International Journal of Transportation Science and Technology xxx (xxxx) xxx

Contents lists available at ScienceDirect

International Journal of Transportation Science and Technology

journal homepage: www.elsevier.com/locate/ijtst

Research Paper

Genetic algorithm optimization and sensitivity analysis of bicycle characteristics for vibration monitoring ☆

Omid Ghaderi [a,*], Salvatore Cafiso [a], Riccardo Caponetto [b], Michele Mangiameli [a], Giuseppina Pappalardo [a]

[a] Department of Civil Engineering & Architecture, University of Catania, Catania, Italy
[b] Department of Engineering, University of Messina, Messina, Italy

ARTICLE INFO

Article history:
Received 20 March 2024
Received in revised form 5 March 2025
Accepted 27 May 2025
Available online xxxx

Keywords:
Road pavement
Bicycle
Smartphone
Vibration
Genetic algorithm

ABSTRACT

It is essential for sustainable cities to promote cycling as one of the modes of urban transport. One significant obstacle to this is the discomfort and safety risks caused by direct impacts of inadequate pavement maintenance. The objective of this study is to investigate the impact of the dynamic characteristics of bicycles on vibration data collected for the purpose of monitoring pavement conditions. To this end, experimental tests and numerical simulations have been employed to gain insight into the influence of these characteristics on vibration measurements.

In the experimental setup, a smartphone-equipped bicycle collected vertical accelerations at various speeds. An advanced laser profiler and Laser Crack Measurement System (LCMS) modeled the pavement surface along the travel path. The data were synchronized and compared with numerical simulations to calibrate the bicycle model using a genetic algorithm, which then calculated changes in vibration output due to variability in the bicycle's primary dynamic characteristics.

This research establishes a comprehensive framework for modeling bicycle dynamics, enhancing the analysis of vibration data across different bicycles. A sensitivity analysis was conducted to explore how variations in bicycle characteristics affect the accuracy of vibration responses. The results indicate that positional variability, speed, and localized high-severity distress along the bicycle path significantly influence the outcome. These findings underscore that while the intrinsic characteristics of bicycles are crucial, the correct selection of bicycle categories and meticulous data cleaning are essential for collecting consistent vibration data through multiple runs, facilitating effective crowdsourcing for pavement condition monitoring.

© 2025 Tongji University and Tongji University Press. Publishing Services by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Peer review under the responsibility of Tongji University and Tongji University Press

* Corresponding author.
E-mail addresses: omid.ghaderi@phd.unict.it (O. Ghaderi), dcafiso@unict.it (S. Cafiso), riccardo.caponetto@unime.it (R. Caponetto), michele.mangiameli@unict.it (M. Mangiameli), giuseppina.pappalardo1@unict.it (G. Pappalardo).

https://doi.org/10.1016/j.ijtst.2025.05.006
2046-0430/© 2025 Tongji University and Tongji University Press. Publishing Services by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Please cite this article as: O. Ghaderi, S. Cafiso, R. Caponetto et al., Genetic algorithm optimization and sensitivity analysis of bicycle characteristics for vibration monitoring, International Journal of Transportation Science and Technology, https://doi.org/10.1016/j.ijtst.2025.05.006

1. Introduction and background

Cycling is recognized as one of the key modes of transport for sustainable urban mobility. Several studies based on user perception (Stamatiadis, Cafiso & Pappalardo 2019; Bourne et al. 2020) have identified the maintenance of cycle lanes and pavement quality as one of the main infrastructural barriers to cycling because vibration caused by the road pavement surface effects comfort, health and safety of bicycle users.

The European Directive 2002/44/EC assumes a health factor of 15–30 min vibration exposure for a reference daily duration of 8 h, corresponding to about 4–10 km cycling distance (European Parliament and of the Council 2002).

International standards like ISO 2631 and ISO 5349 emphasize the significance of vehicle or body vibration to evaluate the human exposure to comfort and health. This preference stems from the direct correlation between driver comfort and the weighted frequency root mean square (RMS) of the acceleration values (ISO 2001; ISO 2018). Human sensitivity to vibration is frequency dependent, and ISO 2631 is concerned with the total weighted RMS acceleration in the frequency range of 0.5–80 Hz, where most of the power is transmitted to the driver for health, comfort, and perception. At normal bicycle speeds, the 0.5–80 Hz range of ISO 2631, covers the full range of pavement surface texture in wavelengths from microtexture (<0.5 mm), macrotexture (0.5–50 mm) to megatexture (50–500 mm) and roughness (> 500 mm) (World Road Association 1987; ASTM International 2012; ISO 2019). Moreover in the ISO standard, the highest weighting factors are in the range 0.5–25 Hz (mainly macro- and megatexture), and frequencies below 1 Hz (mainly roughness) are primarily concerned with motion sickness. Similarly, studies on bicycles dynamics showed that the natural frequencies are influenced by wavelengths in the field of the megatexture (50–500 mm) and roughness until 1000 mm (Niska et al. 2024) with the peak of the response frequencies around 8 Hz (Kotani et al. 2023).

Due to the relevance of vibrations, this section introduces the reference background on pavement condition parameters and bicycle characteristics as factors affecting vehicle vibration.

1.1. Pavement condition and vibration

The condition of pavement surfaces plays a crucial role in the vibration of the bicycle influencing both ride quality and safety while cycling. Key factors such as pavement skid resistance and severe distresses like potholes, large cracks, or bumps can adversely affect bicycle stability and rider safety. Additionally, higher frequency vibrations caused by the Mean Profile Depth (MPD) of the pavement's macrotexture, alongside diffuse cracking and surface unevenness, predominantly impact riding comfort and can pose health risks (López-Molina et al. 2023).

In the realm of quality evaluation of the pavement surface, the International Roughness Index (IRI) (CEN 2006), the Present Serviceability Index (PSI), and the Pavement Condition Index (PCI) (ASTM International 2020) are established metrics used to gauge pavement performance. The IRI is generally measured using mobile equipment equipped with high-precision sensors (i.e. odometer, accelerometers, and laser). Meanwhile, Laser Crack Measurement Systems (LCMSs) represent the cutting edge in technology for detecting, classifying, and measuring surface defects such as cracks, potholes, and ruts, which are essential for calculating PSI and PCI. However, the implementation of these systems is predominantly confined to major rural roads due to their high operational costs and labor-intensive nature, making them less feasible for thorough and regular assessments on urban roads and impractical for use in cycle lanes.

Furthermore, it is recognized that the IRI may not be the most suitable indicator for assessing vibration comfort for low-damped, low-speed vehicles such as bicycles and scooters (Cafiso et al. 2022). Additionally, PSI and PCI are influenced only by certain wavelengths of pavement roughness, limiting their effectiveness in providing a comprehensive picture of surface suitability for all types of road users (Cafiso et al. 2019; Sidess, Ravina & Oged 2021).

Within this framework, while the applications of on-board recording of acceleration and GNSS data to analyze traffic conflicts and riding behavior are more limited (Yaqoob et al. 2023), more studies have focused on distress detection and road surface condition assessment using on-board vertical acceleration data (Allouch et al., 2017; Souza et al., 2017; Souza, Giusti & Batista 2018; Basavaraju et al. 2020; Lekshmipathy, Velayudhan & Mathew 2021; Parastoo et al. 2022).

Instrumented probe bicycles (IPBs) have been used to analyze bicycle comfort (Zhu & Zhu 2019; Ahmad Feizi Jun-Seok Oh & Joo 2020; Miah et al. 2020; Qian, Moore & Niemeier 2020), while high quality instruments are capable of accurate measurements, their installation on all bicycles is not feasible (Mason et al. 2016). Alternatively, smartphone accelerometers are often preferred for pavement surface condition assessment due to their suitability and widespread availability as sensors for collecting vibration and position data (Gao et al. 2018; Litzenberger et al. 2018; Nuñez, Bisconsini & Rodrigues da Silva 2020). Litzenberger et. al (Litzenberger et al. 2018) compared measurements collected by an accelerometer (400 Hz) and acceleration collected by a smartphone (100 Hz) to assess pavement surface quality. The results showed that the accuracy of pavement surface quality by accelerometer was high, but the detection by smartphone accelerometer was also acceptable. A literature review on road surface monitoring using vibration data collected by smartphone sensors showed that they are widely used to detect anomalies (e.g. potholes, bumps) or to derive pavement indices (e.g. IRI, PCI) (Sattar, Li & Chapman 2018). Another significant advantage of smartphones is illustrated by the fact that their penetration rate in the population is almost 100%, especially in developed countries (Newzoo 2019), and the effectiveness of road monitoring could be significantly improved with data crowdsourcing techniques. Moreover, cyclists are hesitant about being equipped with added devices as bicycles are traditionally cheap and simple (Berge et al. 2022) and they are already familiar in sharing crowd-

sourced data from smartphone applications like Strava (Lin & Fan 2020) In such a framework, the absence of standardization in vehicle dynamics and speed, in contrast to the IRI standard, poses new challenges for uniform data collection and analysis from distributed probe vehicles.

Pavement vibration data collected by smartphone resulted suitable to identify medium and high distress severities along road sections (Shtayat et al. 2020) but the authors also stated that different bicycle characteristics need to be evaluated to assess the effectiveness of the classifier on diversified conditions. A recent review paper (Al-Sabaeei, Souliman & Jagadeesh 2024) presents an extensive literature review on smartphone applications for pavement condition monitoring and concludes, among other things, that the effects of different tire pressure, vehicle type, spring stiffness, and vehicle size on acceleration data should be more investigated. Shtayat et al. (Shtayat, Moridpour & Best 2022) compared pavement monitoring using a smartphone on an e-bike and a car, concluded that the pavement vibration data have acceptable accuracy for the estimation of the Present Serviceability Rating, but the results may be subject to various factors that include the vehicle type and quality of the suspensions and tires of the test vehicle. In another recent study published in 2024 the authors stated that the stiffness and damping coefficients of bicycle tires have not been considered in previous studies of road surface measurements made with bicycles (Rizelioğlu & Yazıcı 2024). Similar recommendations for further studies have been suggested by Niska (Niska et al. 2024) to provide better insights to make the readings independent of the type of bicycle.

1.2. Bicycle modelling

Bicycle vibration primarily occurs at the wheel-road contact interface when the vehicle passes over a rough surface. The magnitude of vibration depends on several factors, primarily including road roughness, vehicle speed, suspension system, and bicycle characteristics. For bicycles, the main suspension systems are the air fork in the front axle and the wheel tires. In terms of suspension systems, bicycles mainly fall into two categories: rigid bicycles, which lack a suspension system and bicycles equipped with coil fork systems, utilizing a steel spring and air fork. Without a suspension system, all the absorbed energy would be transferred to bicycle frame and the rider, causing riding discomfort and handling problems. According to Kumar Manilal Sidhant (Kumar Manilal Sidhant 2019), higher stiffness properties result in a decrease in frequency and an increase in amplitude. The amplitude gain of the bicycle depends on the ratio of the excitation frequency to the natural frequency of the bicycle, with the maximum when the ratio is equal to 1. Therefore, increasing the stiffness will result in decreasing the comfort level especially at high speed with higher input frequencies from the road surface.

Several studies have employed bicycle dynamics models to assess rider comfort and optimize vehicle design. Since the first Whipple mathematical model (Whipple 1899) which consisted of four rigid bodies, various methods have been employed to study bicycle dynamics. The most basic models require geometry, mass distributions for the rigid bodies. More complex models require estimates of tire characteristics, human characteristics, friction, stiffness, damping, etc. For the benchmarked Whipple bicycle model, through a laboratory experiment, the authors presented a method to accurately estimate the physical parameters of a bicycle (Moore et al. 2010). In research (Kumar Manilal Sidhant 2019), acceleration data were used to simulate the bicycle response using ADAMS multi-body software and to optimize influential design parameters for comfort improvement. In research (Bulsink et al. 2015), the authors developed a parameterized passive cyclist model to analyze the tire model's influence on stability. Chou et al. (Chou et al. 2015) conducted a research study and evaluated the capability of numerical simulation of bicycle motion. For this purpose, a specialized test bicycle was outfitted with multiple sensors to measure vertical acceleration. Additionally, a lightweight inertial profiler was designed to record the route profile of the bike path, serving as the primary input to the simplified bicycle model. The study's results demonstrated a high degree of agreement between the measured acceleration and those calculated by the model.

These studies have showcased the capabilities of various mathematical and numerical multi-body models to analyze the vertical and pitch motion of the various bodies that make up the bicycle. Comprehensive numerical multi-body models, like ADAMS, are considered the most accurate due to their capability to precisely model each mechanical component of the bicycle as well as the rider's body and pedaling action. They are best suited for the mechanical design and development of bicycle suspensions, frames, and roller chain transmissions. Conversely, the closed form of the numerical solutions must be based on a simplified model because the motion relationships between each part of the bicycle are so complex that some of them could be pages long. However, they have proven effective in simulating the vertical vibrations of the bicycle body induced by the road surface. Moreover, such simplified numerical models can be easily incorporated as part of a more complex simulation framework such as the one used in this paper.

1.3. Study design and contribution

Several studies have examined the relationship between road conditions and bicycle vibrations. These studies have highlighted that the collection of road profiles is a costly survey approach for the extended urban road network. Furthermore, this approach is not suitable for use on bike paths due to the size of the equipment required. Smartphone sensors offer a more cost-effective and widely available solution for monitoring road surfaces, given the correlation between pavement conditions and the root mean square (RMS) of the acceleration values. It is also essential to collect data from multiple users in order to mitigate the impact of uncontrolled factors such as user position and pavement surface variability. This makes crowdsourcing an attractive option for data collection on bicycle vibrations and pavement monitoring. The review of the lit-

erature identified the need for further studies on the effects of variability in bicycle characteristics to evaluate user comfort related to pavement conditions.

In light of the above, the primary research questions of the study were as follows:

"What is the impact of variability in bicycle characteristics on the overall uncertainty when different bicycles are used for crowdsourcing data collection?"

What factors primarily influence the variability in bicycle vibration, and what measures can be implemented to ensure the collected data is more homogeneous?

This study addresses these questions by exploring how the dynamic characteristics of bicycles influence vibration data, employing experimental tests and numerical simulations. Experimental tests were conducted to calibrate the numerical model using a genetic algorithm (GA) and to assess the impact of external factors on vibration outputs, such as wheel path and pavement transversal conditions, and bicycle speed. Once calibrated, the numerical model was used to quantify the changes in vibration output resulting from changes in the bicycle's main dynamic characteristics.

The utilization of a validated numerical model proved value in extrapolating the results to a vast array of bicycle and user characteristics that are not experimentally feasible, as evidenced by the referenced literature. The impact of varying parameters and the necessity for consistent data collection and comparison were assessed by comparing the model outputs with the estimated experimental noise in bicycle vibration. To the best of the authors' knowledge, this paper is the first which analyses the role of variability in vehicle dynamic characteristics on vibration responses in the context of using smartphone applications to monitor and assess pavement conditions using crowdsourced data.

The research development can be described in the following five main tasks, which also represent the organization of this paper.

1) Collection of experimental data on vibration and road surface characteristics, filtering and fusion of data.
2) Modelling of bicycles for numerical simulation of vibration induced by the pavement profile.
3) Use of a genetic algorithm to identify the optimal mechanical configuration of the test bicycle to match the experimental and simulated data.
4) Numerical simulation and sensitivity analysis to investigate how variability in bicycle characteristics contributes to the overall uncertainty in multiple measurements.
5) Presentation and discussion of the results and conclusions.

2. Methodology

In this study, vertical acceleration data with GNSS coordinates, along with pavement longitudinal profiles and distresses, were collected with several runs at different speeds on selected test roads. Various pre-processing steps were then applied to synchronize the data collected by smartphone and mobile laboratory. The vertical acceleration recorded in the field with a smartphone installed on the test bike was filtered and synchronized to ensure comparability with the simulation carried out on the actual profile of the pavement roughness. The road profile was used as input to simulate the vehicle dynamics with a half car model in MATLAB Simulink. The mathematical formulation based on the half-car relaxation model was integrated in a genetic algorithm optimization process to identify the bike parameters which minimize the error between the simulated and in-field RMS of vertical accelerations. The optimization aimed to ensure the model parameters' best consistency with the real dynamic response of bicycle to road profiles. Finally, the sensitivity analysis was carried out by varying the bicycle characteristics in the range of values that can be expected in real world data collection from several bicycle types. This section provides a detailed presentation of the various steps involved in data collection, data preparation and preprocessing, as well as the underlying principles of the modeling approach.

2.1. Data collection

For this research, data were collected from two sources: 1) a 50 Hz accelerometer and 1 Hz GNSS sensors embedded in a smartphone mounted on the bicycle body to record vertical acceleration, speed and position; and 2) the Automatic Road Analyzer (ARAN) with a laser longitudinal profiler and LCMS to collect high-quality pavement longitudinal profiles and pavement distresses. All bicycle tests were repeated on urban road sections, at different speeds in the range $12 \div 25$ km/h. The rider was instructed to maintain, during each test, an almost constant speed and trajectory on the right side of the lane where the ARAN data were collected. Fig. 1 displays the GIS representation of data points collected during one test. All the tests have been carried out on the right lane of three road sections (named A, B, C) with asphalt concrete pavement selected to provide a wide variability of pavement conditions, as shown in Fig. 2 by the distribution of different classes of the Pavement Condition Index (PCI).

Table 1 shows the summary statistic of the experimental dataset which consisted of 153 segments of 10 m in length along the three road sections (A, B, C). This resulted in a final data sample of validated 728 RMS values, with the test being repeated at five different speeds. In this research the test bicycle was a "city e-bike" (model: Atala B-EASY SL 400; electric assisted, front suspension damped, tire 700x40, weight 23 kg) very similar to a common city bike but with the advantage of the electric assistant system that helps to maintain constant the selected speeds during the tests.

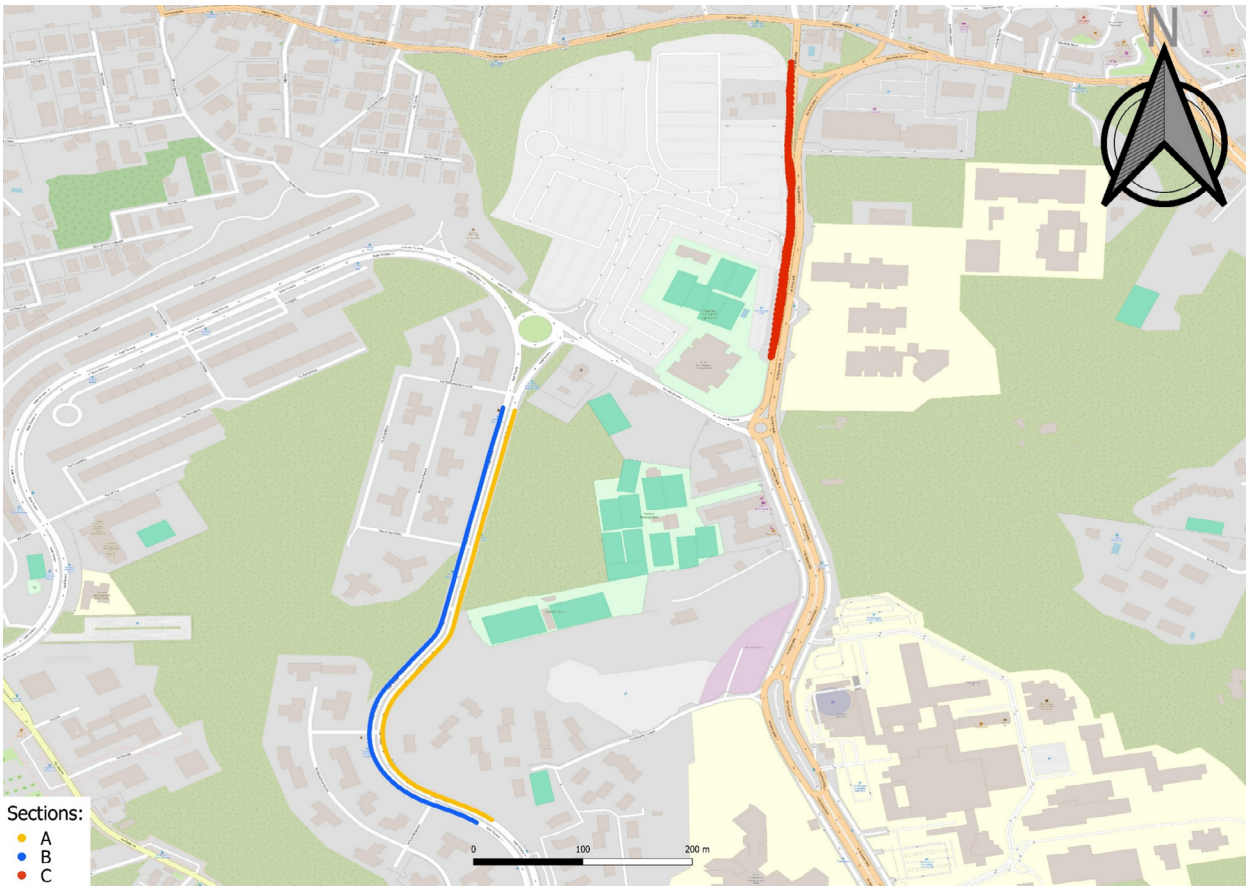


Fig. 1. Visualization of data points in QGIS on OpenStreetMap.

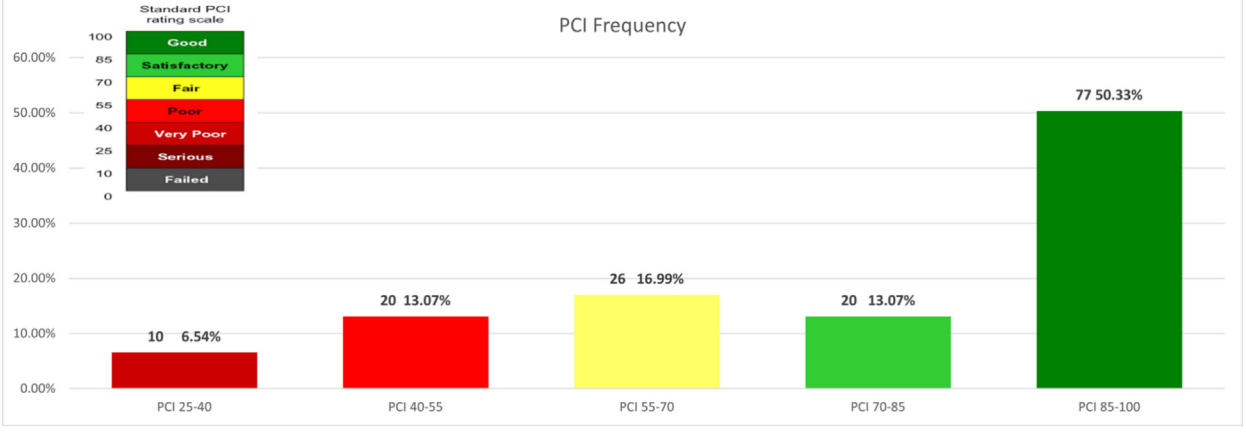


Fig. 2. Frequency distribution of the Pavement Condition Index (PCI).

Data shows a high variability of pavement conditions with IRI ranging from 1.52 to 12.7 mm/m and PCI from 25.93 to 100, while MPD is almost uniform over all the sections. Therefore, the 153 road segments can be considered representative of a wide range of asphalt pavement surfaces from very good to very poor conditions (Fig. 2).

Table 1
Summary statistics of experimental data.

Feature	N samples	Min	Max	Mean	Standard Deviation (SD)	Coefficient of Variation (CV)
Pavement Data						
IRI [mm/m]	153	1.517	12.574	4.116	2.005	49.09%
PCI	153	25.932	100	77.654	21.366	27.60%
MPD [mm]	153	0.286	0.758	0.422	0.093	22.08%
Vibration Data						
Speed [km/h]	728	7.579	26.845	16.252	4.160	25.60%
RMS [m/s^2]	728	0.029	18.601	2.371	2.456	103.63%

2.2. Data preparation and synchronization

2.2.1. Denoising

As previously stated, bicycle dynamics and human comfort are primarily influenced by megatexture and short-wavelength roughness. However, the recorded raw acceleration signal indicates the presence of higher frequency content due to pavement macrotexture (MPD). These surface wavelengths are not included in the pavement elevation profile data collected by ARAN at 0.025 m sampling resolution, which is also the input for the numerical simulation. Furthermore, the paper focuses on vibration caused by pavement surface distress which mainly affects comfort. Therefore, in our study, the recorded higher frequency signals were considered as noise to be removed before comparing the results.

Wavelet decomposition is a powerful signal processing technique used to analyze and denoise signals in both time and frequency domains (Percival & Walden 2000; Mallat, 2008). In this method the signal is decomposed at different scales using a series of high-pass and low-pass filters to reduce or eliminate noise while preserving essential signal features (Donoho 1995). The selection of the denoising level was based on previous studies (Murgano et al. 2021) and a preliminary comparison between recorded and simulated accelerations to minimize the MAPE. As a result, the frequency range for wavelet decomposition was set at level 4, preserving frequencies below 1.56 Hz without thresholding while effectively reducing high-frequency noise in the range of 1.56 to 25 Hz. Fig. 3 shows the comparison between raw and denoised accelerations in one of the test data and Fig. 6 confirms the good agreement between simulated and denoised signals.

2.2.2. Time-space transposition

The main effort in data preparation was to merge the two data sources, which use different acquisition frequencies and references in space (ARAN pavement profile chainage) and time domains (smartphone accelerometer signal), but with a common GNSS positioning (ARAN & Smartphone). Therefore, both datasets were imported into GIS software (Qgis version 3.24) for spatial matching. The positional bias was corrected by applying a re-chainage factor that equals the distance traveled by the bicycle to the length of the road profile recorded by ARAN. Moreover, a time to space transposition was applied to the acceleration signals. To synchronize the two datasets, the time domain of the smartphone sensors was matched to the space domain of ARAN, considering the actual speed of the bicycle:

$$Dist_t = Dist_{t-\Delta t} + \Delta t \times V_{avg} \quad (1)$$

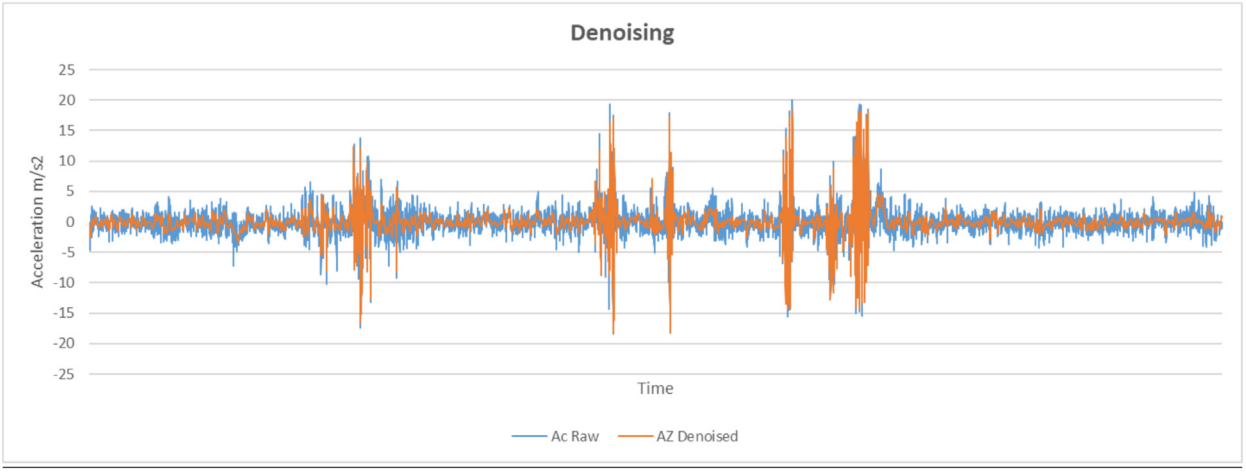


Fig. 3. Comparison of raw and denoised acceleration data.

where: $Dist_t$ is distance at time t; $Dist_{t-\Delta t}$ is distance at time t-Δt; Δt is time step = 1/50 sec; and V_{avg} is average speed between (t-Δt) and (t).

The profile data are acquired with a high spatial resolution of ARAN (0.025 m), which corresponds to 200 Hz in the time domain at the reference bicycle speed of 5 m/s. Therefore, to achieve the highest and most uniform resolution, all acceleration signals were linearly interpolated from the original 50 Hz acquisition rate to 200 Hz frequency. At the end of the process, all the recorded vertical acceleration values were aligned to the linear referencing of the contact point of the front and rear wheel of the bicycle on their actual road profiles. Fig. 4 shows the time–space transposition and synchronization flowchart.

3. Half car model and simulation

While the half-car model is traditionally assumed to be a simplification of the full-car model for four-wheel vehicles, Fig. 5 shows that its structure is well-suitable for bicycles when properly adapted through parameter tuning. In bicycle dynamics modeling, the quarter-car (Rizelioğlu & Yazıcı 2024) and half-car model have been used to study the behavior and performance of a bicycle's body and suspension system (Gillespie 1992; Wong 2022). The basic elements of the half-car model include front and rear sprung and unsprung masses, suspension components wheel springs and dampers, as well as bicycle and rider body mass and inertia (Gillespie 1992). Various resources covering vehicle dynamics, automotive engineering, and suspension systems delve into the concepts and calculations associated with this model, which is also familiar to pavement engineering because it is consistent with the IRI quarter car model (Gillespie 1992; Mahroogi et al. 2019; Wong 2022). Springer et alt. (Springer & Ament 2020) assessed the capability to estimate the IRI utilizing a half-car model parametrized to the bicycle characteristics. Rizelioğlu et alt. (Rizelioğlu & Yazıcı 2024) used an instrumented bike as the test vehicle to calculate IRI. They showed promising results by adjusting the quarter-car model to account for the stiffness and damping coefficients of bicycle tire.

The half-car model for bicycles divides the bicycle into three primary components: (1) – the front half, which represents the front wheel, fork, and associated suspension system, and is typically modeled as a simple spring-damper system. The body includes a sprung mass that represents the mass of the bicycle frame and any additional components at the front of the bicycle. (2) – the rear half, which represents the rear wheel and suspension. (3) – the bicycle frame which connects the front and rear halves and supports the rider. The interaction between the bicycle wheels and the road surface is crucial for modeling vertical vibrations, the road longitudinal profile is considered as imposed vertical position of the wheel contact points (Zr_1, Zr_2) forcing the dynamics of the whole system in the time domain. In this model, bike parameters include spring stiffness, damping coefficients, tire characteristics, rider and bike mass and geometry. The mathematical modeling of the half-car model based on Fig. 5 and using Newton's second law of motion is defined as follows:

For sprung mass C.G:

$$m\ddot{Z} = -K_1(Z - a_1.\theta) - C_1\left(\dot{Z} - a_1.\dot{\theta}\right) + K_1.Zu_1 + C_1.Zu_1 - K_2(Z + a_2.\theta) - C2\left(\dot{Z} + a_2.\dot{\theta}\right) + K_2.Zu_2 + (C_2.\dot{Zu_2}) \quad (2)$$

Pitch axis moment of inertia (Kg.m^2):

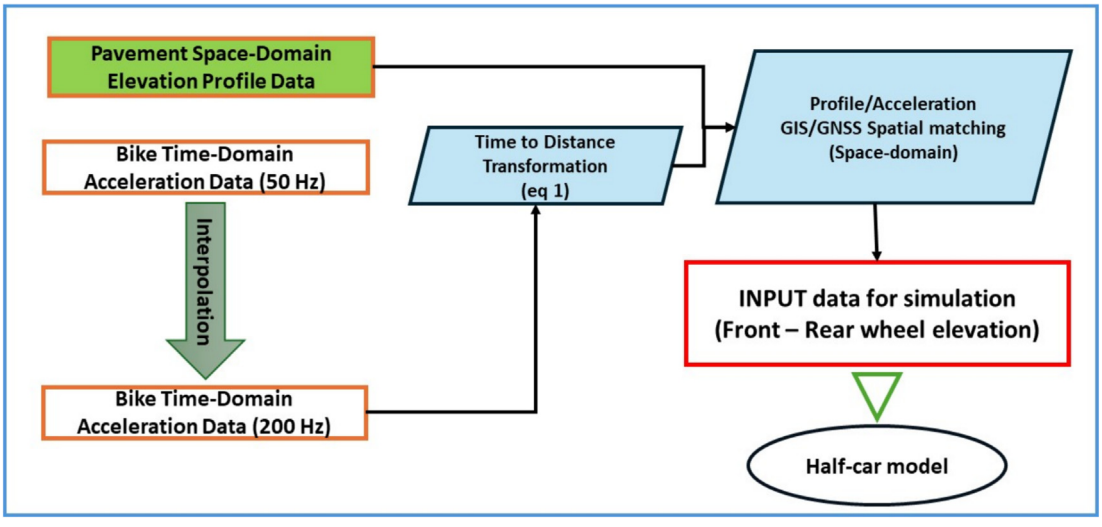


Fig. 4. Time-space transposition and synchronisation flowchart.

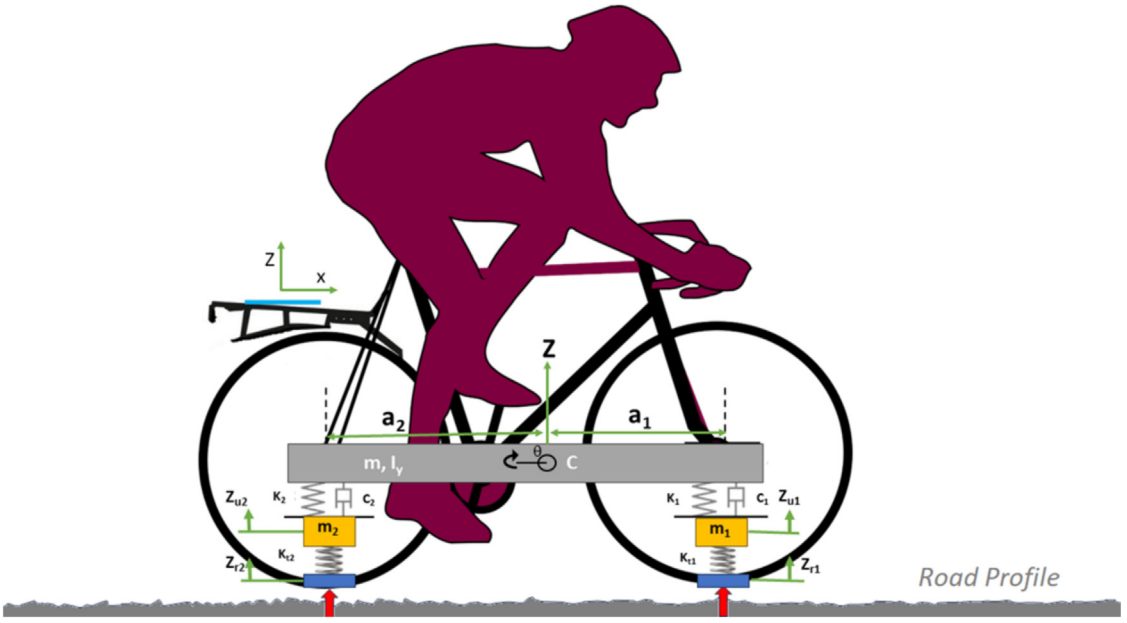


Fig. 5. Schematic of the Half Car Model.

$$J\ddot{\theta} = (-K_1.a).(Z + \theta(-a_1) + C_1.a_1(\dot{Z} + \dot{\theta}(-a_1) + (-K_1.a_1.Zu_1) + (-C_1.a_1).\dot{Zu_1} + (-K_2.a_2).(Z + \theta.a_2) - C_2.a_2\left(\dot{Z} + \dot{\theta}.a_2\right) + K_2.a_2.Zu_2 + C_2.a_2.\dot{Zu_2}) \quad (3)$$

The equation for the front suspension system can be written as:

$$m_1\ddot{Zu_1} = K_1(Z + \theta(-a_1)) + C_1\left(\dot{Z} + \dot{\theta}(-a_1)\right) + (-Kt_1.Zu_1) - K_1.Zu_1 - C_1.\dot{Zu_1} + Kt_1.Zr_1 \quad (4)$$

Similarly, the equation for the rear suspension system employs terms analogous to those in the front suspension equation, differentiated by a subscript of 2 for the rear half:

$$m_2\ddot{Zu_2} = K_2(Z + \theta(-a_2)) + C_2\left(\dot{Z} + \dot{\theta}(-a_2)\right) + (-Kt_2.Zu_2) - K_2.Zu_2 - C_2.\dot{Zu_2} + Kt_2.Zr_2 \quad (5)$$

where: m_1 and m_2 are the masses of the front and rear halves (sprung mass). *C_1 and C_2* are the damping coefficients of the front and rear suspensions. K_1 and K_2 are the stiffness of the front and rear suspensions. Kt_1 and Kt_2 are the tire stiffness of the front and rear suspensions.$\ddot{Z}$, $\dot{Z}$, Z are acceleration velocity and displacement of the sprung mass C.G.$\ddot{\theta}$, $\dot{\theta}$, θ are angular acceleration, angular velocity, and angular displacement.$\ddot{Zu_1}$, $\dot{Zu_1}$, Zu_1 are acceleration, velocity, and displacement of the front unsprung mass.$\ddot{Zu_2}$, $\dot{Zu_2}$, Zu_2 are rear unsprung mass acceleration, velocity, and displacement.

The beam and frame characteristics are also considered to model the interaction between the front and rear halves of the bicycle (Zr, Zu, Z, θ). The joint and frame equation (*Eq. (6)*) typically accounts for the relative motion and forces between the two halves of the bicycle.

$$F_c = \frac{I_y}{Z} m.\ddot{\theta} \quad (6)$$

where: I_y is the moment of inertia, Z is the height of the center of mass, m is mass, $\ddot{\theta}$ is angular acceleration and F_c is the force exerted by the joint or frame.

Fig. 5 depicts the half car model and provides definitions of the parameters for this dynamic model. In this figure, Z is the body's vertical motion coordinate, a_1 and a_2 are the distance of the center of gravity from the front and rear axles, m is the body mass, m_1 and m_2 are the mass of the front and rear unsprung masses, I_y is the body and rider moment of inertia ($kg.m^2$), and theta (θ) is the body's pitch motion.

All the procedure was implemented in Matlab Simulink to calculate the vertical accelerations at the smartphone position due to the pavement profile in the three road sections given as input the actual speed of bicycle in each road segment of 10 m.

3.1. Simulated and measured signals

After simulation, two signals of vertical accelerations can be compared: one is the simulated acceleration in the model and the other is the in-field acceleration signal collected by the sensor. Fig. 6 shows the simulation by half car model for one test and comparison with the denoised measured signal.

As mentioned earlier, these signals have been aligned and synchronized at a frequency of 200 Hz. The root mean square (RMS) of the signal is then used as the reference measure, as also suggested by ISO 2631-1 (ISO 1997):

$$\mathrm{RMS}_{t,u} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left|X_{t,u} - x_{t,u,n}\right|^2} \tag{7}$$

In which: $\mathrm{RMS}_{t,u}$ is the RMS for test "t" in the road section "u"; $X_{t,u}$ is the average of vertical accelerations; $x_{t,u,n}$ is the n^{th} vertical acceleration value; N is the number of vertical acceleration values for each road section of 10 m length at the 200 Hz frequency output (e.g. N = 200/5 × 10 = 400 at a speed of 5 m/s).

3.1.1. Peak alignment

The higher peaks in the RMS values are made by singularities in the pavement surface (e.g., potholes, large cracks, bumps). Despite the effort in synchronization of the data, shifts in peaks between recorded and simulated RMS have been occasionally observed. These shifts can be attributed to the GNSS positioning accuracy in the spatial matching carried out in GIS. For this reason, an automatic peak alignment was performed in the MATLAB environment and the optimal shift parameter was calculated for each test to maximize the cross-correlation between the real RMS signal and the simulated one.

4. Optimization of vehicle dynamic parameter by genetic algorithm (GA)

Genetic algorithms (GAs) are increasingly popular in various engineering fields because they effectively tackle complex optimization challenges (Alsajri & Steiti 2024; Jlifi, Ferjani & Duvallet 2025). This study focuses on using GAs to optimize vehicle dynamic parameters in the half-car model, aiming for the best alignment between simulated and measured RMS values.

Given a road profile and running speed, in the half-car model, the acceleration signal is affected by the 6 bicycle dynamic parameters $kt_{1\text{-}2}$, $k_{1\text{-}2}$ and $C_{1\text{-}2}$. Different settings of the bicycle parameters change the goodness of fitting between measured and simulated RMSs. Due to the complexity of the system and the number of variables, a genetic algorithm (GA) was implemented to identify the model parameters that minimize the difference between simulated and measured RMS (MAPE). Genetic algorithms (GAs) are optimization techniques inspired by the process of natural selection and genetics (Holland 1975). They are used to find solutions to optimization and search problems by mimicking the principles of biological evolution. Haupt and Haupt (Haupt & Haupt 1998) discussed the practical aspects of implementing genetic algorithms, emphasizing algorithm design, parameter tuning, and optimization strategies, and provides insights into the effective use of GAs in real-world scenarios.

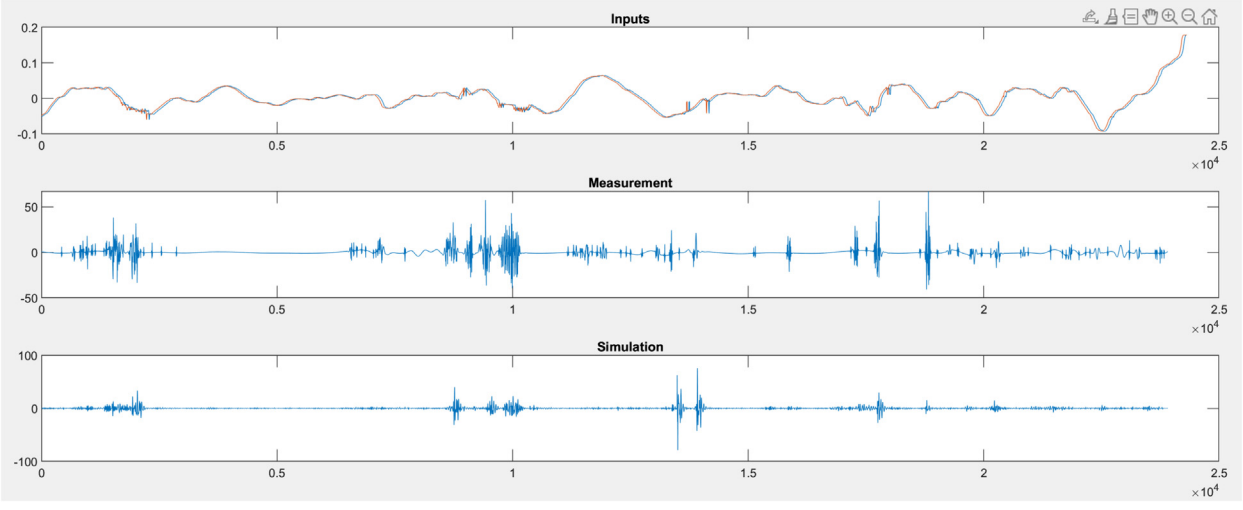


Fig. 6. Simulation by half car model for one test; Inputs: road profiles at the front and rear wheel contact, Measurement) measured acceleration in in the field, Simulation) simulated accelerations by the model.

The process begins with the **initialization** of a population of individuals, which is randomly generated to start the optimization process. The fitness of each individual, indicating the goodness or optimality of the solution, is evaluated using a fitness function specific to the problem being solved (Mitchell 1996). As mentioned by Eiben and Smith (Eiben & Smith 2015), the **selection** process involves selecting individuals from the population based on their fitness. Individuals with higher fitness have a greater chance of being selected for reproduction to contribute to the creation of the next generation. In this research, Stochastic Universal Sampling (SUS) is used as a single-phase sampling algorithm with minimal spread and zero bias. It is a variation of roulette wheel selection designed to address limitations associated with traditional roulette wheel selection, such as the risk of selecting the same individual multiple times. Selected individuals undergo genetic operations such as **crossover**, involving the swapping of genetic material between two parents, and **mutation**, which introduces random changes in the individuals' genome. Single-point crossover is the genetic operator used to combine genetic material from two or more parent individuals to create one or more offspring. Mutation introduces random changes to the genetic information of individuals in the population that contributes to genetic diversity and exploration of the solution space. **Reinsertion** is a strategy used to determine how newly created offspring are introduced into the population after undergoing selection, crossover, and mutation. The purpose of reinsertion is to maintain diversity, incorporate potentially better solutions, and manage population size. Fitness-based insertion was chosen as the reinsertion procedure. These steps are repeated for several generations until a **termination criterion** is met (Goldberg & Holland 1988). To improve the stability and robustness of the results, multiple runs of the GA have been conducted across different initial populations and parameter settings. Table 2 & Table 3 shows the configuration settings and operational details of the GA which provided a fast and accurate convergence. The GA parameters were selected based on predefined values commonly used in engineering optimization, relying on prior knowledge and experience. Since the GA inherently monitors error minimization across generations, as shown in Fig. 8, the stabilized error indicates that the selected parameters are appropriate. Fig. 7 shows the general genetic algorithm flowchart for optimization which is integrated with the bicycle model simulation.

The procedure involved a population of six bicycle parameters which provided the simulated RMS by running the bicycle model. The objective function was to minimize the MAPE between the simulated and recorded RMS values. The six bicycle parameters are:

- **K_1 & K_2:** front and rear suspension spring constant [N/m].
- **Kt_1 & kt_2**: front and rear tire spring constant [N/m].
- **C_1 & C_2:** front and rear suspension spring constant [N*s/m].

The overall weight was fixed at 100 kg which includes bicycle (23 kg), instruments (2 kg) and cyclist (75 kg) weights.

For validation purposes, two scenarios were devised for training and testing, considering the variability in pavement characteristics and speeds. In **Scenario 1**, road sections A and B were selected for training while road section C served as the testing set. Conversely, **Scenario 2** involved a mixture of tests at different speed in road sections A, B and C for both training and validation (Fig. 1). The distribution of data samples for training (**75%**) and validation (**25%**) remained the same across both scenarios, including always the overall range of speeds.

Table 4 shows the combination of dataset for each scenario.

Moreover, given the possibility of multiple minima, a defined range of output values was defined to avoid unrealistic solutions, high number of population size and generations were implemented, and different parameter settings were selected to conduct a more comprehensive analysis of the results.

For the tire spring constant "Kt", a laboratory test was conducted by applying different loads and measuring tire deflections. The results showed a constant and consistent spring value of 119.8 kN/m, which was rounded to 120 kN/m. The damping parameter for the rear wheel (C2) was identified close to 2000 [N*s/m] based on engineering judgment after preliminary tests. It is worth mentioning that the real bicycle is not equipped with damping in the suspension of the rear wheel, except

Table 2
Genetic algorithm configuration settings.

Parameter	Description	Value
Population Size	The number of potential solutions (individuals) in each generation. Higher values increase genetic diversity.	100 Individuals
Number of Generations	Total cycles of selection, crossover, mutation, and insertion processes.	90 Cycles
Generation Gap	Determines the replacement rate of generations. A gap of 1.0 means full replacement each generation, enabling rapid solution space exploration.	1.0
Mutation Probability	Probability of a mutation occurring in any single gene, calculated adaptively as of $(\frac{0.7}{chromosomelength})$. Balances exploration and exploitation.	Adaptive as of $(\frac{0.7}{chromosomelength})$
Crossover Probability	Frequency of crossover between pairs of individuals. Higher rates increase genetic diversity in offspring.	0.7

Table 3
Genetic algorithm operational details.

GA Process	Description	Method / Setting
Initialization	Population created with random binary values to begin the optimization process.	Random binary initialization
Binary to Real Conversion	Converts binary chromosomes to real-number parameters for model use.	Binary to real decoding
Selection	Stochastic Universal Sampling (SUS) selects individuals based on fitness, reducing the risk of selecting the same individual repeatedly.	SUS (Stochastic Universal Sampling)
Recombination (Crossover)	Single-point crossover merges genetic information from two parents to produce offspring, enhancing solution diversity.	Single-point crossover
Mutation	Random changes are introduced in genes to maintain diversity and enable thorough solution exploration.	Adaptive mutation
Reinsertion	Fitness-based insertion retains offspring in the population, prioritizing better solutions while managing population diversity.	Fitness-based insertion

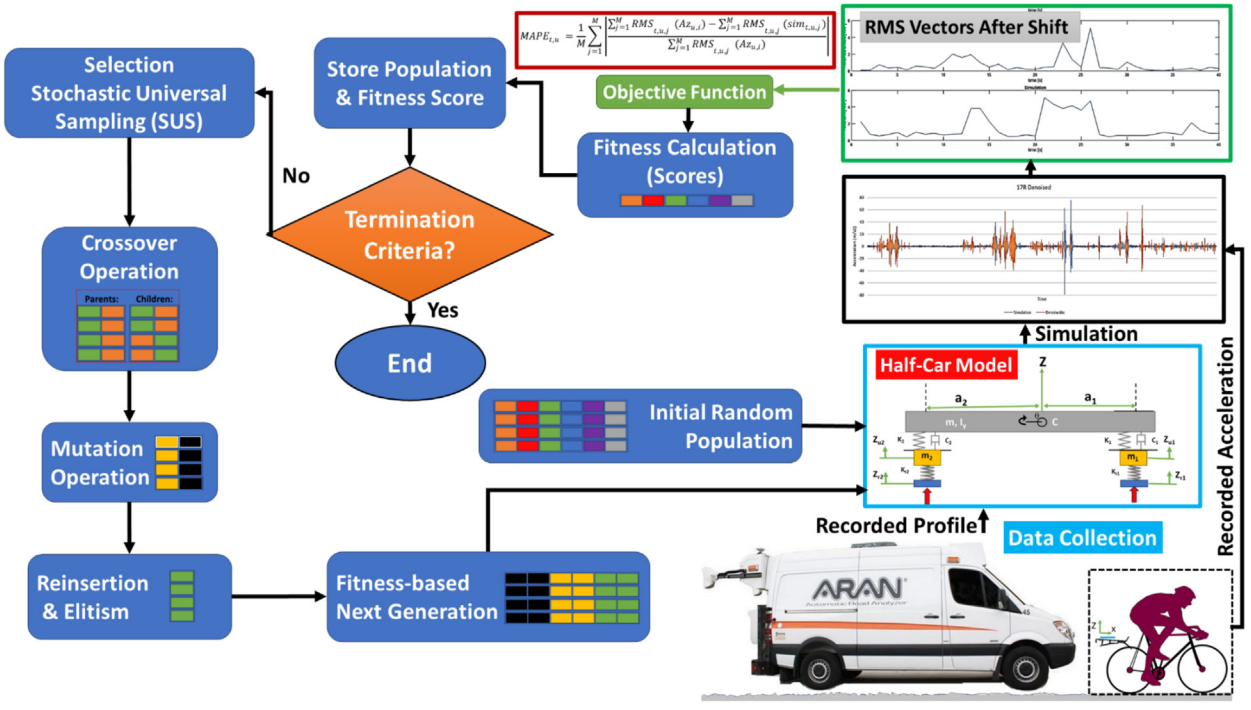


Fig. 7. Conceptual flowchart of general simulation and genetic algorithm integration.

for the damping component of the tire. Therefore, C2 is a model component effective in adjusting the numerical results to the real-word bicycle behavior.

Finally, four different configurations (Con 1, 2, 3, 4) were utilized to investigate the optimization parameters:

Con1: all parameters optimized within the defined ranges,
Con2: fixing Kt1 & Kt2 to 120 [kN/m], based on laboratory test, other parameters optimized.
Con3: fixing C2 to 2000 [N*s/m], based on test results, other parameters optimized.
Con4: fixing Kt1 & Kt2 to 120 [kN/m] and C2 to 2000 [N*s/m], K1, K2 and C1 parameters optimized.

4.1. Optimization and objective function

By choosing an objective function, the fitness of the population can be evaluated. The Mean Absolute Percentage Error (MAPE) between the measured and simulated RMS was chosen as the objective function to be minimized. MAPE is a well-known measure of prediction accuracy in statistics. In the present application, the MAPE of the total RMS of each test was defined by the formula:

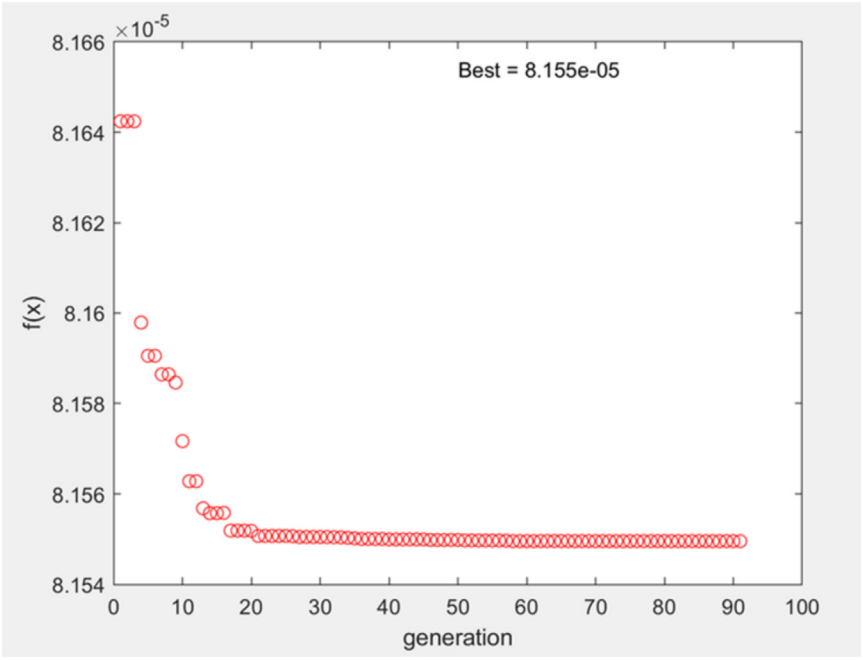


Fig. 8. Error trends in genetic algorithm iterations.

Table 4
Scenario combinations for analysis.

scenario	Training dataset (75%)	Validation dataset (25%)
Sc1	5 speed tests in Section A + 5 speed tests in Section B (Total average sample size 540)	5 speed tests in Section C (Total average sample size 180)
Sc2	mix speed tests from all three road sections (Total average sample size 540)	mix of 5 speed tests from all 3 three sections (Total average sample size 180)

$$MAPE = \frac{1}{Z}\sum_{i=1}^{Z} MAPE_i \tag{8}$$

where:

Z is the total number of tests in the training data set.

$MAPE_i$ is the error in each test and road section, calculated as in the following ($MAPE_{t,u}$):

$$MAPE_{t,u} = \frac{1}{M}\sum_{j=1}^{M}\left|\frac{\sum_{j=1}^{M} RMS_{t,u,j}(Az_{u,i}) - \sum_{j=1}^{M} RMS_{t,u,j}(sim_{t,u,j})}{\sum_{j=1}^{M} RMS_{t,u,j}(Az_{u,i})}\right| \tag{9}$$

$MAPE_{t,u}$: Mean Absolute Percentage Error calculated for each speed test "t" and road section "u" in the training sample. M is the number of 10 m sample units in the simulations for each test "t" and road section "u"; $RMS_{t,u,j}$ ($Az_{u,j}$) is the root mean square of the measured vertical acceleration; $RMS_{u,j}$ ($Sim_{u,j}$) is the RMS of simulated vertical acceleration.

5. GA results and data analysis

After defining these four configurations (Con 1, Con 2, Con 3, Con 4), the optimization was performed based on the two previously defined scenarios in Table 4 as Sc1 (training and testing on different road sections) and Sc2 (training and testing on the same road sections). The genetic algorithm demonstrated excellent performance, enabling the fast optimization of all the six parameters that characterize the bicycle model (Fig. 8).

The bar graphs in Fig. 9, Fig. 10 & Fig. 11 show the optimized parameters according to the four configurations Con1, Con2, Con3 and Con4 in scenarios Sc1 and Sc2.

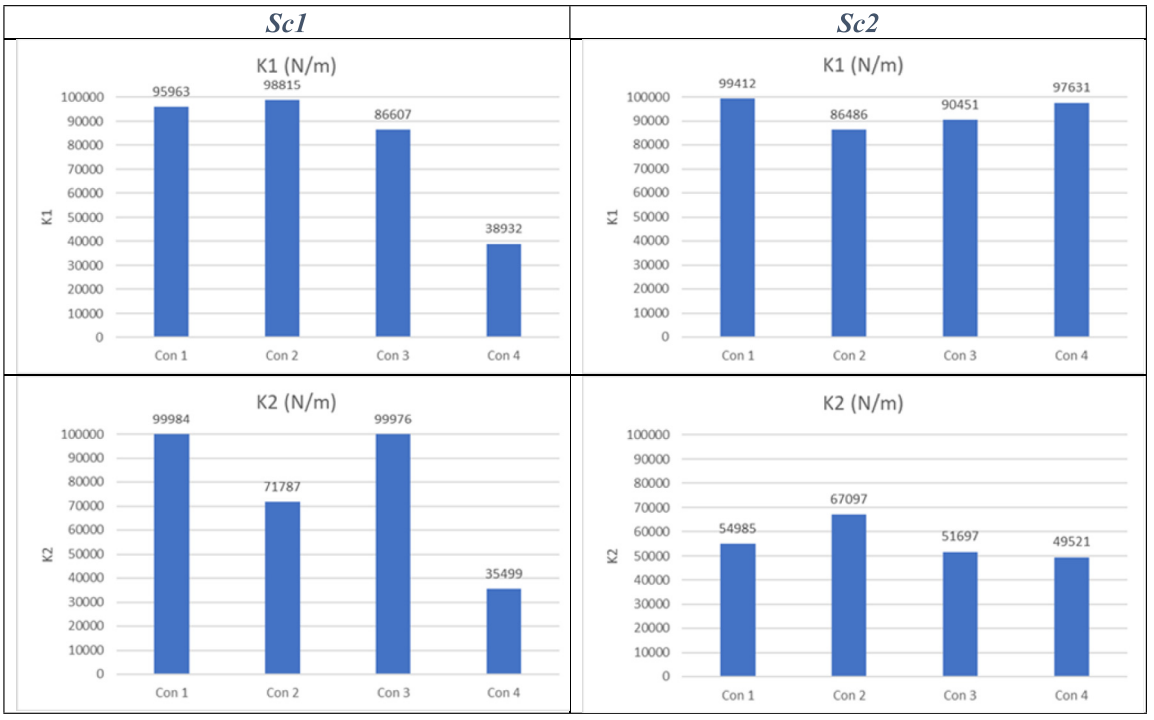


Fig. 9. Optimized K1 and K2 [N/m] for scenario Sc1 and Sc2 analyzed by 4 configurations Con1, 2, 3, 4.

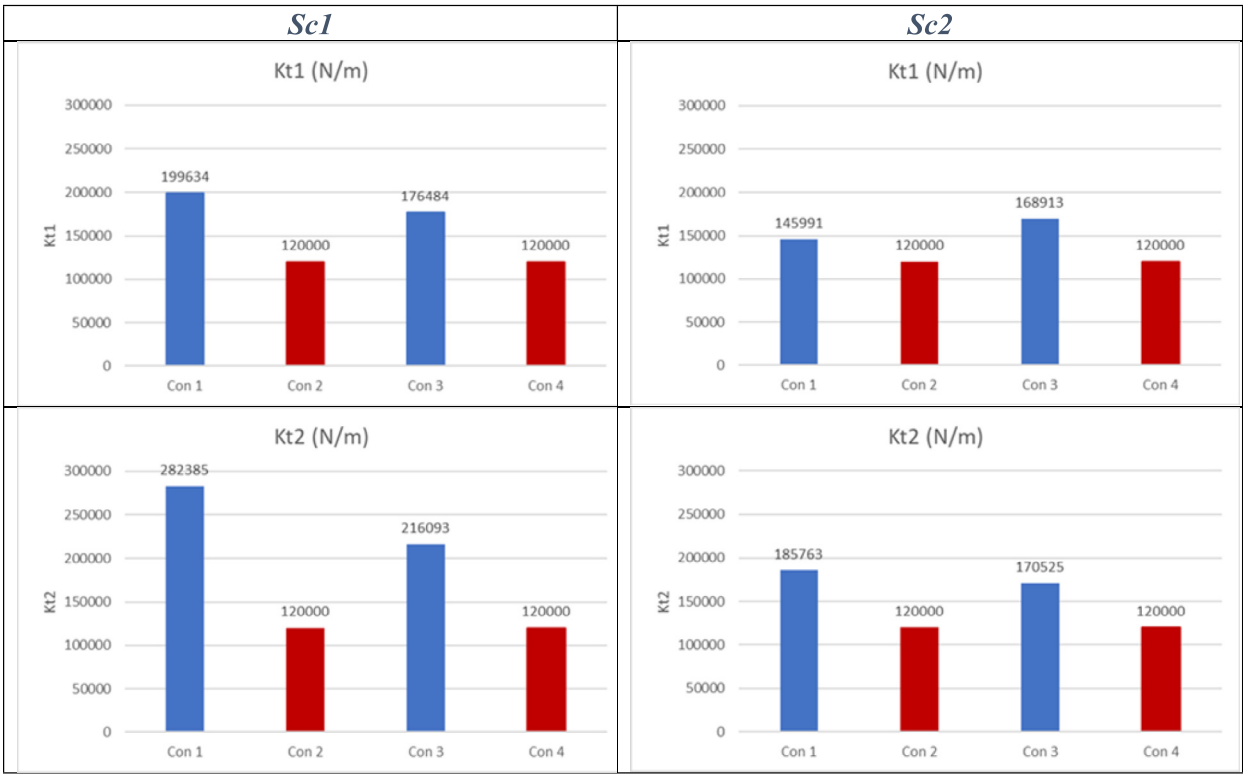


Fig. 10. Optimized Kt1 and Kt2 [N/m] for scenario Sc1 and Sc2 analyzed by 4 configurations Con1, 2, 3, 4. (Red bars: fixed parameters).

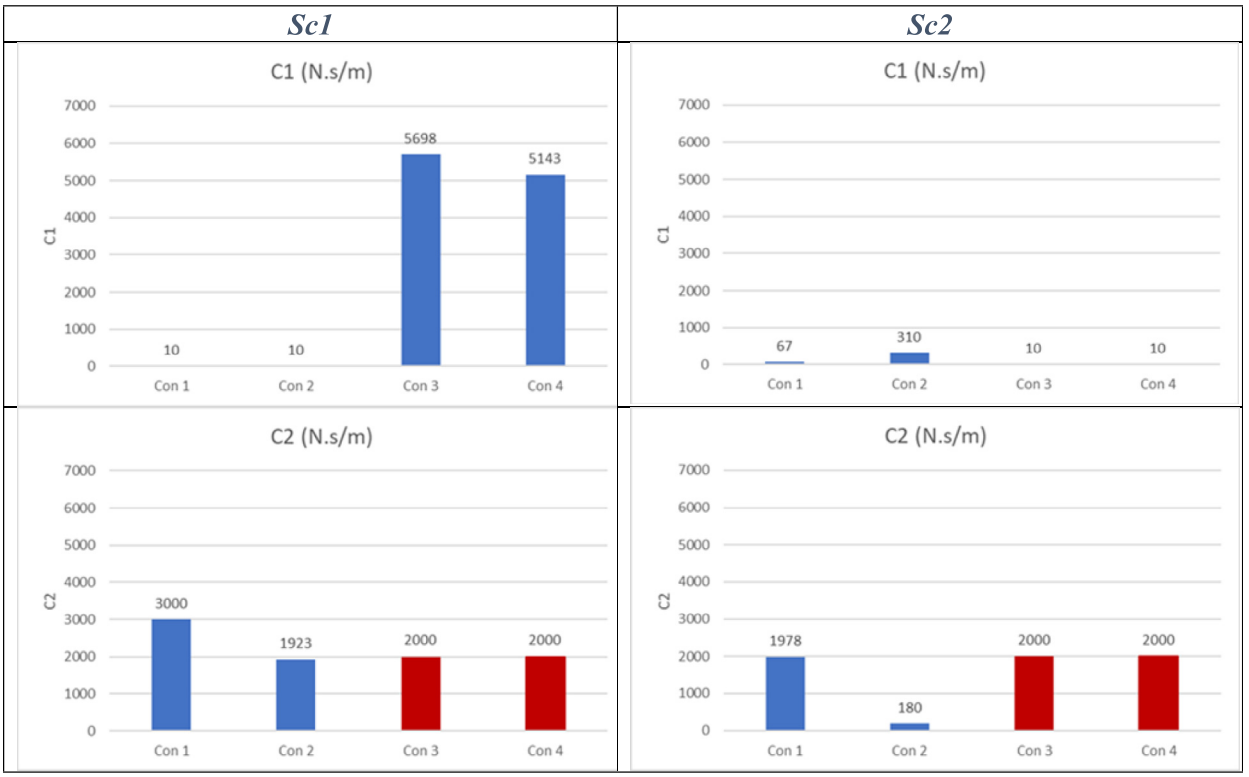


Fig. 11. Optimized C1 and C2 [N*s/m] for scenario Sc1 and Sc2 analyzed by 4 configurations Con1, 2, 3, 4. (Red bars: fixed parameters).

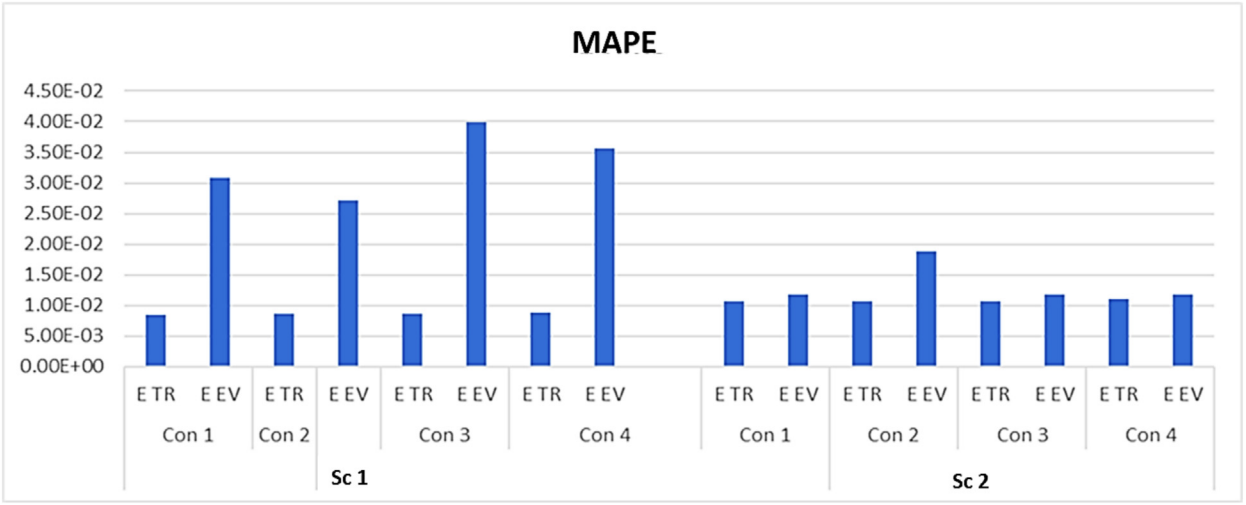


Fig. 12. Error graphs for two scenarios; in this graph, MAPE of RMS that used for objective function, TR and EV subscripts represent training and evaluation.

Examination of the results in Sc2 reveals consistent values for K1, K2, Kt1, and C1 across all configurations (Con 1, 2, 3, 4). Instead, the higher variability of Kt2 and especially of C2 is the consequence of the balancing adaptation of the numerical model to the real bike dynamics without damping in the rear suspension.

Comparison of the results shown in Fig. 9,Fig. 10 & Fig. 11 indicates that Sc2 provided more uniform and consistent values in the GA selection of the model parameters, in all bike configurations (i.e., Con 1, 2, 3, 4). Furthermore, Fig. 12 shows the best performance of GA in Sc2 with lower and consistent values of MAPE in both the training (TR) and validation (EV) samples.

As expected, the results show that Scenario 2, composed of a mix of speeds and road conditions in the training and validation datasets, produced the best results with lower differences between training (ETR) and validation (EEV) MAPEs, com-

pared to Scenario 1, where training was performed on road sections A and B with different distributions of distress type, extension, and severity from section C that was used for validation. No relevant variability in the results was seen among the 4 different bicycle configurations, except Con 2, where only the tires' stiffness was fixed. Moreover, good results achieved by the full set of bicycle characteristics identified in Con 1 showed the good performance of the GA in tuning our multi-parameter and complex model.

A correlation analysis was also tested between the measured and simulated RMS at the optimal GA model parameters in scenario 2. Table 5 shows a significant correlation with P-values less than 0.1% for all configurations, with similar correlation coefficients. Therefore, Scenario 2 (Sc2) and Configuration 1 (Con 1) were chosen as benchmarks for the sensitivity analysis, because they always returned the best results providing a good fitting of the experimental and simulation data for all the bicycle characteristics.

6. Sensitivity analysis

After setting the bicycle model and assessing its ability to mimic real-world bicycle dynamics in relation to pavement conditions, the final step of the research involves conducting a sensitivity analysis. This analysis was deemed significant due to its relevance in practical applications involving crowdsourced data collection from several bicycles for evaluating pavement conditions and riding comfort. Given the different characteristics of bicycles participating in such data collection, the effects of variability in bicycle attributes need consideration. On the other hand, even with the same bicycle, as in this study, variability in vertical accelerations among different runs can be expected due to the trajectories dispersions and the transversal variability of the pavement surface characteristics in the limited tire-pavement contact area (about 2–3 cm tire width) and accuracy in the bicycle longitudinal positioning and speed in repeated measurement, as also reported by (Niska et al. 2024) and (Shtayat et al. 2022). Therefore, the experimental noise was set at the 85th percentile of the MAPE values calculated in each of the 728 samples at the reference bicycle characteristics.

Hence, the numerical model and sensitivity analysis serve to generalize how the variability in bicycle characteristics contributes to the overall uncertainty in estimating vibrations induced by pavement conditions. For this purpose, the optimized parameters from scenario Sc2 and configuration Con 1, which has lower MAPE and consistent results, were set as default and used as the benchmark. One by one, each of the 6 parameters was varied over a wide practical range, while the other 5 were set at the default values. Running the simulation with different parameters is expected to result in an increased MAPE between the simulated and in-field RMSs. This is confirmed in Fig. 13 and Fig. 14, which show the results of the sensitivity analysis. The red triangle, denoting the GA default configuration, is always at the minimum of the blue line, which depicts the rise in MAPE for varying parameter settings. Furthermore, the MAPEt,u in each road element calculated between the in-field RMS and the simulated RMS at the GA-optimized parameters can be attributed both to the numerical simulation accuracy on the reference pavement profile and to the variability in RMS resulting from in-field factors, such as the variability in the wheel path and pavement transversal conditions. In order to account for the aforementioned intrinsic noise in repeated data collections, the orange line represents the 85th upper confidence limit of the 728 MAPEt,u values at the GA reference conditions, calculated across all road elements in Table 1. Accordingly, when the parameter changes result in a MAPE exceeding the statistical 85th upper limit, this indicates that the vibration outcomes are statistically different from the default reference bicycle characteristics. Therefore, it can be assumed that the effect of the variability of bicycle characteristics is statistically higher than the overall variability and experimental noise observed in multiple data collection processes.

As a general comment to the MAPE plots, the sensitivity analysis confirmed that the Genetic Algorithm (GA) consistently identified the optimal values, always placed at the minimum of the MAPE blue curves.

Our analysis revealed that the suspension parameters of the fork where the sensor is placed (i.e., K2, C2) have a significant impact on the RMS results, with the potential to introduce errors that are statistically higher than the general noise observed in RMS variability during repeated runs. In this study, the designation '2′ refers specifically to the rear tire and suspension where the sensor was mounted. Furthermore, our analysis demonstrates that the user's weight, which exhibits a strong negative linear correlation with the RMS (R^2 = 0.95), can result in a moderate rate of RMS variability of approximately 10% across a broad user weight range from 35 kg (W60) to 95 kg (W120). This suggests that the influence of weight variability on RMS estimation is comparatively less significant than other uncontrolled factors.

Additionally, comparisons between infield data and simulated data show differences primarily in the occurrence of RMS peaks, which sometimes are significantly higher in simulations than in actual measurements or vice versa, as shown in Fig. 15.

Table 5
Statistical analysis of various configurations.

Configuration	R	Std error	F	Sig
Con 1	0.552	1.502	310.623	0<.001
Con 2	0.523	2.005	267.712	0<.001
Con 3	0.552	1.512	311.094	0<.001
Con 4	0.538	1.620	289,635	0<.001

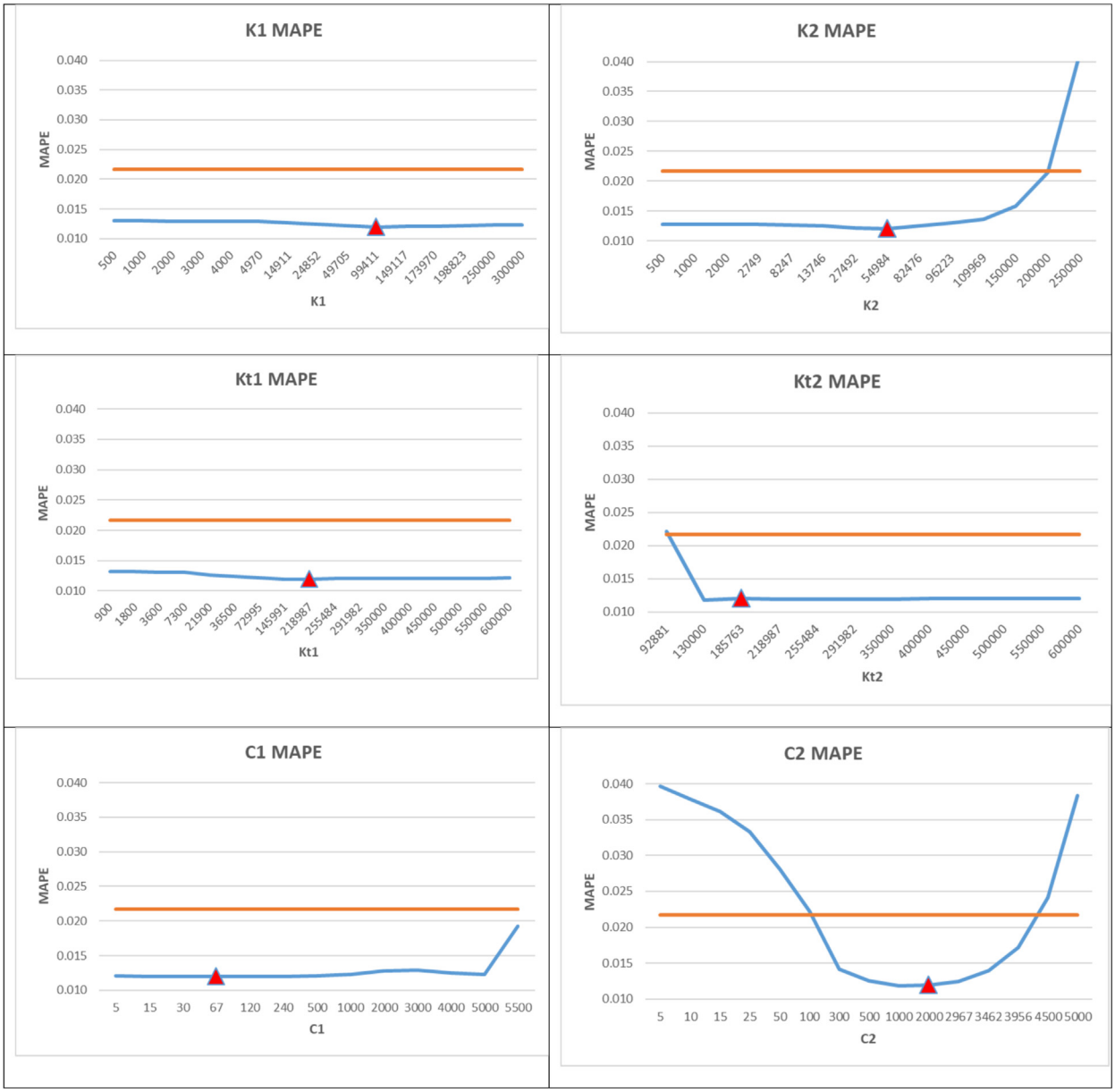


Fig. 13. Sensitivity Analysis of Bicycle Dynamic Parameters (orange line is the 85th confidence limit).

7. Discussion

While the main focus of the study is the variability in bicycle characteristics, the results also offer valuable insights into the various aspects of crowdsourced data collection for pavement monitoring including pavement characteristics and user behavior. They highlight the key issues and considerations for future research and practical applications in this field.

7.1. Bicycle characteristics

The results have confirmed that the suspension characteristics are relevant for RMS estimation (Sattar et al. 2018; Al-Sabaeei et al. 2024). The observed variability in RMS measurements related to the suspension parameters where the sensor is placed, C2 in our study, underscores the potential impact of specific bicycle characteristics and sensor configurations. When considering the placement of smartphones on the handler bar, the impact of the fork characteristic can be significant if a spring suspension system is present (Fig. 13). Noting that standard city bikes are typically built with rigid front and rear

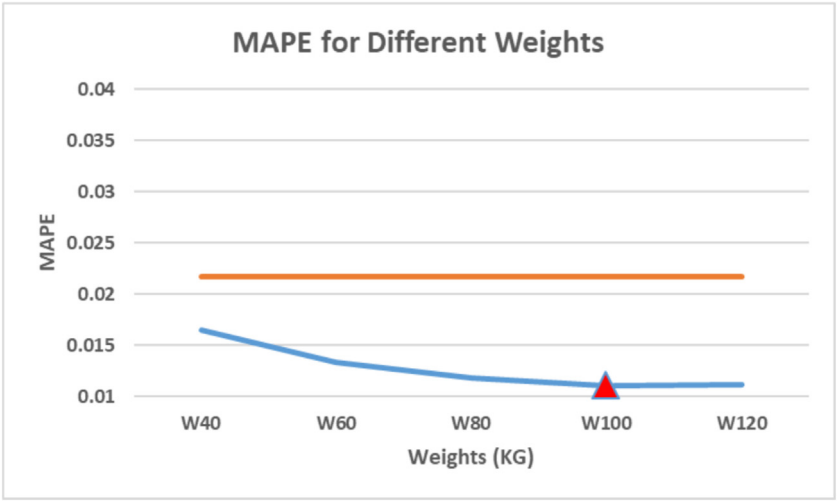


Fig. 14. Sensitivity analysis based on users' weight.

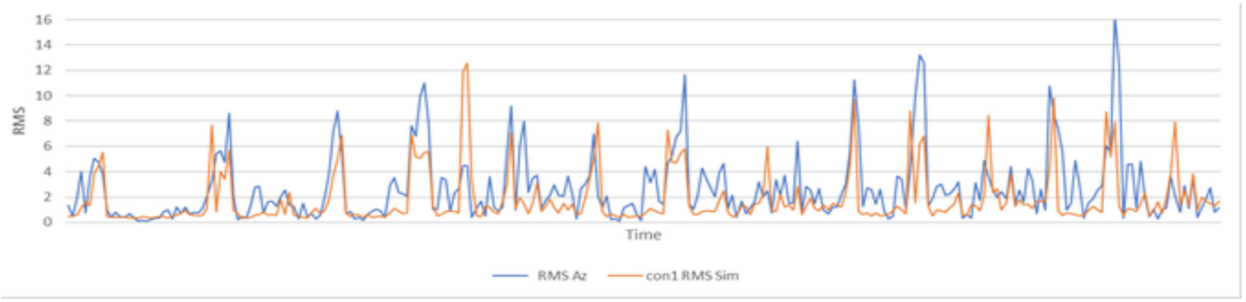


Fig. 15. Example of measured and simulated RMS in one test.

forks, which prioritize lightness and efficiency over shock absorption, e-bikes, trekking bikes, and mountain bikes often feature spring suspensions in the front fork. For the e-bike used in the study, which had a spring front fork, the sensor was placed over the rigid rear fork. However, the results indicated that in the presence of a spring/damped fork, only a significant variability in the damping coefficient of the bicycle suspension can affect the RMS estimation. Ultimately, selecting only bicycles with rigid suspension or positioning the sensor on the rear side of the bicycle frame (e.g., instrumented bike sharing vehicles) is the best solution to obtain more homogeneous RMS data. In other conditions, such as when users place the smartphone on the handlebar of a bicycle with front suspension, the bicycle typology should be known (e.g., at user registration) and considered in the data collection. Another factor to consider is the orientation of the sensor. The study ensured that the sensor was placed so that the z-axis was oriented vertically. If the sensor (e.g., smartphone) is not oriented correctly, the vertical acceleration will be distributed over the three cartesian axes, affecting the accuracy of RMS estimations (Nuñez et al. 2020). The 3 axles accelerometers available in all the smartphones can be used to compute angles through the known relationship with the gravity vector, providing a method accurate enough to compute the vertical components of the accelerations used for RMS estimation (Fisher 2010; McClintock et al. 2024). Lastly, the issue of localization accuracy was crucial in the study, with a great deal of effort put into locating the acceleration data and peak alignment over a short 10 m stretch of road to achieve as accurate a data comparison as possible. However, reporting data on very short sections emphasizes the localization error. Even with the GNSS augmentation already available in the last generation of smartphones (e.g., SBAS, multi-constellation), the GNSS position accuracy in the constrained urban canyon can be around 20 m without multi-sensor fusion. Using a longer road section (e.g., 50–100 m) as a reference base for RMS, can mitigate the errors in the data localization, while maintaining a good mapping distribution of pavement conditions (Miah et al. 2020; Shtayat et al. 2020; Al-Sabaeei et al. 2024).

Additionally, the study addressed the variation in bicycle tire typology (i.e. Kt parameter). The standard width of tires is 23–45 mm, which can be increased to 75 mm for mountain bike tires. Nowadays, even some city e-bikes are fitted with fat tires ranging from 90 to 127 mm wide. The bike used in the study was equipped with a tire 700/45 with a width of 45 mm. This dimension fits well within the range of standard tires; therefore, it can be assumed that the results are suitable to consider the variability of Kt as representative both of different types of tires and of inflation pressure. However, bicycles

equipped with extra-large fat tires may have a different performance (Monz & Kulmatiski 2016; Dressel & Sadauckas 2020) and therefore should be further studied.

The study also observed the effect of user weight on RMS variability during the experimental test, where the overall weight of the bicycle (frame plus cyclist) was about 100 kg. The literature review doesn't report any research on this topic, which is sometimes mentioned as an open question that should be investigate (Sattar et al. 2018; Al-Sabaeei et al. 2024). Our results show that the strong correlation between user weight and RMS values is mitigated by the lower variability it introduces, compared to other factors, within the average weights of the user population and bicycles (Fig. 14). Although the effect of weight on RMS variability was less than the noise from other bicycle factors, knowing the weight of the rider can easily be used to adjust the RMS data.

7.2. Pavement characteristics

Our experiment was carried out on asphaltic concrete pavements. Surface distresses include longitudinal, transverse and alligator cracks of varying extent and severity, potholes and irregularities, giving the 153 road samples a PCI ranging from 25 (very poor) to 100 (very good) (Fig. 2). Therefore, it can be assumed that the pavement characteristics are representative of a wide range of distresses and conditions for asphalt concrete pavements. Although asphalt concrete pavements represent the main pavement typology in the road network, other pavement typologies such as stone or concrete block pavement should be considered in future studies. This type of pavement is expected to have a different surface texture and poorer vibration quality. Some studies have compared vibrations in both asphalt and stone pavements, (Zoccali, Loprencipe & Galoni 2017; Xie et al. 2019) highlighting that abnormal values of vertical accelerations are particularly evident in the latter, and intermittent vibrations caused by the gaps between paving stones are more unpleasant than higher regular vibrations caused by rougher asphalt or gravel surfaces (Niska et al. 2024). In practical applications, different data analysis must be performed for the different pavement typologies in the urban road network, which are usually known and reported in a reference digital map (e.g., Open Street Map) (Zoccali et al. 2017; Xie et al. 2019).

The analysis of RMS peaks (Fig. 15) revealed the potential for uncertainty in the data due to the presence of high-severity localized distresses such as potholes and bumps in repeated runs. This highlights the need for further refinement of the pavement monitoring process, particularly with the introduction of crowdsourced data cleaning (Shtayat et al. 2022) and (Al-Sabaeei et al. 2024) and specific anomaly detection algorithms, as proposed by (Lekshmipathy et al. 2021) and (Sattar et al. 2018).

In terms of IRI prediction, the literature reveals a predominance of studies that report a linear correlation between vertical accelerations and IRI (Múčka 2017). This finding suggests a proportional relationship between the accuracy of IRI prediction and that of the acceleration measures. Utilizing models tailored to bicycles (Cafiso et al. 2022) and (Alatoom et al. 2024), we have determined a variability of + 15–20% in IRI prediction within the range of bicycle parameters below the orange threshold identified in Fig. 13, which is consistent with the accuracy of the models in these studies.

7.3. User behaviour

It should be noted that there is considerable variation in speed and trajectory among bicyclists. In our study each test was simulated and compared at the actual cycling speed in each 10 m segment, without the need for any speed adjustment. While in the practical application to ensure uniformity in the data, it is necessary to normalize the RMS for speed variability (Sattar et al. 2018). In practical application, the speed can be recorded with good accuracy with the GNSS signal, and the RMS should be normalized to a reference speed. However, the RMS variability with speed confirmed a non-perfect linear relationship (Cafiso et al. 2022) which can be accentuated by the variability of bicycle characteristics. Our study indicates that the best agreement between measured and simulated values has been achieved at a moderate speed (i.e. 17 km/h). At high or low speeds, the non-linearity is emphasized by a damping/spring fork. Moreover, to collect suitable data at higher speeds the sampling rate should be increased (e.g. 100 Hz), but with an increase in data storage and power consumption. Therefore, it is suggested to use a reference speed of 5 m/s (18 km/h) for normalization, excluding data collected at extremely high or low speeds (e.g., 4 < speed < 6 m/s), a range of values that is consistent with the results of Shtayat (Shtayat et al. 2022), covering also about 70% of bicycle free flow speeds (Pazdan & Kiec 2023).

Despite the trajectory of the path during the test being constrained to a narrow width to the right side of the lane, the variability of the bicycle's trajectory is a condition that cannot be totally eliminated in these studies. Indeed, it is considered one of the factors contributing to noise in the acceleration recordings. Trajectory dispersion was evaluated as one of the main sources of variance in RMS, noting that the wheel path of a bicycle is very narrow, and the trajectory of the cyclist varies over the pavement surface. Some studies (Shtayat et al. 2022)reported that over 10–20 repetitions there is a good agreement in the average values of RMS, among different users. In the perspective of crowdsourced data collection, it is expected that several tens of records can be aggregated within a limited reference time frame (e.g., 15 days).

8. Conclusion

The authors conducted a comprehensive comparison between measured and simulated vertical accelerations to calibrate a numerical model, which was then used to investigate the influence of bicycle characteristics on the variability of smartphone acceleration data used for the urban network screening of pavement conditions and user comfort. This is the main paper's contribution to the field, addressing a gap identified by many authors and in the most recent literature reviews (e.g., Sattar et al. 2018; Al-Sabaeei et al. 2024). The study is not directly intended to propose a new method to detect pavement anomalies in the extensive literature on this topic. Although it provides novel information for such applications to improve the quality of crowdsourced bicycle vibration measurements for pavement condition evaluation.

Furthermore, the methodological approach, which is based on experimental data and model simulation tuned by a genetic algorithm, offers an original contribution to this field and can serve as a reference for further studies.

The half-car model was chosen to simulate the bicycle's response to the road profile, and experimental accelerations were gathered at various speeds on asphalt pavements with a wide range of surface conditions. This model was integrated with a genetic algorithm on MATLAB Simulink to optimize the six dynamic bicycle parameters. Previous studies such as Gillespie (Gillespie 1992), Wong (Wong 2022), and Rizelioğlu & Yazıcı (Rizelioğlu & Yazıcı 2024), alongside our findings, have demonstrated that this model is well-suited for analyzing vertical accelerations induced by road roughness. The model was also integrated into a more complex optimization procedure, involving multiple iterations and the tuning of model parameters with the Genetic Algorithm (GA). The GA performed well in a complex, non-linear, multi-variable system, aiming to identify the best set of bicycle dynamic parameters that minimized the MAPE of RMS between measured and simulated estimations. The study compared different dataset scenarios for training and validation, as well as different parameter settings, with the objective of achieving a stable optimization result. The most effective approach was identified as selecting GA settings and training the model on a wider range of surface conditions (Scenario A) and optimizing the full set of parameters (Con 1). This approach enhanced the model's adaptability to the experimental data. Following calibration, the model has the capability to generalize results by varying the characteristics of the bicycle. This approach offers valuable insights into the generalizability of the findings on the potential impact of different bicycle configurations on vibration data collections and studies.

Since the main research question was whether bicycle characteristics should be considered in the analysis of vertical accelerations for pavement monitoring, the results of several simulations allow us to conclude that the overall results may be more influenced by the positional accuracy and trajectory dispersion in multiple runs than by the bicycle characteristics and user weights, with some restrictions. Bicycles should be of a normal type with rigid suspension and standard tires and average total weight, including cyclist, in the wide range of 60 and 120 kg. If the bicycle is equipped with front spring suspension or fat tires, and the smartphone is placed on the front handlebar, the suspension and tire characteristics should be considered in the RMS data collection. Additionally, the results of this study cannot be extended to pavement types other than asphalt concrete.

Furthermore, the results confirmed the opportunity to apply different data analyses for the detection of pavement anomalies such as potholes and bumps based on the RMS of the peaks that should be excluded from the overall data analysis of pavement surface. Moreover, as the data is collected from users traveling at different speeds, the RMS adjustment for speed variability must be applied. Based on the results, a reference speed of 5 m/s is suggested, including only data collected in the speed range of $4 \leq$ speed ≤ 6 m/s, in accordance with the findings of Shtayat et al. (Shtayat et al. 2020).

These observations limit the generalization of our results, highlighting the need for further studies and more complex models to extend the analysis to any type of bicycle, pavement, and speed. However, it is important to note that in practical applications, a simple and reliable model based on average RMS estimation on long sections of 50–100 m can be used to monitor asphalt pavement conditions. Despite the specified selections of bicycle types and speeds, it is easy to collect homogeneous data with multiple runs (e.g. more than 20) within a reference period of 15 days. This provides useful data which can be used for the application of crowdsourcing data collection for pavement condition monitoring.

Funding

This work has been partially supported by the European Union – Next Generation EU PNRR, Spoke 9 "Digital Society & SMart Cities" of ICSC – Centro Nazionale di Ricerca in High Performance-Computing, Big Data and Quantum Computing, (PNRR-HPC, CUP: E63C22001000006), and the "Sicilian MicronanoTech Research And Innovation Center – SAMOTHRACE" (PNRR-CUP E63C22000900006).

Authors' contributions

The authors confirm contribution to the paper as follows: study conception and design: Cafiso, Pappalardo; data collection, analysis and interpretation of results, draft manuscript preparation: Cafiso, Ghaderi, Pappalardo; model training and validation, Caponnetto, Ghaderi, Pappalardo; Funding: Cafiso, Pappalardo. All authors reviewed the results and approved the final version of the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Omid Ghaderi: Writing – review & editing, Writing – original draft, Visualization, Validation, Formal analysis, Data curation. **Salvatore Cafiso:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition, Conceptualization. **Riccardo Caponetto:** Validation, Data curation. **Michele Mangiameli:** Writing – review & editing. **Giuseppina Pappalardo:** Writing – original draft, Validation, Methodology, Data curation, Conceptualization.

References

Ahmad Feizi Jun-Seok Oh, V.K., Joo, S., 2020. Cycling environment analysis by bicyclists' skill levels using instrumented probe bicycle (IPB). *Int. J. Sustain. Transp.* 14 (9), 722–732.

Alatoom, Y.I., Zihan, Z.U., Nlenanya, I., Al-Hamdan, A.B., Smadi, O., 2024. A sequence-based hybrid ensemble approach for estimating Trail pavement roughness using Smartphone and bicycle data. *Infrastructures* 9 (10).

Allouch, A., Koubâa, A., Abbes, T., Ammar, A., 2017. RoadSense: Smartphone application to estimate road conditions using accelerometer and gyroscope. *IEEE Sens. J.* 17 (13), 4231–4238.

Al-Sabaeei, A.M., Souliman, M.I., Jagadeesh, A., 2024. Smartphone applications for pavement condition monitoring: a review. *Constr. Build. Mater.* 410, 134207.

Alsajri, A., Steiti, A., 2024. Intrusion detection system based on machine learning algorithms:(SVM and genetic algorithm). *Babylonian Journal of Machine Learning* 2024, 15–29.

ASTM International, 2012, *ASTM E867. Standard Terminology Relating to Vehicle-Pavement Systems.*

ASTM International, 2020, *ASTM D6433-20 Standard Practice for Roads and Parking Lots Pavement Condition Index Surveys.*

Basavaraju, A., Du, J., Zhou, F., Ji, J., 2020. A machine learning approach to road surface anomaly assessment using smartphone sensors. *IEEE Sens. J.* 20 (5), 2635–2647.

Berge, S.H., Hagenzieker, M., Farah, H., de Winter, J., 2022. 'Do cyclists need HMIs in future automated traffic? An Interview Study'. *Transportation Research Part F: Traffic Psychology and Behaviour* 84, 33–52.

Bourne, J., Cooper, A., Kelly, P., Kinnear, F., England, C., Leary, S., Page, A., 2020. The impact of e-cycling on travel behaviour: a scoping review. *J. Transp. Health* 19, 100910.

Bulsink, V.E., Doria, A., Belt, D. Van De and Koopman, B., 2015, 'The effect of tyre and rider properties on the stability of a bicycle'. Adv. Mech. Eng., 7(12).

Cafiso, S., Graziano, A.D., Goulias, D.G., D'Agostino, C., 2019. Distress and profile data analysis for condition assessment in pavement management systems. *Int. J. Pavement Res. Technol.* 12 (5), 527–536.

Cafiso, S., Graziano, A.D., Marchetta, V., Pappalardo, G., 2022. Urban road pavements monitoring and assessment using bike and e-scooter as probe vehicles. *Case Stud. Constr. Mater.* 16.

(CEN), E.C. for S., 2006, 'Road and airfield surface characteristics—Test methods. Part 5: Determination of longitudinal unevenness indices'.

Chou, C.-P., Lee, W.-J., Chen, A.-C., Wang, R.-Z., Tseng, I.-C., Lee, C.-C., 2015. Simulation of bicycle-riding smoothness by bicycle motion analysis model. *J. Transp. Eng.* 141, (12)4015031.

Donoho, D.L., 1995. De-noising by soft-thresholding. *IEEE Trans. Inf. Theory* 41 (3), 613–627.

Dressel, A., Sadauckas, J., 2020. Characterization and modelling of various sized mountain bike tires and the effects of tire tread knobs and inflation pressure. *Appl. Sci.* 10 (9), 3156.

Eiben, A.E., Smith, J.E., 2015. *Introduction to evolutionary computing*. Springer.

European Parliament and of the Council, 2002, Directive 2002/44/EC-the minimum health and safety requirements regarding the exposure of workers to the risks arising from physical agents (vibration) (sixteenth individual Directive within the meaning of Article 16(1) of Directive 89/391/EEC) .

Fisher, C.J., 2010. 'Using an accelerometer for inclination sensing', *AN-1057*. Application note, Analog Devices, pp. 1–8.

Gao, J., Sha, A., Huang, Y., Hu, L., Tong, Z., Jiang, W., 2018. Evaluating the cycling comfort on urban roads based on cyclists' perception of vibration. *J. Clean. Prod.* 192, 531–541.

Gillespie, T.D., 1992, Fundamentals of Vehicle Dynamics.

Goldberg, D.E., Holland, J.H., 1988. Genetic algorithms and machine learning. *Mach. Learn.* 3, 95–99.

Haupt, R.L. and Haupt, S.E., 1998, Practical Genetic Algorithms.

Holland, J.H., 1975. Adaptation in natural and artificial systems.

International Organization for Standardization, 2019. Characterization of pavement texture by use of surface profiles–Part 1: Determination of mean profile depth.

International Organization for Standardization (ISO), 1997, UNE2631_Mechanical Vibration and Shock - Evaluation of Human Exposure to Whole-body Vibration - Part 1: General Requirements, International Organization for Standardization.

International Organization for Standardization (ISO), 2001, ISO 5349- :2001 Mechanical vibration Measurement and evaluation of human exposure to hand-transmitted vibration.

International Organization for Standardization (ISO), 2018, ISO 2631:2018 Mechanical vibration and shock Evaluation of human exposure to whole-body vibration.

Jlifi, B., Ferjani, S., Duvallet, C., 2025. A genetic algorithm based three HyperParameter optimization of deep long short term memory (GA3P-DLSTM) for predicting electric vehicles energy consumption. *Comput. Electr. Eng.* 123, 110185.

Kotani, Y., Tomiyama, K., Nishigai, H., Sasaki, K., Yamaguchi, Y., Moriishi, K., 2023. Characterization of surface roughness on the basis of vibration response corresponding to the ride quality of micromobilities. In: *Roads and Airports Pavement Surface Characteristics*. CRC Press, pp. 355–364.

Kumar Manilal Sidhant, 2019, Bicycle ride comfort evaluation and optimization MRN422 Final Report.

Lekshmipathy, J., Velayudhan, S., Mathew, S., 2021. Effect of combining algorithms in smartphone based pothole detection. *Int. J. Pavement Res. Technol.* 14 (1), 63–72.

Lin, Z., Fan, W. (David), 2020. Modeling bicycle volume using crowdsourced data from strava smartphone application. *Int. J. Transp. Sci. Technol.* 9 (4), 334–343.

Litzenberger, S., Christensen, T., Hofstätter, O., Sabo, A., 2018. Prediction of road surface quality during cycling using smartphone accelerometer data. *Proceedings* 2 (6).

López-Molina, M., Llopis-Castelló, D., Pérez-Zuriaga, A.M., Alonso-Troyano, C., García, A., 2023. Skid resistance analysis of urban bike lane pavements for safe micromobility. *Sustainability* 15 (1).

Mahroogi, F.O., Narayan, S., Abudbaker, S., Sakthivel, R., Kaisan, M.U., Alammari, Y., 2019. *Introduction to automotive engineering*. John Wiley & Sons.

Mallat, S., 2008. *A wavelet tour of signal processing.* The Sparse Way.
Mason, G.S., Larson, M.J., Deng, R., Reed, D., Pahlmeyer, M., Wright, N.C., Wu, Z.P., Yahata, J., 2016. A robust low cost device for measuring road induced vibrations. *Journal of Science and Cycling* 5, 13–17.
McClintock, F.A., Callaway, A.J., Clark, C.J., Williams, J.M., 2024. Assessing the impact of sensor orientation on accelerometer-derived angles: a systematic analysis and proposed error Reduction. *Appl. Sci.* 14 (2), 842.
Miah, S., Milonidis, E., Kaparias, I., Karcanias, N., 2020. An innovative multi-sensor fusion algorithm to enhance positioning Accuracy of an instrumented bicycle. *IEEE Trans. Intell. Transp. Syst.* 21 (3), 1145–1153.
Mitchell, M., 1996, An introduction to genetic algorithms.
Monz, C., Kulmatiski, A., 2016. The emergence of "fat bikes" in the USA: trends, potential consequences and management implications. *J. Outdoor Recreat. Tour.* 15, 20–25.
Moore, J.K., Hubbard, M., Schwab, A.L. & Kooijman, J., 2010. Accurate Measurement of Bicycle Parameters.
Múčka, P., 2017. Road roughness limit values based on measured vehicle vibration. *J. Infrastruct. Syst.* 23, (2)04016029.
Murgano, E., Caponetto, R., Pappalardo, G., Cafiso, S.D., Severino, A., 2021. A novel acceleration signal processing procedure for cycling safety assessment. *Sensors* 21 (12), 4183.
Newzoo, 2019, Newzoo Global Mobile Market Report 2019.
Niska, A., Sjögren, L., Andrén, P., Weber, C., Jong, T.D., Fyhri, A., 2024. Determination of riding comfort on cycleways using a smartphone application. *Journal of Traffic and Transportation Engineering (English Edition)*.
Nuñez, J.Y.M., Bisconsini, D.R., Rodrigues da Silva, A.N., 2020. Combining environmental quality assessment of bicycle infrastructures with vertical acceleration measurements. *Transp. Res. A Policy Pract.* 137, 447–458.
Parastoo, K., David, L., Amarda, S., Shelley, S., 2022. Pavement distress recognition via wavelet-based clustering of Smartphone accelerometer data. *J. Comput. Civ. Eng.* 36, (4)04022007.
Pazdan, S., Kiec, M., 2023. Bicycle free-flow speed estimation based on GPS data – comparison of bikesharing system and strava data. *Archives of Transport* 68 (4), 77–90.
Percival, D.B., Walden, A.T., 2000. *Wavelet Methods for Time Series Analysis*, vol. 4. Cambridge University Press.
Qian, X., Moore, J.K., Niemeier, D., 2020. Predicting bicycle pavement ride quality: sensor-based statistical model. *J. Infrastruct. Syst.* 26, (3)4020033.
Rizelioğlu, M., Yazıcı, M., 2024. New approach to determining the roughness of bicycle roads. *Transp. Res. Rec.* 2678 (1), 781–793.
Sattar, S., Li, S., Chapman, M., 2018. *Road surface monitoring using smartphone sensors: a review*. *Sensors (switzerland)* 18 (11).
Shtayat, A., Moridpour, S., Best, B., Shahriar Rumi, M., 2020. *Using a smartphone software and a regular bicycle to monitor pavement health statues.* In: *Proceedings of the 2020 2nd International Conference on Robotics Systems and Vehicle Technology*, pp. 121–126.
Shtayat, A., Moridpour, S., Best, B., 2022. Using e-bikes and private cars in dynamic road pavement monitoring. *Int. J. Transp. Sci. Technol.* 11 (1), 132–143.
Sidess, A., Ravina, A., Oged, E., 2021. A model for predicting the deterioration of the pavement condition index. *Int. J. Pavement Eng.* 22 (13), 1625–1636.
Souza, V.M.A., Cherman, E.A., R.R.G., & S.R.A. (2017). Towards Automatic Evaluation of Asphalt Irregularity Using Smartphone's Sensors. In N. Adams, A. Tucker, & W.D. Kieran (Eds.), Advances in Intelligent Data Analysis XVI (pp. 322–333). Springer International Publishing, Cham.
Souza, V.M.A., Giusti, R., Batista, A.J.L., 2018. Asfault: a low-cost system to evaluate pavement conditions in real-time using smartphones and machine learning. *Pervasive Mob. Comput.* 51, 121–137.
Springer, M. and Ament, C., 2020, A mobile and modular low-cost sensor system for road surface recognition using a bicycle, 2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), 360–366, IEEE.
Stamatiadis, N., Cafiso, S., Pappalardo, G., 2019. *A Comparison of bicyclist attitudes in two urban areas in USA and Italy*. In: Nathanail, E.G., Karakikes, I.D. (Eds.), *Data Analytics: Paving the Way to Sustainable Urban Mobility*. Springer International Publishing, Cham, pp. 272–279.
Whipple, F.J.W., 1899. 'The stability of the motion of a bicycle.'. *Quart. J. Pure Appl. Math.* 30, 312–348.
Wong, J.Y., 2022. *Theory of ground vehicles*. John Wiley & Sons.
World Road Association, 1987. *Report of the committee on surface characteristics. Proceeding of XVIII World Road Congress*, 13–19.
Xie, N., Li, H., Zhao, W., Ni, Y., Liu, C., Zhang, Y., Xu, Z., 2019. Measurement of dynamic vibration in cycling using portable terminal measurement system. *IET Intel. Transport Syst.* 13 (3), 469–474.
Yaqoob, S., Cafiso, S., Morabito, G., Pappalardo, G., 2023. Deep transfer learning-based anomaly detection for cycling safety. *J. Saf. Res.* 87, 122–131.
Zhu, S., Zhu, F., 2019. Cycling comfort evaluation with instrumented probe bicycle. *Transp. Res. A Policy Pract.* 129, 217–231.
Zoccali, P., Loprencipe, G., Galoni, A., 2017. Sampietrini stone pavements: distress analysis using pavement condition index method. *Appl. Sci.* 7 (7), 669.