Reply to the reviewer's evaluation report

Title of the dissertation: Automotive and mechatronics applications of IMU sensors

Author: Roland Nagy

Reviewer: Dr. Zsolt János Viharos

Dear Reviewer,

I would like to thank you for the thorough, detailed, and constructive evaluation of my dissertation, for the supportive opinion, and for the acceptance of all three of my theses. Below I respond to the weaknesses, open questions, and comments on the theses, in the same order in which they appear in the review. Since I am unable to make any retroactive modifications to the submitted, printed thesis, every clarification, additional reference, and wording refinement is recorded in the current reply.

Yours sincerely,
Roland Nagy

Replies to the points raised in the review

1. Feature Selection techniques

The reviewer is right that the classical Feature Selection techniques do not appear in the thesis as a standalone stage. Instead, the work used post-hoc importance metrics:

- At the maneuver-detection stage, Principal Component Analysis compressed 95.2% of the variable variance into two principal components (Table 11, page 89).
- At the road-quality classification stage, Random Forest native feature importance and SHAP value analysis (Figs. 33–34, pages 102–104).

These evaluate a fixed, manually engineered feature set rather than generating a large pool of candidates and pruning it systematically. With this approach, I aimed to build upon and make use of existing domain knowledge, as the range and impact of influencing factors in this field are relatively well documented thanks to the wealth of literature available. However, I would like to thank you for drawing attention to the inclusion of classical Feature Selection methods, as this could improve the results and open up a new, valid approach in further work.

2. Statements not sufficiently supported by references

Thank you for drawing my attention to the lack of references. In the second half of the dissertation, several statements should have been backed by an explicit literature reference. In order to supplement these statements, I assign the following references to the points flagged in the review:

- "While some studies do not specify their labeling methodology for real-world experiments, and others rely entirely on computer simulations:"
    - The works below are already cited in §2.3 of the thesis, and I designate them here as the concrete sources for the mentioned statement.
    - Praznowski K., Mamala J., Smieja M., et al. Assessment of the road surface condition with longitudinal acceleration signal of the car body. MDPI Sensors, 20(21):1–19, 2020. (cited in the thesis as ref. [37])
    - Mikko P., Oleksiy M., Fengyu C., et al. Distributed road surface condition monitoring using mobile phones. Springer - Verlag, (6905):64–78, 2011. (cited in the thesis as ref. [53])
    - Seraj F., van der Zwaag B. J., Dilo A., et al. Roads: A road pavement monitoring system for anomaly detection using smart phones. Big Data Analytics in the Social and Ubiquitous Context, Springer, pages 128–146, 2016. (cited in the thesis as ref. [58])
- "Furthermore, affordability is critical for widespread deployment, especially across diverse vehicle fleets:"
    - Sattar S., Li S., Chapman M. A. Road Surface Monitoring Using Smartphone Sensors: A Review. Sensors, 18(11):3845, 2018.
    - Bhoraskar R., Vankadhara N., Raman B., Kulkarni P. Wolverine: Traffic and road condition estimation using smartphone sensors. COMSNETS, 2012.

- "x = 0.2 m was selected, with a trade-off between sufficient detail for localized terrain analysis and oversampling:"
    - Ward C. C., Iagnemma K. Speed-independent vibration-based terrain classification for passenger vehicles. Vehicle System Dynamics, 47(9):1095–1113, 2008. (cited in the thesis as ref. [45])
- "Since road vehicles operate within a limited dynamic range, previous studies..."
    - ISO 8608. Mechanical vibration - Road surface profiles.
    - Cantisani G., Loprencipe G. Road Roughness and Whole Body Vibration: Evaluation Tools and Comfort Limits. Journal of Transportation Engineering, 136(9):818–826, 2010.
- "median ... std ... the two core indicators of surface uniformity and variability:"
    - Du Y., Liu C., Wu D., Jiang S. Measurement of International Roughness Index by Using Z-Axis Accelerometers and GPS. Mathematical Problems in Engineering, 2014:928980, 2014.
    - Mednis A., Strazdins G., Zviedris R., Kanonirs G., Selavo L. Real Time Pothole Detection Using Android Smartphones with Accelerometers. DCOSS, 2011.
- "This methodology has already proven successful across various domains:"
    - Xingjiao W., Luwei X., Yixuan S., et al. A survey of human-in-the-loop for machine learning. Future Generation Computer Systems, 135(107299):364–381, 2022. (cited in the thesis as ref. [87])
    - Settles B. Active Learning Literature Survey. 2009.
    - Budd S., Robinson E. C., Kainz B. A Survey on Active Learning and Human-in-the-Loop Deep Learning for Medical Image Analysis. Medical Image Analysis, 71:102062, 2021.
- Reference for BIC and AIC formulas:
    - Schwarz G. Estimating the Dimension of a Model. Annals of Statistics, 6(2):461–464, 1978. (BIC)
    - Akaike H. A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control, 19(6):716–723, 1974. (AIC)
    - McLachlan G., Peel D. Finite Mixture Models. Wiley, 2000.
- Maneuver thresholds (7 / 10 / 5 m/s²):
    - As described in the paper, the exact threshold values were mainly selected based on our research and empirical observations from real-world measurements. This allowed us to create the most accurate labeling in our case, which best suited the available test vehicle. For further clarification, I am also attaching articles below that deal with similar issues and were reviewed during my work.
    - Eboli L., Mazzulla G., Pungillo G. Combining speed and acceleration to define car users' safe or unsafe driving behaviour. Transportation Research Part C, 68:113–125, 2017.
    - Vaiana R., Gallelli V., Iuele T., Minani V. Driving Behavior and Traffic Safety: An Acceleration-Based Safety Evaluation Procedure for Smartphones. Modern Applied Science, 8(1):88–96, 2014.

I agree with the reviewer that placing these citations inline would have improved the transparency of the work.

3. Sensor location selection vs. optimal sensor selection approach

Thank you for bringing attention to the apparent tension between §3.1, where the first research question may suggest an optimal sensor selection approach, and §3.3, which states that the selection of sensor locations was driven by practical engineering constraints rather than an explicit optimization procedure. According to my intentions, the two statements are complementary rather than contradictory:

- §3.1 provides an information-content question, which asks if the suspension-mounted measurement carries useful additional information compared to the cabin measurement, for the subsequent classification task? This is the scientific question to which Thesis #1 provides an answer.
- §3.3 describes the implementation decisions, as the IMU is placed close to the cabin centre of gravity, plus piezoelectric accelerometers on the rear control arms, chosen to provide mechanically stable, flat attachment surfaces and direct exposure to road excitation. This was an engineering decision held fixed across the experiments.
- The phrase "rather than an explicit optimization procedure" refers specifically to the fact that the exact mounting point on the suspension itself (contról arm vs. wheel hub vs. damper top mount) was not the subject of a formal numerical optimization.

I acknowledge that the wording in §3.3 invites the misreading the reviewer noted, and I thank you for the opportunity to clarify it.

4. Infinity-norm normalization

The mentioned concern is valid, as infinity-norm normalization is inherently sensitive to a single outlier value. I would like to thank you for pointing out this potential vulnerability. Below, I detail the rationale behind my choice, which is primarily rooted in the requirements of vibration diagnostics and multivariate signal processing. Despite the valid concern, while infinity-norm normalization works properly under controlled conditions, a different approach should be chosen for a system operating in a real-world environment.

The movement of the car at any given moment is a 3D vector. When hitting a pothole, the vertical acceleration suddenly becomes much larger than the lateral acceleration. The ratios between these axes carry the physical sign of what happened to the wheel.

- Infinity-norm normalization (Row-wise/Sample-wise): This method finds the maximum value in each specific time sample and divides all other axes by this maximum at that exact moment. If the vertical value is divided by 5 and the lateral value is also divided by 5, the ratio between them is perfectly preserved. The physical meaning remains intact.
- Per-axis Z-score normalization (Column-wise/Axis-wise): Z-score evaluates the entire dataset axis by axis, independently. Since a car's suspension experiences orders of magnitude larger impacts vertically due to the road than laterally, the variance of the two axes will be completely different. The Z-score will divide the vertical axis by a large number and the lateral axis by a very small number. This artificially inflates the lateral movements and dampens the vertical ones. The original physical ratio between the

axes is destroyed, and the model will no longer be able to distinguish whether an event was a sharp vertical pothole impact or a lateral cornering maneuver.

Furthermore, to avoid the sensitivity of the infinity-norm, the following aspects were ensured during the development:

- The characteristics and the source of the vibrations are always the same under normal measurement conditions, the impacts caused by potholes belong to the normal, expected operation of the system.
- The range is limited, and the signal can only move within this range due to its "physical" constraints stemming from the measurement limits of the sensor, which has a minimum and a maximum value. Despite this, a significantly large signal value deviating from the rest can indeed occur.
- In the case of the signals measured and processed during the research, it was strictly checked that the signal follows a normal distribution and that significantly large outliers do not occur at the edges of the distribution.

However, as mentioned above, if the system is developed in the future, it is definitely recommended to choose a different and more robust solution that simultaneously preserves the ratios, but ensures that outliers do not cause an issue:

- The best solution would be not to normalize the axes independently.
- Instead, L2 normalization could be applied, which still operates as a row-wise normalization, but the basis of the division is the length of the vector rather than the raw maximum.
- This preserves the ratios similarly to the infinity-norm, but it is mathematically more stable against spikes.

5. Table 5: SMOTE algorithm, where the variable 'x' should be 'x(min)'

I thank you for the observation. In step 4 of Table 5 (page 61) and in Eq. 11, the variable 'x' should be read as 'x(min)', the minority-class instance selected at the start of the iteration.

6. "After comparative analysis..." - the comparisons are not described

The mentioned sentence in §4.5.3 (page 74) is a forward reference. The actual comparison appears in §5.5, Table 16 (page 96), where six classifiers are compared.

Random Forest yielded the best performance both in overall accuracy and in macro-averaged F1, followed by Gradient Boosting and SVM. I accept the reviewer's observation that the §4.5.3 sentence would have been clearer with an explicit forward reference to §5.5, which would have made the reference clearer.

7. Determination of the values of w(t) in the Gaussian-weighted correction mask

The correction mask w(t) is determined through a practical process based on the real-world timing of driving events. First, the exact center of each aggressive maneuver is located in the

signal, and a sample window is placed around it. The value of w is set empirically to the typical distance of an aggressive event, equivalent to 2 driving meters. To prevent any sharp or artificial jumps in the data, a Gaussian filter is applied to smooth the boundaries of this window. Finally, this smoothed shape is inverted to create the weighting function. This ensures that the distorted signal on the affected axes is fully suppressed to zero at the exact peak of the aggressive maneuver, while smoothly returning to one as the event ends, leaving all normal driving sections completely untouched.

8. Argumentation supporting the lowest-BIC GMM model selection

Thank you for drawing attention to the not entirely accurate description. The mentioned sentence in §4.6.1 states the general principle of BIC-based model selection, and it does not describe the concrete decision logic that led to K = 4. The actual decision is given later in §5.4 (pages 91–92), in the elbow plot of Fig. 28, and in Table 15, and is based on the joint evaluation of four criteria:

- Plateau behaviour of BIC and AIC: Both metrics decrease as K increases, but exhibit a plateau-like behaviour beyond K = 5 – 6, where the marginal improvement is small. Statistically, therefore, there is no strong argument for going to a higher K.
- Silhouette score: Maximal between K = 3 and K = 5, declining sharply for K ≥ 6, therefore, cluster separation is best in this band.
- Industry practice: Road management organisations typically use 3 - 7 quality classes (cited in §2.1.4 as ref. [17]), which sets a practical upper bound on usable K.
- MIL manual validation: At K = 4, the four clusters mapped onto easy-to-interpret Poor / Medium / Good / Perfect; at K = 3, the Medium and Good classes merged; at K = 5, the Perfect class split into two physically non-meaningful clusters, according to my experience from the experimentation.

These four criteria jointly identified K = 4 as the common optimum. The selection was therefore not based on the lowest BIC value alone. I acknowledge that the wording in §4.6.1 could have been more precise on this point.

9. Method for the removal of sensor errors in the preprocessing pipeline

The phrase in §5.2 is a summary expression that refers to the operations described in detail in §4.2 (pages 65–67). The three concrete steps are:

- Bias removal: Subtraction of a low-degree polynomial fitted to the accelerometer signals by least-squares regression. This eliminates the static offset due to calibration drift and environmental effects.
- Drift removal: The same polynomial detrending is applied to the gyroscope signals, which typically exhibit low-frequency integration drift.
- Spatial resampling: The signal was then resampled based on spatial distance, thereby reducing the negative influence of varying vehicle speed.

These three steps together are under the umbrella term mentioned in §5.2.

10. Reason for the higher classification error of acceleration and braking

I would like to do a small clarification first, as the figure showing per-class maneuver errors is Fig. 26 (page 90) in the printed thesis, however, the review refers to it as "Fig. 25". Nevertheless, this does not affect the substance of the question.

Three factors explain the higher classification error for the longitudinal maneuvers:

- The main reason stems from differences in signal physics, as aggressive longitudinal events generally blend better into normal driving conditions, often have a less characteristic signal shape, and shorter duration than lateral cornering events.
- A slight class overlap also appears, as a moderate acceleration or braking event on a rough road can produce a longitudinal signal that closely resembles a strong, road-induced vertical excitation. The SVM decision boundary between "aggressive longitudinal" and "normal driving on rough road" is therefore fuzzier than between "aggressive turn" and "normal straight driving", where the signal structure differs much more sharply.
- The suspension dependence was also noticeable during the measurements. The Nissan Leaf hatchback has a softer rear suspension (torsion beam with separated spring–damper elements). Softer pitch behaviour during braking compresses the longitudinal-acceleration peaks, widening the apparent overlap with normal driving. This effect is directly visible in the per-vehicle bars of Fig. 26, as the hatchback's acceleration- and brake-error is meaningfully higher than the BMW's, while the cornering errors of the two vehicles are comparable.

These three factors together explain the wider acceleration- and brake-error bands, and they also pinpoint the possible directions of refinement. Additional discriminative features in the longitudinal domain would probably improve the model accuracy.

11. Meaning of "These observations and experiences are supported by my visual observations"

The sentence at the end of §5.7 (page 107) refers to the manual MIL-validation step described in §4.5.1 (Fig. 18, page 70) and in the "Role of Human Validation" subsection of §4.6.1 (page 77). The concrete procedure was:

- During every measurement, I either drove the test vehicle or carried out the measurement seated in the front passenger seat.
- In both cases, I tried to keep manual notes on the condition of the road, especially highlighting and recording any longer sections with significantly different quality.
- During the evaluation, I also used Google Maps Street View if recent footage of the given route was available.
- During the measurements, I tried to choose sections that I personally drive on frequently, so I basically had a preliminary and deeper knowledge of the given road section.

- Representative road segments classified by the algorithm were subsequently validated against these notes and footage; in ambiguous cases, a repeated on-site visit was carried out.

Visual observations, therefore, refer to a more systematic validation of the measured and classified road segments, performed as part of the MIL framework.

Theses

1. Thesis #1 b) - first bullet

I accept the reviewer's reasoning. The development and construction of the data acquisition platform is engineering work, which, on its own, does not constitute a scientific thesis result. The bullet was included among the theses because the comparative investigation of the information content of cabin and suspension sensors could not have been carried out without the described measurement system. The apparatus is therefore the precondition for the scientific question, not the answer to it.

2. Thesis #1 b) - second bullet

I likewise accept the reviewer's observation. The description of the measurements is a description of experimental work, and is not in itself a scientific claim.

Closing remarks

I would once again like to thank the reviewer for the detailed, thorough, and forward-looking evaluation, and for the constructive open questions posed for the defense. The weaknesses raised, the open questions, and the comments on the structure of Thesis #1 each provide substantive feedback.

Veszprém, 19. June 2026



Roland Nagy