

**Pre-print Manuscript of Article:**

Bridgelall, R., Tolliver, D., “Accuracy enhancement of roadway anomaly localization using connected vehicles,” *International Journal of Pavement Engineering*, pp. 1-7, March 18, 2016.

**Accuracy enhancement of roadway anomaly localization using connected vehicles**

Raj Bridgelall, Denver Tolliver

*Upper Great Plains Transportation Institute, North Dakota State University, Fargo, ND, USA*

Raj Bridgelall (Corresponding Author): P.O. Box 863676, Plano, TX, 75086; Phone 408-607-3214; E-mail: [raj@bridgelall.com](mailto:raj@bridgelall.com)

Denver Tolliver: P.O. Box 6050, Fargo, ND 58108; Phone 701-231-7190; E-mail: [denver.tolliver@ndsu.edu](mailto:denver.tolliver@ndsu.edu)

The University Transportation Centre, a program of the United States Department of Transportation, sponsored this research through its Mountain Plains Consortium (MPC) under grant DTRT12-G-UTC08.

**ABSTRACT**

The timely identification and localization of roadway anomalies that pose hazards to the traveling public is currently a critical but very expensive task. Hence, transportation agencies are evaluating emerging alternatives that use connected vehicles to lower the cost dramatically and to increase simultaneously both the monitoring frequency and the network coverage. Connected vehicle methods use conventional GPS receivers to tag the inertial data stream with geospatial position estimates. In addition to the anticipated GPS trilateration errors, numerous other factors reduce the accuracy of anomaly localization. However, practitioners currently lack information about their characteristics and significance. This study developed error models to characterize the factors in position biases so that practitioners can estimate and remove them. The field studies revealed the typical and relative contributions of each factor, and validated the models by demonstrating agreement of their statistics with the anticipated norms. The results revealed a surprising potential for tagging errors from embedded systems latencies to exceed the typical GPS errors and become dominant at highway speeds.

**Keywords:** accelerometer; GPS; inertial profiler; pavement performance; ride quality; sensor fusion; smartphone

## 1 Introduction

Localized roughness from anomalies such as frost heaves, pavement cracking, potholes, spills, and debris pose serious hazards to the traveling public. Rough roads cause damage to transported goods (Steyn *et al.* 2015), and increase the cost of operating a vehicle (Robbins and Tran 2015). Transportation agencies rely on the regular and accurate reporting of localized roughness to prioritize maintenance needs (Karamihas and Senn 2012). Hence, the inaccurate reporting of anomaly positions could lead to unnecessary and costly decisions or maintenance actions. To enforce measurement precision and accuracy requirements, some transportation agencies penalize contractors for erroneously reporting localized roughness (Chen and Dye 2014).

Existing approaches use heavily instrumented vehicles with laser-based systems and computers to measure the road elevation profile. Post processing transforms the data to the international roughness index (IRI) for a selected spatial resolution. These probe vehicles are relatively expensive to deploy and maintain. In addition, agencies do not use such vehicles on most local and unpaved roads because of numerous technical shortcomings (Papagiannakis 1997) and practical limitations (Ahlin and Granlund 2002). Such constraints also limit monitoring of the critical highways to at most once annually (NCHRP 2015). Hence, important vulnerabilities such as frost heaves, dangerous spills, or debris that appear and disappear between monitoring cycles go undetected. Consequently, agencies are evaluating connected vehicle methods because of their potential to provide more affordable and continuous monitoring for the entire network.

Connected vehicle methods use on-board accelerometers and conventional GPS receivers to report the inertial response and position of vehicles. However, the ability to

transform inertial and geospatial position data from many vehicles into a consistent summary of roughness has remained a significant challenge (Du *et al.* 2014). In related work, we recently addressed this challenge (Bridgelall 2014a) by inventing and demonstrating a signal transform that converts and merges the sensor data stream from many connected vehicles to produce a new roughness index that is directly proportional to the IRI. The road impact factor (RIF) transform produces the RIF-index from direct inertial measurements to characterize the actual roughness that riders experience within a specified speed band. In contrast, the IRI is an estimate of the roughness experienced based on a simulated quarter-car moving at a fixed speed across the elevation profile samples collected.

The IRI procedure computes the damped mass-spring motions of a standard quarter-car model as it responds to traversing the sampled elevation profile, at a precise reference speed of  $80 \text{ km h}^{-1}$  (Papagiannakis 1997). The procedure computes the IRI as the accumulated absolute rate difference between the sprung- and unsprung-mass motions for a specified traversal distance. In contrast, the RIF-transform computes the energy of the longitudinal velocity-modulated vertical acceleration sensed from the vehicle's body. This computation produces a RIF-index per unit of distance travelled (Bridgelall 2014a). Therefore, the RIF-index is the linear energy density of the actual g-forces that riders experience within any speed band. On the other hand, the IRI is an accumulation of the simulated vertical motions of a quarter-car model. Nevertheless, both indices result from the linear time-invariant transformation of motions from equivalent damped mass-spring systems. Hence, the RIF-index and the IRI are directly proportional at any fixed speed.

## Accuracy enhancement of roadway anomaly localization using connected vehicles

Previous work demonstrated this proportionality at both local and highway speeds (Bridgelall 2014a).

A corresponding Time Wavelength-Intensity Transform (TWIT) combines the RIF-indices from all available speed bands to produce a speed-independent summary of roughness. The connected vehicle approach obviates the need for calibration with individual vehicle installations by leveraging the central limit theorem for a large data volume (Bridgelall 2014a). Hence, the average RIF-index across all speed bands reflects the typical ride quality experienced at any speed, and establishes a practical figure-of-merit to trigger specific remediation actions.

One shortcoming of the connected vehicle approach is that the data fusion produces a resultant position tagging error. This aggregate bias accumulates from the position tagging errors of the individual data streams. Therefore, the main idea of this paper is to develop a method to characterize and remove position biases from the individual data streams. This method will enhance the accuracy of localizing anomalies from the merged data. The ever-increasing traversal volume of connected vehicles will yield a corresponding enhancement in the precision of localization.

Related research investigated participatory sensing approaches to identify clusters of roughness from smartphone user reports to estimate the position of possible anomalies such as potholes (Byrne *et al.* 2013). Methods that directly analyse the accelerometer signal stream from individual vehicles used short-time spectral transforms to identify the signatures of anomalies (Ayenu-Prah and Attoh-Okine 2009). Other methods include attempting to recover the road profile by double integration of the accelerometer signal

(Islam *et al.* 2014, Nomura and Shiraishi 2015), and signal classification using machine learning methods (Rajamohan, Gannu and Rajan 2015).

In adjacent fields, researchers report on the use of non-conventional GPS receivers such as differential GPS to measure the static and dynamic behaviours of large structures such as bridges (Yi, Li and Gu 2010), and high-rise buildings (Yi, Li and Gu 2013). Such applications focus on designing or using high performance geospatial positioning systems to maximize the accuracy of structural models used in finite element analysis. Given the high specialization of using connected vehicles to localize roadway anomalies, practitioners currently know very little about how to estimate and remove errors in position tagging of the inertial data stream.

The main objective of this study is to develop models for the position tagging errors. Practitioners will benefit by using the models to estimate and remove bias from the position tags of the RIF-indices derived from each data stream. Therefore, the sensor data fusion will produce minimal bias, and the precision will continue to improve with traversal volume. To evaluate the model, we designed and conducted six field studies to characterize the error distribution and to compare the relative magnitude of each error component. To validate the model, we tested the error distribution against classic distributions to demonstrate agreement with the established norms. Subsequently, the organization of this paper is as follows: the next section develops the error model. The third section describes the field studies conducted to characterize the statistics of the overall position tagging error and to compare the relative contributions from each factor. The final section summarizes the findings and recommends an approach to minimize the localization error.

## 2 Anomaly position estimate

For an isolated anomaly, the position tag  $\hat{\nu}_p$  of the corresponding inertial signal peak is an estimate of the true position  $\nu_p$  of the anomaly. The estimate includes distance biases such that

$$\hat{\nu}_p = \nu_p + \bar{\varepsilon}_{\text{GPS}} + (\bar{\varepsilon}_{\text{DSP}} + \bar{\varepsilon}_i + \bar{\varepsilon}_b + \bar{\varepsilon}_s). \quad (1)$$

This expression groups the biases into two categories: GPS and non-GPS related. The GPS related bias is the average offset in position tags  $\bar{\varepsilon}_{\text{GPS}}$ . The position bias from digital signal processing (DSP) is the expected delay  $\bar{\varepsilon}_{\text{DSP}}$  from digital filtering. The position bias  $\bar{\varepsilon}_i$  is the average error in locating a peak within the interpolation sub-interval. For the typical anomaly type, the vehicle mechanical response delay  $\bar{\varepsilon}_b$  primarily depends on the frequency response of the vehicle suspension and the traversal speed. The sensor position bias  $\bar{\varepsilon}_s$  is the longitudinal distance of the GPS antenna from the first axle that traverses the anomaly.

### 2.1 GPS related errors

The geospatial position tag  $\bar{\varepsilon}_{\text{GPS}}$  reported by the GPS receiver and its associated embedded system has two error components such that the average position bias is

$$\bar{\varepsilon}_{\text{GPS}} = \bar{\varepsilon}_{\text{dGPS}} + \bar{\varepsilon}_{\text{dlag}} \quad (2)$$

The first error component  $\bar{\varepsilon}_{\text{dGPS}}$  is the familiar geospatial position bias from GPS trilateration that the literature has long established to have zero mean and normal distribution (Gade 2010). However, the literature seldom identifies or reports on the second

Accuracy enhancement of roadway anomaly localization using connected vehicles

error component  $\bar{\epsilon}_{\text{dlag}}$ . It arises from latencies in the embedded system that fetches coordinates from the GPS receiver at the selected update interval to assign position tags to the inertial samples. Therefore, unlike the trilateration error, the expected value of position tagging latency will be non-zero, and equation (2) becomes  $\bar{\epsilon}_{\text{GPS}} = \bar{\epsilon}_{\text{dlag}}$ .

From equation (2) the total variance in geospatial position tagging  $\sigma_{\text{GPS}}^2$  is

$$\sigma_{\text{GPS}}^2 = \sigma_{\text{dGPS}}^2 + \sigma_{\text{dlag}}^2 \quad (3)$$

According to GPS system administrators (USDHS 1996), the variance  $\sigma_{\text{dGPS}}^2$  from trilateration errors can be substantial because of random changes in atmospheric effects, multipath propagation, and GPS receiver performance. System administrators expect that the 95% confidence interval for horizontal position precision under direct line-of-sight conditions will be about 6.7 meters. However, this uncertainty could increase to more than 10 meters when multi-path reflections from buildings, large trees, and other tall structures distort the weak satellite signals.

The average latency in position tagging  $\bar{\epsilon}_{\text{dlag}}$  and its variance  $\sigma_{\text{dlag}}^2$  depends on the embedded system implementation, which is often a trade secret. The average lag distance in position tagging is directly proportional to the average lag time  $\bar{\tau}_{\text{lag}}$  from embedded system latencies such that

$$\bar{\epsilon}_{\text{dlag}} = \bar{\tau}_{\text{lag}} \bar{v}. \quad (4)$$

Hence, the corresponding variance in position tagging bias  $\sigma_{\text{dlag}}^2$  is

$$\sigma_{\text{dlag}}^2 = (\bar{v} \sigma_{\tau_{\text{lag}}})^2 + (\bar{\tau}_{\text{lag}} \sigma_v)^2 \quad (5)$$

where  $\sigma_v^2$  is the variance in vehicle velocity. The asynchronous timing of GPS receiver updates with respect to the interrupt response times of the embedded system lead to variances in the lag time  $\sigma_{\tau_{\text{lag}}}^2$ . Such information is not generally available from the equipment manufacturers.

## 2.2 *Non-GPS related errors*

The phase response of the vehicle suspension system causes a predictable delay of the response motion peak (Jazar 2008). This is analogous to the established approach of estimating the time delay of a signal by knowing the phase response of the filter (Young, Freedman and Ford 2011). Researchers have recently demonstrated that it is also possible to estimate the equivalent quarter-car parameters directly from the inertial response signal of any vehicle that traverses a rough spot (Blum 2015). We have also demonstrated similar results in previous work (Bridgelall 2014b) and established that the distance delay of the peak inertial response for typical anomalies is approximately 1.5 cm per  $\text{m s}^{-1}$  change in longitudinal velocity such that

$$\bar{\varepsilon}_b = 0.015 \cdot \bar{v}. \quad (6)$$



Subsequently, the corresponding variance  $\sigma_{\text{eb}}^2$  of the first peak offset is

$$\sigma_{\text{eb}}^2 = (0.015)^2 \cdot \sigma_v^2 \quad (7)$$

The signal processing for the case studies applies a standard digital low-pass filter to suppress noise and improve the precision of detecting the inertial peak. The expected filter delay  $\bar{\varepsilon}_{\text{DSP}}$  is

$$\bar{\varepsilon}_{\text{DSP}} = N_{\text{LP}} \frac{\bar{v}}{\bar{f}_A} \quad (8)$$

where  $N_{\text{LP}}$  is the number of stages in the digital filter cascade, and  $\bar{f}_A$  is the average sample rate, which is the same as the accelerometer sample rate. Hence, the corresponding variance of the DSP filter delay  $\sigma_{\varepsilon_{\text{DSP}}}^2$  is

$$\sigma_{\varepsilon_{\text{DSP}}}^2 = \left( \frac{\partial \bar{\varepsilon}_{\text{DSP}}}{\partial \bar{v}} \sigma_v \right)^2 + \left( \frac{\partial \bar{\varepsilon}_{\text{DSP}}}{\partial \bar{f}_A} \sigma_{f_A} \right)^2 + 2 \cdot \text{cov}[\bar{v}, \bar{f}_A] = N_{\text{FIR}} \left[ \left( \frac{1}{\bar{f}_A} \sigma_v \right)^2 + \left( \frac{\bar{v}}{\bar{f}_A^2} \sigma_{f_A} \right)^2 \right] \quad (9)$$

where  $\sigma_v$  and  $\sigma_{f_A}$  are the standard deviations of the vehicle speed and the accelerometer sample rate, respectively. The covariance factors are zero because the vehicle speed is not a function of the accelerometer sample rate.

The average sample interval is  $\bar{\delta}_v = \bar{v} \cdot \bar{\tau}_A$  where the average sample period of the accelerometer signal is  $\bar{\tau}_A$ . Hence, the error in estimating the position of a signal peak within the sample interval will be at most  $\bar{\delta}_v$ . If the distribution of the peak position is uniform within the sample interval, then the average error is

$$\bar{\varepsilon}_i = \frac{1}{2} \bar{v} \bar{t}_A = \frac{\bar{v}}{2f_A} \quad (10)$$

Therefore, the variance  $\sigma_{ei}^2$  is

$$\sigma_{ei}^2 = \left( \frac{\partial \bar{\varepsilon}_i}{\partial \bar{v}} \sigma_v \right)^2 + \left( \frac{\partial \bar{\varepsilon}_i}{\partial f_A} \sigma_{fA} \right)^2 + 2 \cdot \text{cov}[\bar{v}, f_A] = \left( \frac{1}{2f_A} \sigma_v \right)^2 + \left( \frac{\bar{v}}{2f_A^2} \sigma_{fA} \right)^2. \quad (11)$$

This variance is directly proportional to the DSP variance such that

$$\sigma_{ei}^2 = \frac{\sigma_{eDSP}^2}{N_{FIR}}. \quad (12)$$

### 3 Field studies and results

We designed the field studies to compare the position tags assigned to an isolated inertial peak with its known reference position. Isolated anomalies produce an easily detectable peak response in the samples of the inertial signal. The anomalies were a speed bump on a park road, a raised concrete-to-asphalt pavement joint on an airport access road, and an uneven rail grade crossing a local road. Figure 1 shows street level views of those isolated anomalies. Each anomaly interrupted smoother segments measuring 30-meters on either side to produce a single dominant inertial peak.

[Figure 1 near here].

Even though the areas of uneven pavement joint and the rail grade crossing produced multiple signal peaks, one peak remained dominant for all traversals. A local geographic

information systems (GIS) database (North Dakota State Government 2015) provided the ground truth coordinates of each anomaly, and a GPS receiver validated the position by resting the device on top of the anomaly.

To simulate connected vehicles, we developed a smartphone data logger application (app) called Pavement Analysis Via Vehicle Electronic Telemetry (PAVVET) to collect the accelerometer and GPS data (Bridgelall 2014b). We mounted the smartphone on the dashboard of each vehicle and left it in the same position for all traversals. We also travelled at the speed limit for those facilities. A 2001 Ford Explorer sports utility vehicle (SUV) produced the data from the park bump traversals at three different speeds. A 2007 Toyota Camry LE sedan produced the data from traversing the rail grade crossing at a steady speed. A 2007 Subaru Legacy sedan produced the data from the airport road bump for eastbound (EB) and westbound (WB) traversals at a steady speed. We excluded two data logs from the 30 traversals in each of the six datasets to remove outlier GPS position tags.

### ***3.1 Distribution of peak position tags***

An algorithm identified the position tags of the inertial peak for each traversal by locating the peak of the first signal that exceeded a threshold  $G\sigma$  set at two standard deviations above the signal mean. To demonstrate the variations in position tags among traversals, Figure 2 shows the filtered accelerometer signal output from two traversals of the park bump at a relatively safe speed of approximately  $7 \text{ m s}^{-1}$ .

[Figure 2 near here].

The plots indicate the path distance tags of the first peaks from traversals 1 and 2 at the distance markers  $\epsilon_{p1}$  and  $\epsilon_{p2}$ , respectively. These positions are located at -6.2 and 0.1 meters

relative to the actual position of the anomaly at 30 meters. The position tags of the first peaks of the remaining 26 traversals randomly appear before or after the actual position of the anomaly. Table 1 summarizes for each experiment the parameters related to the signal processing, the vehicle, and the GPS receiver. The average position bias of the first inertial peak  $\bar{d}$  from the true position of the anomaly and the average spread  $\sigma_v$  across all cases was -4.54 and 3.28 meters, respectively. The standard deviation of the spread  $\sigma_v$  across all cases was 1.2 meters, which highlights that environmental differences can significantly affect the degree of uncertainty in GPS position tagging.

[Table 1 near here].

The histograms of Figure 3 provide a visual confirmation of the average delay, and the relative magnitude and differences in the position tag spreads. The number of bins for each histogram is proportional to the typical guideline, which is the square root of the number of traversals available. A least squares fit of the Gaussian, Student-t, and logistic distributions superimpose each histogram. The fit for these three distributions are very similar and almost indistinguishable.

[Figure 3 near here].

Table 2 summarizes the parameters of the distribution fit and the associated chi-squared ( $\chi^2$ ) values for the hypothesis that the data are accordingly distributed. The chi-squared test statistic listed ( $\chi^2$  data) is

$$\chi^2 = \sum_{k=1}^n \frac{(O_k - E_k)^2}{E_k} \quad (13)$$

where  $O_k$  are histogram values observed in bin  $k$  and  $E_k$  are the expected values from the hypothesized distribution (Papoulis 1991). The  $\chi^2$  at  $\alpha = 5\%$  is the value where the cumulative chi-squared distribution  $\chi_{df}^2$  reaches 95%. Hence, the significance  $\alpha$  is the probability of observing a chi-squared value that is at least as large as the  $\chi^2$  statistic from equation (18).

[Table 2 near here].

Statisticians generally reject a null hypothesis that the data follows a tested distribution if the significance of the  $\chi^2$  statistic is less than 5%, or equivalently, if the  $\chi^2$  statistic is larger than the cumulative chi-square distribution value at 5% significance. Table 3 highlights the largest significance levels in bold font for each case study. Hence, the chi-squared method cannot reject the hypothesis that the distributions of the position tags for the first peak follow those classic distributions.

[Table 3 near here].

### **3.2 Digital signal processing related errors**

Equations (8) and (9) provide the mean DSP filter delay  $\bar{\epsilon}_{\text{DSP}}$  and its standard deviation  $\sigma_{\text{eDSP}}$ . The average delay and associated spread across all cases were 56 cm and 4 cm respectively. Equation (6) provides the bias  $\bar{\epsilon}_i$  and equation (7) provides the uncertainty  $\sigma_{\text{ei}}$  in peak position estimation given the sample rate of the accelerometer signal. The average sample rate was approximately 93 Hz, which was the highest rate practically achieved with the smartphone app. Across all cases, the average bias and spread of the peak position in the signal were 31 and 2 millimetres respectively.

### **3.3 Vehicle response related errors**

For each field study, the first two rows of Table 2 list the batch means of the average velocity  $\bar{v}$  and the speed standard deviations  $\bar{\sigma}_v$ , respectively. The average speed variability across all case studies was  $0.38 \text{ m s}^{-1}$  (less than 1 mph). The expected values for the aggregate vehicle suspension response related delay  $\bar{\epsilon}_b$  and delay spread  $\sigma_{\epsilon_b}$  across all case studies were 87 mm and 6 mm, respectively. The spread in suspension delay across all case studies was 27 mm. The actual position delay of an inertial sample is associated with the instantaneous speed at that position. Therefore, any bias minimization algorithm should use the instantaneous speed when applying a distance correction to each position tag based on the estimate. This minimal delay is consistent with the results of previous work (Bridgelall 2014b).

### ***3.4 Estimate of position tag latency***

Removing the non-GPS related biases produces an estimate for the position tagging latency as follows:

$$\bar{\epsilon}_{\text{dtag}} = \tilde{D} - (\bar{\epsilon}_{\text{DSP}} + \bar{\epsilon}_i + \bar{\epsilon}_b + \bar{\epsilon}_s). \quad (14)$$

The residual distance bias from tagging latency was an average of -4.41 meters across all case studies. From equation (4), the equivalent tagging latency was 0.78 seconds at the local road speed limits. For all cases, the mean time lag was 77.8% of the mean update interval (approximately 1 second) for the GPS receiver. This result indicates that if the tag delays are normally distributed, then the embedded system of the GPS receiver biased the delay almost two standard deviations away from the midpoint of the update interval. Hence,

for a normal distribution of the tag latency, equation (5) indicates that the average spread in geospatial position tagging  $\sigma_{\text{dtag}}$  was 1.03 meters across all traversals.

When not disclosed by the manufacturer, practitioners could conduct similar experiments to determine the latency of the actual embedded system used in a connected vehicle. Ideally, emerging standards for connected vehicle data exchanges should include the expected value of this delay. Otherwise, practitioners could simply apply a blind approximation by setting  $\bar{\tau}_{\text{lag}}$  equal to the median of the selected GPS update interval. It became evident in this study that a blind approximation will still achieve substantial improvements in localization accuracy relative to no bias removal. For these experiments, a 50% (versus the full 77.8%) bias time removal would be equivalent to an average correction of 2.83 meters (versus 4.41 meters).

### ***3.5 Estimate of GPS precision***

The chi-squared tests promote a high level of confidence that the position tags distribute normally. Therefore, the residual variance is a good estimator for the actual GPS precision  $\sigma_{\text{dGPS}}$  realized for comparison with the expected norms. The residual variance is

$$\sigma_{\text{dGPS}}^2 = \sigma_v^2 - \left[ \sigma_{\text{eDSP}}^2 \left( 1 + \frac{1}{N_{\text{FIR}}} \right) + \sigma_{\text{eb}}^2 + \sigma_{\text{es}}^2 + \sigma_{\text{dtag}}^2 \right]. \quad (15)$$

## Accuracy enhancement of roadway anomaly localization using connected vehicles

Its average interval of uncertainty was 3.056 meters across all traversals. This result is consistent with the levels that operators expect for conventional GPS receivers (USDHS 1996).

### ***3.6 Relative contribution of errors***

Table 3 summarizes the relative contribution of the individual error factors in the overall spread of peak position tags for the case studies. For normally distributed tag delays, variances in the tagging latency contributed 16.3% on average to the overall uncertainty while errors in GPS trilateration contributed 83.7%. The spread of relative contributions for both the tagging latency and the GPS position uncertainty was approximately 12% across all case studies. The insignificance of delay spreads from vehicle suspension filtering is consistent with the results from previous research (Bridgelall 2014b). Although the dominance from GPS spread is not surprising, the potential for embedded system latencies to contribute relatively large errors points to the importance of being able to estimate and remove them. In particular, the spread in tag latency increases in direct proportion to vehicle speed (equation 4) and, therefore, has the potential to dominate the overall error at high traversal speeds.

## **4 Summary and conclusions**

Affordable and scalable methods of measuring localized roughness enable improved efficiencies and effectiveness in the practice of roadway asset management. The expense and limitations of existing approaches has motivated agencies to evaluate evolving methods that leverage connected vehicles. However, the potential adopters know very little about the practical performance of connected vehicle technologies that use accelerometers and conventional GPS receivers to characterize roadway roughness.



We reviewed the most recent approach called the road impact factor (RIF) transform that precludes the need for calibration with individual vehicle parameters. The RIF-transform leverages the central limit theorem by combining large data volumes from connected vehicles to produce a consistent measure of roughness. We then developed error models to estimate and remove position bias from the RIF-indices. With six field studies, we characterized the error distribution and compared the relative magnitude of each component. The dominant error from GPS position estimation (84%) and the relative insignificance of errors from suspension system filtering was not surprising. However, we did not anticipate the potential for position tagging errors from embedded systems latency (16% at local road speed limits) to dominate at highway speeds. The statistical distribution of position tagging errors agreed with the expected classic distributions, thereby validating the model. Furthermore, strong agreement with the normal distribution indicates that the precision of the approach will continue to improve with the ever-increasing volume of connected vehicles.

This model and method of error characterization is broadly applicable to mobile applications that tag sensor data with position coordinates derived from GPS receivers. Future work will examine the utility of the error models for applications involving the use of unmanned aircraft systems (UAS) to monitor a variety of other transportation infrastructure issues.

## References

- Ahlin, K. and Granlund, N.J., 2002. Relating road roughness and vehicle speeds to human whole body vibration and exposure limits. *Int. J. Pavement Eng.*, 3 (4), 207-216.
- Ayenu-Prah, A.Y. and Attoh-Okine, N.O., 2009. Comparative study of Hilbert–Huang transform, Fourier transform and wavelet transform in pavement profile analysis. *Vehicle Syst. Dyn.*, 47 (4), 437-456.

## Accuracy enhancement of roadway anomaly localization using connected vehicles

- Blum N.C., 2015. *System Identification of vehicle dynamics and road conditions using wireless sensors*. College Park: University of Maryland.
- Bridgelall R., 2014a. A participatory sensing approach to characterize ride quality. In: J.P. Lynch, K-W. Wang, and H. Sohn, ed. *Proc. SPIE 9061, Sensors and smart structures technologies for civil, mechanical, and aerospace systems*, 8 March 2014, San Diego. Bellingham: International Society for Optics and Photonics (SPIE).
- Bridgelall R., 2014b. Connected vehicle approach for pavement roughness evaluation. *J. Infrastruct. Syst.*, 20 (1), 1-6.
- Bridgelall R., 2015. *Pavement Performance Evaluation Using Connected Vehicles*. Fargo: North Dakota State University.
- Bridgelall R., 2015. Precision bounds of pavement distress localization with connected vehicle sensors. *J. Infrastruct. Syst.*, 21 (1), 1-7.
- Byrne, M., Parry, T., Isola, R., and Dawson, A., 2013. Identifying road defect information from smartphones. *Road Transp. Res.*, 22 (1), 39-50.
- Chen, D., and Dye, M., 2014. Evaluation of initial IRI values as acceptance criteria for flexible pavements. *Int. J. Eng. Res. Innov.*, 6 (2), 55-61.
- 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.
- Gade, K., 2010. A non-singular horizontal position representation. *J. Navigation*, 63 (3), 365-417.
- Islam, S., Buttler, W.G., Aldunate, R.G., and Vavrik, W.R., 2014. Measurement of pavement roughness using Android-based smartphone application. *Transport Res. Rec.*, (2457), 30-38.
- Jazar, R.N., 2008. *Vehicle dynamics: theory and applications*. New York: Springer.
- Karamihas, S.M., and Senn, K., 2012. *Curl and warp analysis of the LTPP SPS-2 site in Arizona*. Washington, D.C.: Federal Highway Administration.
- NCHRP, 2013. *Measuring, characterizing, and reporting pavement roughness of low-speed and urban roads*. Washington, D.C.: National Cooperative Highway Research Program.
- Nomura, T. and Shiraishi, Y., 2015. A method for estimating road surface conditions with a smartphone. *Int. J. Informatics Society*, 7 (1), 29-36.
- North Dakota State Government, 2015. *North Dakota Geographic Information Systems* [online]. Available from: <http://www.nd.gov/gis> [Accessed 29 September 2015].
- Papagiannakis, A.T., 1997. *The need for a new pavement roughness index; RIDE*. Washington, D.C.: Society of Automotive Engineers International.
- Papoulis, A., 1991. *Probability, random variables, and stochastic processes*. New York: McGraw-Hill.
- Rajamohan, D., Gannu, B., and Rajan, K.S., 2015. MAARGHA: a prototype system for road condition and surface type estimation by fusing multi-sensor data. *ISPRS Int. J. Geo-Information*, 4 (3), 1225-1245.
- Robbins, M.M. and Tran, N., 2015. *Literature review: the effect of pavement roughness on vehicle operating costs*. Report 15-02. Auburn: National Center for Asphalt Technology.

## Accuracy enhancement of roadway anomaly localization using connected vehicles

- Steyn, W.J. vdM., Nokes, B., du Plessis, L., Agacer, R., Burmas, N., and Popescu, L., 2015. Evaluation of the effect of rural road condition on agricultural produce transportation. *Transport Res. Rec.*, (2473), 33-41.
- USDHS, 1996. *NAVSTAR GPS user equipment introduction*. Washington, D.C.: U.S. Department of Homeland Security.
- Yi, T., Li, H-N., and Gu, M., 2010. Recent research and applications of GPS based technology for bridge health monitoring. *Sci. China Technol. Sc.*, 53 (10), 2597-2610.
- Yi, T., Li, H-N., and Gu, M., 2013. Recent research and applications of GPS-based monitoring technology for high-rise structures. *Struct. Control Hlth.*, 20 (5), 649-670.
- Young, H.D., Freedman, R.A., and Ford, A.L., 2011. *University physics with Modern physics*. 13<sup>th</sup> ed. Boston: Addison-Wesley.
- Ahlin, Kjella, and N.O. Johan Granlund. "Relating Road Roughness and Vehicle Speeds to Human Whole Body Vibration and Exposure Limits." *International Journal of Pavement Engineering* (Taylor & Francis) 3, no. 4 (2002): 207-216.
- Ayenu-Prah, A. Y., and N. O. Attoh-Okine. "Comparative study of Hilbert–Huang transform, Fourier transform and wavelet transform in pavement profile analysis." *Vehicle System Dynamics* (Taylor & Francis) 47, no. 4 (2009): 437-456.
- Blum, Noah Corey. *System Identification of Vehicle Dynamics and Road Conditions Using Wireless Sensors*. Doctoral Dissertation, Civil and Environmental Engineering, University of Maryland, College Park: University of Maryland, 2015, 232.
- Bridgelall, Raj. "A participatory sensing approach to characterize ride quality." *Proceedings of SPIE Volume 9061, Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems*. Bellingham, Washington: International Society for Optics and Photonics (SPIE), 2014.
- Bridgelall, Raj. "Connected Vehicle Approach for Pavement Roughness Evaluation." *Journal of Infrastructure Systems* (American Society of Civil Engineers) 20, no. 1 (April 2014): 04013001.1-04013001.6.
- Byrne, Matthew, Tony Parry, Riccardo Isola, and Andrew Dawson. "Identifying Road Defect Information from Smartphones." *Road & Transport Research: A Journal of Australian and New Zealand Research and Practice* 22, no. 1 (2013): 39-50.
- Chen, Don, and Michael Dye. "Evaluation of Initial IRI Values as Acceptance Criteria for Flexible Pavements." *International Journal of Engineering Research & Innovation* (International Association of Journals and Conferences) 6, no. 2 (2014): 55-61.
- Du, Yuchuan, Chenglong Liu, Difei Wu, and Shengchuan Jiang. "Measurement of International Roughness Index by Using Z-Axis Accelerometers and GPS." *Mathematical Problems in Engineering* (Hindawi Publishing Corporation) 2014 (June 2014): 1-10.
- Gade, Kenneth. "A Non-singular Horizontal Position Representation." *The Journal of Navigation* (The Royal Institute of Navigation) 63, no. 3 (2010): 365-417.
- Islam, Shahidul, William G. Buttlar, Roberto G. Aldunate, and William R. Vavrik. "Measurement of Pavement Roughness Using Android-Based Smartphone Application." *Transportation Research Record: Journal of the Transportation Research Board* 2457, no. 1 (2014): 30-38.
- Jazar, Reza N. *Vehicle Dynamics: Theory and Applications*. New York: Springer, 2008.

## Accuracy enhancement of roadway anomaly localization using connected vehicles

- Karamihas, Steven M., and Kevin Senn. *Curl and Warp Analysis of The LTPP SPS-2 Site in Arizona*. FHWA-HRT-12-068, Washington, D.C.: Federal Highway Administration (FHWA), 2012, 114.
- NCHRP. "Measuring, Characterizing, and Reporting Pavement Roughness of Low-Speed and Urban Roads." *Research in Progress*. National Cooperative Highway Research Program (NCHRP). Vols. 10-93. Edited by Amir N. Hanna. Washington, D.C.: Transportation Research Board of the National Academies, October 1, 2015.
- Nomura, Tomohiro, and Yoh Shiraishi. "A Method for Estimating Road Surface Conditions with a Smartphone." *International Journal of Informatics Society* 7, no. 1 (2015): 29-36.
- North Dakota State Government. *North Dakota Geographic Information Systems*. September 1, 2015. [www.nd.gov/gis](http://www.nd.gov/gis) (accessed September 18, 2015).
- Papagiannakis, A. Tom. "The Need for a New Pavement Roughness Index; RIDE." *International Truck & Bus Meeting & Exposition*. Washington, D.C.: Society of Automotive Engineers International, 1997.
- Papoulis, Athanasios. *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1991.
- Rajamohan, Deepak, Bhavana Gannu, and Krishnan Sundara Rajan. "MAARGHA: A Prototype System for Road Condition and Surface Type Estimation by Fusing Multi-Sensor Data." *ISPRS International Journal of Geo-Information* (Multidisciplinary Digital Publishing Institute) 4, no. 3 (2015): 1225-1245.
- Robbins, Mary M., and Nam Tran. "Literature Review: The Effect of Pavement Roughness On Vehicle Operating Costs." *National Center for Asphalt Technology Report 15-02* (National Center for Asphalt Technology Report) 15-02 (2015): 22.
- Steyn, W. J. vdM., B Nokes, L. du Plessis, R. Agacer, N. Burmas, and L. Popescu. "Evaluation of the effect of rural road condition on agricultural produce transportation." *Transportation Research Record: Journal of the Transportation Research Board* 2473 (2015): 33-41.
- USDHS. *NAVSTAR GPS User Equipment Introduction*. Technical Guide, United States Coast Guard, Navigation Center, Washington, D.C.: U.S. Department of Homeland Security, 1996, 215.
- Yi, TingHua, Hong-Nan Li, and Ming Gu. "Recent research and applications of GPS based technology for bridge health monitoring." *Science China Technological Sciences* 53, no. 10 (2010): 2597-2610.
- Yi, TingHua, Hong-Nan Li, and Ming Gu. "Recent research and applications of GPS-based monitoring technology for high-rise structures." *Structural Control and Health Monitoring* 20, no. 5 (2013): 649-670.
- Young, Hugh D., Roger A. Freedman, and A. Lewis Ford. *University Physics with Modern Physics*. 13th. Addison-Wesley, 2011.

Table 1. Parameters derived from the data of the six case studies.

Parameters	Park Bump			Airport Road Joint		Tracks
	2.5 m s <sup>-1</sup>	5 m s <sup>-1</sup>	7 m s <sup>-1</sup>	EB	WB	NB
Batch mean speed, $\bar{v}$ (m s <sup>-1</sup> )	2.552	4.983	7.187	6.769	6.715	6.685
Batch mean spread, $\bar{\sigma}_v$ (m s <sup>-1</sup> )	0.204	0.570	0.428	0.428	0.251	0.416
Suspension delay, $\bar{\epsilon}_b$ (m)	0.038	0.075	0.108	0.102	0.101	0.100
Suspension delay, $\sigma_{\epsilon_b}$ (m)	0.003	0.009	0.006	0.006	0.004	0.006
DSP latency mean, $\bar{\epsilon}_{\text{DSP}}$ (m)	0.246	0.481	0.693	0.654	0.648	0.649
DSP latency spread, $\sigma_{\epsilon_{\text{DSP}}}$ (m)	0.020	0.055	0.041	0.041	0.024	0.040
Interpolation, $\bar{\epsilon}_i$ (m)	0.014	0.027	0.038	0.036	0.036	0.036
Interpolation, $\sigma_{\epsilon_i}$ (m)	0.001	0.003	0.002	0.002	0.001	0.002
Sensor position, $\bar{\epsilon}_s$ (m)	-0.920	-0.920	-0.920	-0.710	-0.710	-0.710
Inertial sample rate, $f_A$ (Hz)	93.227	93.277	93.341	93.169	93.198	92.669
Inertial rate spread, $\sigma_{f_A}$ (Hz)	0.097	0.070	0.101	0.073	0.103	0.077
GPS update rate, $T_{\mu\text{GPS}}$ (s)	1.009	1.004	0.993	0.986	0.988	1.015
GPS update spread, $\sigma_{\tau_{\text{lag}}}$ (s)	0.168	0.167	0.166	0.164	0.165	0.169
GPS tag spread, $\sigma_{\text{dlag}}$ (m)	0.460	1.011	1.212	1.162	1.109	1.230
First peak offset, $\tilde{v}$ (m)	-2.712	-5.337	-3.944	-5.227	-2.334	-7.708
GPS tag lag, $\bar{\epsilon}_{\text{lag}}$ (m)	-2.091	-4.999	-3.864	-5.309	-2.409	-7.784
GPS tag lag, $\bar{\tau}_{\text{lag}}$ (s)	-0.819	-1.003	-0.538	-0.784	-0.359	-1.164
First peak spread, $\sigma_v$ (m)	5.013	4.361	3.349	2.637	2.010	2.294
GPS spread, $\sigma_{\text{dGPS}}$ (m)	4.992	4.242	3.122	2.367	1.676	1.936

Table 2. Best fit parameters for distributions of the first peak position.

Parameters	Park Bump			Airport Road Joint		Tracks
	2.5 m s <sup>-1</sup>	5 m s <sup>-1</sup>	7 m s <sup>-1</sup>	EB	WB	NB
<b>Gaussian</b>						
<i>df</i>	2	2	2	2	2	2
$\chi^2, \alpha = 5\%$	5.991	5.991	5.991	5.991	5.991	5.991
$\chi^2$ Data	1.511	2.152	2.075	0.873	1.004	0.347
Significance $\alpha$ (%)	46.986	34.103	35.431	64.628	60.527	84.062
Amplitude	141.611	122.735	107.874	303.516	69.35	67.407
Mean	28.173	25.67	25.784	18.684	27.345	25.169
Standard Dev.	6.467	4.665	3.14	5.238	2.838	2.345
<b>Student-t</b>						
<i>df</i>	2	2	2	2	2	2
$\chi^2, \alpha = 5\%$	5.991	5.991	5.991	5.991	5.991	5.991
$\chi^2$ Data	3.071	3.574	1.534	1.064	2.626	2.135
Significance $\alpha$ (%)	21.534	16.744	46.438	58.745	26.901	34.391
Amplitude	166.519	143.144	120.988	130.934	81.315	77.353
Mean	28.59	26.102	25.542	23.254	27.262	25.191
Standard Dev.	6.415	4.658	2.839	2.661	2.835	2.255
<b>Logistic</b>						
<i>df</i>	2	2	2	2	2	2
$\chi^2, \alpha = 5\%$	5.991	5.991	5.991	5.991	5.991	5.991
$\chi^2$ Data	1.73	2.216	0.576	0.561	1.317	0.701
Significance $\alpha$ (%)	42.111	33.026	74.974	75.53	51.763	70.441
Amplitude	148.791	128.291	111.262	171.415	72.27	69.95
Mean	28.373	25.841	25.703	21.741	27.305	25.164
Scale	4.173	3.016	1.968	2.405	1.813	1.488

Table 3. Summary of relative error contributions.

Parameters	Park Bump			Airport Road Joint		Tracks
	2.5 m s <sup>-1</sup>	5 m s <sup>-1</sup>	7 m s <sup>-1</sup>	EB	WB	NB
First peak spread, $\sigma_v$ (m)	5.013	4.361	3.349	2.637	2.010	2.294
Ratio, $\sigma_{eb}^2 / \sigma_v^2$	0.000%	0.000%	0.000%	0.001%	0.000%	0.001%
Ratio, $\sigma_{eDSP}^2 / \sigma_v^2$	0.002%	0.016%	0.015%	0.024%	0.014%	0.030%
Ratio, $\sigma_{dlag}^2 / \sigma_v^2$	0.842%	5.374%	13.097%	19.417%	30.442%	28.749%
Ratio, $\sigma_{dGPS}^2 / \sigma_v^2$	99.164%	94.617%	86.903%	80.571%	69.527%	71.224%



Figure 1. Isolated anomalies traversed for the case study measurements.

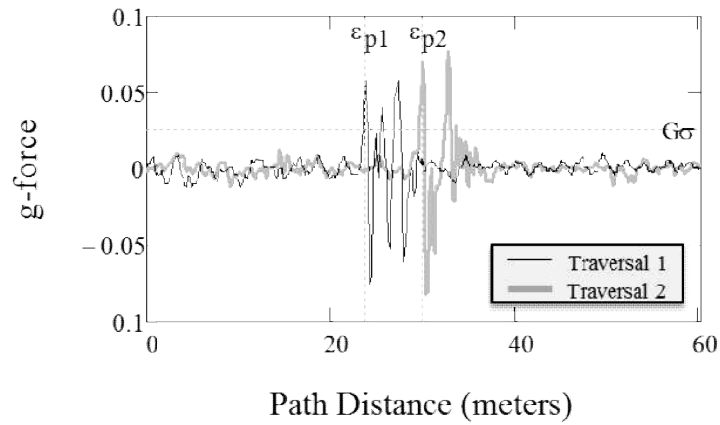


Figure 2. Accelerometer signal for two traversals of the same anomaly.

Accuracy enhancement of roadway anomaly localization using connected vehicles

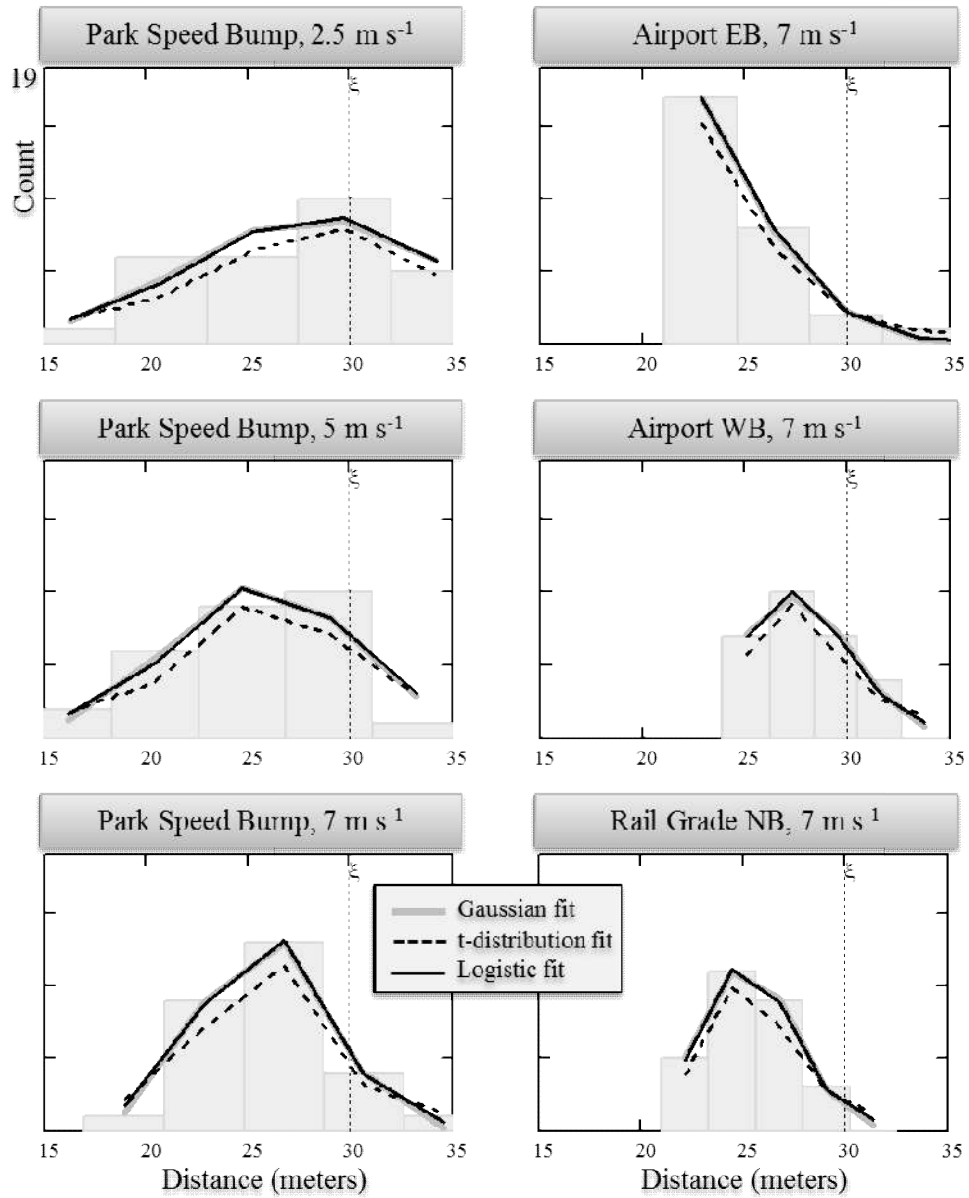


Figure 3. Distribution of peak position tags from the case study.