

Research Article

Forecasting Urban Rail Transit Vehicle Interior Noise and Its Applications in Railway Alignment Design

Yifeng Wang,¹ Ping Wang,¹ Zihan Li,¹ Zhengxing Chen,¹ and Qing He ^{1,2}

¹Key Laboratory of High-Speed Railway Engineering of the Ministry of Education, School of Civil Engineering, Southwest Jiaotong University, Chengdu 610031, China

²Department of Civil, Structural and Environmental Engineering and Department of Industrial and Systems Engineering, University at Buffalo, The State University of New York, 313 Bell Hall, Buffalo, NY 14260, USA

Correspondence should be addressed to Qing He; qinghe@buffalo.edu

Received 12 December 2019; Revised 3 May 2020; Accepted 29 May 2020; Published 23 June 2020

Academic Editor: Xingju Wang

Copyright © 2020 Yifeng Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this study, a data-driven interior noise prediction model is developed for vehicles on an urban rail transit system based on random forest (RF) and a vehicle/track coupling dynamic model (VTCDM). The proposed prediction model can evaluate and optimize the sustainability of railway alignment from the perspective of interior noise. First, a data collection framework via embedded sensors of onboard smartphones was developed. Then, for establishing the mapping relationship between the dynamic responses of the car body and interior noise, the collected dataset was fed to the RF. Parameter, error distribution, and feature importance analyses were conducted for evaluating and optimizing the performance of the RF. With the optimized parameters, the probability of prediction errors being within 5 dB was 86.9%. Next, the VTCDM was established using an existing industry multibody simulation tool and verified through a comparison between the simulated and field dynamic responses. Finally, a case study that extends the application of this interior noise prediction model to railway alignment design is presented.

1. Introduction

By the end of 2018, 35 cities in mainland China had opened 180 urban rail transit (URT) lines, with a total operation mileage of more than 5700 km. Moreover, the length of subway lines was approximately 4350 km, accounting for 75.6% of all URT lines. The rapidly expanded URT networks offered great convenience to the citizen and were highly advantageous in solving traffic congestions. As a critical URT component, the subway is crucial in people's daily life. However, owing to the unreasonable design of railway alignments and the degradation of track infrastructures, the running quality of trains worsens with the increase in service time. Furthermore, these issues caused abnormal interior noise and vibration, which severely affected passengers' ride comfort. The interior noise of subway vehicles significantly affected passengers' experience; therefore, the design, construction, and operation of subway lines must be emphasized [1].

Generally, train noise can be divided into two categories: external and interior noises [2]. The interior noise is a complex sound field resulting from external acoustics and mechanical excitation sources transmitting, attenuating, and radiating inward through the carriage structure [3]. The interior noise of trains primarily comes from the electrical equipment, aerodynamic, and wheel-rail noises of the train [2]. The operation speed of subway trains is typically within the range of 30–80 km/h, and at such speeds, wheel-rail noises such as rolling noise, squeal, and impact noise are dominant in the generation of interior noise [4]. Curve squeal noise is an intense noise generated when railway vehicles pass sharp curves [5]. Owing to the limitations in cities, many sharp curves appear in subway lines; therefore, the interior noise caused by squeal is serious. Furthermore, defects on the treads of wheels and rails, such as rail corrugation [4], spall of railheads, and polygonal wear of wheels [6], significantly affect the generation of interior noise.

Therefore, wheel-rail noise can be considered a main source of vehicle interior noise.

Wheel-rail noise, including rolling noise, impact noise, and curve squeal, is primarily affected by the wheel-rail relationship when trains move along a track. However, the wheel-rail interaction not only affects the generation of noise but also affects the dynamic responses of vehicles. For example, when rail corrugation occurs, a series of dynamic problems as well as serious train interior noise are likely to appear. Hence, onboard devices to detect track faults from cabin vibrations and interior noise have been developed [7–9]. Because noise and dynamic responses are both related to wheel-rail relationships, it is assumed that interior noise can be forecasted with the dynamic responses of the car body. As for subways, most trains move in a tunnel, the sound field inside of which is not easily interfered by external environmental factors. Therefore, we propose an idea to establish a mapping relationship between vehicle interior noise and the dynamic responses of trains with RF, which is suitable for nonlinear problems. The RF algorithm was first proposed by [10]; it is widely used owing to its high accuracy and strong robustness. Although the dynamic responses and interior noise between individual vehicles differed even through the same section of the alignment, the strong robustness rendered it capable of learning from complex data.

Because RF is a supervised learning algorithm, a large amount of training datasets is required when modeling. For acquiring these datasets, an onboard smartphone application was developed to collect the dynamic responses and interior noise of the subway vehicle. With the development of microelectromechanical system technologies, smartphones with built-in sensors exhibit excellent properties in terms of portability and practicability. Moreover, several applications regarding transport infrastructure condition monitoring and vehicle motion pattern recognition have been reported [11–14]. These applications and our previous studies indicate that the sensing capabilities of the smartphone can satisfy the requirement of this study [15, 16].

The goal of this study was to develop a data-driven method for forecasting vehicle interior noise for the optimization of URT railway alignments. Hence, a large number of train operation status data were collected using the embedded sensors of an Android smartphone. Subsequently, the relationship between car body dynamic responses and interior noise of subway trains was established using RF. For obtaining the simulated dynamic responses of the car body under specific railway alignments, we built a vehicle/track coupling dynamic model (VTCDM) via the Universal Mechanism (UM) software.

Finally, based on the mapping relationship learning from data and dynamic responses from the simulated model, the interior noise under specific railway alignments can be predicted. Further, the prediction results can be applied for optimizing railway alignments at the design stage. The key contributions of this study are as follows:

- (1) Propose a data-driven prediction framework for subway vehicle interior noise based on the smartphone-collected dataset, RF, and VTCDM
- (2) Establish a VTCDM via the UM software based on multibody system dynamics
- (3) Validate the VTCDM and the proposed interior noise prediction model using the field test data from Chengdu Subway Line 7
- (4) Develop an application scenario for optimizing railway alignments with the prediction model of vehicle interior noise

The remainder of this study is organized as follows. Section 2 briefly reviews previous studies regarding methods to predict interior noise in the railway domain. In Section 3, a comprehensive overview of the proposed methodology is provided. Furthermore, in Section 4, the data and features used in our model are introduced. Section 5 describes the modeling approach of RF and VTCDM in detail. The analysis results and discussion are elaborated in Section 6. Finally, conclusions are presented in Section 7.

2. Literature Review

Because train interior noise directly affects passengers' ride comfort, noise control becomes a challenging technology in the design and construction of railway lines. Many factors have an impact on vehicle interior noise. The potential sources of interior noise can be categorized into two groups according to the characteristics of sound: the airborne and structure-borne noise sources [17]. The airborne noise denotes the noise generated by the wheel-rail track system and aerodynamic noise that transmit into the cabin via air, and the structure-borne one refers to the noise radiated from the vibration of vehicle structures inside the cabin. The results of partial coherence analysis of the vibroacoustical signals indicated that the structural vibrations, the cause of structure-borne noise, contribute more to interior noise than airborne noise [18]. Figure 1 describes the composition of vehicle interior noise and some critical influencing factors. The wheel-rail contact force, originating from the roughness of the wheel-rail contact surface, excites the vibration of the wheel and rail to radiate noise. On the other hand, it transmits along with the bogie and suspension system and finally becomes the suspension force to excite vibration of the carriage cabin bottom, generating noise at low frequency [19]. Travel speed of the train, having a relationship with both the wheel-rail force and aerodynamic noise, is one of the most pivotal influencing factors of vehicle interior noise. Besides, the roughness of the tread of rail and wheel, parameters of the wheel-rail and suspension system, and structures of the carriage cabin all have an impact on the characteristics of vehicle interior noise.

Currently, numerous investigations have been conducted regarding the analysis of noise characteristics [20], sound quality evaluations [21], and noise level prediction [22]. For forecasting vehicle interior noise, typically used methods include the finite element method (FEM) [23],

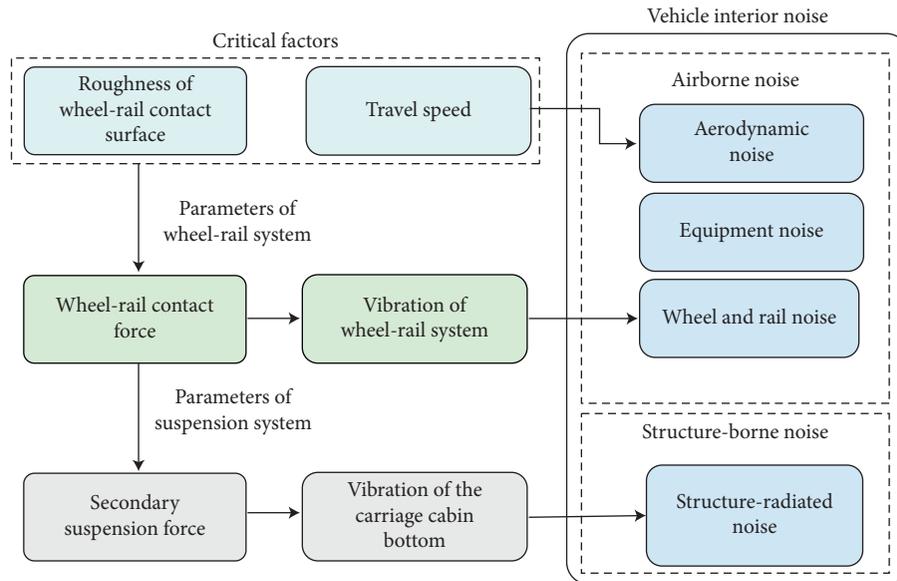


FIGURE 1: The composition of vehicle interior noise and some critical influencing factors.

boundary element method (BEM) [24], and statistical energy analysis method (SEAM) [3]. As each method has its limitations, these approaches can only be applied to specific situations. To achieve higher prediction accuracy and more exceptional generalization ability, hybrid models comprising two or three of the above numerical methods have been widely implemented in the prediction of vehicle interior noise.

In 1966, Gladwell et al. first proposed formulating acoustic problems in terms of displacements, similar to a continuous elastic structure [25], which is the beginning of applying the FEM for acoustic analysis. The FEM was first used to predict the interior noise of a car supported by General Motors Corporation. Owing to the low running speed of the train, the vehicle interior noise was dominated by low-frequency structural noise, and the FEM performed well [26]. However, when Lucas et al. attempted to predict high-frequency aerodynamic noise using the FEM, satisfactory results were unavailable. They indicated that the FEM could not predict high-frequency noise; that is, as the frequency increased, the modes of the structure became denser and could not be identified by the current identification technologies [27]. Therefore, the FEM was primarily used to predict the structural and wheel-rail noise of the car body.

Unlike the FEM, the BEM does not require the acoustic space to be discretized; only the boundary conditions require discretization, which can reduce the calculation time and minimize the calculation error. Owing to the superiority of the boundary element method, it has been widely promoted recently. Similar to the FEM, the BEM was also first used to predict the interior noise of cars [28]. Subsequently, Letourneaux established a train interior noise within a low-frequency range prediction model based on the BEM, and the high-frequency noise prediction problem was investigated [29]. For a long time, the BEM has only been applied

for predicting the vibration and radiated noise of components. In 2011, Soltani established an entire train vehicle model using the BEM and systematically studied the structural noise of the car body [30]. However, the fundamentals of the BEM and the FEM are the same, both of which are based on the method of modal analysis. Therefore, the BEM yields a significant prediction error in the high-frequency range, which has been verified by other researchers [31]. In summary, although the BEM is improved compared with the FEM, it still cannot accurately predict the full-spectrum train interior noise.

The SEAM is entirely different from the two methods above. It utilizes energy as a variable to describe the state of a system, and its core is energy conservation. By constructing the energy flow equation between the cavity subsystems, the energy stored in the relevant acoustic cavity subsystem is available, which can be used to predict the acoustic and dynamic responses. In 1997, Radcliff used the SEAM to predict the interior noise caused by automobile engines. In 2003, James predicted the aerodynamic noise of high-speed trains based on the SEAM, and the feasibility and accuracy of the SEAM for predicting the high-frequency train interior noise were verified by comparison with wind tunnel test results [32]. Givargis and Forssén used the SEAM to predict the high-speed train interior noise in their respective countries and proved the validity of the SEAM in predicting interior noise in high-speed trains [33, 34]. However, the accuracy of the SEAM depends on the damping coefficient matrix, which is difficult to obtain.

The methods above are numerical methods. Although they can yield more accurate results under specific conditions, the requirements for complex model parameters and a large amount of computing capacity render them challenging for practical applications. Recently, data-driven approaches have provided insight for different areas in railway transportation fields [35]. However, to our best

knowledge, studies regarding interior noise prediction based on data-driven methods are few.

3. Research Methodology

Herein, a data-driven interior noise prediction method is proposed, which is different from the traditional numerical model, as shown in Figure 2. Vehicle dynamics responses are key parameters to be considered in the design of URT alignments [36]. Generally, engineers evaluate the designed railway alignments with the simulated vehicle dynamic responses from the VTCDM. Subsequently, the interior noise can serve as a reference in the design of railway alignments using the forecasting model for the vehicle interior noise of the train.

4. Data Collection and Description

Figure 3 shows the setup of the field test when collecting data with an Android smartphone (Huawei Honor FRD-AL00). The smartphone was placed on the cabin floor above the bogie to measure the condition of the vehicle-track system when trains move along the track. The embedded inertia measurement unit (LSM6DS3, manufactured by STMicroelectronics) and microphone sensor (MP34DB02, manufactured by STMicroelectronics) of the smartphone were used to acquire the vibration acceleration, angular velocity, and audio data during the test. More detailed information about the smartphone and sensors is shown in Table 1. According to the comparative experiment, laying the smartphone on the floor directly without securement slightly affects the amplitude of the vertical acceleration, but it performs well in other directions and frequency domains, which have no significant impact on our analysis [16]. Additionally, in this study, the measure accelerations were transformed from the coordinate system of the smartphone to the vehicle for eliminating errors caused by different coordinate systems [37]. However, the accelerations in different positions within the car were different; therefore, in this study, the smartphone was placed in the same position to overcome this issue.

By reading the embedded sensors of the smartphone with the developed application, the data required, that is, seven types of signals, were readily available. The audio signal was acquired using the microphone sensors of the smartphone, and the dynamic responses of the car body including the vibration accelerations (horizontal, vertical, and longitudinal accelerations) and rotational angular velocities (pitch angular, yaw angular, and roll angular velocities) were available with the built-in inertial sensors. In our study, the sampling frequency of the audio signal was 22,050 Hz and that of all the inertial sensors was set as 100 Hz.

The sound pressure recorded by using the microphone differed from that perceived by the human ear (even in the same field) owing to various factors, such as psychological effects, presence of outer ear, and cochlear health status. To objectively reflect the passenger's hearing experience, we employed the A-weighting sound pressure level (SPL(A)) in

this study to evaluate the interior noise. The SPL(A) can be calculated as follows:

$$\text{SPL(A)} = 20 * \log\left(\frac{P_A}{P_{\text{ref}}}\right), \quad (1)$$

where P_A is the A-weighting sound pressure (in Pa) and P_{ref} is the reference pressure (in Pa), which is typically set as 2×10^{-6} Pa. The data, including horizontal, vertical, and longitudinal accelerations and pitch angular, yaw angular, and roll angular velocities collected by the smartphone, were used to describe the dynamic responses of the car body. However, using only the data collected by the vehicle-carried smartphone may not be sufficient to explain the causes of the interior noise. For better prediction results, we considered the rotational angular accelerations of the car body obtained by calculating the first derivative of the rotational angular velocity. The effects of the trains' running speed on the generation of interior noise were nonnegligible. As a crucial parameter, the train speed was introduced to the prediction model. Because the subway tunnels are a GPS-free environment, the location and running speed information were not available through the GPS module embedded in the smartphones. The first-order integration of the longitudinal acceleration was used to overcome the problem. The running velocity of the train can be calculated by the following equation [37]:

$$v(t) = \int_0^t a_L dx + v_0, \quad (2)$$

in which $v(t)$ corresponds to the running velocity of the train at time t , a_L is the longitudinal acceleration of the car body, and v_0 is the initial velocity of the train. Because the stopping state of the train can be easily recognized using a smartphone, v_0 is typically regarded as 0. Because the velocity cannot be validated directly, by regarding the interval length between adjacent stations as the ground truth, the error between the integral displacement and real interval length is 9.5%. [37]. The data collected and used in this model are presented in Table 2.

A disadvantage of using raw signals collected by smartphones is that they show track quality problems in only single points, which complicates the generation of a reasonable section length for optimization. Additionally, the sampling frequencies of the inertial sensors and microphone are different, which renders it challenging to build a point-to-point relationship directly. Hence, a moving time window method was employed in this study. For confirming the optimal time window, the size of the time window was varied from 0.5 to 10 s, and the hop length of the window was half the size of the time window. Additionally, a series of features was selected to reflect the characteristics of the signals inside the window. We selected the root mean square (RMS) of the SPL(A) in the windows as the index of the interior noise. As for the dynamic responses of the car body, the following features were used: (1) mean value, (2) RMS, (3) variance, (4) standard deviation, (5) peak value, (6) skewness, (7) kurtosis, (8) shape factor, (9) crest factor, and (10) clearance factor. In each frame of the dynamic response signals of the car body,

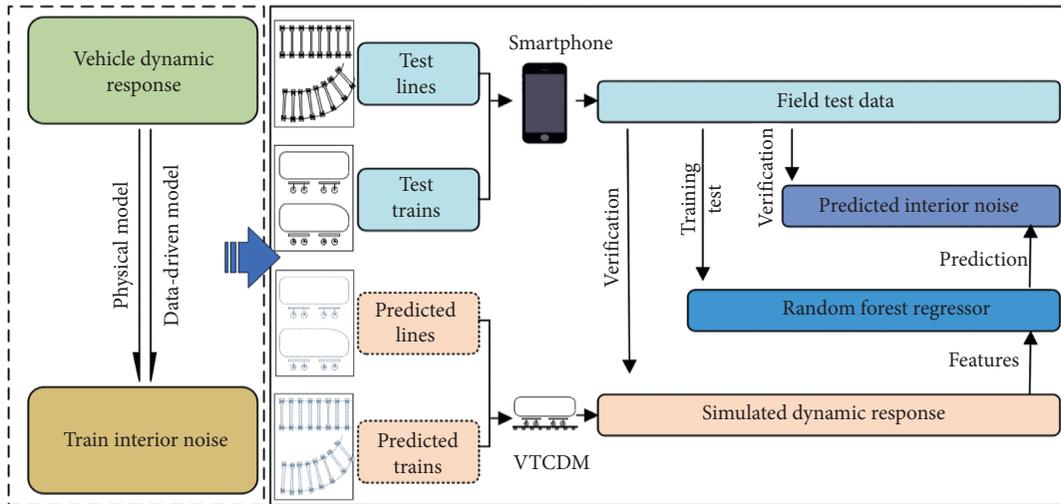


FIGURE 2: Research methodology.

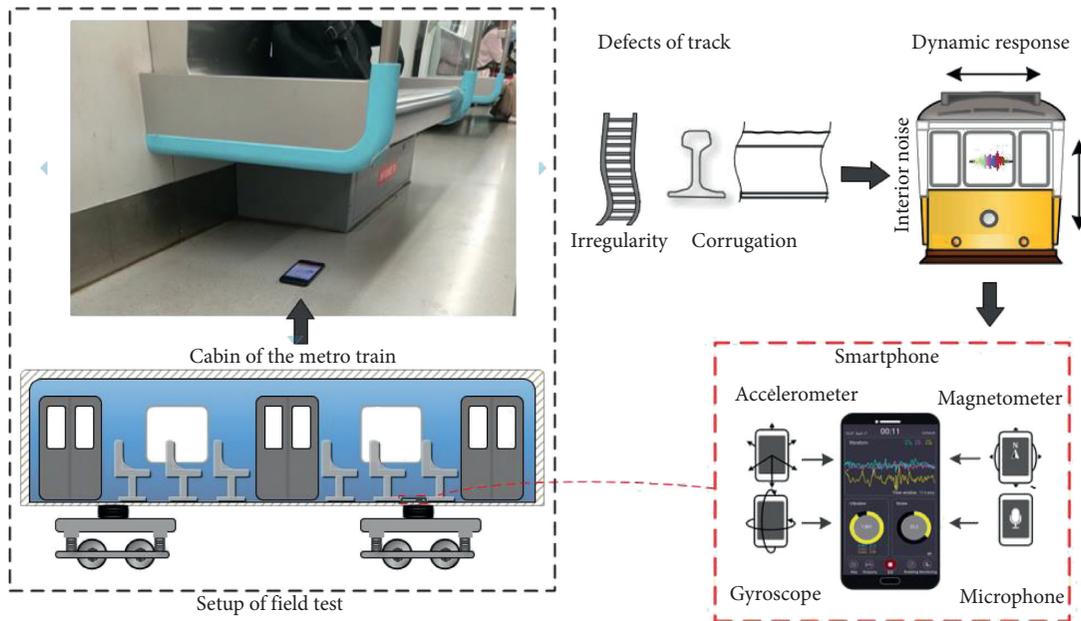


FIGURE 3: Data collection framework for this study.

TABLE 1: Parameters about the smartphone and sensors.

Device	Item	Parameter
Smartphone	Model	FRD-AL00
	Operating system	Android 8.0
	Chipset	HiSilicon Kirin 950
	CPU	Octa-core (4×2.3 GHz and 4×1.8 GHz)
	GPU	Mali-T880 MP4
	Memory	4 GB RAM
Inertial measurement unit	Model	LSM6DS3
	Acceleration range	±2/±4/±8/±16 g
	Angular rate range	±25/±250/±500 dps
	Analog supply voltage	1.71 V to 3.6 V
	Power consumption	1.25 mA

TABLE 1: Continued.

Device	Item	Parameter
Microphone sensors	Model	MP34DB02
	Frequency range	20 Hz to 20 kHz
	Supply voltage range	1.64 V to 3.6 V
	Acoustic overload point	120 dB SPL
	Signal-to-noise ratio	62.6 dB
	Sensitivity	-26 dBFS

TABLE 2: Features involved in the prediction model.

	Collected signal	Used signal	Selected features
1	Audio signal	SPL(A)	RMS
2	Horizontal acceleration	Horizontal acceleration	Mean value
3	Vertical acceleration	Vertical acceleration	RMS
4	Longitudinal acceleration	Longitudinal acceleration	Variance
5	Pitch angular velocity	Pitch angular velocity	Standard deviation
6	Yaw angular velocity	Yaw angular velocity	Peak value
7	Roll angular velocity	Roll angular velocity	Skewness
8		Pitch angular acceleration	Kurtosis
9		Yaw angular acceleration	Shape factor
10		Roll angular acceleration	Crest factor
11		Operating velocity of trains	Clearance factor

these features were calculated as input parameters for our prediction model.

5. Modeling Approach

5.1. RF. RF is an ensemble learning approach comprising hundreds of decision trees for performing classification or regression tasks independently. Using the average of all the trees' results as the final output can significantly improve the predictive accuracy. RF was developed based on the decision tree structure; however, two additional characteristics, including bagging and random subspace methods, were added to improve the accuracy and robustness. First, by creating a series of bootstrap samples, the bagging algorithm is fundamental for improving the accuracy and stability of machine-learning models. The introduction of the bagging method helps reduce variance and control overfitting. Next, the random subspace method is designed to increase tree independence by generating trees with a random sample of features rather than the entire feature set [38]. Additionally, the collected raw data can be directly fed into the RF model without any preprocessing [39], which renders the RF method easier to implement. Furthermore, the RF method can be used for feature importance analysis.

As a data-driven method, RF relies more on a tremendous amount of data to perform interior noise prediction. Compared with mathematical and physical models, data-driven prediction models do not require complex parameters, strict conditions, or hypotheses. Furthermore, this method can potentially reduce computational time and require less computer memory compared with numerical models. The mapping relationship between the dynamic responses of the car body and interior noise exhibits a strong nonlinear characteristic. Because the RF is a combination of

a series of decision trees, it exhibits an outstanding performance for fitting nonlinear relationships. Therefore, RF was selected in this study to fit the mapping relationship between the dynamic responses of the car body and the interior noise.

In RF, one dependent variable exists, that is, the RMS of SPL(A) for each moving window. Moreover, 100 independent variables (10 dynamic response signals \times 10 selected features; 10 dynamical response signals are items 2 to 11 of Column 3, Table 2; 10 selected features are the items 2 to 11 of Column 4, Table 2) were input to the model. In the model study, the effects of the number of decision trees, window size, and modes of the maximum amount of features were investigated. The computational time of different max feature modes was also considered. Using the feature selection method provided by RF, we performed the feature importance analysis.

5.2. VTCDM. Using RF, we established the mapping relationship between the dynamic response of the car body and the subway vehicle interior noise. However, it is difficult to exert its value using only such a mapping relationship. To expand its application scenario, we propose applying the mapping relationship for evaluating the design of the URT railway alignments based on the simulated dynamic responses of the vehicles. Hence, a VTCDM was constructed, with which the required simulated dynamic response signals could be obtained.

Vibrations of the vehicle can be transmitted to the track via the wheel-rail contact and cause vibrations of the track structure. However, the vibrations of the vehicle are also affected by the track in reverse [40]. Therefore, the vibrations of the track and the vehicle couple with each other, which allows us to evaluate the track state with the dynamic

responses of the train. Using the parameters of a type A subway train, we established the vehicle submodel in UM, which comprised 15 rigid bodies, such as the car body, bogies, wheelsets, and axle boxes. Moreover, the detailed parameters of the type A subway train used in our model are shown in Table 3. In the model, the flexibility and mass of the rail were ignored. Additionally, the rail was regarded as a massless block connected to the foundation with spring and damping force elements. Such a track model could be used to analyze the dynamic response in a low-frequency range. The effect of high-frequency vibration on the wheel-rail contact behavior was neglected. However, this method provided a high calculation efficiency that satisfied the general design requirements. In this study, the fifth-grade American spectrum was adapted as an excitation to input the model. Finally, a nonelliptical contact model was used to couple the two subsystems.

6. Results and Discussion

After RF and the VTCDM were established, a series of analyses were implemented. First, the performance of the RF regression model was analyzed in terms of parameter tuning, feature importance, and error distribution. Next, the VTCDM and the interior noise prediction model were verified through the field test data. Subsequently, an application scenario for evaluating the railway alignments from the perspective of interior noise was presented.

6.1. Analysis of the RF Regression Model

6.1.1. Parametric Analysis. To obtain the optimal parameters of RF, the effects of the number of decision trees, size of time windows, and mode of the max feature numbers on the performance were studied. It is noteworthy that the Out-of-Bag R^2 (OOB R^2) was employed as a critical parameter to evaluate the performance of the RF-based regression model [41–43]. Figure 4(a) presents the effects of the number of decision trees on the performance of this forecasting model with different time windows. As one can see, OOB R^2 increases gradually as the number of trees increases. However, when there are more than 150 trees, OOB R^2 remains stable. The relationship between OOB R^2 and the size of time windows is presented in Figure 4(b). It is indicated that, as the window size changes from 0.5 to 10 s, OOB R^2 shows an overall downward trend, and when the window size is 0.5 s, OOB R^2 reaches its maximum value of 0.776. Meanwhile, local fluctuations occur when the window size is 4 s.

The number of features considered when forming the random forests significantly affects the goodness of fit. Three max feature modes, “auto,” “sqrt,” and “log2,” are discussed from the perspective of computation time and goodness of fit, as shown in Figure 4(c). If the “auto” mode was selected, then the number of features to consider when forming the trees was equal to the number of all features input to the model. If “sqrt” was selected, the number of features considered was the square root of the number of all features. When mode “log2” was selected, the number of features considered was the base-two logarithm of the number of total features. In terms of

TABLE 3: Key parameters of type A subway vehicle.

	Parameters	Value
1	Length between bogie pivot centers (m)	15.7
2	Wheelbase (m)	2.5
3	Lateral distance of the rolling circle (m)	1.493
4	Nominal rolling circle radius (m)	0.84
5	Mass of car body (kg)	50 877
6	Mass of bogie (kg)	2 721.5
7	Mass of wheel (kg)	1900
8	Wheel tread type	LMA

goodness of fit, the “auto” mode performed the best, followed by “sqrt” and “log2.” Furthermore, the calculation time of the three methods increased linearly with the increase in the number of decision trees. Under the same condition, the “auto” mode consumed the most time, whereas the “sqrt” mode consumed slightly more time than the “log2” mode. It is clear from the analysis that although the “auto” mode consumed the most time, its fitting effect was the best; the “log2” mode computed faster than the other two methods but yielded the worst performance.

From the analysis above, the optimal parameters of the RF were confirmed. The RF with 200 trees and “auto” max feature mode was selected in our study to obtain better fitting results with less computing capacity. Figure 4(d) shows the comparison between the predicted interior noise by RF and the field test results. This figure shows that the interior noise by RF agrees well with the field test results. However, when the train stopped at the stations, significant errors occurred therein. This was because when the train stopped or started, the primary source of interior noise of the subway vehicles changed from wheel-rail noise to electro-mechanical equipment noise and broadcast sound.

6.1.2. Prediction Error Distribution Analysis. Additionally, we studied the distribution characteristics of prediction errors of the RF regression model (RFRM). The prediction error distribution of the RFRM was compared with that of the linear regression model (LRM), radial-basis-function-based support vector regression model (rbf-SVRM), and gradient boosting regression model (GBRM). Figure 5(a) shows the density function curves of prediction errors of different models. Figure 5(b) shows the cumulative distribution function curves of different models. First, the figure shows that the prediction errors of the RF have the highest occurrence probability in the area nearby 0, which indicates that the RFRM is more accurate than the other models, that is, GBRM, rbf-SVRM, and LRM, in that order. For the RFRM, prediction errors within 5 and 10 dB(A) constitute 86.9% and 98.1% of the 13,000 test samples, respectively. It is noteworthy that the accuracy of this prediction model requires further improvement.

6.1.3. Feature Importance Analysis. The essence of feature importance assessment in RF is to calculate the average of each feature’s contribution to all trees in the forest and subsequently compare the contributions among those

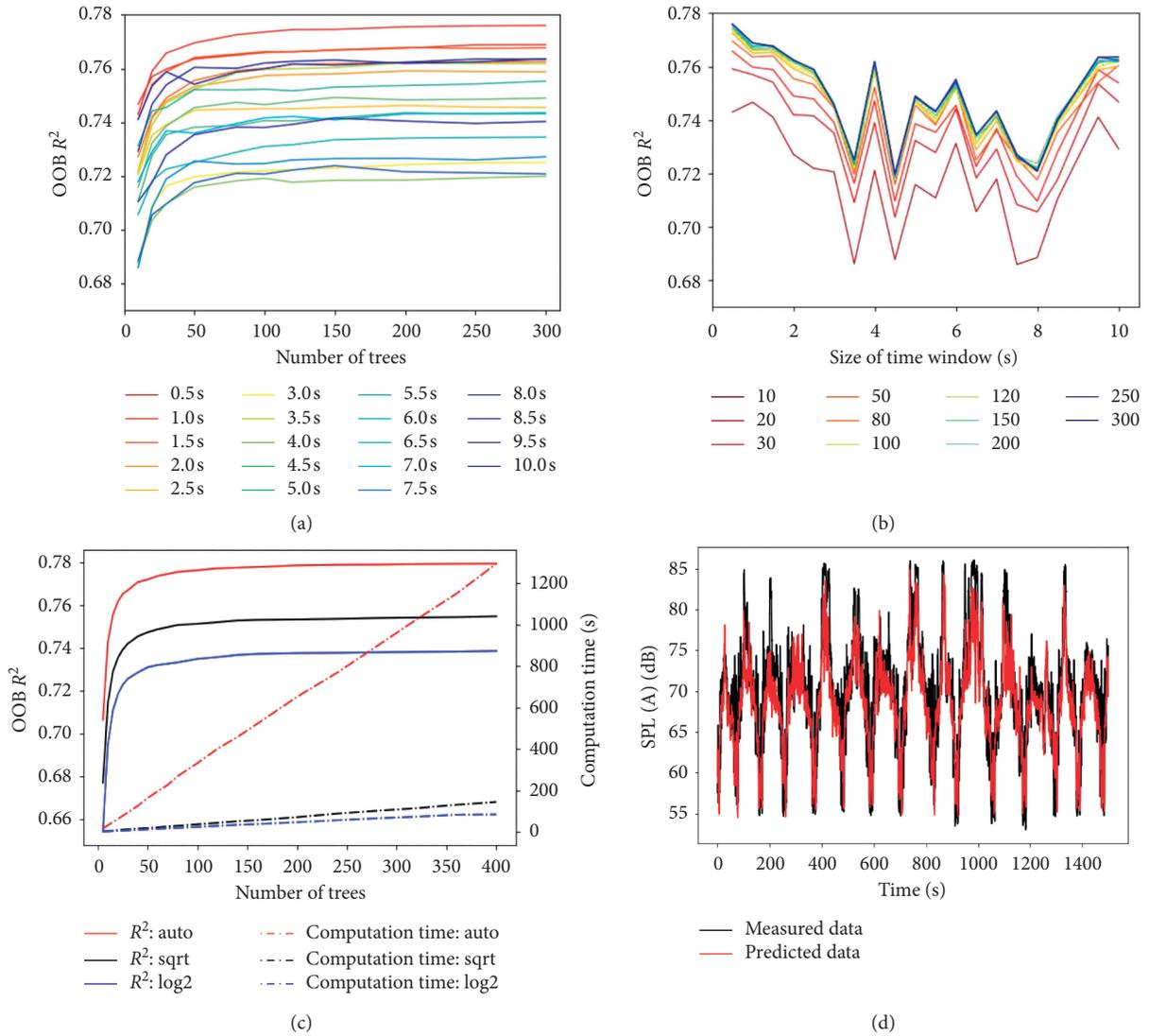


FIGURE 4: Effect of number of estimators on OOB R^2 .

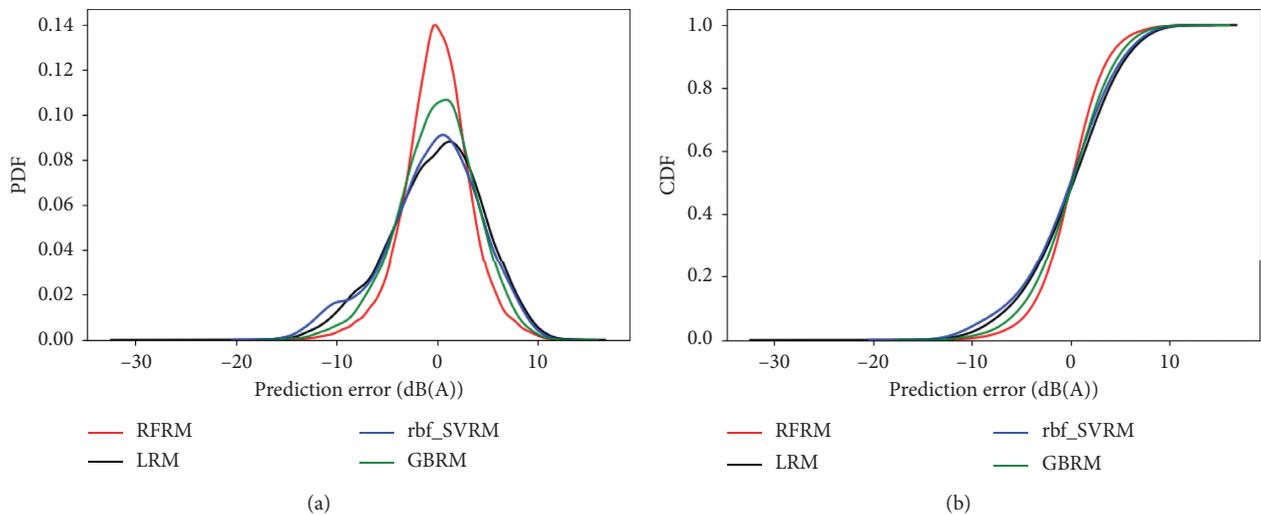


FIGURE 5: Comparison of the prediction error distributions.

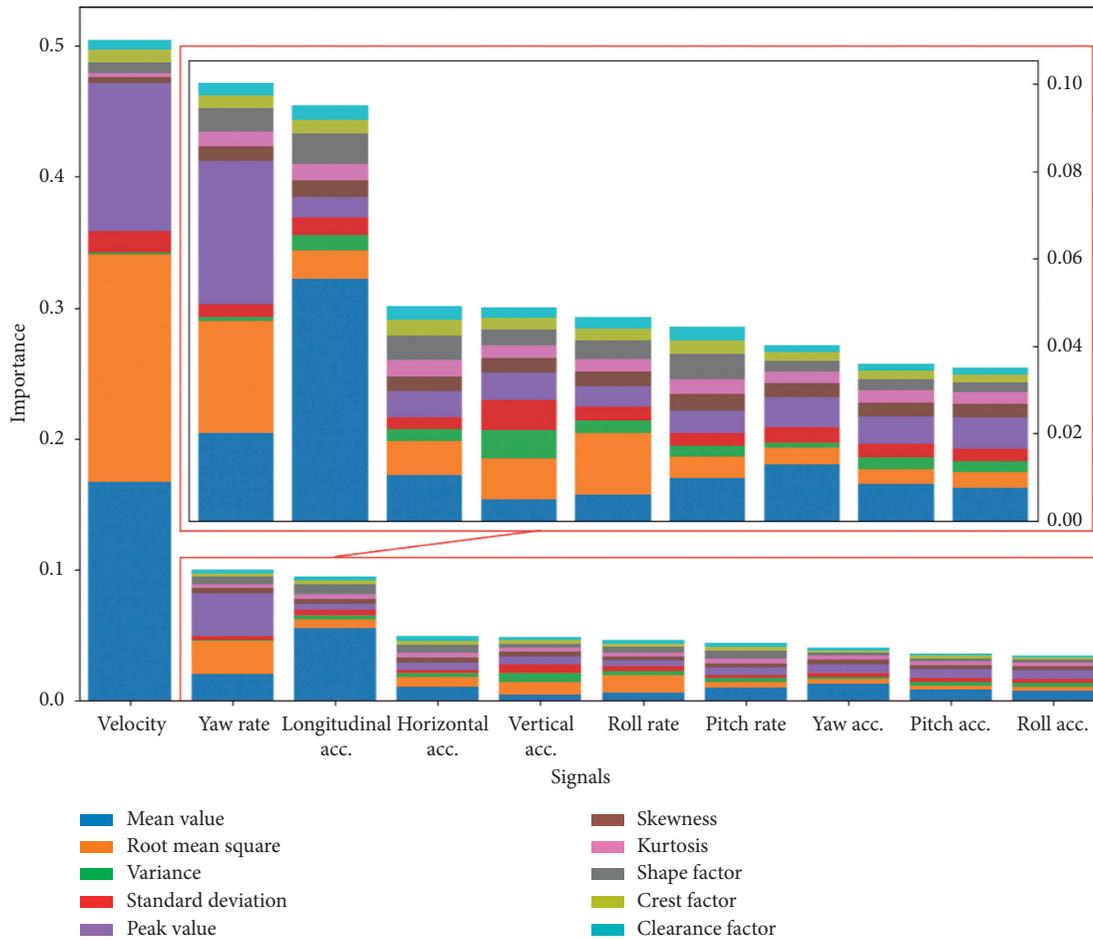


FIGURE 6: Feature importance analysis.

features. The mean decrease impurity method was adopted in this study to perform feature importance analysis and variance was used as the impurity measure [44]. The results of the feature importance analysis are presented in Figure 6. For showing the specific details clearly, an enlarged view of the feature importance of the partial signals is included. A total of 100 features (10 signals \times 10 indexes) are presented, and the summary of these features' importance coefficient is 1. Figure 6 shows that the importance of velocity is the highest, reaching nearly 0.5. This suggests that the running speed is the most crucial factor affecting vehicle interior noise. The importance of yaw rate and longitudinal acceleration ranks second and third with 0.10 and 0.096, respectively. This is because the yaw rate relates well with the curve radius, where rail corrugation and squeal often occur. Among all the indexes, the mean value, RMS, and peak value are the most important ones.

6.2. Verifications of VTCDM and the Interior Noise Prediction Model. Not only the RF but also the VTCDM significantly affects the performance of the interior noise prediction model. In Figure 7(a), the simulated dynamic responses, including the horizontal acceleration (H. acc.), vertical acceleration (V. acc.), longitudinal acceleration (L.

acc.), pitch rate, yaw rate, and roll rate, are compared with the field measured data. The figure shows that the simulated data agree well with the measured data. However, the simulated vertical acceleration of the car body is lower than the measured data, which may be caused by the simplification of the car body into a rigid body. The comparison between the measured and forecasted vehicle interior noise is shown in Figure 7(b). The predicted SPL(A) by our model is similar to the measured data. However, our model cannot predict the fluctuation of the interior noise caused by the broadcast sound when the train starts and brakes.

6.3. Case Study for Optimizing Railway Alignment Designs. We developed an application case of the interior noise prediction model for evaluating railway alignments in the design stage. In this study, we designed three types of railway alignments, as shown in Figure 8(a). These three railway alignments have the same parameters except the radius of the circular curve. The radii of the circular curve were 400, 800, and 2000 m, which are typical in subway lines. Subsequently, the designed railway alignments were input to the VTCDM for obtaining the simulated dynamic responses of the car body. The simulated train moved at a constant speed of 60 km/h under the three alignments above. Subsequently,

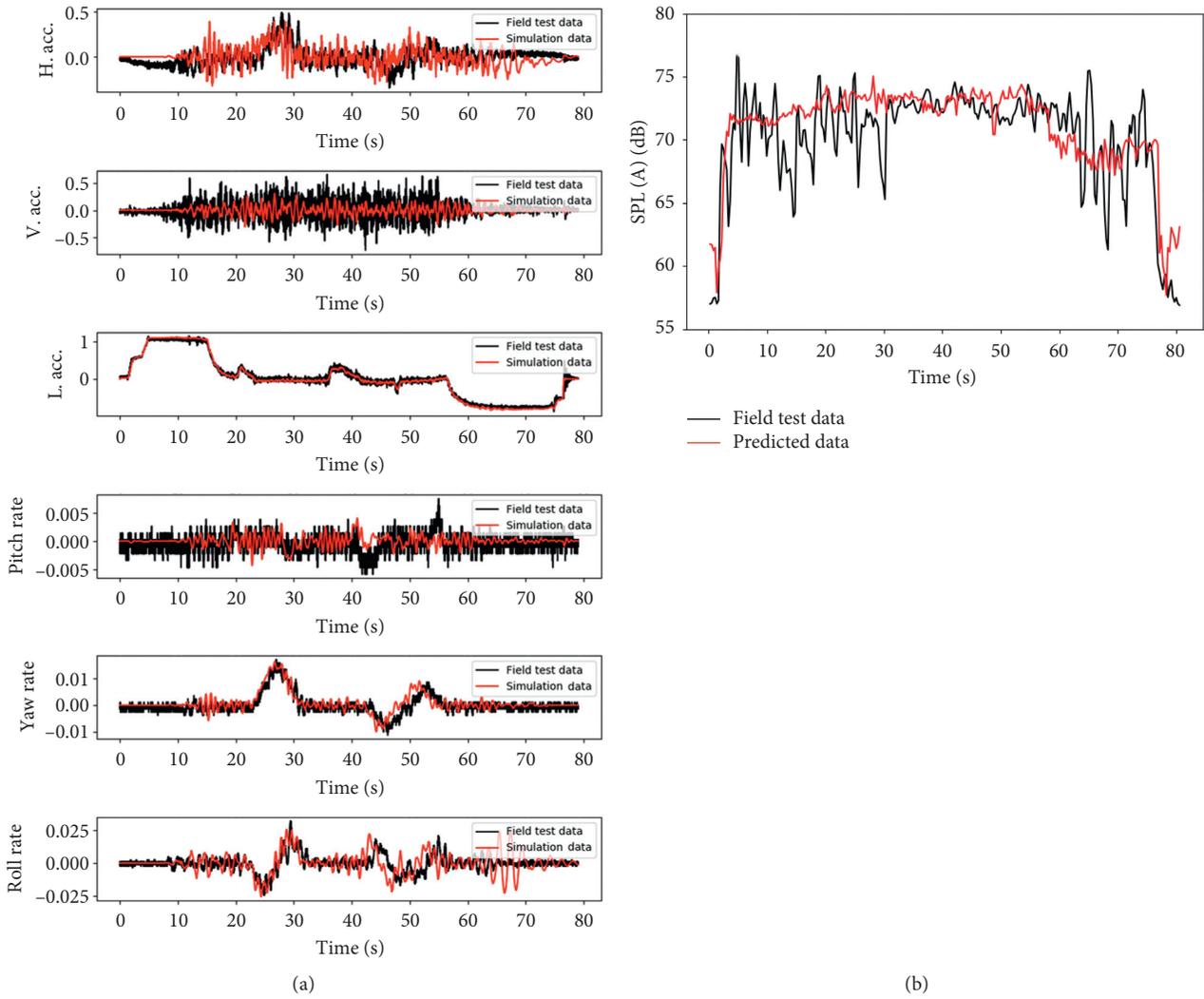


FIGURE 7: Verification of the VTCDM and interior noise prediction model.

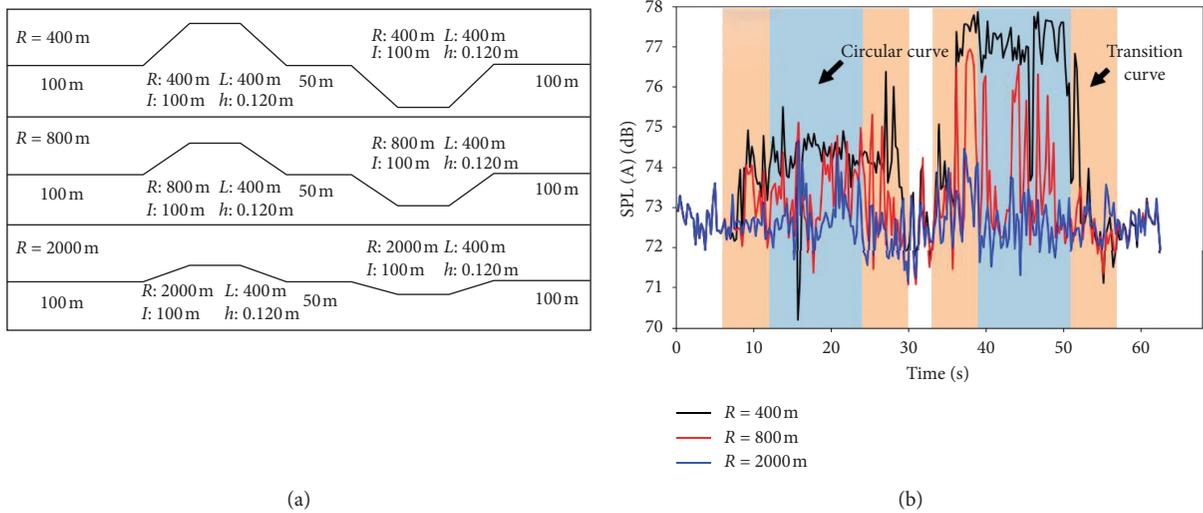


FIGURE 8: Evaluation of railway alignments considering interior noise.

the predicted interior noise (SPL(A)) was obtained by inputting the simulated dynamic responses to the trained RF.

In Figure 8(b), the predicted interior noise under different railway alignments is presented (the black line refers to the SPL(A) with circular curves of radius 400 m; the red line denotes the SPL(A) with circular curves of radius 800 m; the blue line represents the SPL(A) with circular curves of radius 2000 m). The findings obtained from the figure are as follows: first, the predicted SPL(A) in the curved segments, especially in the circular sections, is higher than that in the straight-line segments; next, the predicted SPL(A) increases as the radius of the curve decreases. Therefore, sharp curves must be avoided in the design of railway alignments for reducing the interior noise of the train. The purpose of this case study is to verify the feasibility of using this interior noise prediction model for evaluating railway alignments. Regulations for designing rail alignments consider more about the economy and the running stability of trains, such as the lateral acceleration [45–47], whereas, as one of the critical indices for the ride comfort, the noise level is not sufficiently considered in designing rail alignments. These regulations usually provide a general limit value of noise level, which cannot be directly applied to designing rail alignments. Comparing the predicted noise levels corresponding to different alignments with those limit values allows us to evaluate rail alignments from the perspective of noise. The combination of the predicted noise level and regulations in force is able to make the designed rail alignments more reasonable.

In this case, we did not consider different vehicles, as only one type of vehicle was selected for a specific subway line in general. For simplifying the model of the cases, the impact of travel speeds of the train on the noise level was not considered, whereas in practical operation, trains usually repeatedly go through the process of starting-speeding up—keeping a high travel speed—speeding down—stopping. In the section with small-radius curves, the higher the speed of the train is, the more likely it causes noise issues, such as squeal. Therefore, the sections where trains have a high travel speed should avoid small-radius curves or reverse curves as much as possible in the design of rail alignments.

7. Conclusions

The interior noise of subway vehicles based on RF and VTCDM was forecasted herein using data collected from onboard smartphones. Through parameter study, we confirmed the optimal parameters of the RF model: number of trees, 200; size of time window, 0.5 s; “auto” max feature mode selected. With such parameters, R^2 reached its highest value, 0.78. Error analysis showed that RF had a higher prediction accuracy than the other models. The probability of a prediction error within 5 dB was 86.9%.

Feature importance analysis demonstrated that the running velocity of the train affected the interior noise the most. Additionally, the yaw rate and longitudinal acceleration of the car body affected the vehicle interior noise significantly. The effectiveness of the VTCDM was verified through comparisons between the simulated and measured

dynamic responses of the car body. From the comparison with the field-test vehicle interior noise, it was evident that the proposed prediction method could accurately predict the trend of subway train interior noise. However, the local fluctuations of the interior noise caused by some nonwheel-rail noise, such as the sounds of broadcast and electrical devices, could not be forecasted correctly. The case study demonstrated that the proposed interior noise forecasting method could be used for evaluating and optimizing railway alignment designs in the early stage from the perspective of interior noise.

Next, to improve the performance of the prediction model, we could categorize the training data in detail according to different characteristics, such as track type, train type, and service time of the track system. Because the simplification of the vehicle track into a multirigid-body system in the current model reduced the forecasting accuracy, the flexibility of the car body and track shall be considered in future studies.

Data Availability

Noise data will be available upon request to the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was funded by the China Scholarship Council (ID: 201907000077) and National Natural Science Foundation of China (Grant no. 51878576).

References

- [1] J. Zhang, X. Xiao, X. Sheng, Z. Li, and X. Jin, “A systematic approach to identify sources of abnormal interior noise for a high-speed train,” *Shock and Vibration*, vol. 2018, Article ID 5085847, 12 pages, 2018.
- [2] B. T. Atmaja, M. F. Puabdillah, M. N. Farid, and W. A. Asmoro, “Prediction and simulation of internal train noise resulted by different speed and air conditioning unit,” *Journal Of Physics*, vol. 1075, Article ID 012038, 2018.
- [3] W. Dai, X. Zheng, L. Luo, Z. Hao, and Y. Qiu, “Prediction of high-speed train full-spectrum interior noise using statistical vibration and acoustic energy flow,” *Applied Acoustics*, vol. 145, pp. 205–219, 2019.
- [4] J. Han, X. Xiao, Y. Wu, Z. Wen, and G. Zhao, “Effect of rail corrugation on metro interior noise and its control,” *Applied Acoustics*, vol. 130, pp. 63–70, 2018.
- [5] D. J. Thompson, G. Squicciarini, B. Ding, and L. Baeza, “A state-of-the-art review of curve squeal noise: phenomena, mechanisms, modelling and mitigation,” *Noise and Vibration Mitigation for Rail Transportation Systems*, vol. 3–41, 2018.
- [6] J. Zhang, G. Han, X. Xiao, R. Wang, Y. Zhao, and X. Jin, “Influence of wheel polygonal wear on interior noise of high-speed trains,” *Journal of Zhejiang University Science*, vol. 15, no. 12, pp. 1002–1018, 2014.
- [7] H. Tsunashima, Y. Naganuma, A. Matsumoto, T. Mizuma, and H. Mori, “Condition monitoring of railway track using

- in-service vehicle,” *Reliability and Safety in Railway*, vol. 12, pp. 334–356, 2012.
- [8] Y. Hayashi, T. Kojima, H. Tsunashima, and Y. Marumo, “Real time fault detection of railway vehicles and tracks,” in *Proceedings of the Institution of Engineering & Technology International Conference on Railway Condition Monitoring*, Birmingham, UK, 2007.
- [9] M. Odashima, S. Azami, Y. Naganuma, H. Mori, and H. Tsunashima, “Track geometry estimation of a conventional railway from car-body acceleration measurement,” *Mechanical Engineering Journal*, vol. 16, Article ID 00498, 2017.
- [10] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [11] Y. Yu, R. Han, X. Zhao et al., “Initial validation of mobile-structural health monitoring method using smartphones,” *International Journal of Distributed Sensor Networks*, vol. 11, no. 2, Article ID 274391, 2015.
- [12] S. H. de Frutos and M. Castro, “Using smartphones as a very low-cost tool for road inventories,” *Transportation Research Part C: Emerging Technologies*, vol. 38, pp. 136–145, 2014.
- [13] A. Paixão, E. Fortunato, and C. Rui, “Smartphone’s sensing capabilities for on-board railway track monitoring: structural performance and geometrical degradation assessment,” *Advances in Civil Engineering*, vol. 2019, Article ID 1729153, 13 pages, 2019.
- [14] S. Kanarachos, S.-R. G. Christopoulos, and A. Chroneos, “Smartphones as an integrated platform for monitoring driver behaviour: the role of sensor fusion and connectivity,” *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 867–882, 2018.
- [15] A. Ghose, P. Biswas, C. Bhaumik, M. Sharma, A. Pal, and A. Jha, “Road condition monitoring and alert application: using in-vehicle smartphone as internet-connected sensor,” in *Proceedings of the 2012 IEEE International Conference on Pervasive Computing and Communications Workshops*, IEEE, Lugano, Switzerland, pp. 489–491, March 2012.
- [16] P. Wang, Y. Wang, Li Wang, R. Chen, and J. Xiao, “Measurement of carbody vibration in urban rail transit using smartphones,” 2017.
- [17] H.-M. Noh, “Contribution analysis of interior noise and floor vibration in high-speed trains by operational transfer path analysis,” *Advances in Mechanical Engineering*, vol. 9, no. 8, Article ID 1687814017714986, 2017.
- [18] H. B. Huang, X. R. Huang, M. L. Yang, T. C. Lim, and W. P. Ding, “Identification of vehicle interior noise sources based on wavelet transform and partial coherence analysis,” *Mechanical Systems and Signal Processing*, vol. 109, pp. 247–267, 2018.
- [19] J. Zhang, X. B. Xiao, X. Z. Sheng, R. Fu, D. Yao, and X. S. Jin, “Characteristics of interior noise of a Chinese high-speed train under a variety of conditions,” *Journal of Zhejiang University-Science A*, vol. 18, no. 8, pp. 617–630, 2017.
- [20] Y. Sun and Y. Zhao, “Characteristics of interior noise in MonoRail and noise control,” in *Proceedings of the Inter-Noise and Noise-Con Congress and Conference*, vol. 258, Institute of Noise Control Engineering, Chicago, IL, USA, pp. 1461–1467, August 2018.
- [21] K. Hu, Y. Wang, H. Guo, and H. Chen, “Sound quality evaluation and optimization for interior noise of rail vehicle,” *Advances in Mechanical Engineering*, vol. 6, p. 820875, 2014.
- [22] L. G. Kurzweil and R. Lotz, *Prediction and Control of Noise and Vibration in Rail Transit Systems*, Urban Mass Transportation Administration, Washington, DC, USA, 1978.
- [23] W. U. Di and J. Ge, “Analysis of the influence of racks on high speed train interior noise using finite element method,” *Applied Mechanics & Materials*, vol. 675–677, pp. 257–260, 2014.
- [24] L. Franzoni, J. Rouse, and T. Duvall, “A broadband energy based boundary element method for predicting vehicle interior noise,” *The Journal of the Acoustical Society of America*, vol. 115, no. 5, p. 2538, 2004.
- [25] G. M. L. Gladwell and G. Zimmermann, “On energy and complementary energy formulations of acoustic and structural vibration problems,” *Journal of Sound and Vibration*, vol. 3, no. 3, pp. 233–241, 1966.
- [26] S. H. Sung and D. J. Nefske, “A coupled structural-acoustic finite element model for vehicle interior noise analysis,” *Journal of Vibration and Acoustics*, vol. 106, no. 2, pp. 314–318, 1984.
- [27] C. Lucas, D. Shahmirzadi, and M. N. Bahrami, “A FEM-based quasi-static neuro-model for acoustic noise in switch reluctance Motors,” *Journal of Computational Acoustics*, vol. 12, no. 1, pp. 85–97.
- [28] E. Ghafoorifard, M. Claessens, and X. Hu, “An analytical-experimental method for analysing the low-frequency interior acoustics of a passenger car,” in *Proceedings of the International Seminar on Modal Analysis*, vol. 3, KU Leuven, Orlando, FL, USA, pp. 1331–1338, February 1998.
- [29] F. Létourneaux, S. Guerrand, and F. Poisson, “Low-frequency acoustic transmission of high-speed trains: simplified vibroacoustic model,” *Journal of Sound and Vibration*, vol. 231, no. 3, pp. 847–851, 2000.
- [30] A. Soltani and M. K. Demmeh, “Analyzing of noise inside a simple vehicle cabin using boundary element method,” *World Academy of Science, Engineering and Technology*, vol. 49, pp. 630–634, 2011.
- [31] R. H. Lyon and G. Maidanik, “Power flow between linearly coupled oscillators,” *The Journal of the Acoustical Society of America*, vol. 34, no. 5, pp. 623–639, 1962.
- [32] B. James, H. Andrew, and D. Sam, “Aero-acoustic tests and analysis to predict internal and external railway vehicle aerodynamically-generated noise,” in *Proceedings of the Tenth International Congress on Sound and Vibration*, Stockholm International Institute of Acoustics and Vibration, Stockholm, Sweden, pp. 1697–1704, July 2003.
- [33] S. Givargis and H. Karimi, “Mathematical, statistical and neural models capable of predicting LA,max for the Tehran-Karaj express train,” *Applied Acoustics*, vol. 70, no. 7, pp. 1015–1020, 2009.
- [34] J. Forssén, S. Tober, A. C. Corakci, and A. Frid, “Modelling the interior sound field of a railway vehicle using statistical energy analysis,” *Applied Acoustics*, vol. 73, no. 4, pp. 307–311, 2012.
- [35] F. Kropp, Q. He, R. M. P. Goverde, and X. Liu, “Recent applications of big data analytics in railway transportation systems: a survey,” *Transportation Research Part C: Emerging Technologies*, vol. 90, pp. 226–246, 2018.
- [36] X. Lai and S. Paul, “Optimization of rail transit alignments considering vehicle dynamics,” *Transportation Research Record*, vol. 2275, no. 2275, pp. 77–87, 2012.
- [37] Y. Wang, J. Cong, H. Tang, X. Liu, T. Gao, and P. Wang, “A data fusion approach for speed estimation and location calibration of a metro train in underground environment based on low-cost sensors in smartphones,” *IEEE Sensors Journal*, vol. 19, no. 22, pp. 10744–10752, 2019.
- [38] A. Falamarzi, S. Moridpour, M. Nazem, and S. Cheraghi, “Development of random forests regression model to predict track degradation index: melbourne case study,” in

Proceedings of the Australasian Transport Research Forum (ATRF), 40th, 2018, Darwin, Australia, October 2018.

- [39] Y. Li, C. Zou, M. Bercibar et al., "Random forest regression for online capacity estimation of lithium-ion batteries," *Applied Energy*, vol. 232, pp. 197–210, 2018.
- [40] W. Zhai, K. Wang, and C. Cai, "Fundamentals of vehicle-track coupled dynamics," *Vehicle System Dynamics*, vol. 47, no. 11, pp. 1349–1376, 2009.
- [41] U. Grömping, "Variable importance assessment in regression: linear regression versus random forest," *The American Statistician*, vol. 63, no. 4, pp. 308–319, 2009.
- [42] D. Liakhovitski, Y. Bryukhov, and M. Conklin, "Relative importance of predictors: comparison of random forests with johnson's relative weights," *Model Assisted Statistics and Applications*, vol. 5, no. 4, pp. 235–249, 2010.
- [43] M. Vreugdenhil, W. Wagner, B. Marschallinger, I. Pfeil, I. Teubner, and C. Rüdiger, "Sensitivity of sentinel-1 backscatter to vegetation dynamics: an Austrian case study," *Remote Sensing*, vol. 10, no. 9, p. 1396, 2018.
- [44] G. Strauss, L. Wehenkel, A. Sutera, and P. Geurts, "Understanding variable importances in forests of randomized trees," *Advances In Neural Information Processing Systems*, vol. 1, pp. 431–439, 2013.
- [45] EU Commission, "Commission regulation (EU) no 1304/2014," 2014, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32014R1304>.
- [46] GB 50157-2013 and Mohurd, *Code for Design of Metro*, Scientific Research, Wuhan, China, 2013.
- [47] GB 14892-2006 and Mohurd, *Noise Limit and Measurement for Train of Urban Rail Transit Standards*, Scientific Research, Wuhan, China, 2006.