



Contents lists available at ScienceDirect

Transportation Research Part C

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



Large-scale pavement roughness measurements with vehicle crowdsourced data using semi-supervised learning

Chenglong Liu ^{a,b}, Difei Wu ^a, Yishun Li ^a, Yuchuan Du ^{a,b,*}

^a Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, 4800 Cao'an Road, Shanghai 201804, China

^b Shanghai Engineering Research Center of Urban Infrastructure Renewal, Shanghai 200032, China



ARTICLE INFO

Keywords:

Pavement roughness
Large-scale measurement
Semi-supervised learning
Power spectral density
Full-car simulation

ABSTRACT

Rapid measurements of large-scale pavement roughness have long been a hot topic in pavement condition evaluation and maintenance. Most traditional methods rely on dedicated devices, such as laser, Lidar and so on, which should be set up on customized vehicles. With the rapid development of sensing technology, vehicles owned by the general public are empowered with the ability to collect vibration measurements themselves. This crowdsourced dataset is convenient, extensive coverage, inexpensive, and has high sampling frequency, making it a suitable source for large-scale pavement roughness evaluation. However, vehicle information is missing for these data due to privacy protection, which renders them quite difficult to directly use with traditional model-based methods. Thus, in this paper, we propose a semi-supervised learning (SSL) model to deal with the problem of incomplete data and multi-vehicle data fusion. A mathematical derivation of the ‘international roughness index’ (IRI) using in-car vibrations is established. Furthermore, given the multi-vehicle scenario, a self-training model is designed to iteratively estimate IRIs in a roadway network. Both the confidences of the vehicle parameters and IRI estimation are considered in the algorithm to improve its reliability and robustness. A full-car simulation model is constructed to verify the effectiveness of the proposed model. The results show that the overall relative error is less than 10% for 50 road sections in the network, which is a significant improvement compared to traditional multi-vehicle average models. The errors of the SSL model are found to be significantly dependent on the iteration order. Based on the proposed model, the coupled impact of the sampling rate and vehicle quantity on the model’s accuracy is further discussed. The proposed approach provides new insights into large-scale pavement roughness measurements.

1. Introduction

Pavement roughness is a fundamental property of roads that significantly influences driving safety, comfort and the service life of roads. Sayer et al. proposed the ‘international roughness index’ (IRI) to quantify this property in 1986 (Sayer et al., 1986), which has since become well recognized. The measurement of roughness is challenging because it relies on the characteristics of the vehicle in addition to the conditions of the road (Prasad et al., 2013). Many devices have been developed to measure the IRI, such as leveling instruments, three-meter-long beams, and the General Motors research profilometer (Du et al., 2014). More advanced devices that use

* Corresponding author.

E-mail address: yedu@tongji.edu.cn (Y. Du).

precise lasers, such as laser surface testers and 3D lidar testers, have also been manufactured (Chang et al., 2006). These devices can be divided into three types: I) direct profile measurements, II) response-type road roughness measurements (RTRRMS) and III) subjective rating panels. Type I devices measure the pavement longitudinal profile directly using laser rangefinder (Wang et al., 2011). Type II devices measure the IRI by correlating the relationship between vehicular vibration and pavement roughness (Bil et al., 2015; Hudson et al., 1985). Type III devices evaluate pavement riding quality by professional panel assessment according to their subjective experience (Janoff, 1988). Table 1 shows the features of the common pavement roughness measurements. However, most traditional devices are either time-consuming, labor-intensive, and/or expensive, rendering them unsuitable for measurements over extensive ranges and at high sampling frequencies.

Advanced pavement maintenance techniques place higher demands on pavement roughness data. Traditional sampling measurement methods cannot meet the current maintenance requirement. The large-scale and high-frequency pavement condition evaluation has become a predominant trend (Mahmoudzadeh et al., 2019). Rapid and economical measurements of pavement roughness have long been the subject of research. Several inexpensive alternatives to lasers or Lidar devices have been developed. In the 1990s, researchers began developing indirect methods to measure pavement roughness (Gillespie, 2001). A vehicle's vibration is the most intuitive response to pavement roughness, and several devices, such as the bump integrator (Jordan and Young, 1980) and the Bureau of Public Roads (BPR) roughometer (Hudson and Hain, 1961), have been proposed to evaluate it using in-car vibrations. Such devices measure the vertical acceleration of the vehicle and calculate the double integral of acceleration to derive the change in elevation (Islam et al., 2014). However, this integration is sensitive to sudden changes, and pavement roughness is discontinuous when determined using a discrete sample rate. Apart from the direct integration of vibrations, frequency-domain analyses are another commonly used methodology. As a vehicle's acceleration has the feature of a stationary random signal, patterns appearing in the frequency domain are more stable than those in the time domain. For example, Douangphachanh et al. proposed a fast Fourier transform (FFT)-based method to estimate IRIs at different sample rates (Douangphachanh and Oneyama, 2013) using vehicle-mounted smartphone data. Chen et al. developed a viscoelastic model to estimate IRIs based on acceleration measured using a smartphone. They found a quasi-linear correlation between the outputs of the model and the IRIs (Chen et al., 2017). Sandamal and Pasindu defined an estimated IRI (eIRI) based on the peak and root mean square (RMS) vibration collected by smartphones (Sandamal and Pasindu, 2020). Although the correlation between eIRI and IRI is solid, the RMS is significantly affected by vehicle parameters, mobile phone positions and driving speed, etc., resulting in huge calibration pressure. To address this issue, Jeong et al. proposed a convolutional neural network for pavement roughness assessment with calibration-free vehicle dynamics, which provides a new insight for large-scale IRI estimation (Jeong et al., 2020). The algorithm performs well within the experimental conditions. However, the types of vehicles, the distribution of IRIs in real case are quite complicated, making it hard to use an end-to-end neural network. Zhao et al. calculated the RMS of a vehicle's sprung mass by multiplying a transfer function by the original vibrations (Zhao et al., 2017). Du et al. proposed a power spectral density (PSD)-based algorithm to calculate IRIs using in-car vibrations at multiple positions (Du et al., 2016). Wei et al. used wavelet analysis to identify the features of the profile of pavement roughness in both the frequency and distance domains (Wei et al., 2005). Bridgelall created a new index named the 'road impact factor' (RIF) to predict IRIs using connected vehicles (Bridgelall, 2013). These response-based measurement methods are more economical and less time-consuming than laser-based devices. However, most still require the installation of dedicated devices and the vehicle parameters need to be calibrated via complex experiments, which makes them hard to apply to large-scale pavement roughness detection.

With the rapid development of sensing technology, vehicles used by the general public are empowered with the ability to perceive and collect vibration data themselves. For instance, most vehicles have a vibration diagnosis system to monitor their conditions, while some vehicles use in-car accelerometers to improve comfort and switch between different driving modes (Du et al., 2018). These vibration data do not rely on dedicated devices and can be easily acquired by the vehicles. There are thousands of GB of data generated by vehicles every day (Nelson, 2016). These crowdsourced data are convenient, extensive coverage, inexpensive and have high sampling frequencies, which makes them a potentially suitable tool for large-scale pavement roughness measurements. Mirtabar et al. developed a crowdsourcing-based system for computing the IRI, which requires obtaining vehicular parameters in advance. The pavement profile is obtained by double integration of Z-axis acceleration data (Mirtabar et al., 2020). Medina et al. designed a Monte

Table 1
Features of different pavement roughness measurement methods.

Measurement	Type	ACC	EFF	STA	LAB	PRI	CAL
3-Meter long beam (Janoff, 1988)	I	+++	+	++	+++	+	×
Level (Janoff, 1988)	I	+++	+	++	+++	+	×
Bump integrator (Nguyen et al., 2019)	I/II	++	++	+	+	+	✓
Laser road surface tester (Bitelli et al., 2012)	I	+++	+++	+++	+	+++	✓
GMR profilometer (McCann and Nguyen, 2007)	I	++	++	++	+	++	✓
LiDAR (Alhasan et al., 2017)	I	+++	+++	+++	+	+++	✓
Depth Camera (Mahmoudzadeh et al., 2019)	I	++	+	+++	+	++	✓
Mays Meter (Gillespie, 1980)	II	+	++	++	++	++	✓
Road Unevenness Recorder (Srivastava and Nanda, 2001)	II	++	+++	++	+	+++	✓
ARRB Roughometer (Al-Rousan et al., 2010)	II	++	+++	++	+	+	✓
Smartphone Probe Car (Yi et al., 2015)	II	++	+++	++	+	+	✓

*+ indicates the level of different indexes; +(bold +) means this method performs the best in terms of the corresponding index.

*ACC = Accuracy, EFF = Efficiency, STA = Stability, LAB = Labor-cost, PRI = Price, CAL = Calibration.

Carlo and probabilistic point estimate test to simulate pavement roughness, which proved that the crowdsourced smartphone-based roughness measurement can be an alternative method to evaluate pavement conditions (Medina et al., 2020). Yi et al. developed a crowdsourced-based pavement monitoring system named smartphone probe car (SPC), it can detect and assess the road surface anomalies such as potholes and bumps (Yi et al., 2015). This method provides a very effective tool to evaluate pavement performance. However, it is not able to estimate the IRI as the parameters of the vehicles equipped with smartphones are missing. Similarly, Bose et al. invented a D&RSense system to collect the vehicle's driving patterns and road anomalies, but cannot quantitatively estimate the IRI as well (Bose et al., 2018). Chen et al. further explored the correlation between crowdsourced vehicle vibration data and IRI using the Gaussian Mixture Model and created the CRSM system (Chen et al., 2013). This system assesses the pavement roughness level instead of the exact values. Alessandroni et al proposed a collaborative road surface condition monitoring method: Smartroadsense. The pavement roughness is denoted by the road roughness (RI), which is calculated by averaging the power of the prediction error for the three axial components (x, y, z) (Alessandroni et al., 2014). The RI indicates the road conditions, but it is difficult to establish a direct connection with IRI. Due to data privacy laws, most crowdsourced data do not contain user information, such as vehicle type, model, physical parameters, etc. Therefore, traditional IRI-estimation models are hard to be directly applied to these unlabeled vehicular vibration data.

As the vehicular parameters and road roughness are both unknown quantities, it is very challenging to apply supervised and unsupervised learning methods to the crowdsourced data. Therefore, this paper proposes a semi-supervised learning (SSL) method. SSL is a class of machine-learning tasks and techniques that make use of unlabeled data for training, typically a small amount of labeled data, with a large amount of unlabeled data (Chapelle et al., 2009). Many SSL algorithms have been invented to deal with unlabeled data, such as generative mixture models, self-training, co-training, transductive support vector machine, and graph-based methods (Zhu, 2005). For this specific problem, we select a small portion of the roads in the network as labeled data, and measure the IRIs of these roads in advance. Then, by combining these labeled roads with many unlabeled roads, a self-training model is established to predict the IRIs for the whole network. The proposed method makes most of the crowdsourced data and has the potential to be applied to large-scale pavement roughness evaluation.

The remainder of this paper is organized as follows: Section 2 presents the large-scale pavement roughness measurement based on SSL, and introduces the full-car simulation model. Section 3 describes the numerical case study and verifies the effectiveness of the proposed model. Finally, Section 4 presents the conclusions of this study.

2. Methodology

2.1. Relationship between the IRI and vehicular vibrations

2.1.1. Kinetic model of a quarter car

The suspension of a vehicle is a complex system of tires, springs, dampers, and linkages that connects the body of the vehicle to its wheels and allows for the relative motion of different components. A mathematical model of the coordinates of the suspension system represents the vehicle's characteristics and simplifies its physical features. The most commonly employed model for suspension system is the quarter-car model. Although the quarter-car model does not adequately reflect the entire information about the vehicle, it contains the main features of the suspension system with respect to vertical vibrations, and is simple. Fig. 1 shows a schematic diagram of the quarter-car model, where the dynamic equation of the quarter-car model can be expressed as shown in Eqs. (1)–(4):

$$m_s \ddot{z}_s + C_s (\dot{z}_s - \dot{z}_u) + k_s (z_s - z_u) = 0 \quad (1)$$

$$m_u \ddot{z}_u + C_s (\dot{z}_u - \dot{z}_s) + k_s (z_u - z_s) + k_t (z_u - Y) = 0 \quad (2)$$

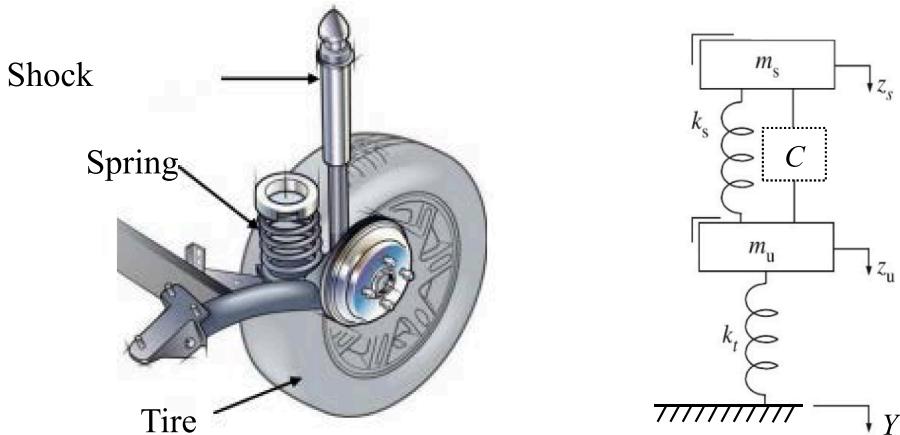


Fig. 1. Quarter-car model.

$$\ddot{z}_s = 1/m_s \cdot \left[C_s(\dot{z}_u - \dot{z}_s) + k_s(z_u - z_s) \right] \quad (3)$$

$$\ddot{z}_u = 1/m_u \cdot \left[C_s(\dot{z}_s - \dot{z}_u) + k_s(z_s - z_u) + k_t(Y - z_u) \right] \quad (4)$$

where m_s and m_u denote the sprung and the unsprung masses, respectively, k_s indicates the stiffness of the spring, c_s the damping coefficient of the shock absorber, k_t represents the stiffness of the tire, Y is the surface roughness, and z_s and z_u are the displacements of the sprung and the unsprung masses, respectively.

Although in-car vibrations are induced by changes in the elevation of the road surface, they are also influenced by the body's mass, the elasticity of the tires, and features of the suspension during the transmission process. According to the 1986 World Bank report (Sayer et al., 1986), the IRI is defined as the cumulative vertical displacement of the dynamic reaction of the suspension system within the traveling distance L and at a driving speed of 80 km/h. This is determined using the quarter-car model, as shown in Eq. (5):

$$IRI = \frac{1}{L} \int_0^L |Z_s - Z_u| dx \quad (5)$$

As shown in the above equations, the IRI is obtained by combining the impacts of the pavement roughness and the dynamic characteristics of the vehicle. However, due to such features, the IRI can neither be directly calculated by measuring the absolute elevation of the road surface nor can it eliminate the deviations caused by the natural vibrations of the vehicle.

2.1.2. Systematic excitation due to pavement roughness and suspension vibrations

Given the definition of the IRI, a quarter-car model satisfies the requirements of a linear time-invariant (LTI) system. Changes in the elevation of the road surface are regarded as system excitations, and the vibrations of the sprung and unsprung masses are considered as systematic responses. A Laplace transform (LT) is used to solve Eqs. (1) and (2). As the sprung and unsprung displacements, velocities, and accelerations are all zero under the initial conditions, the following equation holds:

$$\begin{pmatrix} m_s s^2 + C_s s + k_s & -C_s s - k_s \\ -C_s s - k_s & m_u s^2 + C_s s + k_s + k_t \end{pmatrix} \begin{pmatrix} F_s(s) \\ F_u(s) \end{pmatrix} = \begin{pmatrix} 0 \\ k_t \ell(Y) \end{pmatrix} \quad (6)$$

where $F_s(s)$ and $F_u(s)$ are the LT functions for the movement of the sprung and the unsprung masses, respectively, $s = \sigma + j\omega$, and $\ell(Y)$ represents the results of the roughness determined by the LT. The steady-state response of the suspension's vibration can be obtained by:

$$\ell(z_s(t) - z_u(t)) = (H_s(s) - H_u(s)) \cdot \ell(Y) \quad (7)$$

$$H_s(s) = \frac{k_t m_s s^2 + k_t C_s s + k_t k_s}{m_s m_u s^4 + C_s(m_s + m_u)s^3 + (k_s m_s + k_t m_s + k_s m_u)s^2 + C_s k_t s + k_t k_s} \quad (8)$$

$$H_u(s) = \frac{k_t C_s s + k_t k_s}{m_s m_u s^4 + C_s(m_s + m_u)s^3 + (k_s m_s + k_t m_s + k_s m_u)s^2 + C_s k_t s + k_t k_s} \quad (9)$$

where $H_s(s)$ and $H_u(s)$ are systematic response functions, which can be solved using Eq. (6).

Next, let $H(s) = H_s(s) - H_u(s)$; then, the vibrational response in the time domain becomes:

$$z(t) = |z_s(t) - z_u(t)| = \ell^{-1}|H(s)| \cdot Y(t) \quad (10)$$

where $\ell^{-1}|H(s)|$ is the inverse LT function.

The PSD is a measure of the mean square value of a random signal. It is frequently used to solve the excitation and response of an LTI system. LTs are performed on both sides of Eq. (10) to obtain the PSD function:

$$S_z(s) = \ell(\ell^{-1}|H(-s)|) \cdot S_Y(s) \cdot \ell(\ell^{-1}|H^T(-s)|) = |H(s)|^2 S_Y(s) \quad (11)$$

Note that a prerequisite to the above formula is that the input signal must satisfy the features of a stationary steady state. It is well-recognized that the elevation distribution of a road surface satisfies a stationary stochastic process. To simplify the calculation, the real part of the s -domain of the LT is often taken as zero. Then, $s = \omega j \omega = 2\pi f$, where ω represents the angular frequency. If $j = 1$, the LT turns into a Fourier transform. Therefore, we can obtain the relationship between the PSD of the pavement roughness based on the angular frequency, as shown in Eq. (12):

$$S_z(\omega) = |H(\omega)|^2 S_Y(\omega) \quad (12)$$

$$H(\omega) = \frac{k_t m_s \omega^2}{m_s m_u \omega^4 + C_s(m_s + m_u)\omega^3 + (k_s m_s + k_t m_s + k_s m_u)\omega^2 + C_s k_t \omega + k_t k_s} \quad (13)$$

The sequence of road roughness is randomly distributed, which follows a Gaussian probability distribution with a zero mean ergodic stochastic field. According to the differentiation property of the LT, the Laplace function $F_a(s)$ of the in-car acceleration is

$F_a(\omega) = \omega^2 F_s(\omega)$. Therefore, the response function of the acceleration is shown in Eq. (14):

$$H_a(\omega) = \omega^2 \frac{F_s(\omega)}{\ell(Y)} = \omega^2 \cdot H_s(\omega) = H(\omega) \cdot \left(\omega^2 + \frac{C_s}{m_s} \omega + \frac{k_s}{m_s} \right) \quad (14)$$

where $H_a(s)$ indicates the corresponding transfer function and $S_a(\omega)$ represents the PSD of the vehicular acceleration. Based on the definition of the transfer function, the PSD of the pavement roughness can be rewritten as:

$$S_Y(\omega) = S_a(\omega) / |H_a(\omega)|^2 = S_a(\omega) / \left| H(\omega)^2 \cdot \left(\omega^2 + \frac{C_s}{m_s} \omega + \frac{k_s}{m_s} \right)^2 \right| \quad (15)$$

According to Parseval's theorem (Arfken et al., 1999), the sum of the square of the signal in the time domain represents its energy, which is equal to its Fourier-transformed integral. The total power of the PSD represents its total average signal power, which is equal to its autocorrelation value at zero, as shown in Eq. (16):

$$\varphi_z^2 = R_z(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(\omega)|^2 / |H_a(\omega)|^2 \cdot S_a(\omega) d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{S_a(\omega)}{\left(\omega^2 + \frac{C_s}{m_s} \omega + \frac{k_s}{m_s} \right)^2} d\omega \quad (16)$$

The average power of the signal is a measure of the mean squared value of the signal. It is generally linear with its mean value for pavement roughness (Du et al., 2014):

$$E(|z_s(t) - z_u(t)|) = k \cdot \varphi_z - b \quad (17)$$

where k, b are regression coefficients. In practice, it is impossible to continuously detect the road surface elevation or vehicular vibrations; instead such data are obtained at specified sampling rates. The discrete expression of Eq. (5) is shown in Eq. (18):

$$IRI = \frac{1}{L} \int_0^L |Z_s - Z_u| dx = \frac{1}{N} \sum_{i=1}^N |z_s(i) - z_u(i)| \quad (18)$$

where N refers to the sample size. Using Eqs. (17) and (18), we can derive the final formulation:

$$IRI = k \cdot \sqrt{\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{S_a(\omega)}{\left(\omega^2 + \frac{C_s}{m_s} \omega + \frac{k_s}{m_s} \right)^2} d\omega} - b \quad (19)$$

where C_s/m_s and k_s/m_s are constants reflecting the systematic features of the suspension. The PSD of the acceleration can be calculated based on in-car acceleration data. The parameters k, b are calibrated by field tests.

Eq. (19) describes the vehicle's responses to pavement roughness using a quarter-car model. However, it remains challenging to distinguish whether the vibration is caused by the engine or the pavement roughness in practice. Therefore, the effect of vehicular vibration must be considered when evaluating the IRI. The engine excitation can be regarded as a harmonic process, the steady-state response of the suspension is also harmonic, whose frequency is identical to the excitation frequency. The amplitude of the vibration depends only on the physical parameters of the vehicle, then the impact caused by engine self-vibration can be expressed as Eq. (20) (Liu et al., 2020, working paper).

$$E|z_s - z_u| = \frac{2}{\pi} \left| \frac{(-\gamma^2 m_s + k_t) F_0}{\sqrt{\hat{A}^2 + \hat{B}^2}} \right| \quad (20)$$

where γ indicates the angular velocity of the cylinder in the engine, which is approximately constant under the uniform driving speed. F_0 indicates the vertical excitation force caused by engine, including the inertial force generated by the piston rod and the centrifugal force generated by the flywheels of the crankshaft. \hat{A} and \hat{B} are the contracted indexes, which are only determined by the vehicular suspension parameters and γ .

When the vehicle is driving at a constant speed, Eq. (20) can be regarded as a constant during a fixed distance. Therefore, we also reckon the impact of the engine on IRI evaluation as a part of intercept b in Eq. (19).

2.2. SSL model for large-scale roughness evaluation

The previous section introduced the relationship between the suspension vibrations and the IRI of a single vehicle. Although this method can be used to evaluate the roughness level of the pavement, it still requires an accurate calibration of the vehicular parameters, which is a time-consuming and inefficient process. Meanwhile, with the rapid development of sensing technology, vehicles have become empowered with the ability to collect enormous amounts of data (such as vibrations, speed, GPS location, etc.) during driving. These data directly reflect the performance of the pavement, and thus can be used to calculate the IRI. However, due to user privacy protection, most vibration data are quite hard to match with practical vehicles; thus, applying supervised regression may be appropriate to evaluate IRIs based on Eq. (19). On the other hand, since vehicular vibrations are easily affected by the speed, and they vary between different types of vehicle, it is also difficult to completely estimate the IRI via unsupervised clustering. Therefore, this

paper proposes a SSL model to estimate the IRIs for large-scale road sections, which enables us to utilize all the unlabeled data based on only a small dataset of labeled data. The framework of using multiple-vehicle crowdsourced data to coordinately estimate the network roughness is illustrated in Fig. 2.

A small quantity of the roads was selected as labeled roads in advance, whose IRIs were pre-measured, as shown in Fig. 2 (red lines). When vehicles pass through these roads, their parameters (as determined using Eq. (19)) can be estimated from their vibration data and the IRIs of the roads. These vehicles are defined as ‘calibrated vehicles’ (CbVs); in contrast, the ones whose parameters are unknown are defined as ‘uncalibrated vehicles’ (uCVs). When a CbV drives on an unlabeled road, the IRI of the road can be estimated using Eq. (19).

There are two essential problems in this framework:

- (1) The confidence level of the parameter estimation for each vehicle varies: some uCVs may only pass through one or two labeled roads, yet others may pass through five or more. Therefore, the parameter estimation results for the latter are obviously more reliable.
- (2) The confidence level of the IRI estimation for each unlabeled road varies: some unlabeled roads may be traveled by one or two CbVs, yet the others may be traveled by a hundred or more CbVs. The IRI evaluation results of the latter are inclined to be more reliable.

Therefore, we selected vehicles and roads whose confidence level exceeded 85% of all samples as an effective estimation to lower the error caused by randomness.

The first step of the framework is to calibrate vehicle parameters. There are five parameters that need to be estimated in Eq. (19), which contains three inherent parameters and two fitting coefficients, and it is a nonlinear expression, which makes it complicated to calibrate these parameters. Therefore, we discretized the IRI formula as:

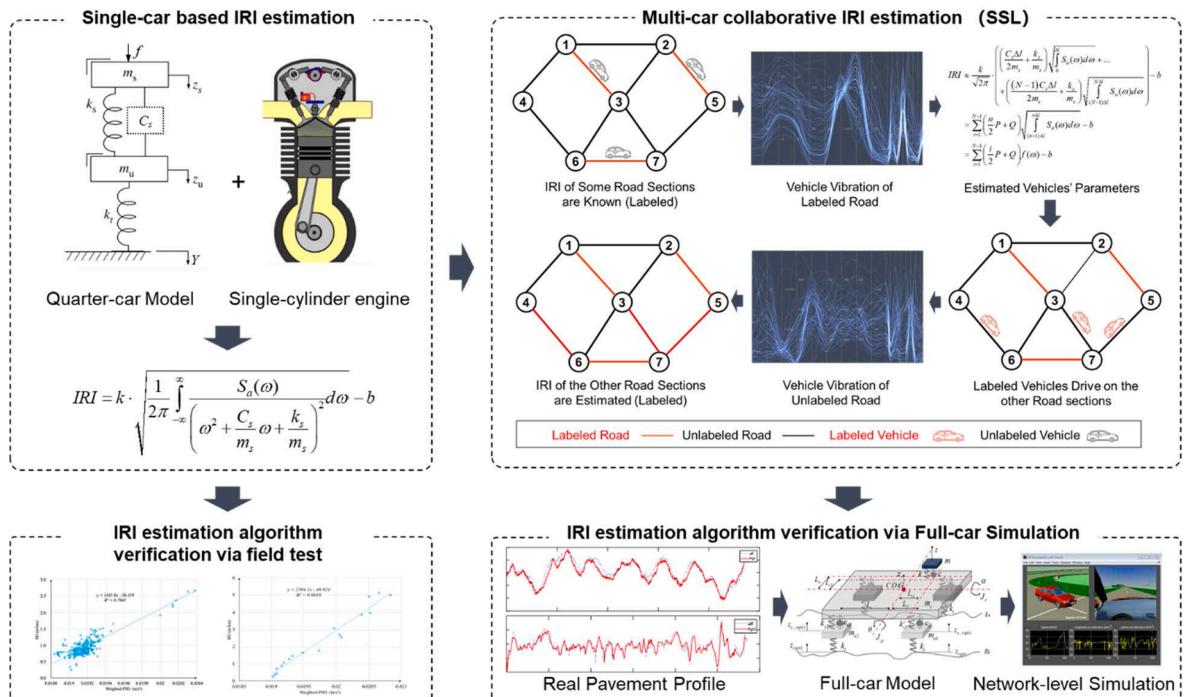


Fig. 2. The framework of large-scale pavement roughness estimation.

$$IRI = k \cdot \sqrt{\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{S_a(\omega)}{\left(\frac{C_s}{m_s}\omega + \frac{k_s}{m_s}\right)^2} d\omega} - b$$

$$\approx \frac{k}{\sqrt{2\pi}} \cdot \left(\begin{array}{l} \sqrt{\int_0^{l_1} \frac{S_a(\omega)}{\left(\frac{C_s l_1}{2m_s} + \frac{k_s}{m_s}\right)^2} d\omega} + \sqrt{\int_{l_1}^{l_2} \frac{S_a(\omega)}{\left(\frac{C_s(l_2 + l_1)}{2m_s} + \frac{k_s}{m_s}\right)^2} d\omega} + \dots \\ + \sqrt{\int_{l_{N-1}}^{l_N} \frac{S_a(\omega)}{\left(\frac{C_s(l_N + l_{N-1})}{2m_s} + \frac{k_s}{m_s}\right)^2} d\omega} \end{array} \right) - b \quad (21)$$

where N is the upper limit of the frequency of the power spectral density function (PSD). Considering the symmetry of the PSD bilateral spectrum, only an interval with a positive frequency is calculated. If the interval of the frequency segment is equally divided, then the above equation can be expressed as:

$$IRI \approx \frac{k}{\sqrt{2\pi}} \cdot \left(\begin{array}{l} \left(\frac{C_s \Delta l}{2m_s} + \frac{k_s}{m_s}\right) \sqrt{\int_0^{\Delta l} S_a(\omega) d\omega} + \left(\frac{C_s \Delta l}{m_s} + \frac{k_s}{m_s}\right) \sqrt{\int_{\Delta l}^{2\Delta l} S_a(\omega) d\omega} + \dots \\ + \left(\frac{(N-1)C_s \Delta l}{2m_s} + \frac{k_s}{m_s}\right) \sqrt{\int_{(N-1)\Delta l}^{N\Delta l} S_a(\omega) d\omega} \end{array} \right) - b$$

$$= \sum_{i=1}^{N-1} \left(\frac{n}{2} P + Q \right) \sqrt{\int_{(n-1)\Delta l}^{n\Delta l} S_a(\omega) d\omega} - b = \sum_{i=1}^{N-1} \left(\frac{i}{2} P + Q \right) f(\omega) - b \quad (22)$$

where P, Q, b are model fitting parameters, which are related to the fixed parameters of the vehicle, and $f(\omega)$ represents the integral of the mean square value of the PSD within each segment. In this approach, a nonlinear problem with five parameters can be transformed into a linear problem with only three parameters, which can be solved more conveniently. The maximum likelihood and least square method are used to estimate P and Q , as shown in Eqs. (23) & (24). After estimating P and Q , the interpret b can be directly calculated by Eq. (22) based on the labelled IRI.

$$\hat{P} = \frac{\sum_i \left(\sum_{n=1}^{N-1} f(\omega)^{(i)} \cdot IRI^{(i)} \right) \cdot \sum_i \left(\sum_{n=1}^{N-1} f(\omega)^{(i)} \cdot \sum_{n=1}^{N-1} \left(\frac{n}{2} \cdot f(\omega)^{(i)} \right) \right)}{-\sum_i \left(\sum_{n=1}^{N-1} \left(\frac{n}{2} \cdot f(\omega)^{(i)} \right) \cdot IRI^{(i)} \right) \cdot \left(\sum_i \sum_{n=1}^{N-1} \left(\frac{n}{2} \cdot f(\omega)^{(i)} \right) \right)^2} \quad (23)$$

$$\hat{Q} = \frac{\sum_i \left(\sum_{n=1}^{N-1} f(\omega)^{(i)} \cdot IRI^{(i)} \right) \cdot \left(\sum_i \sum_{n=1}^{N-1} \left(\frac{n}{2} \cdot f(\omega)^{(i)} \right) \right)^2}{-\sum_i \left(\sum_{n=1}^{N-1} \left(\frac{n}{2} \cdot f(\omega)^{(i)} \right) \cdot IRI^{(i)} \right) \cdot \sum_i \left(\sum_{n=1}^{N-1} f(\omega)^{(i)} \cdot \sum_{n=1}^{N-1} \left(\frac{n}{2} \cdot f(\omega)^{(i)} \right) \right)} - \left(\sum_i \left(\sum_{n=1}^{N-1} f(\omega)^{(i)} \cdot \sum_{n=1}^{N-1} \left(\frac{n}{2} \cdot f(\omega)^{(i)} \right) \right) \right)^2 + \left(\sum_i \sum_{n=1}^{N-1} \left(\frac{n}{2} \cdot f(\omega)^{(i)} \right) \right)^2 \quad (24)$$

where n indicates the n^{th} frequency band, and i represents the i^{th} sample. As mentioned above, the number of samples affects the confidence degree of the vehicle's parameter estimations and the pavement roughness evaluation. Therefore, we established an iterative model based on SSL to predict P, Q and the pavement roughness with relatively high confidence.

Let \mathbf{R} denote a set that contains roads with known IRIs, and \mathbf{U} denotes its complementary set. \mathbf{V} represents the vehicle set. The algorithm processes the following steps:

Step 1 - Select and calibrate the CbVs: For each vehicle $v \in \mathbf{V}$, count the number of roads in \mathbf{R} that are in the trajectory records of v . Rank \mathbf{V} by the counts. The top 15% of the vehicles in \mathbf{V} are recognized as reliable probes and selected as CbVs. For every CbV, its corresponding parameters are calibrated based on Eqs. (23)–(24).

Step 2 - Label the most reliable unlabeled road: Estimate the IRIs for the roads in \mathbf{U} using Eq. (22) with all the selected CbVs. Note that each CbV will calculate one value for each road it has traveled, while the value will be left null if a CbV did not go through the specific road. For each road u in \mathbf{U} , count the number of CbVs that passed through it. Select the road $\bar{r} \in \mathbf{U}$ with the highest count and

lowest IRI variance (Eq. (25)) as the most reliable estimation. Transfer \bar{r} from \mathbf{U} to \mathbf{R} , and label \bar{r} with the average value of IRI_r :

$$\bar{r} = \min(\sigma(\max(\text{size}(IRI_r)))) \quad (25)$$

where r is the road number, IRI_r represents the set of estimated IRIs for road r , $\text{size}(IRI_r)$ means the size of set IRI_r , and σ represents its variance.

Step 3 – Repeat the labelling: Use a new pair of \mathbf{U} and \mathbf{R} and repeat steps 1–2 until $U = \emptyset$, whereby the IRIs of all the roads have been estimated.

The greatest strength of this SSL-based approach is that each iteration only labels one road that shares the highest confidence in its predicted IRI value, and \mathbf{R} is continuously expanded to better calibrate the parameters for \mathbf{V} . Increasing the parameter estimation confidence, in turn, enhances the IRI prediction for \mathbf{U} until the whole network is traversed.

3. Results and discussion

3.1. Field tests for validation of IRI estimation model

The relationship between IRI and PSD has been proved to be strongly correlated in many previous studies. However, few of them shed light on the mathematical derivations between them. Even though PSD and IRI are described quantitatively in the literature, such as (Chen et al., 2011; Du et al., 2014; Pawar et al., 2018; Sun, 2003), etc., the vehicle parameters are assumed to be “known” in their models, and the errors caused by vehicle vibration are not considered. This paper considers the impact of the suspension and engine together to improve the reliability of the single-car based IRI estimation model. More than 200 km highways and urban roads in Ningbo and Shanghai, China are selected to conduct field tests. The testing vehicle is Nissan Xterra 2013, the acceleration of which was collected using vibration sensors right above the vehicle suspension. Note that the regression coefficients shown in Fig. 3 are only effective to the specific testing vehicle, which is a Nissan Xterra, 2013 for this case. However, the derived mathematical formulations are supposed to be universal for most vehicles as long as they meet the features of a quarter car model. Driving speed affects the results of the IRI evaluation significantly. Previous studies have proved a quadratic relationship between speed and the PSD values (Du et al., 2018; Liu et al., 2020). A speed calibration model is proposed to consider the impact of the speed on the PSD and the result accuracy, as shown in Eq. (26) (Du et al., 2014).

$$f(\omega, v) = 0.0263v^2 + 0.6027v \quad (26)$$

For the large-scale pavement roughness evaluation, it's assumed that the network's traffic status is stable. Therefore, the driving speed on each road section can be represented by the space mean speed of the traffic, which is constant for a short time. The vehicular paraments $P(v)$ and $Q(v)$ (in Eqs. (23)–(24)) estimated by the proposed SSL algorithm are case-specific and unique regarding the speed v . As the IRI is defined by a quarter-car with a constant speed 80 km/h, the parameters $P(v)$ and $Q(v)$ for a certain vehicle with the speed of v are converted to their equivalent values $P(v = 80 \text{ km/h})$ and $Q(v = 80 \text{ km/h})$. For simplicity, the driving speed is set as 80 km/h. The sampling rate is 200 Hz. The standard IRI values (ground truth data) were measured by a laser surface tester. Fig. 3 illustrates the testing results.

As shown in Fig. 3(a), the tested data yield a positive correlation between IRI and weighted PSD (Eq. (19)). Compared with the vehicle-self vibration, the vibration caused by the roughness is relatively small for low-IRI roads ($IRI < 2 \text{ m/km}$), whose PSD mainly shows at the low-frequency bands. As the low-frequency bands contain lots of vibration components generated by the engine, the PSD resolution is relatively low. On the contrary, the roughness-induced vibration is predominant for high-IRI roads. The proposed algorithm has the capability of high-pass filtering to eliminate the high-frequency noise effectively. Therefore, the weighted PSD

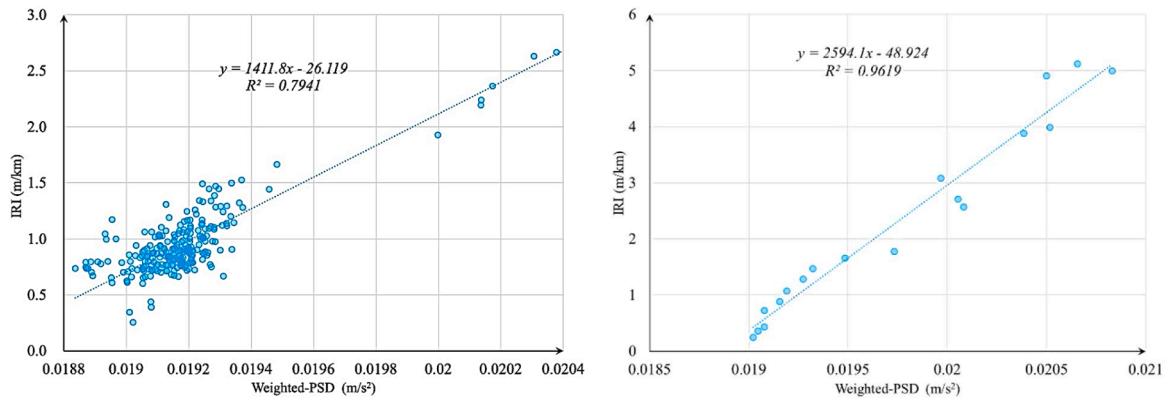


Fig. 3. Field Tests Results.

performs better and more stable on thigh-IRI roads. Given the error's randomness, we clustered the data into ten sub-groups and utilized their mean values as the kernel of each cluster. Fig. 3(b) shows the regression results after clustering. The R-square value is over 0.9, which proves the effectiveness of the proposed model (Eqs. (19) & (20)).

3.2. Simulation tests for validation of SSL model

In order to verify the reliability and effectiveness of the proposed model for large-scale IRI estimation, 50 road sections were selected in the testing network, of which 1/5 were roads with known IRIs. We also incorporated pavement elevation data, which was collected by a laser detection vehicle in Baoshan District and Jiading District of Shanghai. The IRI distribution of the selected 50 sections basically followed a normal distribution with a mean value of 4.5 m/km. The IRI values of the roads with known labels were normally distributed to improve the generalization of the parameter estimation for the vehicles. The sampling frequency was 200 Hz. Table 2 shows the IRI values of the 50 designed roads.

The bold blue numbers in Table 1 indicate roads with known IRIs, and the remaining roads were unlabeled data. To simulate the most common vehicle types, the coefficients of the testing vehicles were selected from the CARSIM vehicle library (Johansson et al., 2004), including the CS-A/B/C/D/E/F car models. The basic parameters of these vehicles are shown in Table 3. The vehicles were evenly distributed in the traffic flow. The total number of test vehicles was 500.

This paper proposed a full-car model to generate vibration data due to lack of the practical crowdsourced vibration in the current condition. As a car's body is a rigid structure, when it drives over the continuous surface of pavement, if a change is produced on one side, a corresponding change will inevitably appear on the other side. Thus, although the tires and the suspension system can reduce the effects of the two wheels on each other, the quarter-car model used in the model described previously (see Section 2.1.1) does not provide an accurate description of reality. In addition, there is an insufficient number of connected vehicles to conduct field tests in the current stage. Therefore, we constructed a full-car simulation environment following the work of Cantisani and Loprencipe (Cantisani and Loprencipe, 2010) to better describe the practical scenarios and verify the effectiveness of the SSL model, as shown in Fig. 4. The dynamic vibration of a vehicle can be predicted based on the full-car model. In this paper, we established a simulation platform using Simulink version 2019b in MATLAB.

The following assumptions were made when designing the testing network: (1) The physical topology is not included in the road network design since in this research we only consider the accuracy of the roughness prediction, we do not consider the impact of the topology and traffic assignment; (2) the number of roads traveled by the vehicles is in accordance with a normal distribution; (3) the probability of each road to be selected by the vehicles is equal; and (4) the driving speed of each vehicle on the same road is basically consistent. Based on the above assumptions, we simulated the vehicles' routes in the network using a Weibull distribution.

During the actual tests, the parameters of the suspension quality (considering the driver), tire spring rigidity and dynamic suspension damping varied depending on the weight and tire pressure of each vehicle, and other conditions. To simulate randomness of the three basic parameters, random tail terms were considered in the simulations, where the parameters of a given vehicle varied by no more than 10%.

Seat damping was set to 2200 Ns/m, the elasticity of which was 64140 N/m, the distance between the center of gravity and the edge of the seats was 0.4 m, and the moment of inertia for a sprung mass was 2700 Nms². Fig. 5 illustrates the results of the proposed algorithm, where the average IRI estimation relative error of SSL model is 9.71%. The highest relative prediction error is 46.5% on road 40, whose absolute error is about 1.15 m/km, as shown in Fig. 5(a). The potential reason for this large error may lie in the distribution of the training set, which follows a normal distribution with a mean of 4.5 m/km, so the model performance on lower IRIs is relatively bad. It can be seen from the figure that the SSL performs well on most roads, and basically conforms to the actual IRI distribution of the pavement. Most of the relative errors were kept within 10%, and the errors were inclined to occur on roads with high or low IRIs.

Fig. 5(b) shows another large-scale IRI estimation algorithm named 'multi-vehicle average estimation' (MAE). The core of MAE algorithm is to superimpose the IRI results of different vehicles on the same road section without considering the IRI prediction confidence of each vehicle. The IRI of a road section is calculated by the average value of multi-vehicle prediction results. The advantage of this method is that it can fully consider the data of all vehicles on the road section, so that it is not affected by the number of iterations, while the disadvantage is that it is very vulnerable to outliers. The following steps are taken in the MAE algorithm:

Step 1 - Select and calibrate CbVs: For each vehicle $v \in V$, select all vehicles that travel through roads with known IRIs (R) as CbVs. Calibrate the vehicles' parameters using Eqs. (23)–(24).

Table 2

The ground truth pavement roughness of the simulation network.

No.	1	2	3	4	5	6	7	8	9	10
IRI	3.898	3.673	3.787	3.987	3.498	4.447	3.538	4.016	4.911	4.690
No.	11	12	13	14	15	16	17	18	19	20
IRI	3.346	3.289	4.221	3.685	3.322	4.152	5.652	4.482	4.389	4.439
No.	21	22	23	24	25	26	27	28	29	30
IRI	2.982	4.982	5.608	5.864	5.304	6.354	6.353	2.917	5.264	5.913
No.	31	32	33	34	35	36	37	38	39	40
IRI	5.914	6.147	5.896	5.144	6.539	5.394	7.409	7.200	9.860	2.477
No.	41	42	43	44	45	46	47	48	49	50
IRI	2.048	3.071	3.130	2.762	2.859	3.084	6.103	3.398	3.342	2.422

Table 3

Fundamental parameters of the simulated vehicles.

Vehicle Type	l (m)	m_s (kg)	m_{u1} (kg)	m_{u2} (kg)	c_{s1} (Ns/m)	k_{s1} (N/m)	k_{s2} (N/m)	k_r (N/m)
P1	2.59	1300	40	35	3100	38,889	35,000	200,000
C-Class, Hatchback	2.59	1274	71	71	3551	27,000	30,000	228,000
F-Class, Sedan	3.17	1823	100	100	3551	83,000	44,000	230,000
D-Class, SUV	2.62	1429	80	100	5189	130,000	40,000	230,000
B-Class, Sports	2.33	1020	60	60	3551	130,500	91,000	230,000
B-Class, Hatchback	2.6	1111	60	60	3551	28,000	35,000	230,000
A-Class, Hatchback	2.35	747	41.5	41.5	1876	14,000	18,000	200,000
CS Stock car	2.68	1820	70	140	5189	40,000	40,000	230,000
European Van	2.58	1100	100	100	2301	31,000	31,000	200,000
D-Class, Sedan	2.78	1370	80	80	3551	153,000	82,000	230,000

* The above data were collected from the CARSIM software database.

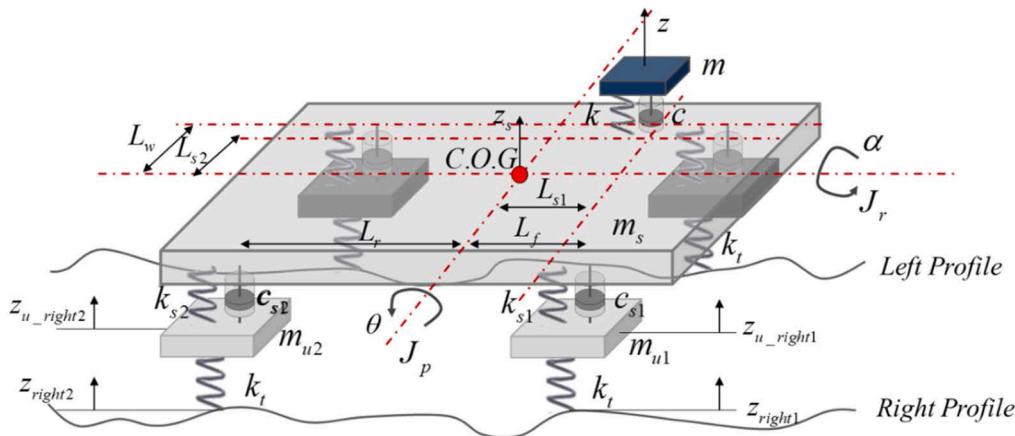


Fig. 4. Eight degrees-of-freedom vehicle vibration model.

Step 2 – Estimate the IRI_s for unlabeled roads: Estimate IRI_s for roads in U using Eq. (22) with all the selected CbVs, as shown in Fig. 4b (blue dots).

Step 3 – Use the average value as the predicted IRI: The IRI of each road in U is calculated using the mean value of the IRI_s, which are estimated by multiple the CbVs on the road, as shown in Fig. 4b (black line).

The zero point was removed during the calculation, and all outliers with IRI greater than 10 m/km were excluded. The MAE model prediction results generally conform to the overall distribution of the IRI. The average relative error is 18.01%, which is almost double the error from the SSL. The highest prediction error is 35.13% on road 20, whose absolute error is about -1.56 m/km. The overall model predictions are smaller than the actual IRI values, and only roads with relatively low IRI values, such as road 40, are slightly larger than the real IRI value.

Compared with the SSL model, the error distribution of the MAE model is more discrete, indicating the impact of the random errors. The key difference between the error sources of the two models is that the errors of the SSL can be mainly attributed to the fact that each time an unknown label road is added to the known set R, the uncertainty of R increases, resulting in the accumulation of errors during each iteration. For the MAE algorithm, the errors mainly come from the inaccurate estimation of a vehicles' parameters and an insufficient sample size during the calibration process, which thus accumulate and affect the unknown roads' IRI predictions.

Fig. 6(a) shows the relative errors of the two models. Overall, the SSL model has a higher prediction accuracy, where its relative error is 8.70% smaller than that of the MAE method. Most of the MAE model errors are negative, and the actual IRI values are underestimated. However, for the same specific road, the MAE model performs better, such as for roads 38, 40, and 41. The MAE model tends to underestimate the real IRI, and the IRI_s of these roads are quite small, which makes MAE relatively accurate. It is seen that the errors of the SSL model are significantly affected by the iteration order. For example, the IRI values of roads 38, 40, and 41 were only determined after a large number of iterations. As mentioned before, error accumulation increases with the number of iterations, which results in the high relative errors of the IRI values estimated in later iterations. Fig. 6(b) shows the effect of the iteration order on the model accuracy. Since the SSL algorithm assigns values to the road section and vehicle with the highest estimation confidence in each iteration, the estimated error will inevitably increase as the number of iterations increases. Therefore, the overall IRI error of the whole network calculated by SSL is relatively low from the macro-perspective. As for some road sections with later iterations, the MAE method can be considered for the IRI estimation.

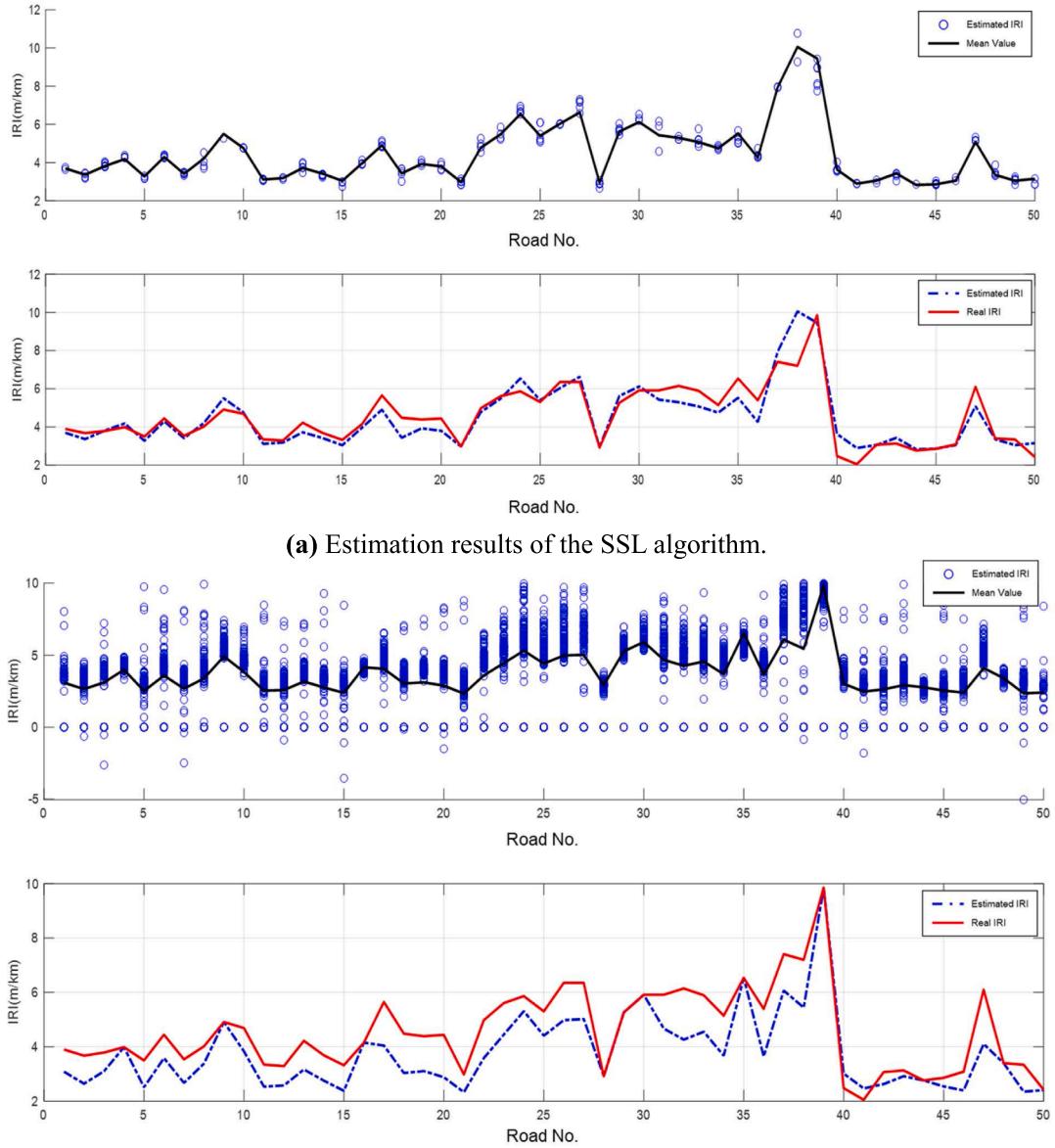


Fig. 5. Simulation results.

3.3. Sensitivity analysis

The total number of vehicles and the vibration sampling rate are the two main factors that affect the prediction ability of the proposed model. High-frequency collection provides more detailed roughness information of roads, which improves the confidence level of each prediction. Fig. 7 reveals the coupling effect between the sampling frequency and the total number of test vehicles on the accuracy of the SSL model. The 3D image was linearly interpolated using griddata with an interpolation rate of 80. The results show that the average prediction error decreases as the sampling rate and vehicle number increase. The relative estimation error varies significantly in the sampling rate range of 0–50 Hz, and the effect of the declining errors is evident within this range. When the sampling frequency exceeds 50 Hz, the impact of the frequency is not significant. The potential reason may lie in the PSD calculation (Eq. (19)): a lower sampling range (for example, 50 Hz) results in poorer frequency resolution (25 Hz), which barely provides enough vibration information for the subsequent pavement roughness estimation. Therefore, the effectiveness of increasing the sampling frequency is evident. The most sensitive range of the vehicle quantity is 0–200. However, the relative errors do not monotonically decrease with vehicle amount, which may be due to the impact of the error accumulation. Too many vehicles can exacerbate this

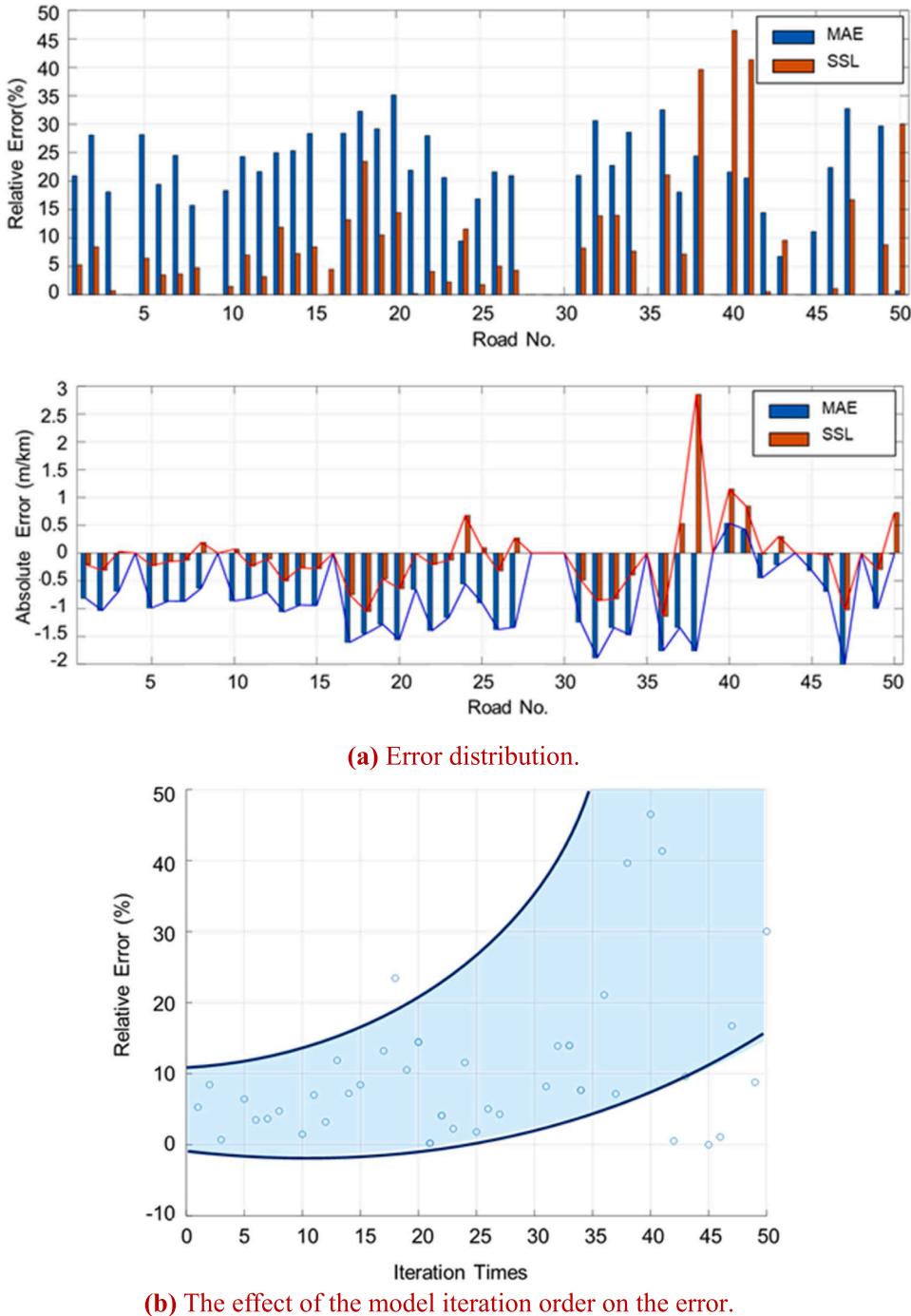


Fig. 6. Model error analysis.

negative effect. Note that the sensitivity range of the sampling rate and vehicle quantity change according to the size of the network. Therefore, in practice, the suitable threshold for these two factors should be analyzed for specific cases. In addition to the total number of vehicle and the vibration sampling rate, the lateral trajectory of the vehicle is also a significant factor affecting the accuracy of the algorithm. Most vehicles drive through the two wheel-tracks on the roads. Since the SSL algorithm selects the vehicle and road with the highest confidence for calculation in each iteration, it reduces the negative impact of the vehicles not on the wheel track.

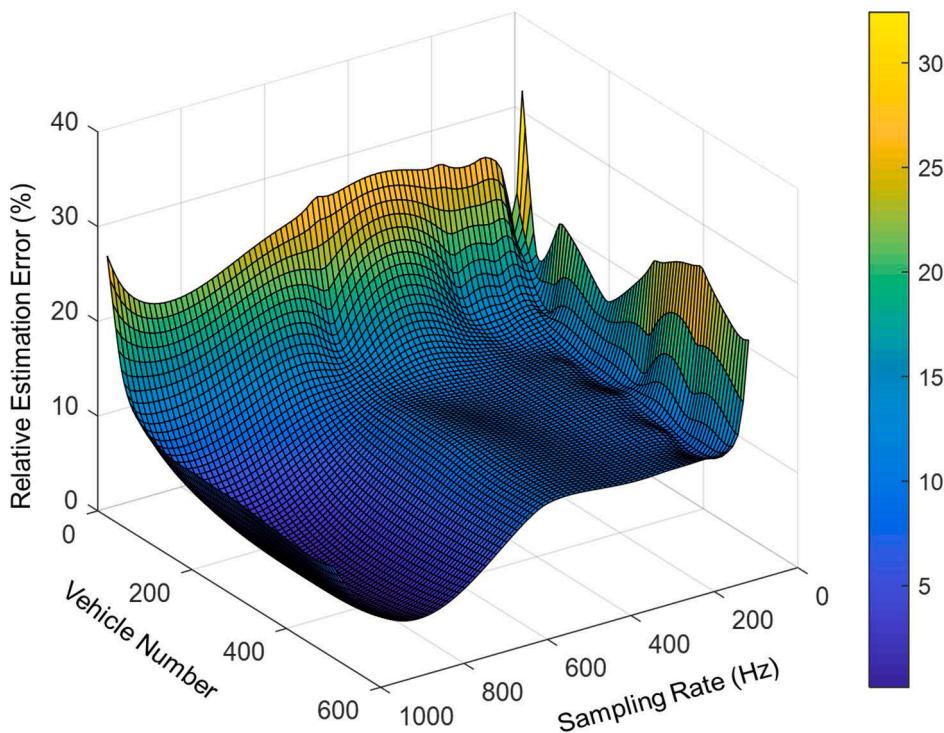


Fig. 7. Dependence of the relative estimation error on the sampling rate and vehicle quantity.

4. Conclusions

This study proposed an SSL model to rapidly evaluate large-scale pavement roughness based on crowdsourced data of multiple vehicles. The relationship between pavement roughness and in-car vibrations was mathematically derived using a PSD analysis and an LTI system. Based on the multi-vehicle vibration data, a self-training algorithm was devised that used both labeled and unlabeled data to comprehensively evaluate the IRIs. The confidences of the vehicular parameters and the IRI estimations were considered to ensure that we obtained the most reliable results in each iteration. A full-car simulation environment with eight degrees-of-freedom was constructed to verify the effectiveness of the proposed model. The results showed the high accuracy of the SSL model, where the overall relative error remained less than 10%. Compared with the MAE model, SSL had, on average, 8.3% higher accuracy, and its error distribution was steadier. However, the errors of the SSL model were significantly affected by the iteration order. Error accumulation increased with the iteration order, which led to large relative errors being produced in later iterations. A sensitivity analysis illustrated that the sensitivity ranges of the sampling rate and vehicle quantity were 0–50 Hz and 0–200 cars, respectively, for this specific case. Future work will focus on the impact of the network's topology and traffic volume on the performance of the proposed model. Meanwhile, the way of combining the SSL and MAE algorithms to solve the problem of IRI estimation for road sections with the later iterations is also an important research direction in the future.

CRediT authorship contribution statement

Chenglong Liu: Conceptualization, Methodology, Writing - original draft. **Difei Wu:** Data curation, Software, Writing - original draft. **Yishun Li:** Investigation. **Yuchuan Du:** Supervision, Validation, Writing - review & editing.

Acknowledgments

This study was jointly supported by National Natural Science Foundation of China (NSFC-51978519, NSFC-52008309), Chinese Postdoctoral Science Foundation (2020M671221), and Scientific Research Project of Shanghai Science and Technology Commission (19DZ1209102). The authors are responsible for all views and opinions expressed in this paper.

References

- Alessandroni, G., Klopfenstein, L.C., Delpriori, S., Dromedari, M., Luchetti, G., Paolini, B., Seraghiti, A., Lattanzi, E., Freschi, V., Carini, A., 2014. Smartroadsense: Collaborative road surface condition monitoring. In: Proceedings of the UBICOMM, pp. 210–215.
 Alhasan, A., White, D.J., De Brabanter, K., 2017. Spatial pavement roughness from stationary laser scanning. Int. J. Pavement Eng. 18, 83–96.

- Al-Rousan, T., Asi, I., Baker, A.A., 2010. Roughness evaluation of Jordan highway network. In: 24th ARRB Conference Melbourne, Australia.
- Arfken, G.B., Weber, H.J., Spector, D., 1999. Mathematical Methods for Physicists. Am. J. Phys. 67, 165–169. <https://doi.org/10.1119/1.19217>.
- Bil, M., Andrasik, R., Kubecek, J., 2015. How comfortable are your cycling tracks? A new method for objective bicycle vibration measurement. Transport. Res. Part C: Emerg. Technol. 56, 415–425.
- Bitelli, G., Simone, A., Girardi, F., Lantieri, C., 2012. Laser scanning on road pavements: A new approach for characterizing surface texture. Sensors 12, 9110–9128.
- Bose, B., Dutta, J., Ghosh, S., Pramanick, P., Roy, S., 2018. D&RSense: Detection of driving patterns and road anomalies. In: in: 2018 3rd International Conference On Internet of Things: Smart Innovation and Usages (IoT-SIU). IEEE, pp. 1–7.
- Bridgelall, R., 2013. Connected vehicle approach for pavement roughness evaluation. J. Infrastruct. Syst. 20, 04013001.
- Cantisani, G., Loprencipe, G., 2010. Road roughness and whole body vibration: Evaluation tools and comfort limits. J. Transp. Eng. 136, 818–826.
- Chang, J.-R., Chang, K.-T., Chen, D.-H., 2006. Application of 3D laser scanning on measuring pavement roughness. J. Test. Eval. 34, 83–91.
- Chapelle, O., Scholkopf, B., Zien, A., 2009. Semi-supervised learning (Chapelle, O. et al., Eds.; 2006)[book reviews]. IEEE Transactions on Neural Networks 20, 542–542.
- Chen, C.-S., Chou, C.-P., Chen, A.-C., 2017. Viscoelastic Model for Estimating the International Roughness Index by Smartphone Sensors. Presented at the Transportation Research Board 96th Annual MeetingTransportation Research Board.
- Chen, K., Lu, M., Fan, X., Wei, M., Wu, J., 2011. Road condition monitoring using on-board three-axis accelerometer and GPS sensor.
- Chen, K., Lu, M., Tan, G., Wu, J., 2013. CRSM: Crowdsourcing based road surface monitoring. In: 2013 IEEE 10th International Conference on High Performance Computing and Communications & 2013 IEEE International Conference on Embedded and Ubiquitous Computing. IEEE, pp. 2151–2158.
- Douangphachanh, V., Onayama, H., 2013. Estimation of road roughness condition from smartphones under realistic settings. In: 2013 13th International Conference on ITS Telecommunications (ITST). Presented at the 2013 13th International Conference on ITS Telecommunications (ITST), pp. 433–439. <https://doi.org/10.1109/ITST.2013.6685585>.
- Du, Y., Liu, C., Li, Y., 2018. Velocity Control Strategies to Improve Automated Vehicle Driving Comfort. IEEE Intell. Transp. Syst. Mag. 10, 8–18.
- Du, Y., Liu, C., Wu, D., Jiang, S., 2014. Measurement of International Roughness Index by Using-Axis Accelerometers and GPS. Mathematical Problems in Engineering 2014.
- Du, Y., Liu, C., Wu, D., Li, S., 2016. Application of vehicle mounted accelerometers to measure pavement roughness. Int. J. Distrib. Sens. Netw. 12, 8413146.
- Gillespie, T.D., 2001. Everything You Always Wanted to Know about the IRI. But Were Afraid to Ask, Road Profile Users Group Meeting.
- Gillespie, T.D., 1980. Calibration of response-type road roughness measuring systems.
- Hudson, W.R., Hain, R.C., 1961. Calibration and use of BPR Roughometer at the AASHO Road Test. Highway Research Board Special. Report.
- Hudson, W.R., Halbach, D., Zaniewski, J.P., Moser, L., 1985. Root-mean-square vertical acceleration as a summary roughness statistic. Measuring Road Roughness and Its Effects on User Cost and Comfort. ASTM International.
- Islam, S., Buttlar, W., Aldunate, R., Vavrik, W., 2014. Measurement of pavement roughness using android-based smartphone application. Transport. Res. Rec.: J. Transport. Res. Board 30–38.
- Janoff, M.S., 1988. National Cooperative Highway Research Program (NCHRP) Report 308: Pavement Roughness and Rideability Field Evaluation. Transportation Research Board, National Research Council, Washington, DC.
- Jeong, J.-H., Jo, H., Ditzler, G., 2020. Convolutional neural networks for pavement roughness assessment using calibration-free vehicle dynamics. Comput.-Aided Civ. Infrastruct. Eng.
- Johansson, R., Williams, D., Berglund, A., Nogués, P., 2004. Carsim: a system to visualize written road accident reports as animated 3D scenes, in. Proceedings of the 2nd Workshop on Text Meaning and Interpretation.
- Jordan, P.G., Young, J.C., 1980. Developments in the Calibration and Use of the Bump-integrator for Ride Assessment.
- Liu, C., Wu, D., Li, Y., Jiang, S., Du, Y., 2020. Mathematical insights into the relationship between pavement roughness and vehicle vibration. Int. J. Pavement Eng. 1–13.
- Mahmoudzadeh, A., Golroo, A., Jahanshahi, M.R., Firooz Yeganeh, S., 2019. Estimating Pavement Roughness by Fusing Color and Depth Data Obtained from an Inexpensive RGB-D Sensor. Sensors 19, 1655.
- McCann, R., Nguyen, S., 2007. System identification for a model-based observer of a road roughness profiler. In: 2007 IEEE Region 5 Technical Conference. IEEE, pp. 336–343.
- Medina, J.R., Noorvand, H., Shane Underwood, B., Kaloush, K., 2020. Statistical Validation of Crowdsourced Pavement Ride Quality Measurements from Smartphones. J. Comput. Civil Eng. 34, 04020009.
- Mirtabar, Z., Golroo, A., Mahmoudzadeh, A., Barazandeh, F., 2020. Development of a crowdsourcing-based system for computing the international roughness index. Int. J. Pavement Eng. 1–10.
- Nelson, P., 2016. One autonomous car will use 4,000 GB of data per day [WWW Document]. Network World. URL <https://www.networkworld.com/article/3147892/one-autonomous-car-will-use-4000-gb-of-data.html> (accessed 7.25.19).
- Nguyen, T., Lechner, B., Wong, Y.D., 2019. Response-based methods to measure road surface irregularity: a state-of-the-art review. Eur. Transp. Res. Rev. 11, 43.
- Pawar, P.R., Mathew, A.T., Saraf, M.R., 2018. Iri (international roughness index): An indicator of vehicle response. Mater. Today.: Proc. 5, 11738–11750.
- Prasad, J.R., Kanuganti, S., Bhangaonkar, P.N., Sarkar, A.K., Arkatkar, S., 2013. Development of relationship between roughness (IRI) and visible surface distresses: a study on PMGSY roads. Proc.-Soc. Behav. Sci. 104, 322–331.
- Sandamal, R.M.K., Pasindu, H.R., 2020. Applicability of smartphone-based roughness data for rural road pavement condition evaluation. Int. J. Pavement Eng. 1–10.
- Sayer, M.W., Gillespie, T.D., Queiros, C.A.V., 1986. International road roughness experiment. World Bank Tech. Rep. p. 45.
- Srivastava, R., Nanda, P., 2000. Recent Techniques in Pavement Engineering, In: Proceedings of The Silver Jubilee Conference on Road Infrastructure Management (RIM) in the New Millennium, Held March 10-11, 2000, Bangalore.
- Sun, L., 2003. Simulation of pavement roughness and IRI based on power spectral density. Math. Comput. Simul. 61, 77–88.
- Wang, W., Yan, X., Huang, H., Chu, X., Abdel-Aty, M., 2011. Design and verification of a laser based device for pavement macrotexture measurement. Transport. Res. Part C: Emerg. Technol. 19, 682–694.
- Wei, L., Fwa, T.F., Zhe, Z., 2005. Wavelet analysis and interpretation of road roughness. J. Transp. Eng. 131, 120–130.
- Yi, C.-W., Chuang, Y.-T., Nian, C.-S., 2015. Toward crowdsourcing-based road pavement monitoring by mobile sensing technologies. IEEE Trans. Intell. Transp. Syst. 16, 1905–1917.
- Zhao, B., Nagayama, T., Toyoda, M., Makihata, N., Takahashi, M., Ieiri, M., 2017. Vehicle Model Calibration in the Frequency Domain and its Application to Large-Scale IRI Estimation. J. Disaster Res. 12, 446–455. <https://doi.org/10.20965/jdr.2017.p0446>.
- Zhu, X.J., 2005. Semi-supervised learning literature survey. University of Wisconsin-Madison Department of Computer Sciences.