments.

[Figure 3 near here].

4.1 Reliability of measurements

The authors use the iterative Levenberg-Marquardt nonlinear least squares method of curve fitting to identify the parameters of the best-fit distributions (Gill *et al.* 1981). Table 2 lists those amplitude, mean, and standard deviation (γ, μ, σ) parameters. For each set of measurements, the χ^2-statistic does not exceed the χ^2-critical value for either the normal or the t-distributions. The table includes the p-values corresponding to the χ^2-statistic in percentages. These results indicate that the chi-squared tests cannot reject the null hypothesis in any of the cases. Therefore, the statistical tests find no significant departure of the observed distribution from the normal or the t-distributions.

[Table 2 near here].

Table 2 lists the MOE within the 95% confidence interval (MOE_{95}). The average MOE_{95} is below 6% for all the measurements. This indicates that additional measurements will likely yield even further reductions in the MOE_{95}. Hence, tendency towards a central value per the Gaussian, coupled with relatively low margins of error, indicate a high reliability of the measurements. Therefore, the mean value of the RIF-indices is an unbiased estimate of the average roughness experienced. Furthermore, the precision of measurements will increase as the number of connected vehicle traversals increases.

4.2 Distinguishability of measurements

Table 3 indicates that without merging the data for the common routes, the F-tests reject the null hypothesis H_0:V of equal variances for three of the tested segment pairs. This result prompted the use of Cochran t-tests because it does not assume equal sample sizes or equal variances.

[Table 3 near here].

As summarized in Table 3, the t-tests cannot reject the null hypothesis H_0:M of indistinguishable statistics for two of the comparisons. This result correctly indicates that the measurements overlapped for at least one of the data set pairs. Repeating the tests after merging the data for the common SG and UG segments yielded the expected result. As summarized in Table 4, the t-tests rejected the null hypothesis for each of the comparisons, thereby affirming that measurements of RIF-indices are distinguishable among different road segments.

[Table 4 near here].

The minimum difference in roughness among the three road segments was only about 11%. Even so, the level of distinguishability will increase further with an ever-increasing precision by combining measurements from a larger number of connected vehicles.

5 Summary and conclusions

Agencies typically do not evaluate the IRI for local and unpaved roads because of the technical and practical limitations of existing specialized equipment. CV methods are evolving to bridge the gap by enabling continuous roughness characterizations of the entire roadway network. However, speed variability and differences in suspension behavior require data segmentation to achieve some level of desired precision and accuracy with relatively few measurements.

The contribution of this research was to evaluate a CV method that uses the RIF-transform to determine the extent to which it could reliably estimate the average roughness that riders experience when traveling at non-uniform speeds, in different vehicles of the same type. The second contribution was to determine the extent to which the CV method could distinguish among the roughness differences of roadway segments, when measured under the same conditions of speed and suspension variability. The statistical testing determined that under those conditions, the RIF-transform yields an ever-increasing accuracy and precision of roughness characterizations as the number of measurements increase. The RIF-transform distinguished the average roughness difference among the road segments of the case study where the minimum difference was 11%. The power to distinguish even smaller differences in roughness will increase as the MOE_{95} decreases further with larger measurement volume.

In future research, the authors plan to examine properties of the CV method when characterizing the ride quality for unpaved roads under the same conditions of speed and suspension variability. Experiments will include different vehicle types, loading conditions, and roughness averaging distances. Those will be some of the factors used to estimate a model for the standard error as a function of traversal volume.

Acknowledgements

A grant from the National Center for Transit Research (NCTR) and the Small Urban and Rural Transit Center (SURTC) of the Upper Great Plains Transportation Institute, North Dakota State University supported this research. The authors also express their sincere appreciation to Julie Bommelman (Transit Administrator of the City of Fargo), James Gilmour (Planning Director of the City of Fargo), and Gregg Schildberger (Senior Transit Planner for the City of Fargo) for their support in accessing the buses for roughness measurements.

References

Agresti, A. and Finlay, B., 2009. *Statistical Methods for the Social Sciences*. 4th. Upper Saddle River: Pearson Prentice Hall.

Alessandroni, G., Carini, A., Lattanzi, E., Freschi, V., and Bogliolo, A., 2017. A study on the influence of speed on road roughness sensing: the SmartRoadSense case. *Sensors* 17 (2): 305.

Bilodeau, J. -P., Gagnon, L., and Doré, G., 2015. Assessment of the relationship between the international roughness index and dynamic loading of heavy vehicles. *Int. J. Pavement Eng*., DOI: 10.1080/10298436.2015.1121780, 1-9.

Bridgelall, R., 2014a. Connected Vehicle Approach for Pavement Roughness Evaluation. *J. Infrastruct. Syst*., 20 (1), 1-6.

Bridgelall, R., 2014b. Inertial Sensor Sample Rate Selection for Ride Quality Measures. *J. Infrastruct. Syst*., 21 (2), 1-5.

Bridgelall, R., 2015. Precision Bounds of Pavement Distress Localization with Connected Vehicle Sensors. *J. Infrastruct. Syst*. 21 (3), 1-7.

Bridgelall, R. and Tolliver, D. D., 2016. Accuracy enhancement of roadway anomaly localization using connected vehicles. *Int. J. Pavement Eng*., DOI: 10.1080/10298436.2016.1162306, 1-7.

Bridgelall, R., Rahman, M. T., Tolliver, D. D., and Daleiden, J. F., 2016a. Ride quality characterizations using connected vehicles. In: *Proc. 95th Annual Meeting of the Transportation Research Board*. Washington, D.C.

Bridgelall, R., Rahman, M. T., Daleiden, J. F., and Tolliver, D. T., 2016b. Error sensitivity of the connected vehicle approach to pavement performance evaluations. *Int. J. Pavement Eng*., DOI: 10.1080/10298436.2016.1162307, 1-6.

Bridgelall, R., Rahman, M. T., Tolliver, D. D., and Daleiden, J. F., 2017. Wavelength sensitivity of roughness measurements using connected vehicles. *Int. J. Pavement Eng*., DOI: 10.1080/10298436.2017.1316645, 1-7.

Cochran, W. G., and Cox, G. M., 1957. *Experimental Designs*. 2nd. New York: John Wiley and Sons.

Cruz, J. F., and Castro, J. T., 2015. Estimating road roughness conditions using ubiquitous smartphones and geographic information systems and its application to road network planning in the Philippines. In: Proceedings of the *11th International Conference of the Eastern Asia Society for Transportation Studies (EASTS 2015)*, University of the Philippines, Cebu City, Philippines.

Dawkins, J., Bevly, D., Powell, B., and Bishop, R., 2011. *Investigation of Pavement Maintenance Applications of Intellidrive*. Pooled Fund Study, Charlottesville, Virginia: University of Virginia.

Du, Y., Liu, C., Wu, D., and Jiang, S., 2014. Measurement of International Roughness Index by Using Z-Axis Accelerometers and GPS. *Math. Probl. Eng.* 2014: 1-10.

Gill, P. E., Murray, W., and Wright, M. H., 1981. *Practical optimization*. United Kingdom: Emerald Group Publishing Ltd.

Gillespie, T. D., Sayers, M. W., and Queiroz, C. A. V., 1986. *The International Road Roughness Experiment: Establishing Correlation and Calibration Standard for Measurement*. Technical Report No. 45, Washington, D.C.: The World Bank, 464.

HPMS. 2016. *Highway Performance Monitoring System (HPMS) Field Manual*. Office of Highway Policy Information, Office of Management and Budget (OMB), Washington, DC: Federal Highway Administration, 295.

Hughes, W. J., 2016. *Global Positioning System (GPS) Standard Positioning Service (SPS) Performance Analysis Report. Performance Analysis Report 94*, T and E Team, Federal Aviation Administration (FAA), Wide Area Augmentation System (WAAS), Washington, D.C.: Technical Center.

Islam, S., Buttlar, W. G., Aldunate, R. G., and Vavrik, W. R., 2014. Measurement of Pavement Roughness Using Android-Based Smartphone Application. *Transport Res. Rec.* 2457 (1): 30-38.

Karamihas, S. M., 2017. *Measuring, Characterizing, and Reporting Pavement Roughness of Low-Speed and Urban Roads*. Research in Progress, Washington, D.C.: Transportation Research Board of the National Academies. Accessed June 10, 2017. http://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=3404.

Dennis, E. P., and Spulber, A., 2016. *Performance-based Planning and Programming for Pavement Management*, Ann Arbor, MI: Michigan Department of Transportation (MDOT) and Center for Automotive Research (CAR), 63.

Nomura, T., and Shiraishi Y., 2015. A Method for Estimating Road Surface Conditions with a Smartphone. International Journal of Informatics Society 7 (1): 29-36.

Papagiannakis, A. T., 1997. The Need for a New Pavement Roughness Index; RIDE. *International Truck & Bus Meeting & Exposition*. Washington, D.C.: Society of Automotive Engineers International.

Papoulis, A., 1991. *Probalility, Random Variables, and Stochastic Processes*. New York: McGraw-Hill.

Steel R.G.D. and Torrie, J.H., 1980. *Principles and Procedures of Statistics*, Second

Edition, New York: McGraw-Hill.

Stribling, J. W., 2016. *Use of smartphones to measure pavement roughness across multiple vehicle types at different speeds*. Masters Thesis, Civil Engineering, Urbana: University of Illinois at Urbana-Champaign.

Wermers, L. G, 1962. *Analysis of Road Roughness of Flexible Pavements Using the Kentucky Accelerometer*. FHWA/IN/JHRP-62/16 Project No. C-36-54BB, Lafayette, Indiana: Purdue University, 20.

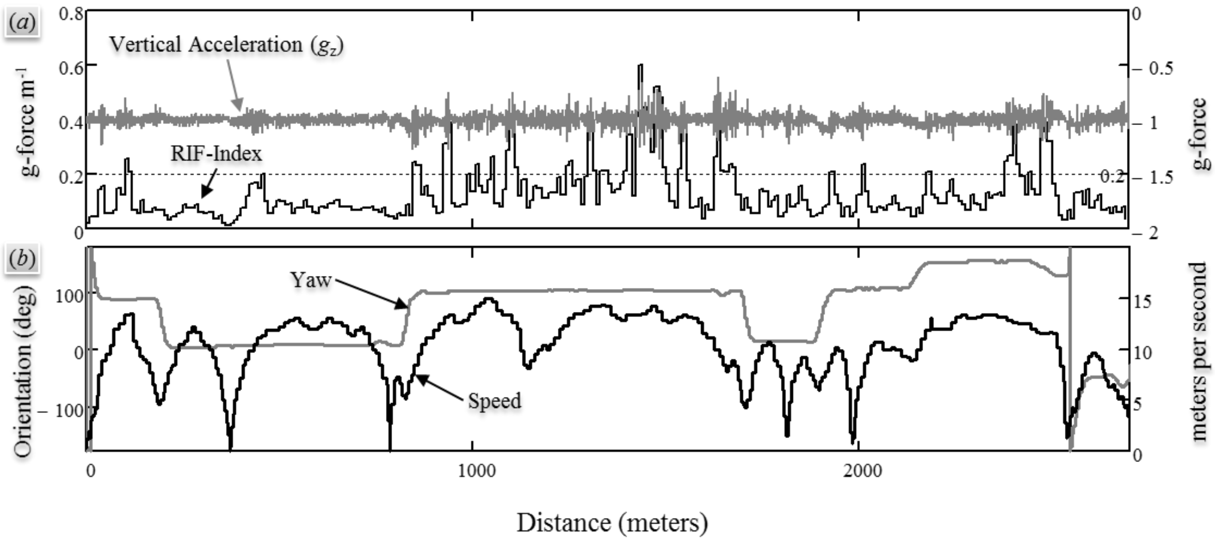

Figure 1. Sensor output of (a) inertial data (b) speed and orientation.

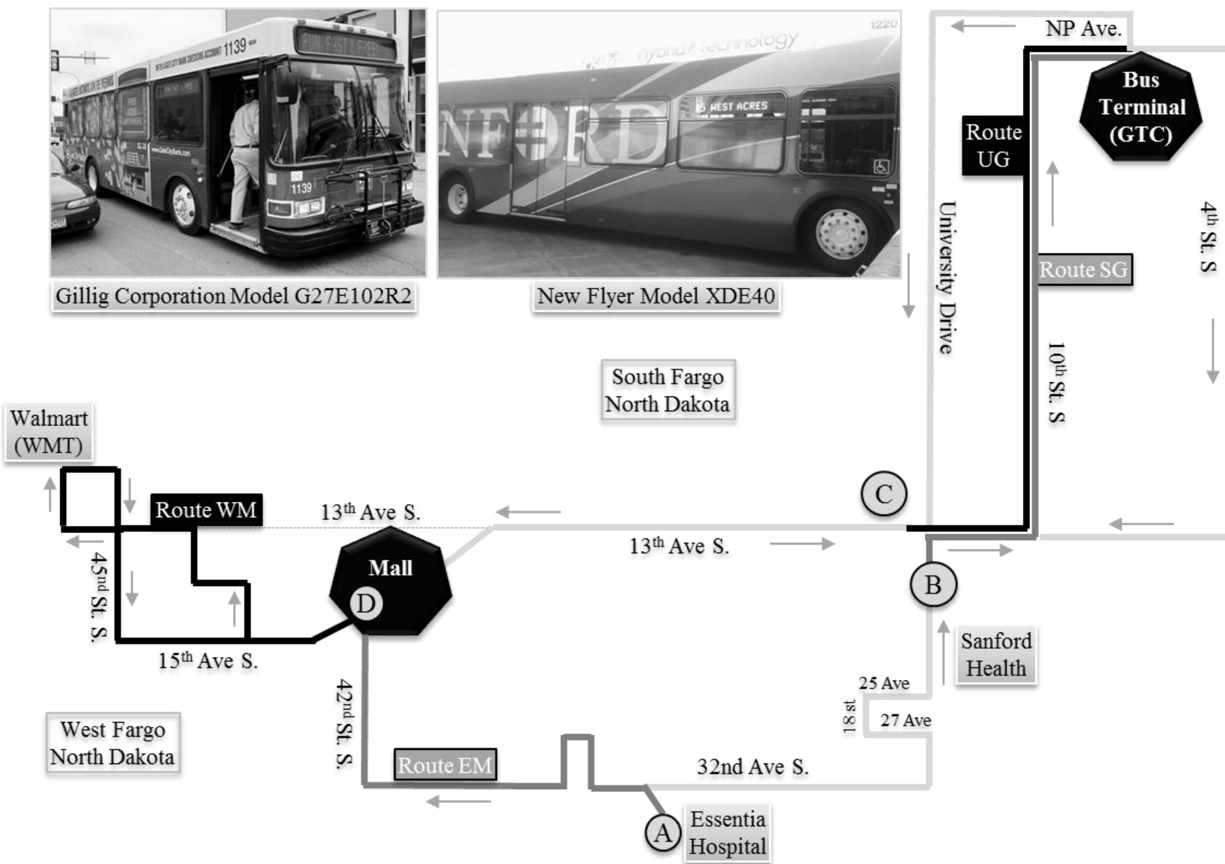

Figure 2. Buses and road segments of the case study.

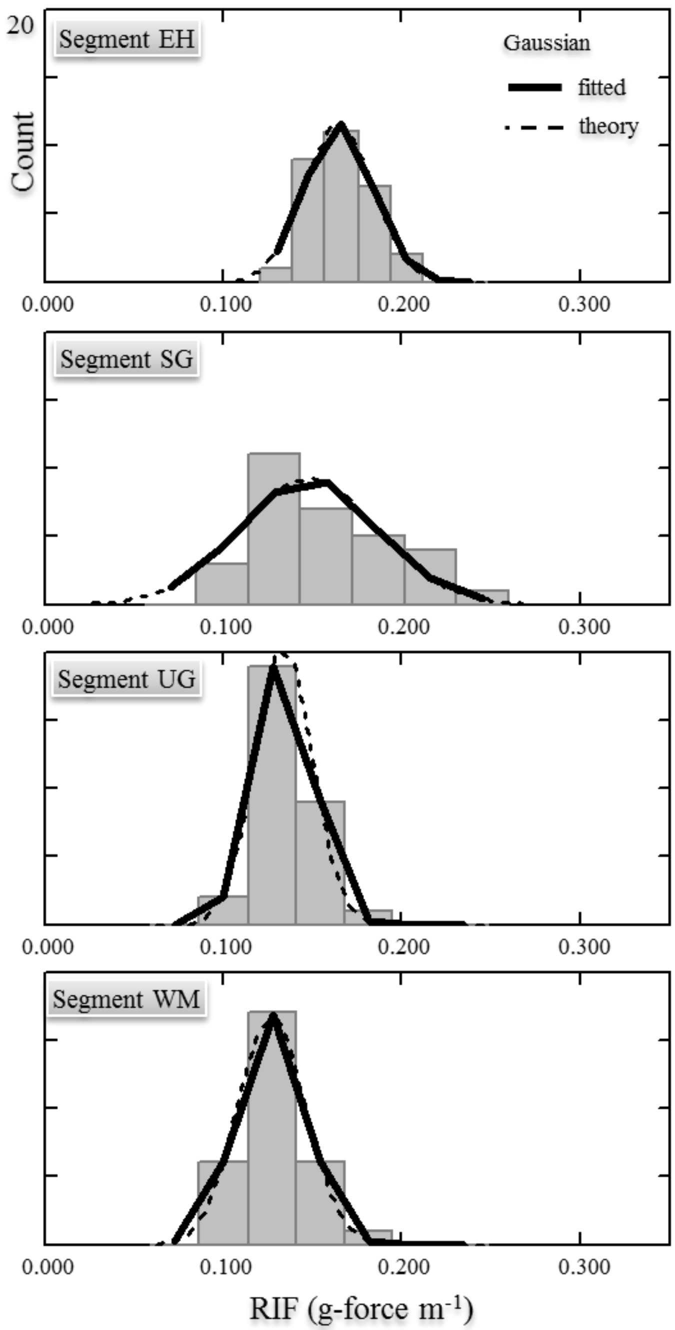

Figure 3. Comparative distribution of RIF-indices for four road segments of four routes.

Table 1. Trips by bus number for each route.

	Measurement-Trips by Bus Number																		
Segment	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Trips
EM	0	0	2	0	2	5	2	0	7	4	1	4	0	3	0	0	0	0	30
SG	0	0	5	3	3	1	1	2	6	3	1	1	2	3	0	0	0	0	31
UG	2	2	3	0	0	0	0	0	0	2	0	0	3	0	10	2	3	4	31
WM	2	2	3	0	0	0	0	0	0	1	0	0	3	0	10	2	2	5	30
Trips	4	4	13	3	5	6	3	2	13	10	2	5	8	6	20	4	5	9	123

Table 2. Chi-squared tests for normality without merging the common route data.

	Road Segments			
Parameters	EM	SG	UG	WM
RIF-Indices	30	31	31	30
Mean	0.166	0.157	0.135	0.130
STD	0.017	0.038	0.018	0.018
MOE_{95} (%)	3.86	8.908	4.828	5.211
Normal distribution				
χ^2 *DF*	4	4	4	4
γ_g	0.554	0.894	0.82	0.799
μ_g	0.164	0.147	0.134	0.127
σ_g	0.019	0.039	0.016	0.019
χ^2 Critical	9.5	9.5	9.5	9.5
χ^2 Statistic	1.1	6.4	2.7	2.2
p-values (%)	89.4	17.2	61.8	70.5
Student's t-distribution				
χ2 DF	4	4	4	4
γ_t	0.593	0.925	0.86	0.839
μ_t	0.164	0.144	0.135	0.127
σ_t	0.019	0.035	0.014	0.018
χ^2 Critical	9.5	9.5	9.5	9.5
χ^2 Statistic	2.2	6.8	0.4	1.2
p-values (%)	70.0	14.7	98.0	87.1

Table 3. F- and t-tests without merging data for common route.

Test Pair	DF	F-value	Pr > F	Reject H0:V?	t-value	Pr > |t|	Reject H_0:M?
EM-SG	30	4.94	<.0001	Yes	1.18	0.2483	No
EM-UG	30	1.07	0.8504	No	6.91	0.0001	Yes
EM-WM	29	1.12	0.7639	No	7.88	<.0001	Yes
SG-UG	30	4.61	<0.0001	Yes	2.92	0.0067	Yes
SG-WM	30	4.43	0.0010	Yes	3.55	0.0013	Yes
UG-WM	29	1.04	0.9155	No	1.09	0.2849	No

Table 4. F- and t-tests after merging data for common route.

Route Pair	DF	F-value	Pr > F	Reject H_0:V?	t-value	Pr > |t|	Reject H_0:M
EM-(SG+UG)	61	3.38	0.0006	Yes	3.91	0.0003	Yes
EM-WM	30	1.12	0.7639	No	7.88	<.0001	Yes
(SG+UG)-WM	61	3.02	0.0017	Yes	3.08	0.0036	Yes