Sound & Vibration 2025, 59(1), 2022.
https://doi.org/10.59400/sv2022

Article

Vehicle structural road noise prediction based on an improved Long Short-Term Memory method

Xiongying Yu[1,3], Ruxue Dai[2], Jian Zhang[1], Yingqi Yin[2], Sha Li[1], Peisong Dai[2,*], Haibo Huang[2,*]

[1] Changan Auto Global R&D Center, Chongqing Changan Automobile Co., Ltd, Chongqing 400023, China
[2] School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, China
[3] State Key Laboratory of Vehicle Noise, Vibration and Harshness (NVH) and Safety Technology, Chongqing 401120, China
*** Corresponding authors:** Peisong Dai, dps@my.swjtu.edu.cn; Haibo Huang, huanghaibo214@swjtu.edu.cn

CITATION

Yu X, Dai R, Zhang J, et al. Vehicle structural road noise prediction based on an improved Long Short-Term Memory method. Sound & Vibration. 2025; 59(1): 2022. https://doi.org/10.59400/sv2022

ARTICLE INFO

Received: 31 July 2024
Accepted: 12 November 2024
Available online: 9 January 2025

Abstract: The control of vehicle interior noise has become a critical metric for assessing noise, vibration, and harshness (NVH) in vehicles. During the initial phases of vehicle development, accurately predicting the impact of road noise on interior noise is essential for reducing noise levels and expediting the product development cycle. In recent years, data-driven methods based on machine learning have gained significant attention due to their robust capability in navigating complex data mapping relationships. Notably, surrogate models have demonstrated exceptional performance in this domain. Numerous researchers have integrated diverse intelligent algorithms into the study of vehicle noise, leveraging advantages such as the elimination of precise modeling requirements, extensive solution space exploration, continuous learning from data, and robust algorithmic versatility. However, in NVH engineering applications, data-driven models face inherent limitations, particularly in interpretability and stability. To address these issues, this paper introduces an improved Long Short-Term Memory (LSTM) network that combines knowledge and data. Inspired by the physical information neural network concept, this approach incorporates values calculated through empirical formulas into the neural network as constraints. Comparative assessments with traditional LSTM networks highlight the advantages of this deep learning model. By integrating empirical formulas constraints, the model not only enhances interpretability but also achieves robust generalization with fewer data samples. The proposed method is validated on a specific vehicle model, showing significant improvements in prediction accuracy and efficiency.

Keywords: NVH; road noise; empirical knowledge; LSTM; neural network

1. Introduction

With rapid scientific and technological advancements and higher living standards, the demand for automobiles has expanded beyond simple mobility to include greater comfort. The Noise, Vibration, and Harshness (NVH) performance of vehicles significantly affects both drivers and passengers psychologically and physiologically [1–3]. Consequently, companies are increasingly prioritizing NVH optimization. Traditional gasoline vehicles produce noise from the engine, transmission, intake and exhaust systems, road, and wind [4–7]. In particular, the road surface and tires play a significant role in traffic noise. The structural noise generated by the interaction of road surface roughness excites the tire's internal cavity and the hub coupling system, which also transmits noise [8–10]. Tires are a major source of noise, especially their tread patterns [11]. The noise response of different tread patterns varies when driving on the road surface. In contrast, electric vehicles, which lack engine noise, emphasize road noise and wind noise [12,13], especially in the process

of driving at medium speed on good urban roads, road noise has become a key research object.

In the early stages, conducting experiments to study acoustic comfort and overall noise in different scenarios is a common approach [14,15]. The composition of traffic and traffic volume have a significant impact on noise generation on typical roads. Alves et al. [16] studied noise generation by measuring the percentage of heavy vehicles in the total traffic volume on typical roads. Praticò et al. [17] emphasized the necessity of reducing traffic noise through the use of quiet road surfaces. The study of automobile road noise has evolved with advances in computer technology and acoustic theory, progressing from early single tests to Computer-Aided Engineering (CAE) simulations [18,19]. Yin et al. [20] predicted the acoustic performance of newly laid low-noise road surfaces by developing two models and using different surface mix parameters across various frequency ranges as inputs for the models. The integration of test technology and CAE is now essential for major automotive research and development. Various methodologies have emerged, such as Huang et al.'s [21] proposed an adaptive parallel filter method for fast response and suppression of different in-vehicle road noise, achieves fast noise reduction response and low steady-state error in attenuating varying in-vehicle road noise. Fan et al. [22] designed a magnetic levitation actuator for vehicle suspension to control road noise, demonstrating excellent vibration isolation in the 20–300 Hz range using a sport utility vehicle. Park et al. [23] utilized a component-level transfer path analysis based on the hysteresis force method to achieve the road noise target, and based on this interior noise prediction, the interior noise and vibration levels were reduced by a hysteresis force-based method. Despite advancements, challenges in predicting vehicle interior noise persist due to long test periods and high costs associated with complex analysis mechanisms.

Research on noise and vibration issues across various industries using experimental methods and CAE approaches often faces challenges such as long-time consumption and complex modeling [24–26]. However, in recent years, with the rapid development of computer technology, these challenges are being addressed. Data-driven approaches based on machine learning show promise for addressing complex data mapping relationships through surrogate models. Scholars have applied these methods to vehicle NVH analysis to improve research efficiency [27,28]. For instance, Wysocki et al. [29] created training data for an artificial neural network by deforming an initial component finite element model to find component design optimization parameters for the frequency response function target curve. Li et al. [30] predicted the noise value of the vehicle by means of Elman neural network and established a body structure optimization method with comprehensive consideration of NVH performance and side impact safety. Huang et al. [31] used empirical modal decomposition and sample entropy to extract noise features, established a wavelet neural network for sound quality prediction, and proposed a modified Long Short-Term Memory (LSTM) model based on adaptive learning rate forests. Yu et al. [32] analyzed that compared to Recurrent Neural Networks (RNN), LSTM can solve the long-term dependency problem well and give better predictions.

With the increase in automotive consumers, acoustic comfort has become an important issue. The design quality of sound within the vehicle cabin [33] can directly

impact driving safety and the overall riding experience [34]. Despite achievements in some NVH engineering applications, predicting vehicle interior noise effectively remains challenging for two main reasons [35,36]. First, neural network models require substantial training data, especially with complex structures. Insufficient training data can lead to overfitting, causing model failure. Second, designers often hesitate to adopt complex data-driven systems due to difficulties in understanding their internal mechanisms, including the interpretability of decision-making and evaluation results.

This paper introduces an improved Long Short-Term Memory (IMP-LSTM) method to address the limitations of existing methodologies. First, it proposes a novel approach by redesigning the neural network's loss function to incorporate local "empirical knowledge" as constraints. This innovation leverages empirical data to guide network learning, creating a hybrid model driven by both knowledge and data. This methodology addresses the shortcomings of traditional LSTM, which heavily rely on precise samples, especially for complex systems. Second, the study introduces normalized adaptive weights to enhance the model's predictive capability. This strategy allows the network to dynamically adjust the weights of loss terms, ensuring numerical comparability and preventing issues such as gradient disappearance or explosion, thereby enhancing training stability and efficiency.

The paper is structured as follows: Section 2 provides an overview of LSTM models, and introduces the proposed IMP-LSTM model theoretically; Section 3 details the experimental design and analyzes the results; Section 4 implements the proposed prediction method and validates the results; Finally, Section 5 summarizes the findings and insights of the study.

2. Methodology

2.1. Introduction to LSTM network structure

Predicting road noise encompasses numerous components, making it a multi-dimensional, nonlinear problem characterized by a small sample size derived from engineering practice. The LSTM network is particularly suitable for addressing this issue due to its superior time series data processing capabilities [37,38].

LSTM is a specialized type of RNN designed to address gradient disappearance and explosion issues common in traditional RNNs when handling long sequences [39]. It employs a unique 'gate' structure to control the flow of information, ensuring stable parameter updates. This structure determines which information to retain and which to discard based on input data characteristics, thus avoiding unnecessary information accumulation while preserving long-term dependencies [40]. As illustrated in **Figure 1**, the LSTM network topology includes an input layer, two hidden layers, and an output layer. The input layer receives external data and transmits it to the first hidden layer. Each hidden layer consists of multiple LSTM units interconnected to facilitate data flow. The second hidden layer passes the data to the output layer, which provides the final prediction results. In practice, the number of hidden layers in the LSTM network can be adjusted based on the complexity of specific tasks and data.

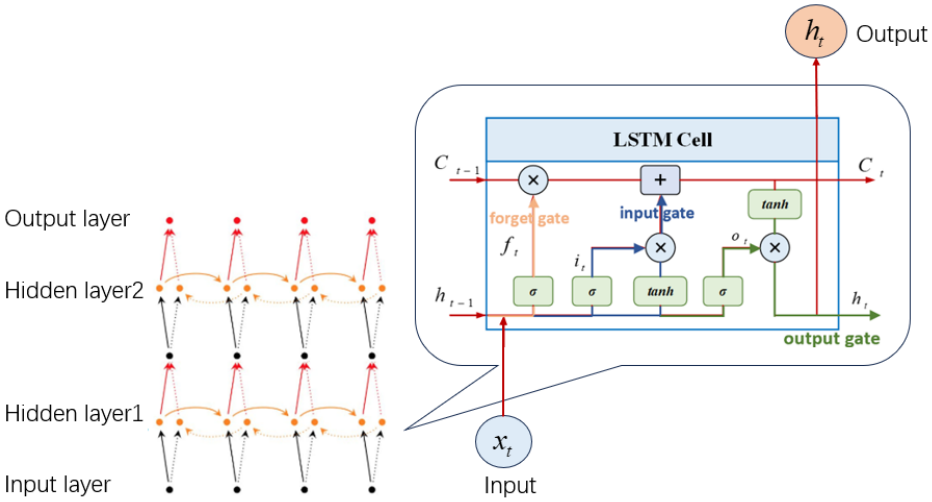

Figure 1. LSTM cell internal structure.

The right side of **Figure 1** illustrates the detailed cellular structure of the LSTM cell. Within the LSTM cell, there are three key "gate" structures: the forget gate, input gate, and output gate. These gates control the flow of information within the cell, enabling it to learn and remember long-term sequential dependencies. The LSTM cell first determines what information to remove from the cell state through the forget gate, which is achieved via a layer with a sigmoid activation function, as shown in Equation (1).

$$f = \sigma(x_t W_f + h_{t-1} W_f + b_f) \tag{1}$$

where the output of the forget gate is derived from the input x_t at moment t, the state h_{t-1} at moment $t-1$, and the bias term b. This output is then passed through the sigmoid function, represented as σ in the equation. The sigmoid function, an S-shaped curve, maps the input variables between 0 and 1, commonly serving as an activation function for neural networks. Each value of f ranges from 0 (complete forgetting) to 1 (complete remembering).

Next, the input gate decides what new information will be stored in the cell state. This process involves two parts: First, a sigmoid layer called the "input gate layer" determines which values will be updated, with a probability i_t. Then, a tanh layer creates a vector of candidate values $\tilde{C}_t$ to be added to the cell state. These two components combine to update the state, as shown in Equation (2). Notably, the tanh function is used instead of the sigmoid function as the activation function because its role is to add new information to the memory cell, not to gate it.

$$i_t = \sigma(x_t W_i + h_{t-1} W_i + b_i), \tilde{C}_t = tanh(W_C \times [h_{t-1}, x_t] + b_C \tag{2}$$

where the old cell state $\tilde{C}_{t-1}$ is then updated to the new cell state $\tilde{C}_t$. The new cell state is a hybrid that combines the old state and the new candidate, depending on the previously determined amount of forgetting and updating, as shown in Equation (3).

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{3}$$

For the design of the output mechanism, the output of the LSTM depends not only on the current input and hidden state but also on the cell state. This process is

facilitated by an output gate. The output gate's primary task is to determine which parts of the cell state should be output. Specifically, it applies a sigmoid layer to interrogate the current input x_t and the current hidden state h_{t-1}, generating a value between 0 and 1 for each element of the cell state. This value determines the amount of information retained in the corresponding section, as shown in Equation (4).

$$o_t = \sigma(x_t W_t + h_{t-1} W_t + b_t) \tag{4}$$

Subsequently, the cell state C_t is passed through a tanh layer that compresses its value to between −1 and 1. The output of the tanh is then multiplied by the output of the output gate O_t to obtain the final output h_t. This process ensures that only the filtered portion of the information is passed to the next layer or used as the current layer's output, as shown in Equation (5).

$$h_t = o_t \times tanh(C_t) \tag{5}$$

Through this mechanism, the LSTM can flexibly extract useful information from its long-term memory while considering the current context. This capability not only enhances the model's predictive accuracy but also improves its ability to handle time-series data, making it well-suited for complex tasks such as road noise prediction.

The training process of neural network methods, such as LSTM, involves calculating errors based on the loss function and then backpropagating to update parameters. The loss function is the core metric that determines the direction of the network's training process. For road noise prediction tasks, in addition to focusing on the overall trend prediction accuracy, it is also necessary to pay attention to the prominent peaks caused by local resonances in the overall noise. Improving the prediction accuracy of these prominent peaks is a current challenge in road noise prediction.

2.2. Improvement of the LSTM model

This study draws on the concept of physics-informed neural networks [41] to refine and expand the loss function. Empirical formula values are introduced as constraints into the network based on the LSTM infrastructure, making it more suitable for the specific needs of road noise prediction. Consequently, the empirical model of the IMP-LSTM network is proposed, with its topology illustrated in **Figure 2**.

The loss function of the IMP-LSTM network is divided into three parts. The first part is the traditional data-driven model loss term, which focuses on the difference between the predicted value and the actual value, as shown in Equation (6).

$$E_d = \sum_{i=1}^{N_d} \left[u_{pred}\left(t_i^d, x_i^d; \theta,\right) - u_{real}\left(t_i^d, x_i^d\right) \right]^2 \tag{6}$$

where N_d is the number of data points, $u_{pred}(t_i^d, x_i^d; \theta,)$ represents the neural network output value and $u_{real}(t_i^d, x_i^d)$ denotes the labeled value. This loss term ensures that the model output remains close to the training data overall, providing a baseline learning objective for the model.

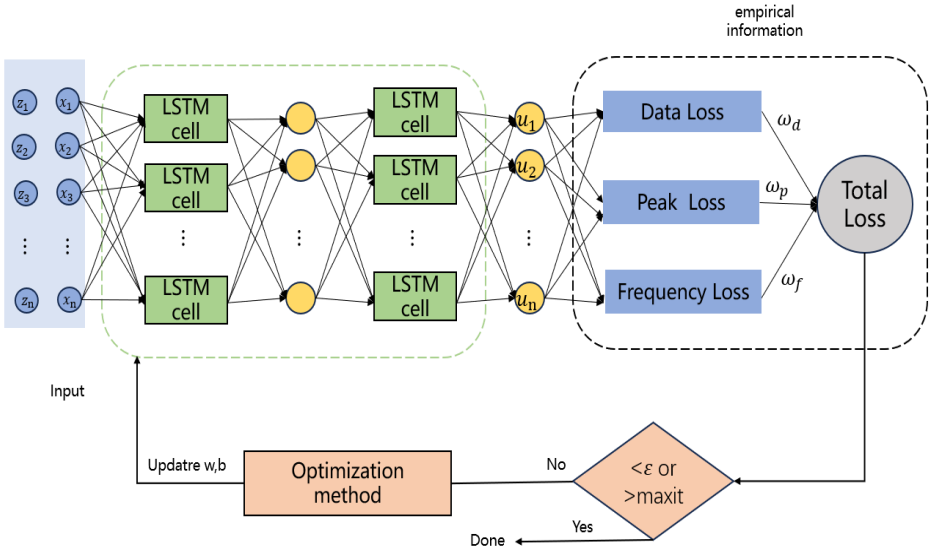

Figure 2. IMP-LSTM network structure.

The second component is the peak error loss term of the predicted sequence features, as shown in Equation (7).

$$MSE_p = \frac{1}{N_p}\sum_{i=1}^{N_p}\left[u_{pred}(0, x_p^{peak}; \theta,) - u_{real}(0, x_p^{peak})\right]^2 \tag{7}$$

where N_p is the number of data points near the peak of the feature, $u_{pred}(0, x_p^{peak}; \theta,)$ is the prediction of the neural network for the pth initial data point x_p^{peak} under parameter θ; $u_{pred}(0, x_p^{peak})$ is the true value of the pth initial data point. This loss term calculates the mean squared error between the neural network output and the true value near the peak location, ensuring that the neural network output matches the true value at the peak of the feature.

The third component is the correlation loss, which is calculated using Pearson's correlation coefficient, as shown in Equation (8). The loss function is defined as 1 minus the correlation coefficient, where the loss is 0 when the output is completely correlated with the labels and 1 when the output is not correlated with the labels, effectively guiding the network training.

$$MSE_r = 1 - \frac{\frac{1}{N_r}\sum_{i=1}^{N_r}\left[(u_{pred}^r - \bar{u}_{pred})(u_{real}^r - \bar{u}_{real})\right]}{\sqrt{\sum_{i=1}^{N_r}\left[(u_{pred}^r - \bar{u}_{pred})^2\right]\sum_{i=1}^{N_r}\left[(u_{real}^r - \bar{u}_{real})^2\right]}} \tag{8}$$

where N_r is the number of data points, $\bar{u}_{pred}$ is the average of network output features, $\bar{u}_{real}$ is the average of labeled values. This loss term measures the correlation between the predicted output of the neural network and the true labeled values. By calculating the Pearson's correlation coefficient between the network output sequence u_{pred} and the labeled value sequence u_{real}^r, the strength of the linear relationship between them is determined.

The total loss of the network is shown in Equation (9).

$$\text{Loss}_{IMP-LSTM} = \omega_d MSE_d + \omega_p MSE_p + \omega_r MSE_r \tag{9}$$

where w_d is the loss weight in the traditional neural network model, w_p is the local loss weight of the feature peaks, and w_r is the serial correlation error loss weight.

As with physics-informed neural networks, the individual weights of the loss function are fixed, and the training efficiency of the network depends on the weights associated with different loss terms. However, the general method of adjusting the loss weights is time-consuming, laborious, and prone to errors and omissions. To address this, an adaptive weighting algorithm based on a normalization method is proposed, which normalizes each weight term so that they contribute equal weights to the total loss, ensuring that no single loss term is biased during model training. This approach improves the generalization ability of the model and usually speeds up convergence by avoiding training instability due to overemphasis on a single loss term.

Specifically, the average value of each loss term is computed by forward propagation before training begins. Initial normalization weights are set to 0.5, as the same in [42]. Using these calculated average loss values and those weights balance the contribution of each loss term to the total loss, as shown in Equation (10).

$$\omega_d^{'} = \frac{1}{\sigma_d\sqrt{N_d}}, \omega_p^{'} = \frac{1}{\sigma_p\sqrt{N_p}}, \omega_r^{'} = \frac{1}{\sigma_r\sqrt{N_r}} \tag{10}$$

where the symbols σ_d, σ_p, σ_r represent the standard deviations of the data loss term, peak error loss term, and correlation error loss term, respectively, during the initial training phase.N_d, N_P, N_r represent the number of samples in the data loss term, peak error loss term, and correlation error loss term, respectively. During each forward propagation, the total loss is calculated using normalized weights, as shown in Equation (11).

$$\text{Loss}_{IMP-LSTM}^{'} = \omega_d^{'} MSE_d + \omega_p^{'} MSE_p + \omega_r^{'} MSE_r \tag{11}$$

Before training begins, calculate the initial average and standard deviation of each loss term. This step serves as a reference for the subsequent weight updates during training. During each training iteration, Equations (10) and (11) are used to dynamically update the weights of the loss terms. This ensures that the model can adaptively adjust the contribution of each loss term based on their changes, which continues throughout the training process. This dynamic normalization process ensures that the model quickly adapts to the data and empirical formulas in the early stages of learning while automatically adjusting in the later stages to avoid overfitting of local loss terms.

3. Vehicle road noise data collection and modeling

3.1. Vehicle Road noise road test

In structural vibration and noise analysis, any system, including vehicles, follows the ‘source-path-receiver’ model. During driving, the tire excites the road surface of the suspension system, transmitting vibration to the body, which generates structural radiation sound ultimately perceived as noise by the human ear [43,44]. Understanding this transmission mechanism involves examining road excitation, tire characteristics, and the properties of each elastic connection in the suspension. In practice, obtaining

and rectifying tire stiffness and damping parameters is challenging [45,46]. To simplify the road noise prediction model, steering knuckle excitation is used as an input kinematic parameter, bypassing the need to study tire and road surface interactions. Within the suspension system, the helical spring's stiffness primarily affects system frequency with minimal impact on road noise. However, the bushing and shock absorber, the main vibration isolation components, significantly influence road noise. Thus, dynamic stiffness of the bushing and shock absorber damping are included as input dynamic parameters [47]. This study focuses on a car model with 'front McPherson, rear multi-link' suspension.

To collect vibration data from the steering knuckle, suspension, and body, as well as noise data from the driver's right ear, a real vehicle road test was conducted [48]. Data were collected synchronously using a 24-channel LMS SCADAS Mobile system from LMS Company and analyzed with the Signature Testing-Advanced module in LMS Test.Lab18. Noise signals were captured with a BSWA sound pressure sensor (model: MPA201-550507), while vibration signals were recorded using PCB three-axis vibration sensors (models: BW13510-J0810/BW13510-J0812). Sampling time was set to 10 seconds, with a frequency of 6400 Hz and a resolution of 1 Hz. To ensure the consistency of the road input during the test, a rough asphalt road is selected and the vehicle speed is kept at 60 km/h at a constant speed. During the test, the external noise interference should be minimized; the windows, air inlet, and outlet should be closed, and the air conditioner and fan should be closed to ensure that no abnormal noise is generated. In addition, the necessary test personnel and drivers try to put the vehicle in a no-load state. Also, to ensure the quality of the test data, each designed data sample is tested four times, and the results with better consistency are screened as the subsequent training samples of the prediction model. The road noise tests were conducted under stable weather conditions to minimize the influence of external environmental factors on the noise data. Specifically, all tests were performed on days with mild weather, avoiding extreme conditions such as heavy rain, strong winds, or extreme temperatures. The ambient temperature during the tests ranged between 25 °C and 28 °C, with minimal wind interference. These controlled conditions helped ensure the reliability and consistency of the collected noise data, allowing the model to focus primarily on the road-induced noise without significant interference from environmental noise sources.

Rough asphalt pavement was chosen for the road noise test for two main reasons: it effectively stimulates the vehicle's structural sound while minimizing other noise contributions, and it is a typical urban passenger car surface, providing relevant real-world noise levels. The sensor layout is illustrated in **Figure 3**. Two PCB three-axis vibration acceleration sensors measured the vibration of the vehicle's left-side steering knuckles. A sound pressure sensor was placed near the driver's right ear in a forward-facing direction, adhering to ISO 5128:2023: Acoustics—Measurement of interior vehicle noise [49]. The driver's right ear noise and the Z-direction vibration acceleration of the front suspension steering knuckle collected in the test are shown in **Figure 4a,b**, respectively. Combined with the transmission characteristics of road structure noise, its frequency band range is mainly concentrated in 50–300 Hz, because this paper focuses on the noise characteristics in the sub-frequency band.

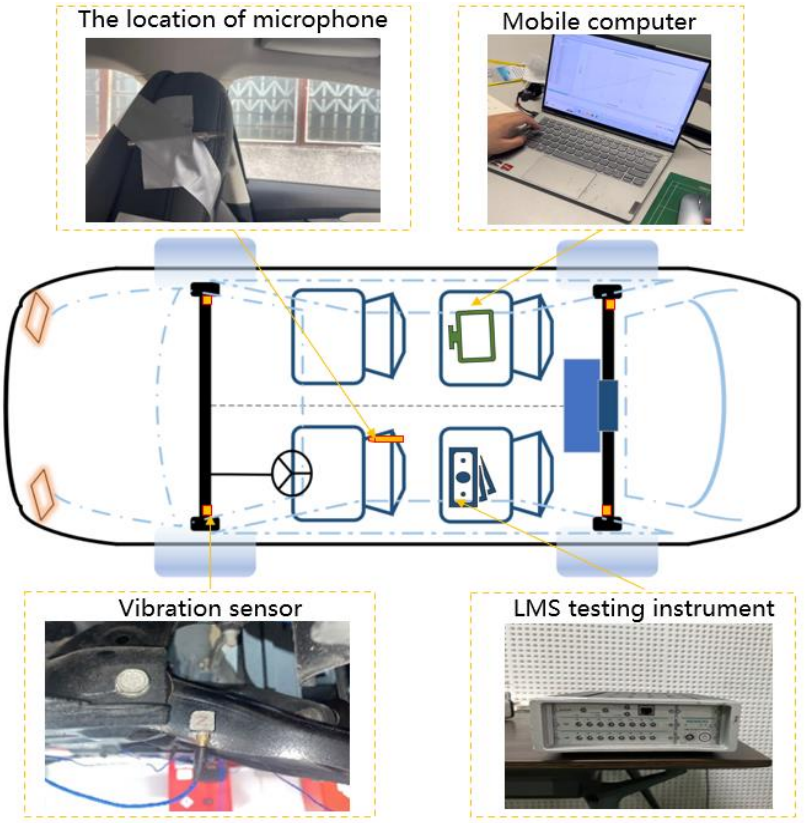

Figure 3. Sensor placement.
The vibration sensor is located at the front steering knuckle.

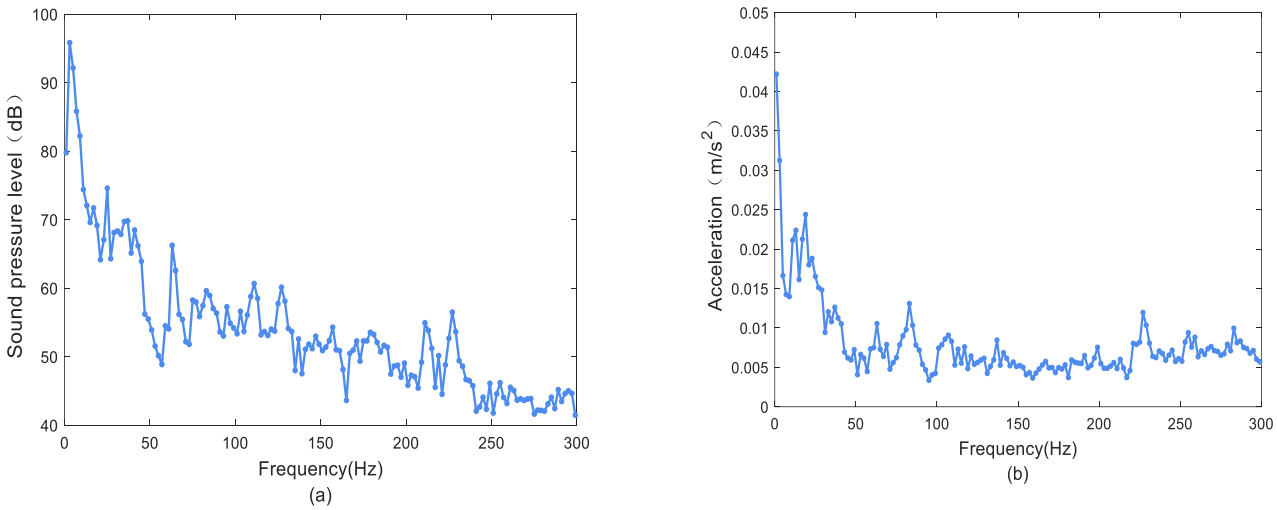

Figure 4. Driver's right ear noise and front steering knuckle vibration acceleration test data, **(a)** driver's right ear noise; **(b)** front suspension steering knuckle acceleration.

3.2. Data pre-processing

A total of 60 sample data points were obtained from the pilot tests using both the original and comparison models. To mitigate the impact of data relative value on modeling, normalization of the dataset was performed [50,51]. Normalization converts the sample eigenvalues to a common scale within the [0,1] interval. The Min-Max

scaler function was used for normalizing the input data and performing inverse normalization on the output results, as shown in Equation (12) and Equation (13).

$$X_{std} = \frac{X - X.min(axis = 0)}{X.max(axis = 0) - X.min(axis = 0)} \tag{12}$$

$$X_{scaled} = X_{std} * (max - min()min) \tag{13}$$

where X is the data to be normalized, usually a two-dimensional matrix; $X.min(axis = 0)$is the row vector consisting of the minimum value in each column; $X.max(axis = 0)$ is the row vector consisting of the maximum value in each column; max is the maximum value of the interval to be mapped to, the default is 1; min is the minimum value of the interval to be mapped to, the default is 0; X_{std} is the result of normalization; X_{std} is the result of anti-normalization; X_{scaled} is the result of anti-normalization.

4. Method application and validation

4.1. Modeling of road noise prediction based on IMP-LSTM

We introduce a new regression-based IMP-LSTM model for predicting in-vehicle noise. First, we compare the model's convergence speed and final loss level with and without the adaptive weighting strategy. Then, we assess the new model's effectiveness and efficiency against traditional machine learning models like LSTMs. The overall model-building process is illustrated in **Figure 5**. First, the input consists of a 250 × 24 feature matrix, which includes 24 sequence features. These features comprise the vibration acceleration of the front and rear suspension steering knuckles, as well as the dynamic stiffness of several components within the suspension system. Specifically, the dynamic stiffness features involve the front and rear shock absorber bushings, the front suspension swing arm (both front and rear bushings), the rear lateral tie rod bushing, the rear lateral control arm bushing, and the rear lower swing arm bushing. Each dynamic stiffness feature is measured in the x, y, and z directions. Each sequence has a length of 141, with a frequency range of 50–300Hz and a frequency resolution of 1 Hz. The output is a 250 × 1 matrix representing the noise at the driver's right ear in decibels (dB).

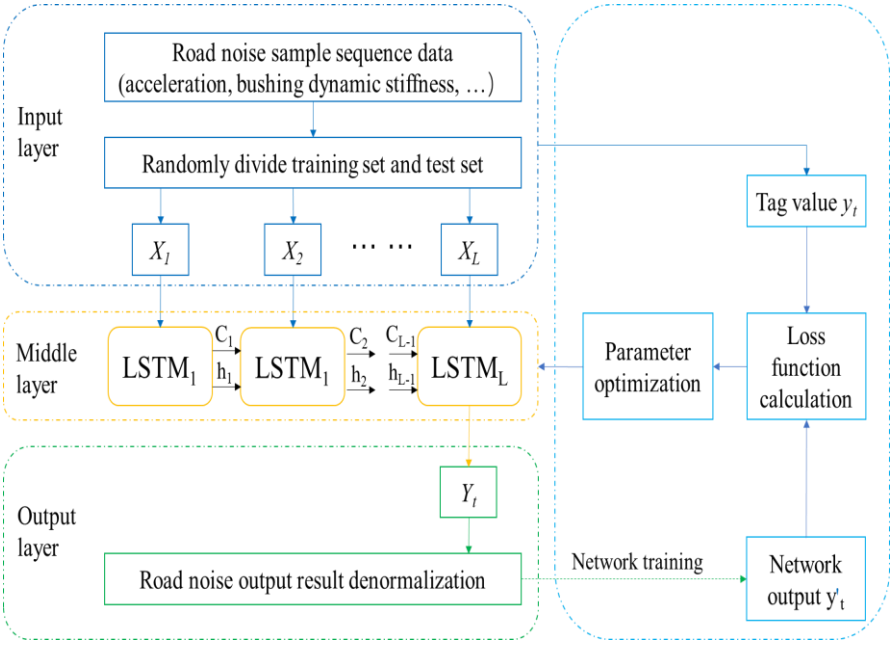

Figure 5. IMP-LSTM algorithm prediction process.

The road noise sample dataset is divided into training, test, and validation sets in an 8:1:1 ratio. The LSTM sequence architecture is used for forward propagation to obtain the network output *y*. The loss function is calculated as per the formula in Section 2.3, and the loss is iteratively backpropagated to adjust network parameters using the training set, aiming to reduce the loss function value. A unified evaluation metric is determined before model training to compare different models' prediction effects. This regression prediction problem uses evaluation indices such as mean square error (MSE), coefficient of determination (R^2), root mean square error (RMSE) [52], peak error, and training time. These metrics characterize the numerical deviation of the prediction results from the actual results.

The modeling software used is Pycharm2022.3, running on an RTX 2080 Ti with 64G RAM. A grid search method is employed to optimize the IMP-LSTM network parameters, focusing on the learning rate and the number of neurons in the hidden layer. The learning rate is tested in the range e ∈ [0.001, 0.01, 0.1], and the number of hidden layer nodes is tested in the range n ∈ [16, 32, 48, 64]. Initial iterations are set to 150, and 12 tests (3 learning rates × 4 node counts) are performed. The optimal parameters are identified based on model accuracy. The highest accuracy obtained is 0.93, achieved with a learning rate of 0.01 and 64 hidden layer nodes. The results of the parameter optimization are presented in **Table 1**.

Table 1. Network search parameter optimization results.

Number of hidden layer nodes	**learning rate**		
	0.00001	0.001	0.01
16	0.68	0.77	0.83
32	0.76	0.84	0.86
48	0.88	0.90	0.89
64	0.89	0.91	0.93

The parameters of the model are set as follows: the learning rate is 0.01, the attenuation coefficient is 0.9, the batch size is 10, and the number of hidden layer nodes is 64. On the premise of the highest accuracy of the model, the parameters of the IMP-LSTM network are optimized. The ADAM algorithm is used as the optimizer, combining Momentum and Adam for efficient parameter space search. To investigate the effect of loss weights on network accuracy, initial weight terms are set as ω_d = 1, ω_p = 0.5, ω_r = 0.1. The performance of the network with fixed weights is compared against that with normalized weights through experiments. The changes in the three loss terms during training are shown in **Figures 6**–**8**.

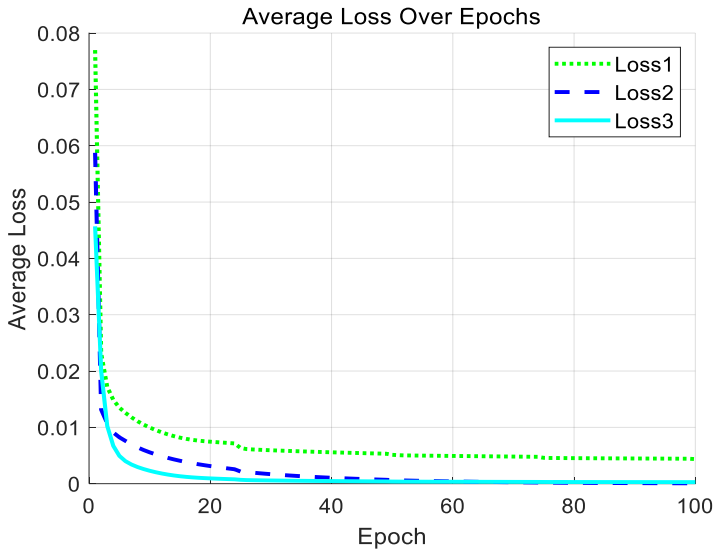

Figure 6. Changes in each loss item of the model under fixed weights.

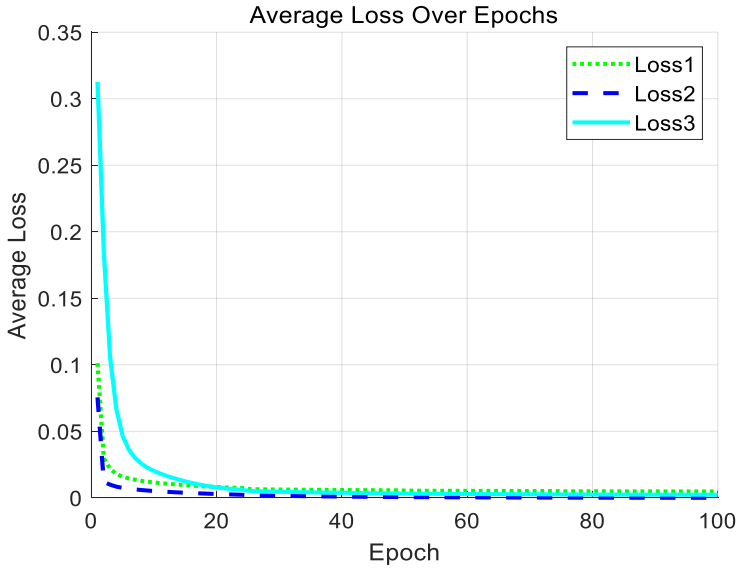

Figure 7. Changes in each loss item of the model under adaptive weights.

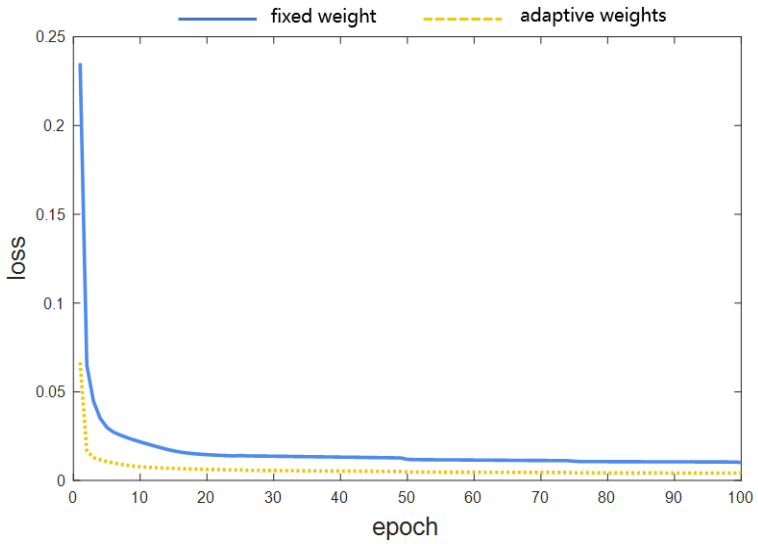

Figure 8. Changes in the total loss of the model.

From the loss weight change diagram shown in **Figure 6** and **7**, it can be seen that Loss1, Loss2 and Loss3 all decrease rapidly and then tend to be stable. In contrast, using normalized adaptive weights, the loss terms of Loss1 and Loss2 decrease faster, and the loss term of normalized adaptive weight Loss3 is smaller than that of fixed weight Loss3. From **Figure 8**, it can be concluded that the introduction of the normalized adaptive weight method in the LSTM network helps to optimize the loss term more evenly, prevent the model from paying too much attention to a specific loss term, learn the overall characteristics of the data more comprehensively, improve the convergence speed and reduce the final loss level.

Based on the validation set, the hyperparameters of the model are adjusted, and the model is optimized during the training process to obtain the IMP-LSTM model with more accurate prediction results on the validation set. At the same time, the performance of the obtained model on the unseen data is evaluated on the test set, and the generalization ability of the model is tested. The IMP-LSTM model predicts the sound pressure level of the driver's right ear using validation set and test set data, as shown in **Figure 9a**,**b**, respectively. It can be seen from the figure that the predicted values and real values of the test set and the validation set of the model have high accuracy, and the predicted results are roughly consistent with the real trend. Among them, the results of the test set in **Figure 9b** prove that the IMP-LSTM model has good generalization. In summary, the results show that the proposed model and method effectively capture the key data features and sound features, which can provide some help for the study of road noise problems.

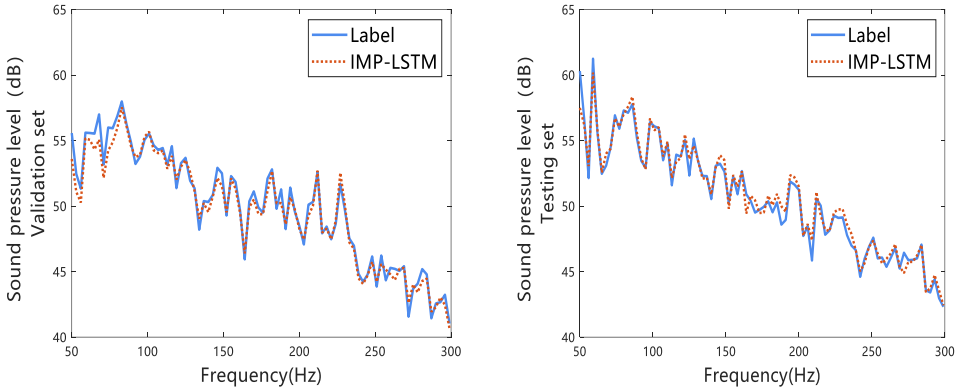

Figure 9. Model total loss weight changes, **(a)** validation set; **(b)** testing set.

4.2. Comparison of IMP-LSTM models

To validate the effectiveness of the IMP-LSTM model in predicting in-vehicle noise, this study introduces a nonlinear fitting LSTM neural network for comparison. Both models use consistent network parameters to ensure fairness. The results on the validation and test sets are shown in **Figures 10**, with performance metrics detailed in **Table 2**, focusing on the prediction accuracy at the driver's right ear.

The IMP-LSTM model, with its focus on loss weight variation, shows significant improvement over the traditional LSTM model. On both test and validation sets, the IMP-LSTM model fits the actual noise labels better at peak frequencies, demonstrating superior local accuracy. While the LSTM model captures the overall noise trend, it exhibits larger prediction biases at certain peaks. This suggests that IMP-LSTM are better suited for handling data with prominent features, crucial for accurately predicting frequency-specific noise impacting passenger comfort.

Overall, the IMP-LSTM model aligns more closely with actual noise levels on both sets, indicating its proficiency in learning the data distribution. In contrast, the LSTM model, despite capturing general noise variations, shows more deviations at specific frequencies. This highlights that the IMP-LSTM model more effectively integrates empirical formulas information with data-driven learning, ensuring higher prediction accuracy. In summary, the IMP-LSTM model offers a better fit for road noise sequence features.

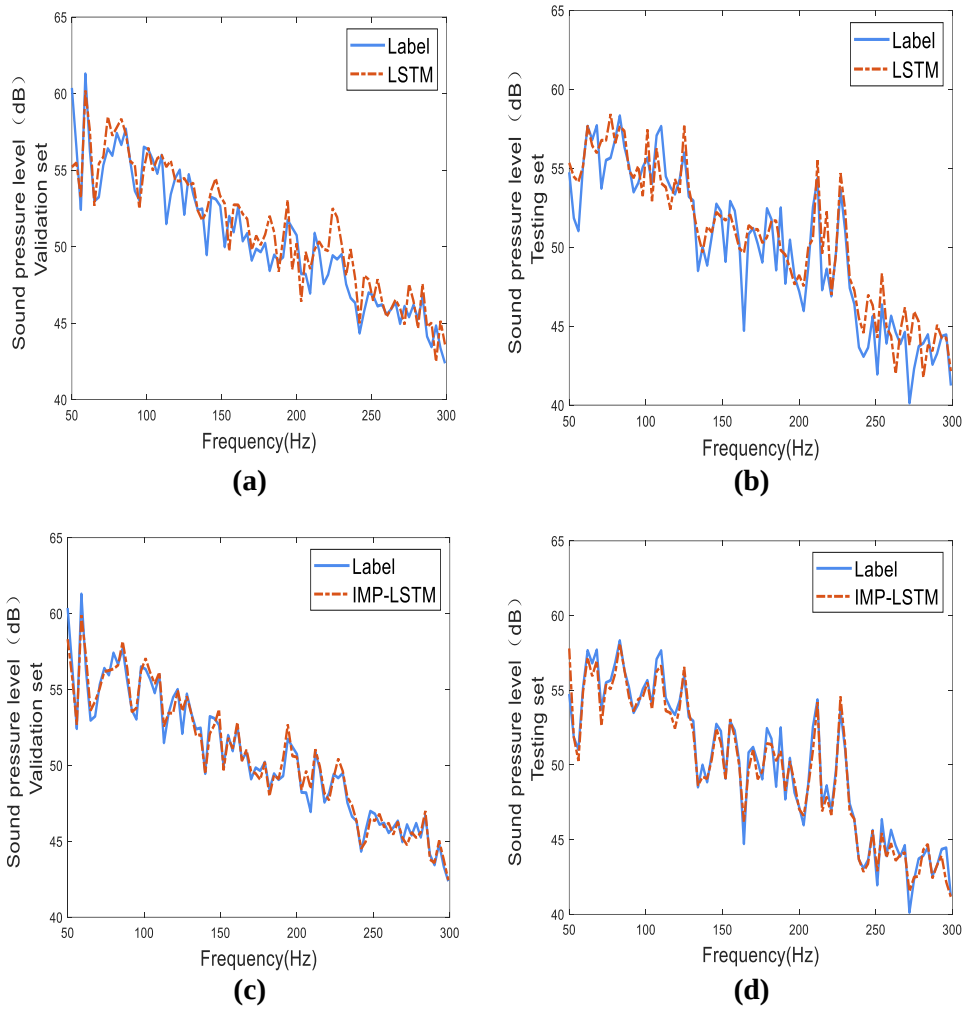

Figure 10. Driver's right ear noise validation and testing results, **(a)** LSTM validation results; **(b)** LSTM testing results; **(c)** IMP-LSTM validation results; **(d)** IMP-LSTM testing results.

Table 2. Comparison of IMP-LSTM model and LSTM model test results (Testing set).

	MSE loss	**69Hz peak error (dB)**	**Training time (150 epoch/min)**	R^2
LSTM data model	0.49	1.2	4.2	0.76
IMP-LSTM data model	0.21	0.1	3.0	0.89

5. Conclusion

In this study, we proposed an improved Long Short-Term Memory (IMP-LSTM) method for predicting vehicle road noise, focusing on the 50 to 300 Hz frequency range. The method integrates knowledge-based constraints with a data-driven approach, enhancing both prediction accuracy and interpretability.

The IMP-LSTM model demonstrated high prediction accuracy with an impressive R^2 value of 0.89, showcasing a substantial improvement over traditional LSTM models. It effectively predicted key sound pressure peaks caused by local

resonances, which are essential in road noise analysis. Additionally, the model's efficiency was evident as it completed 150 training rounds in just 3 minutes, significantly reducing training time compared to conventional methods. This efficiency makes the IMP-LSTM particularly suitable for practical applications in vehicle design, where rapid iterations are often needed. Furthermore, the model showed strong generalization capabilities, as it accurately predicted sound pressure levels on unseen test data, especially in critical peak regions.

Despite the promising results of the IMP-LSTM model, certain limitations must be addressed in future work. Firstly, the model's training was based on a limited dataset, which affects its robustness and generalization. Future studies should incorporate a wider variety of road conditions, vehicle models, and noise sources to enhance its performance. Additionally, while the use of empirical knowledge improved prediction accuracy, further refinement of the physical constraints could help capture more nuanced road noise characteristics, particularly at lower frequencies. To build upon the current research, future investigations should focus on expanding the dataset by including additional real-world road noise data from various vehicle types and road surfaces, which could further enhance the model's predictive capabilities. Moreover, exploring hybrid approaches that combine the strengths of physical modeling and advanced machine learning techniques, such as integrating physics-informed neural networks with deep learning, could lead to even more accurate predictions. Finally, optimizing the model for real-time noise prediction and control during vehicle operation is another critical avenue for development, as it would provide immediate insights into noise issues in dynamic environments.

In conclusion, the IMP-LSTM model provides an effective and efficient solution for vehicle road noise prediction. By combining data-driven methods with empirical knowledge, it significantly improves both prediction accuracy and training efficiency. With further development, this approach could be instrumental in enhancing noise prediction methodologies and vehicle NVH performance analysis.

Author contributions: Resources, RD and JZ; writing—original draft preparation, XY; data curation, YY and SL; writing—review and editing, RD; funding acquisition and Conceptualization, PD and HH. All authors have read and agreed to the published version of the manuscript.

Acknowledgments: The authors would like to acknowledge the support from the Institute of Energy and Power Research for the experimental research.

Funding: This work was supported by the Open Fund of State Key Laboratory of Vehicle NVH and Safety Technology (No. NVHSKL-202101).

Conflict of interest: The authors declare no conflict of interest.

References

1. Zhu, H. L., Zhao, J., Wang, Y. W., Ding, W. P., Pang, J. et al. (2024). Improving of pure electric vehicle sound and vibration comfort using a multi-task learning with task-dependent weighting method. Measurement, 233, 114752. DOI: 10.1016/J.MEASUREMENT.2024.114752.

2. Căşeriu, B., and Petruţa B. (2022). Automotive comfort: State of the art and challenges. International Conference Interdisciplinarity in Engineering. Cham: Springer International Publishing, pp. 375-387. DOI: 10.1007/978-3-031-22375-4_30.
3. Wu, Y. R., Liu, X. B., Huang, H. B., Wu, Y. D., Ding, W. P. et al. (2023). Multi-objective prediction and optimization of vehicle acoustic package based on ResNet neural network. Sound & Vibration, 57(1), 73-95. DOI: 10.32604/sv.2023.044601.
4. Praticò, F. G. (2014). On the dependence of acoustic performance on pavement characteristics. Transportation Research Part D, 79-87. DOI： 10.1016/j.trd.2014.04.004.
5. Hazra, S. and Reddy, J. (2022). An Aspect of Noise, Vibration, and Harshness Issues in Electric Vehicles. SAE International Journal of Vehicle Dynamics, Stability, and NVH, 6(1), 23-33. DOI: 10.4271/10-06-01-0002.
6. SheikhMozafari, M. J. (2024). Enhancing Sound Absorption in Micro-Perforated Panel and Porous Material Composite in Low Frequencies: A Numerical Study Using FEM. Sound & Vibration, 58(1), 81-100. DOI: 10.32604/sv.2024.048897.
7. Huang, H. B., Lim, T. C., Wu, J., Ding, W. P., Pang, J. (2023). Multitarget prediction and optimization of pure electric vehicle tire/road airborne noise sound quality based on a knowledge-and data-driven method. Mechanical Systems and Signal Processing, 197, 110361. DOI: 10.1016/j.ymssp.2023.110361.
8. Bianco, F., Fredianelli, L., Castro, F. L., Paolo, G., Gagliardi, P. et al. (2020). Stabilization of a p-u Sensor Mounted on a Vehicle for Measuring the Acoustic Impedance of Road Surfaces. Sensors, 1239(5). DOI: 10.3390/s20051239.
9. Praticò, F. G., Fedele, R., Pellicano, G. (2021). Monitoring Road Acoustic and Mechanical Performance. European Workshop on Structural Health Monitoring, 127. DOI: 10.1007/978-3-030-64594-6_58.
10. RamosRomero, C., Asensio, C., Moreno, R., Arcas, G. (2022). Urban Road Surface Discrimination by Tire-Road Noise Analysis and Data Clustering. Sensors, 24, 9686-9686. DOI：10.3390/S22249686.
11. DelPizzo, L. G., Teti, L., Moro, A., Bianco, F., Fredianelli, L. et al. (2020). Influence of texture on tyre road noise spectra in rubberized pavements. Applied Acoustics, 107080-107080. DOI： 10.1016/j.apacoust.2019.107080.
12. Sun, P., Dai, R. X., Li, H. Q., Zheng, Z. W., Wu, Y. D. et al. (2024). Multi-Objective Prediction of the Sound Insulation Performance of a Vehicle Body System Using Multiple Kernel Learning–Support Vector Regression. Electronics, 13(3), 538. DOI: 10.3390/electronics13030538.
13. Lan, Z., Yuan, M., Shao, S., Li, F. (2023). Noise Emission Models of Electric Vehicles Considering Speed, Acceleration, and Motion State. Int. J. Environ. Res. Public Health, 20, 3531. DOI: 10.3390/ijerph20043531.
14. Kanka, S., Fredianelli, L., Artuso, F., Fidecaro, F., Licitra, G. (2023). Evaluation of Acoustic Comfort and Sound Energy Transmission in a Yacht. Energies, 2, 808-808. DOI： 10.3390/EN16020808.
15. Borelli, D., Gaggero, T., Rizzuto, E., Schenone, C. (2021). Onboard ship noise: Acoustic comfort in cabins. Applied Acoustics, 177, DOI：10.1016/J.APACOUST.2021.107912.
16. AlvesFilho, J. M., Lenzi, A., Zannin, P. H. T. (2003). Effects of traffic composition on road noise: a case study. Transportation Research Part D, 1, 75-80. DOI： 10.1016/j.trd.2003.08.001.
17. Praticò, F. G., Anfosso-Lédée, F. (2012). Trends and Issues in Mitigating Traffic Noise through Quiet Pavements. Procedia - Social and Behavioral Sciences, 203-212. DOI： 10.1016/j.sbspro.2012.09.873.
18. Xia, X., Ye, S., Hou, L., Wu, D., Lin, Q. (2024). Instantaneous flow area calculation for modeling an axial piston pump: A point cloud-based technique. Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy. DOI: 10.1177/09576509241256499.
19. Teti, L., León, G. D., DelPizzo, L. G, Moro, A., Bianco, F. et al. (2020). Modelling the acoustic performance of newly laid low-noise pavements. Construction and Building Materials, 118509-118509. DOI： 10.1016/j.conbuildmat.2020.118509.
20. Yin, L., Zhang, Z. Q., Wu, M., Wang, Z. L., Ma, C. et al. (2023). Adaptive parallel filter method for active cancellation of road noise inside vehicles. Mechanical Systems and Signal Processing, 193, 110274. DOI: 10.1016/j.ymssp.2023.110274.
21. Huang, H. B., Wang, Y. W., Wu, J. H., Ding, W. P. Pang, J. et al. (2024). Prediction and optimization of pure electric vehicle tire/road structure-borne noise based on knowledge graph and multi-task ResNet.Expert Systems With Applications(PB), 124536-124536. DOI: 10.1016/J.ESWA.2024.124536.
22. Fan, D., Dai, P., Yang, M., Jia, W., Jia, X. et al. (2022). Research on Maglev Vibration Isolation Technology for Vehicle Road Noise Control. SAE International Journal of Vehicle Dynamics, Stability, and NVH, 6(10-06-03-0016), 233-245. DOI: 10.4271/10-06-03-0016.

23. Park, J., Park, S. (2024). Roadnoise Reduction through Component-TPA with Test and Simulation Convergence Using Blocked Force. SAE Technical Paper, 1(2952). DOI: 10.4271/2024-01-2952.
24. Fredianelli, L., Pedrini, G., Bolognese, M., Bernardini, M., Fidecaro, F. et al. (2024). Features for Evaluating Source Localization Effectiveness in Sound Maps from Acoustic Cameras. Sensors, 14, 4696-4696. DOI： 10.3390/S24144696.
25. Licitra, G., Artuso, F., Bernardini, M., Moro, A., Fidecaro, F. et al. (2023). Acoustic Beamforming Algorithms and Their Applications in Environmental Noise. Current Pollution Reports, 3, 486-509. DOI： 10.1007/S40726-023-00264-9.
26. Huang, H., Huang, X., Ding, W., Yang, M., Fan, D., et.al. (2022). Uncertainty optimization of pure electric vehicle interior tire/road noise comfort based on data-driven. Mechanical Systems and Signal Processing, 165, 108300. DOI: 10.1016/j.ymssp.2021.108300
27. Peng, C., Cheng, S., Sun, M., Ren, C., Song, J. et al. (2024). Prediction of Sound Transmission Loss of Vehicle Floor System Based on 1D-Convolutional Neural Networks. Sound & Vibration, 58, 81-100. DOI: 10.32604/sv.2024.046940.
28. Park, U., Kang, Y. J. (2024). Operational transfer path analysis based on neural network. Journal of Sound and Vibration, 579, 118364. DOI: 10.1016/j.jsv.2024.118364.
29. Wysocki, T. V., Rieger, F., Tsokaktsidis, D. E., Gauterin, F. (2021). Generating Component Designs for an Improved NVH Performance by Using an Artificial Neural Network as an Optimization Metamodel. Designs, 5, 36. DOI: 10.3390/designs5020036.
30. Li, M., Zhou, W., Liu, J., Zhang, X., Pan, F. et al. (2021). Vehicle Interior Noise Prediction Based on Elman Neural Network. Appllied Sciences, 11(17), 8029. DOI:10.3390/app11178029.
31. Huang, H., Huang, X., Ding, W., Zhang, S., Pang, J. (2023). Optimization of electric vehicle sound package based on LSTM with an adaptive learning rate forest and multiple-level multiple-object method. Mechanical Systems and Signal Processing, 187, 109932. DOI: 10.1016/j.ymssp.2022.109932.
32. Yu, Y., Si, X. S., Hu, C. H., Zhang, J. X. (2019). A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. Neural Comput, 31(7), 1235–1270. DOI: 10.1162/neco_a_01199.
33. Back, J., Lee, S., Min, L. S., An, K., Kwon, D. H. et al. (2021). Design and implementation of comfort-quality HVAC sound inside a vehicle cabin. Applied Acoustics. DOI： 10.1016/J.APACOUST.2021.107940.
34. Soeta, Y., Sakamoto, Y. (2018). An Exploratory Analysis of Sound Field Characteristics using the Impulse Response in a Car Cabin. Environments, 4, 44. DOI： 10.3390/environments5040044.
35. Deniz, O., Pedraza, A., Vallez, N., Salido, J., Bueno, G. (2020). Robustness to adversarial examples can be improved with overfitting. International Journal of Machine Learning and Cybernetics, 11, 935-944. DOI: 10.1007/s13042-020-01097-4.
36. Taherdoost, H. (2023). Deep Learning and Neural Networks: Decision-Making Implications. Symmetry, 15, 1723. DOI: 10.3390/sym15091723.
37. Al-Selwi, S. M., Hassan, M. F., Abdulkadir, S. J., Muneer, A., Sumiea, E. H. et al. (2024). RNN-LSTM: From applications to modeling techniques and beyond—Systematic review. Journal of King Saud University-Computer and Information Sciences, 102068. DOI: 10.1016/j.jksuci.2024.102068.
38. Mamdouh, M., Ezzat, M., Hefny, H. (2024). Improving flight delays prediction by developing attention-based bidirectional LSTM network. Expert Systems with Applications, 238, 121747. DOI: 10.1016/j.eswa.2023.121747.
39. Staudemeyer, R. C., Morris, E. R. (2019). Understanding LSTM--a tutorial into Long Short-Term Memory recurrent neural networks. arXiv preprint. DOI: 10.48550/arXiv.1909.09586.
40. Huang, Z., Xu, W., Yu, K. (2015). Bidirectional LSTM-CRF models for sequence tagging. arXiv preprint. DOI: 10.48550/arXiv.1508.01991.
41. Raissi, M., Perdikaris, P., Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, Journal of Computational Physics, 378, 686-707. DOI: 10.1016/j.jcp.2018.10.045.
42. Heydari, A. A., Thompson, C. A., Mehmood, A. (2019). Softadapt: Techniques for adaptive loss weighting of neural networks with multi-part loss functions. arXiv preprint. DOI: 10.48550/arXiv.1912.12355.
43. Elliott, A. S., Moorhouse, A. T., Huntley, T., Tate, S. (2013). In-situ source path contribution analysis of structure borne road noise. Journal of Sound and Vibration, 332(24), 6276-6295. DOI: 10.1016/j.jsv.2013.05.031.
44. Zheng, X., Wan, B., Jia, Z., Li, R., Liu, X. et al. (2024). A study on active road noise control based on operational transfer path analysis and selective subband adaptive filtering. Applied Acoustics, 224, 110147. DOI: 10.1016/j.apacoust.2024.110147.

45. Huang, H. B., Huang, X. R., Li, R. X., Lim, T. C., Ding, W. P. (2016) Sound quality prediction of vehicle interior noise using deep belief networks. Applied Acoustics, 113, 149-161. DOI: 10.1016/j.apacoust.2016.06.021.
46. Pu, Z. Q., Yi, J. Q., Liu, Z. (2022). A review of research on group intelligent decision-making methods driven by knowledge and data synergy. Journal of Automation, 48(3), 1-17.7. DOI: 10.16383/j.aas.c210118.
47. Huang, H., Huang, X., Ding, W., Yang, M., Yu, X. et al. (2023). Vehicle vibro-acoustical comfort optimization using a multi-objective interval analysis method. Expert Systems with Applications, 213, 119001. DOI: 10.1016/j.apacoust.2016.06.021
48. Marotta, R., Strano, R., Terzo, M., Tordela, C. (2024). Multi-Output Physically Analyzed Neural Network for the Prediction of Tire–Road Interaction Forces. SAE, 8(2), 285-308, DOI:10.4271/10-08-02-0016.
49. ISO, ISO. (2023). Acoustics — Measurement of interior vehicle noise. Geneva, Switzerland: International Organization for Standardization, 5128:2023.
50. Reyad, M., Sarhan, A. M., Arafa, M. (2023). A modified Adam algorithm for deep neural network optimization. Neural Computing and Applications, 35(23), 17095-17112. DOI: 10.1007/s00521-023-08568-z.
51. Huang, H. B., Li, R. X., Yang, M. L., Lim, T. C., Ding, W. P. (2017) Evaluation of vehicle interior sound quality using a continuous restricted Boltzmann machine-based DBN. Mechanical Systems and Signal Processing, 84, 245-267. DOI: 10.1016/j.ymssp.2016.07.014
52. Wang, W. J., Lu, Y. M. (2018). Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model. IOP Conference Series: Materials Science and Engineering, 324, 012049. DOI:10.1088/1757-899X/324/1/012049.