Article:
Barton, D orcid.org/0000-0003-4986-5817, Wang, S, Zhong, L et al. (5 more authors) (2021) Prediction of frictional braking noise based on brake dynamometer test and artificial intelligent algorithms. Proceedings of the Institution of Mechanical Engineers Part D: Journal of Automobile Engineering. ISSN 0954-4070

https://doi.org/10.1177/09544070211062276

Prediction of Frictional Braking Noise Based on Brake Dynamometer Test and Artificial Intelligent Algorithms

Shuwen Wang[1], Liangwei Zhong[1], Yayun Niu[1], Shuangxia Liu[1], Ke Li[2], Lijing Wang[2], Shaofan Wang[2], David Barton[3]

[1] College of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

[2] Institute of Artificial Intelligence, Beihang University, Beijing 100083, China

[3] School of Mechanical Engineering, University of Leeds, LS2 9JT, United Kingdom

Corresponding authors: Shuwen Wang (e-mail: wsw@ usst.edu.cn; shuwenwang66@163.com); Shaofan Wang (e-mail: Wsf_1993@qq.com).

This work was supported by the Science and Technology Committee of Shanghai Municipal Research Fund under Grant 18060502400.

Abstract

Based on the LINK3900 brake dynamometer test data, combined with the artificial intelligent algorithms, frictional braking noise is quantitatively analyzed and predicted in this study. To achieve this goal, a frictional braking noise prediction method is innovatively proposed, which consists of two main parts: first, based on the experimental data obtained from the LINK3900 brake dynamometer tests, and combining with the improved Long-Short-Term Memory (LSTM) algorithm, the coefficients of friction (COFs) are predicted under various braking test conditions. Then, based on the predicted braking COFs and other selected critical braking parameters, the quantitative prediction of frictional braking noise is obtained by means of the optimized eXtreme Gradient Boosting (XGBoost) algorithm. Finally, the inherent features of the XGBoost algorithm are employed to qualitatively analyze the importance of the main factors affecting the frictional braking noise. The prediction algorithms of COFs and frictional braking noise are validated by the brake dynamometer test data, and the R^2 (R square) scores of both the LSTM and XGBoost prediction algorithms are 0.9, which verifies the feasibility of both algorithms. The main contribution of this work is to propose a novel approach to quantitatively predict the braking noise based on large tested data and combined with the LSTM and XGBoost artificial intelligent algorithms,

which can significantly save time for the brake system development and braking performance testing, and has significance to the rapid prediction of braking frictional noise and fast NVH (noise, vibration, and harshness) optimal design of frictional braking systems.

Keywords: Braking noise, Friction coefficient, Long-Short-Term Memory algorithm, XGBoost model, Noise prediction

1 Introduction

With the rapid development of the automobile industry and the increase of automobile popularity, people have put forward higher requirements for the safety and comfort of automobiles. The braking noise is an important index to judge the performance of braking systems, and the performance of braking systems plays a decisive role in the safety and quality of automobiles. Meanwhile, the comfort of an automobile is highly depended on the braking noise and vibrations that are mainly caused by the friction between the brake discs and brake pads. Through quantitative prediction and qualitative analysis of experimental data of frictional braking noise, the performance and quality of braking systems can be analyzed, the test and development period of brake components can be greatly reduced, and the mechanism of braking noise can be further studied, which has significance to the development, optimization, and improvement of braking systems.

The mechanism of frictional vibrations and noise of automobile braking systems

has been studied for nearly a hundred years, and so far there are at least five theories or mechanisms of the generation of frictional vibrations and noise, including stick-slip theory, sprag-slip theory, hammering theory, modal coupling theory, and surface dynamics theory. The brake noise generation mechanisms were reviewed by a number of researchers, such as North[1], Crolla and Lang[2], Yang and Gibson[3], Kinkaid et al.[4], Oberst and Lai[5], and Ouyang et al.[6].

Based on the brake noise generation theories, scholars developed mathematical models of brake systems with various degrees of freedom (DOFs) for the analysis of the instability of disc brake systems and the prediction of braking squeal propensity. For examples, North[7] proposed a "binary flutter" model with eight degrees of freedom (DOFs) that represented the rigid body motion of the pads, caliper, and a section of disc in contact with the pads; Millner[8] developed a 6-DOF model of such a disc brake system; Murakami et al.[9] developed a lumped mass model of disc brake with more DOFs and correlated the predicted squeal propensities to the occurrence of experimental braking squeal. More recently, Ahmed[10] developed a ten-DOF mathematical model to study the effect of different brake components parameters on the degree of instability and squeal index of the brake system. In addition, numbers of researchers have numerically and experimentally investigated the effect factors on the generation of braking noise and vibrations[11-17].

However, because the braking system itself involves many and complex structural

parameters, it is difficult to develop the relationship between the braking noise and the parameters of brake components, and the braking noise is still lacking a complete mathematical prediction model. Therefore, a large number of experiments and/or complex finite element analysis are still needed to be carried out in order to investigate and predict the generation of braking noise.

Thanks to the rapid development of computing technology, in recent years, scholars have applied the finite element analysis (FEA) and statistical analysis methods to the prediction of braking noise[18~23]. Liles[18] carried out parametric study on the effect of design parameters on the generation of brake squeal via complex eigenvalue analysis of finite element models of disc brake components coupled with friction between the brake rotor and linings. The numerical results show that larger COFs and longer linings tended to produce more brake squeals, and the squeal could be decreased using a softer rotor and larger structural damping. Lee et al.[19,20] developed a predictive tool to evaluate disc brake squeal propensity, used an integrated finite element approach to link the dynamic performance of a disc brake system to the non-linear contact behaviour at the friction interface. Ouyang et al.[21] developed a coupled method for solving the dynamic instability of the car disc brake as a nonlinear eigenvalue problem, in which one component (the brake disc) was analytically treated while the other components (the pads, caliper and mounting) were dealt with by the finite element method. Denimal et al.[22] carried out a stability study of a brake system via a finite element model

considering the influence of several contacts between the caliper, the bracket, the pad and the piston. They found that the eigenvalues of the brake system were significantly changed, indicating strong modifications of the stability and thus of the brake squeal propensity was observed when different contact conditions were considered in the FE analysis. Oberst and Lai[23] statistically analysed brake squeal noise based on population distributions and a correlation analysis, to gain greater insight into the functional dependency between the time-averaged COF and pressure level. The results highlight the nonlinear character of brake squeal and indicate the potential of using nonlinear statistical analysis tools to analyze disc brake squeal.

To establish a more advanced finite element model of a braking system by means of modern finite element software (e.g. ABAQUS), the complex geometry and material features of the braking system can be considered in the simulation, and the noise phenomenon of the braking system can be studied. However, it is difficult to consider the different conditions of the whole test environment in a comprehensive manner. Moreover, the computing time is much longer when the FE model is more complex, and the authenticity and reliability of the simulation results are still needed to be verified by experiments.

Therefore, in automotive industries, brake dynamometer tests are widely used for the development of new brake components and validation of the performance of brake systems. Currently, most of the brake dynamometer tests are based on the European AK

dynamometer noise test procedure and the SAE J2521 procedure, by which the most credibility and closest braking noise data to the actual braking noise can be obtained by measuring the braking noise under the different simulated actual braking conditions. However, this approach is inefficient because the dynamometer tests are expensive and time consuming, and therefore does not meet the needs for the fast and optimal development of braking systems.

With the rapid development of machine learning theories, scholars have tried to adopt machine learning methods to solve the problems in industrial applications, and these methods have been successfully used in many different areas, including in prediction of disc brake performance[24], braking intensity classification and quantitative recognition in various deceleration scenarios[25], electric vehicle brake pressure probability estimation[26], and brake noise characterization and detection[27].

In the present work, machine learning methods are used to quantitatively predict the COFs and braking noise based on the experimental data and testing conditions, and qualitatively analyze the main effect factors on braking noise. The main contents of this article can be summarized as the followings:

(1) In this study, the LINK3900 braking test dynamometer is employed for the braking noise tests of five brake discs based on the SAE J2521 brake noise test procedure, and a large number of reliable test data of braking noise are obtained.

(2) For each brake disc, the LSTM algorithm is used to predict the COFs with

different operation conditions, and the predicted COFs are compared with experimental results. The validity and stability of the LSTM model are verified after optimization.

(3) Combining the brake dynamometer test data with the COFs predicted by LSTM, the linear regression (LiR), decision tree (DT), k-nearest neighbor (KNN), support vector machine (SVM) and XGBoost algorithms are employed to predict the braking noise, and the performance of the five methods are compared; in addition, the feasibility, stability, and practicability of the XGBoost algorithm are verified after hyper parameter optimization.

(4) The hyper parameter optimized XGBoost algorithm is used to qualitatively analyze the main effect factors on braking noise.

The remainder of this paper is arranged as follows: section II briefly introduces the principle of LSTM and XGBoost algorithms, which are used in the prediction of COFs and frictional braking noise, respectively. Section III introduces the process of the prediction and analysis of the frictional braking noise. Section IV introduces and discusses the accuracy of the prediction results and the main factors affecting the accuracy of the method, and section V gives conclusions and outlooks.

2 Related Works

In this section, the machine learning algorithms used in the prediction and analysis of braking frictional noise are briefly introduced. In addition, the reasons for selecting these algorithms are explained.

2.1 Principle of the LSTM model

The frictional noise produced by braking is a very typical frictional vibration and noise phenomenon, which is closely related to the COFs between the brake disc and pads[18, 23]. From the viewpoint of data type, the change of the COFs of the brake disc surface during the dynamometer test can be regarded as a time series-based data. Moreover, in the tests, the sequential operations of the brake discs can be summarized as the braking process of several basic modes at different speeds. Therefore, after learning the change of COFs with the braking process, the machine learning model can be used to predict the change of the COFs by the sequential operations in the braking process. To achieve this, the LSTM algorithm proposed by Sepp Hochreiter in 1997[28] is adopted in the present work. The LSTM has been widely used in various time series data-related tasks, such as analysis of working conditions and prediction of working status and running data of power equipment[29], prediction of the health evolution trend of aero-engine[30], fault diagnosis of rolling bearing[31]. In the principle, to overcome the inherent gradient disappear problem in recurrent neural network (RNN) algorithm, LSTM improved RNN by designing a special structure of duplicate “cell”[32], as shown in Figure 1.

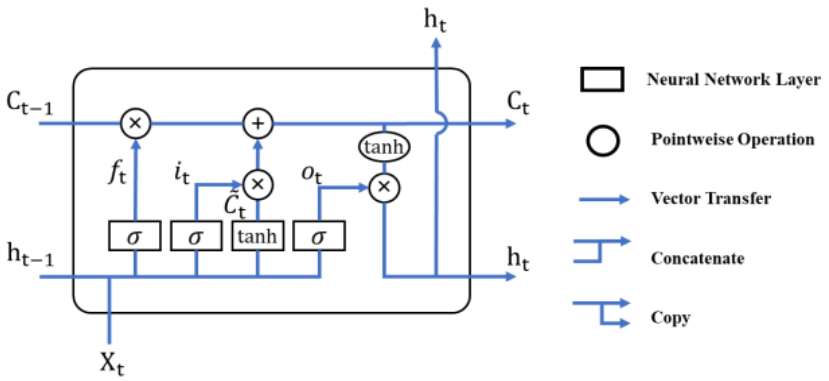


Figure 1. The structure of duplicate "cell" in LSTM

The duplicate structure of LSTM ("memory cell") consists of three different door structures[33]("forget gate", "input gate", "output gate"). In Figure 1, h_{t-1} is the output vector of the previous cell and X_t is the input vector of the current cell. The inference of LSTM consists of the following four steps:

The first step in the LSTM is to decide what information is going to be thrown away from the cell state, which can be expressed as:

$$f_t = \sigma\left(W_f\left[h_{(t-1)}, X_t\right] + b_f\right) \quad (1)$$

Where f_t is the sigmoid function; W_f is the weight; σ is the activate function; b_f is the nodal increment.

The second step is to decide what new information is going to be stored in the cell state. It includes two parts that can be defined as:

$$i_t = \sigma\left(W_i\left[h_{(t-1)}, X_t\right] + b_i\right) \quad (2)$$

$$\overline{C_t} = \tanh\left(W_C\left[h_{(t-1)}, X_t\right] + b_C\right) \quad (3)$$

Where it is the renew, $\overline{C_t}$ is the new candidate vector.

Third, the state of cell will be updated based on the two previous steps, as shown in Eq. (4):

$$C_t = f_t C_{t-1} + i_t \overline{C_t} \tag{4}$$

In Eq. (4), $f_t C_{t-1}$ decides which parts of the states from previous cell will be deleted, while $i_t \overline{C_t}$ decides what new information will be included in the state of the current cell. Then, the final output state of the current cell C_t is calculated.

Initially, O_t is the output gate at t moment and h_t is the output of the LSTM that is calculated as follows:

$$O_t = \sigma\left(W_0\left[h_{(t-1)}, X_t\right] + b_0\right) \tag{5}$$

$$h_t = O_t \tanh\left(C_t\right) \tag{6}$$

In the present work, X_t is the filtered characteristic data under a certain condition, and h_t is the prediction of the COF corresponding to this condition.

2.2 Principle of the XGBoost model

In the present work, the prediction of the frictional braking noise by different operating parameters is essentially a multi variables regression problem in data analysis. In machine learning, there are many algorithms that can solve this problem, such as

LiR, DT, KNN and SVM. However, a single basic learner usually has obvious defects and especially, most of the learners are less explanatory, which is not conducive to analyze the effect of the input feature data on the regression results.

The present work adopts the XGBoost algorithm developed by Chen and Guestrin 34, which is designed to be highly efficient, flexible, portable, and it has been widely used in many industrial fields, such as fault detection of wind turbines[35], prediction of post-fault transient stability status of power systems[36], prediction of the link quality of wireless sensor networks[37]. Figure 2 shows the working principle of the XGBoost algorithm.

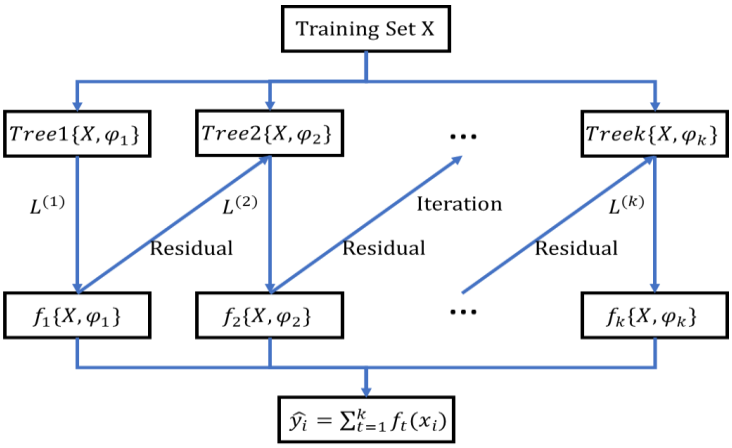


Figure 2. Principle of the XGBoost

Assuming that a dataset is: $D = \{(x_i, y_i) : i = 1, \cdots, n \; x_i \in R_m \; y_i \in R_m\}$

and the objective function of XGBoost is defined as:

Where, k is the number of decision trees; n is the total number of samples;

i represents the i^{th} sample; y_i is the predicted value based on k decision trees by the i^{th} sample, and it can be written as:

$$L = \sum_{i=1}^{n} l\left(y_i \ , \ y_i\right) + \sum_{t=1}^{K} \Omega\left(f_{\mathrm{t}}\right) \tag{7}$$

$$y_i = \sum\nolimits_{k=1}^{k} f_{\mathrm{k}}\left(x_i\right) \tag{8}$$

Here, $\sum_t \Omega\left(f_t\right)$ is the regularization term that represents the complexity of the structure, which is defined as:

$$\Omega\left(f_{\mathrm{t}}\right) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} W_{\mathrm{j}}^2 \tag{9}$$

Both γ and λ are hyper parameters. The role of $\Omega\left(f_{\mathrm{t}}\right)$ is to prevent over fitting, the smaller the value, the lower complexity and the stronger the generalization capability of the structure.

For the optimization, the objective function that adopts additive manner is defined as:

$$L^{(t)} = \sum_{i=1}^{n} l\left(y_i \ , \ y_i^{(t-1)} + f_{\mathrm{t}}\left(x_i\right)\right) + \Omega\left(f_{\mathrm{t}}\right) \tag{10}$$

In contrast to the GDBT (Gradient Boosting Decision Tree)[38], in the XGBoost, the Taylor second order expansion is applied:

$$
\begin{aligned}
L^{(t)} &= \sum_{i=1}^{n} l\left(y_i \ , \ y_i^{(t-1)} + f_t(x_i) \right) + \Omega(f_t) \\
&\approx \sum_{i=1}^{n} \left[l\left(y_i \ , \ y_i^{(t-1)} \right) + g_i f_t(x_i) \right] \\
&+ \sum_{i=1}^{n} \left[\frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t)
\end{aligned}
\tag{11}
$$

Where $g_i = \partial_{y_i^{(t-1)}} l\left(y_i, y_i^{(t-1)} \right)$, $h_i = \partial^2_{y_i^{(t-1)}} l\left(y_i, y_i^{(t-1)} \right)$, g_i and h_i is $l\left(y_i, y_i^{(t-1)} \right)$ to $y_i^{(t-1)}$ the first-order derivative and second-order derivative, respectively; while $L^{(t)}$ is the objective function.

By removing the constant term, the final objective function is simplified as follows:

$$
L^{(t)} = -\frac{1}{2} \sum_{j=1}^{T} \frac{G_j^2}{H_j + \lambda} + \gamma T \tag{12}
$$

Where $G_j = \sum_{i \in I_j} g_i$, $H_j = \sum_{i \in I_j} h_i$

Finally, by calculating the gain after divide the leaf nodes, the gain (Gini) formula before and after the split is defined as follows:

$$
\begin{aligned}
G_{ain} = & \frac{1}{2} \left(\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} \right) \\
& - \frac{1}{2} \frac{(G_L + G_R)^2}{H_R + H_L + \lambda} - \gamma
\end{aligned}
\tag{13}
$$

Where $G_L = \sum_{i \in I_L} g_i$, $H_R = \sum_{i \in I_R} h_i$, I_L and I_R are the set of instances belong to

the left and right nodes after the split. When a leaf node is splitting, the Gain for all candidate features is calculated, and the feature that can contribute to the largest Gain is selected for current splitting.

In summary, the XGBoost has many advantages. First, from the view of structural variety, it contains various of different types of base learners, such as LiR and DT; second, from the view of mathematical theory, it uses the information of secondary derivatives; third, from the view of model training, it adds the complexity of the model structure as a regularization item in the objective function to optimize the model and adds cutting processing in order to make the model difficult to over fit. In addition, it proposes a method of feature secondary sampling.

Based on these features, the XGBoost can combine a series of different weak learners to build a strong regression model with better performance than other base learners. This is helpful to the questions raised in the present work. In industrial scenarios of machine learning applications, data are often obtained through expensive equipment and complex operations, so the amount of data is relatively limited. Based on the limited data, by integrating different weak learners, the final algorithm can have the ability to analyze the data from different viewpoints and further improve the performance of the braking noise prediction algorithm.

3 Approach

This paper proposes a prediction algorithm based on the optimized XGBoost

model to predict frictional braking noise from the selected features obtained by the LINK3900 brake dynamometer tests and the COFs predicted by the optimized LSTM algorithm. The overall flow of the algorithm in the present work is shown in Figure 3.

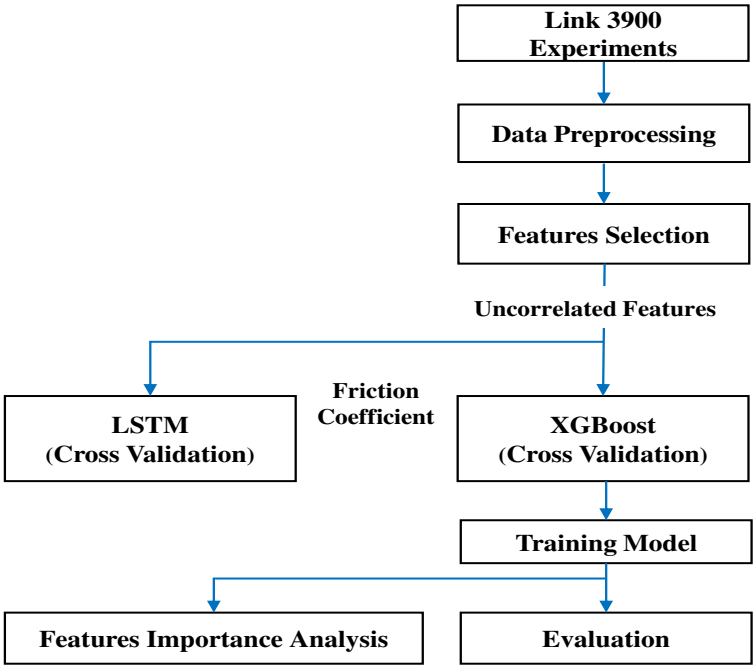


Figure 3. The flow chart of proposed approach

3.1 Data preprocessing and features selection

Normally, because of the instability of different sensors, data obtained directly from engineering tests contain missing data and outliers. Due to the different properties of data from various sensors and the large difference in numerical levels, there are usually different units and orders of magnitude among data from different sensors. To make the prediction results of the machine learning model more reliable and the training more efficient, data normalization and data cleaning methods are employed in the

present work to pre-process the original experimental data.

Data normalization: different algorithms often require different normalization methods to improve their computational accuracy and efficiency. In this paper, in order to avoid over fitting, the Min-Max normalization method is used for the input data of the LSTM algorithm; while for braking noise prediction, the Z-Score standardization method is applied in the LiR, SVM, and other three algorithms to reduce the amount of computation and improve the efficiency of the model.

Data cleaning: In this work, the 'fillna' function in Python language is used to realize the automatic detection and filling of abnormal data. By checking all the test data of each brake disc in turn, using the mean fill method for the missing data, means the missing data is filled with the average value of all the data in the column in which it belongs; after processing the missing data, the quantile-quantile plot is used in combination with the box plot to find the abnormal data.

Feature engineering: Data analysis and visualization are helpful for better understanding the data characteristics, selecting the right machine learning algorithms and removing the redundant features. At the same time, due to a large amount of information and complexity of the test data from the brake dynamometer tests, the data set contains both discrete and continuous data with different operation parameters (features), and possible there are many redundant features. If there are many redundant features in the samples, the complexity of the analysis and the computational cost of the

training process will be increased. Therefore, in the present work, for the training set of each brake disc, data are analyzed by the scatter plot visualization, the Pearson product-moment correlation coefficient (PPMCC), and the maximal information coefficient (MIC).

In addition, the heatmap of the correlation coefficient matrix is also visualized to show the degree of correlation between the features quantitatively and intuitively. Through the comprehensive analysis of the results of the above three methods, only one of the strong correlation features will be retained. By feature engineering the complexity of data can be reduced and the over fitting of the model can be avoided. Thus, this improves the reliability and accuracy of the prediction algorithm.

As discussed above, in order to analysis the correlation between features, at first, the scatter plots between each pair of 12 features are analyzed, as shown in Figure 4, the scatter plots of the data in two features pairs (mean torque f4 and maximum torque f5, mean pressure f6 and maximum pressure f7) are illustrated. Based on the experience, f4 and f5, f6 and f7 should have a strong correlation respectively, which can be verified by the dispersion of data points of these features as shown in Figure 4. To further quantitatively analyze the linear and nonlinear correlation between the features, the PPMCC and the MIC between the features are calculated, and the heat-maps of the PPMCC and MIC correlation coefficient matrixes are presented in Figure 5, to enhance the visualization of information.

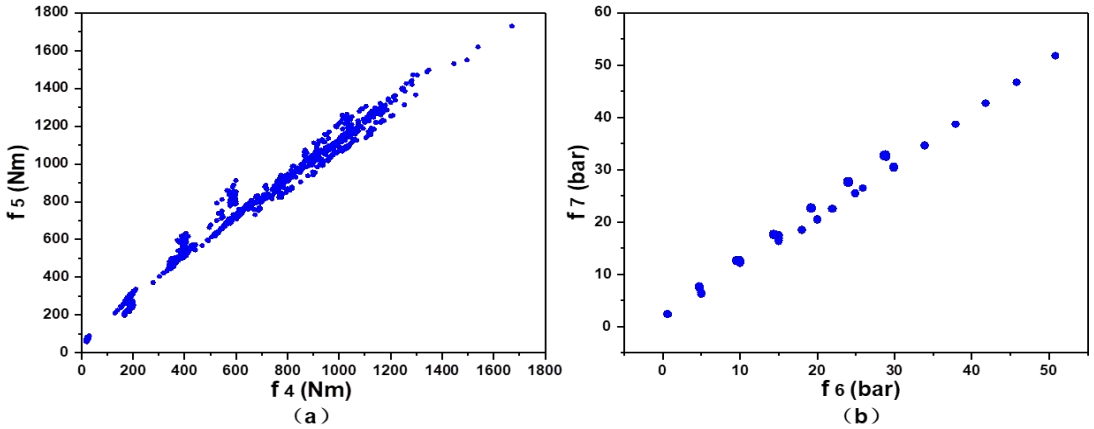
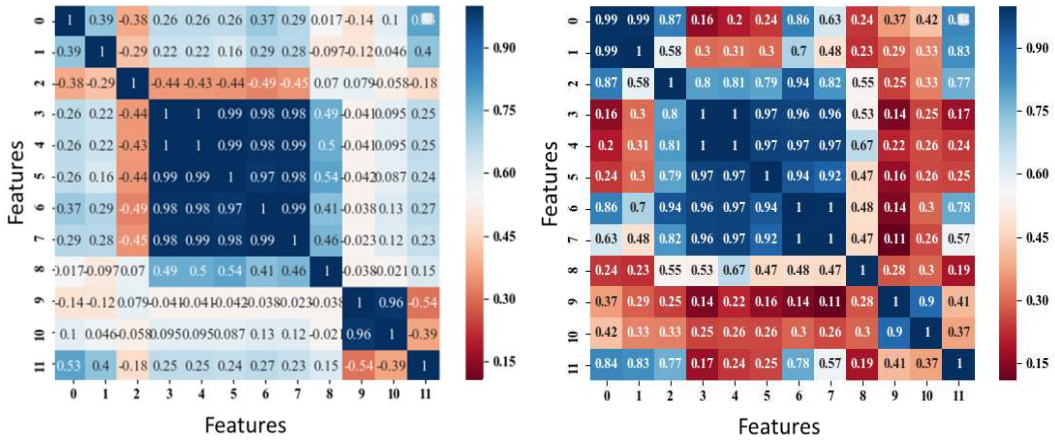


Figure 4. Scatter plots of disc A: (a) correlation between the mean braking pressure and maximum pressure; (b) correlation between the mean torque and maximum torque



(a)PPMCC (b) MIC

Figure 5. Heatmap of correlation metrics of disc A

The PPMCC between variables is defined as 39

$$r = \frac{\sum_{i=1}^{n}(x_i - x)(y_i - y)}{\sqrt{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}\sqrt{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}} \tag{14}$$

Where n is the number of test data; the larger is the PPMCC, the stronger is the correlation; in contrast, the smaller is the PPMCC, the weaker is the correlation between

the variables.

The MIC is defined as 40

$$MIC(D) = \max_{xy \le B(n)} \frac{I(D, x, y)}{\log \min\{x, y\}} \quad (15)$$

Where $I(D, x, y) = \max I(D|G)$ is the MIC under different grid partitioning, and D is the set of all data points. When $MIC(X, Y) = 0$, there is no correlation between two variables (X and Y); while when $MIC(X, Y) = 1$, here is an explicit linear or non-linear correlation between two variables.

In general, the conclusions drawn by the introduced three methods are consistent, which can validate each other. MIC can identify both the linear and non-linear correlation between variables, but the PPMCC cannot. After the analysis of the correlation between features, five parameters are selected as features for training the prediction model, namely, the initial braking speed, release braking speed, mean braking pressure, initial braking teAmperature, and COF of the brake disc.

3.2 Prediction of COFs

In the present work, the LSTM algorithm model is constructed by Tensor flow. The optimal model after hyper parameter optimization is obtained based on the training, inference, and evaluation with the objective of minimizing prediction error by preprocessed test data. In addition, the predicted COFs based on the model are compared with the experimental data.In order to further improve the model

performance, the present work improves the normal LSTM algorithm by means of the Adam optimization[41] and the Xavier parameter initialization [42].

The results are shown in Table 1. In the training, the loss is calculated by the L2 loss (least squared error) method. Because of the performance and computational limitations of the algorithm, the batch training is adopted, and the COFs are predicted iteratively. Finally, the model inference is realized by using the trained LSTM, the accuracy and stability of the model are evaluated by the R^2, mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE) values.

Table 1 shows that LSTM has a system dependent optimal set of hyperparameters. The reason is that because the five brake discs have different structural characters, the measured friction coefficients and brake noise from different discs under the same experimental conditions are significantly different. Five models are trained for the five discs based on the tested data respectively and the hyperparameters for each disc are optimized with corresponding model and data. Therefore, the optimal set of hyperparameters is system dependent.

Table 1. Optimal hyperparameters combinations of discs of LSTM model

Disc	RNN Unit	Learning Rate	Iteration Times
A	10	0.001	10000
B	18	0.0001	10000
C	5	0.0001	10000
D	20	0.0001	10000
E	10	0.0001	10000

3.3 Prediction of braking noise

Integrated with the predicted COFs based on LSTM, the whole datasets for noise prediction contains five different features, namely the initial braking speed, release braking speed, mean braking pressure, initial brake disc temperature, and mean friction coefficient.

Based on the three libraries (TensorFlow, Scikit-learn, and Numpy) in Python, the models of LiR, KNN, DT, SVM, and XGBoost are developed and their initial parameters are set by default. The five models are trained by the processed training datasets respectively and evaluated by the test datasets. Especially, the training process adopts the k-fold cross-validation method, which can effectively avoid the phenomenon of under and over fitting. Calculating the R^2 values of each model based on the comparison between the predicted and measured values allows evaluation of the performance of the five models and the selection of the model with the highest R^2 values.

In the present work, because the XGBoost method gives the best R^2 values, grid searching is applied to the XGBoost method to find the optimal combination of hyper parameters to further improve the model performance, as shown in Table 2. Finally, the XGBoost model is trained by the whole training dataset and to evaluate the accuracy and reliability of the trained model, the R^2, MAE, MSE and RMSE values are calculated based on the trained XGBoost model by the predicted and measured braking noise.

Table 2. Optimal hyper parameters combinations of discs of XGBoost model

Disc	Learning Rate	Depth of Layers	Amount of Estimators	R^2
A	0.3	2	100	0.898
B	0.3	2	50	0.834
C	0.1	5	100	0.706
D	0.1	3	50	0.934
E	0.1	6	50	0.722

3.4 Metrics for model evaluation

For classification problems, there are several different evaluation metrics that can be used to evaluate the accuracy of the regression problem. In general, people prefer to choose multiple metrics because single metric cannot fully evaluate the performance of a model. In the present work, four metrics, R^2, MSE, RMSE, and MAE, are adopted to evaluate the accuracy of the models. At first, the performance of LSTM or XGBoost is evaluated by the above mentioned four indexes; then, the R^2 is used to calculate the improvement of the model's performance before and after the hyper parameter optimization.

The definitions of the four indexes are as follows. The MAE represents the difference between the original and predicted values extracted by averaging the absolute difference over the dataset, which is defined as:

$$MAE = \frac{1}{m}\sum_{i=1}^{m}\left|\left(y_i - \hat{y}_i\right)\right| \qquad (16)$$

Where m is the amount of test data, y_i is the measured value, and $\hat{y}_i$ is the predicted value.

The MSE represents the difference between the original and predicted values extracted by squaring the average difference over the dataset, which is defined as:

$$MSE = \frac{1}{m}\sum_{i=1}^{m}\left(y_i - \hat{y}_i\right)^2 \quad (17)$$

The RMSE is the error rate given by the square root of MSE.

$$RMSE = \left(\frac{1}{m}\sum_{i=1}^{m}\left(y_i - y_i\right)^2\right)^{0.5} \quad (18)$$

The R^2 (R-squared) represents the coefficient of how well the predicted values fit the original values, which is defined as:

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum_{i=1}^{n}\left(y_i - y_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \overline{y_i}\right)^2} \quad (19)$$

Where RSS is the explained variation and TSS is the total variation. The value of R^2 is from 0 to 1 interpreted as a percentage. The higher the R^2 value, the better the model.

3.5 Analysis of the feature importance

There are many different variables that can affect frictional braking noise; these can be summarized as friction pair factors, brake structure factors, environmental factors, and braking conditions, which is difficult to quantify. Based on the perspective

of the study of quantifiable factors, this paper narrows the scope of the study, focuses on the critical factors that are closely related to the braking noise, and makes full use of the advantages of the XGBoost algorithm to sort the importance of these factors, by means of three common metrics, namely weight, gain, and cover. In this context, the weight refers to the number of times the feature is used as a splitting tree node in all trees or the number of times it is used to split samples; the gain refers to the average gain generated when the feature is used to split tree nodes in all decision trees, and the cover refers to the average coverage when the feature is used in the tree or the average coverage that it is applied within the tree structure.

In this study, the weight is chosen as the metric to calculate the importance scores of different features. First, the XGBoost model corresponding to the five brake discs and the datasets are used to calculate the importance index of each feature, and then multiply the weight of each brake disc, and finally calculate the sorting average of each feature, in which the weight of each disc is proportional to the R^2 value of the corresponding trained XGBoost model.

4 Results and analysis

In this section, the braking dynamometer test equipment and the test procedure are firstly introduced; then, the performance of the prediction algorithms of the COFs and the frictional braking noise proposed for the first time in this paper are tested by comparison with the experimental data; in addition, the noise effect factors are analyzed

based on the XGBoost model; finally, some factors closely related to the performance of the algorithm framework proposed in this study are discussed.

5 Experiments setup and methodology

The LINK3900 brake test dynamometer was employed in the present work to carry out the tests, as shown in Figure 6. In order to prove that the method proposed in this study is effective for the brake discs of different structures, five brake discs with different surface structures were tested. Figure 7 shows the five brake discs before testing. More details about the surface structures on the five brake discs can be found in[43]. The brake dynamometer can test the braking performance of the brake discs and friction pads at different speeds, temperatures, and contact pressures by running various test procedures (e.g. QC/T564, ISO2667, SAE-J2521, etc.). This is the most convincing way to obtain and measure the important performance data of the brake discs and pads, such as wear, COF, DTV (disc thickness variation), and braking noise, etc.

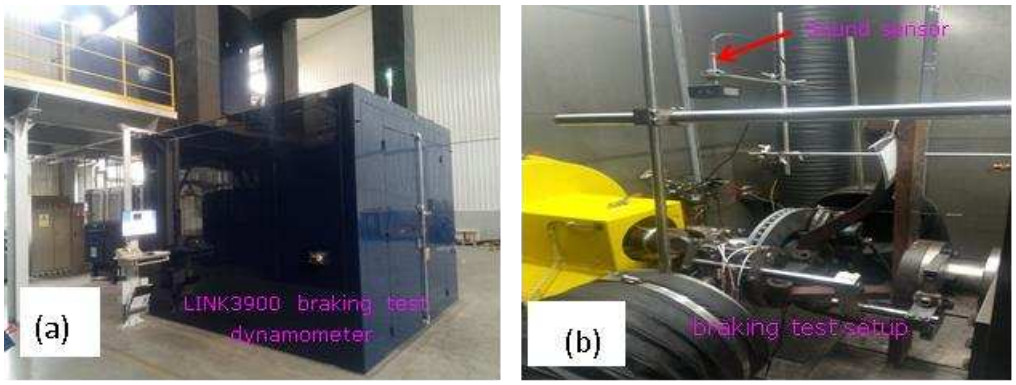


Figure 6. Illustration of experimental equipment and setup

Figure 7. Five brake discs before testing

In this study, the standard test procedure of SAE-J2521 for braking noise was strictly followed. This procedure was developed by the International Society of Automotive Engineers and is recognized by the global automotive industry as the standard experimental procedure to simulate various braking conditions in which vehicle braking systems may produce noise such as brake squeals. The SAE-J2521 standard braking test procedure consists mainly of the following basic brake conditions: Snub Brake, Brake, Deceleration Brake, Cold Brake, and Fade Brake, which includes a total of 2321 braking stops in 31 different test modules with various braking conditions.

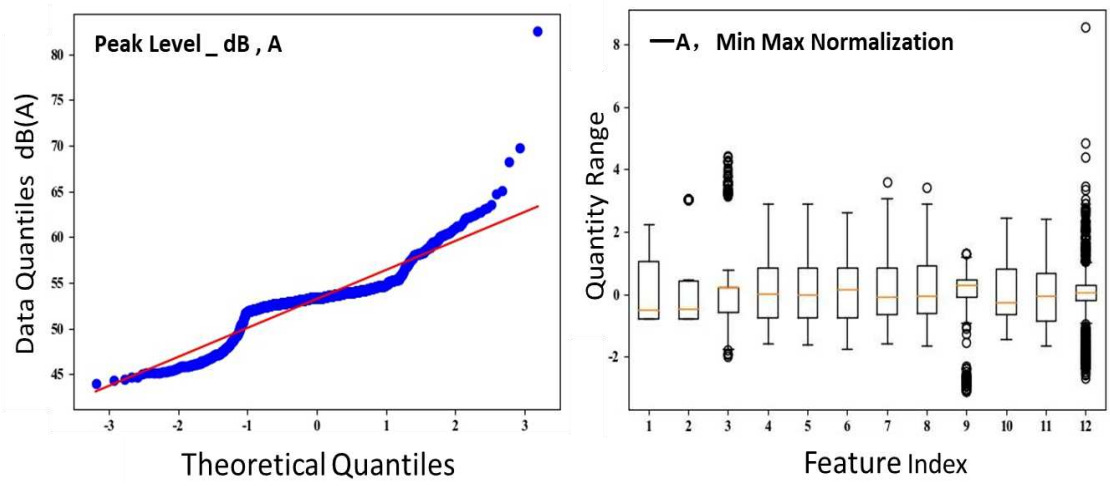
In the present work, nearly 1000 braking noise data obtained from braking tests of each brake disc under various braking conditions were used for the model training and prediction of brake noise. The feature sets and descriptions are listed in Table 3.

The data used in this study are peak noise levels in dB(A), which were measured in a frequency range of 0.9-17 kHz using a B&K microphone. The installation of the sound sensor can be seen in Figure 6. In order to effectively predict the braking noise, data cleaning and feature engineering are firstly employed to remove outliers and redundant features. Then, the remaining data are used to generate the complete training data sets in combination with the COFs predicted by the LSTM.

In this work, the Q-Q quantity plot division method combined with the Box plot method are employed to clean the data. Figure 8 (a) and (b) demonstrates the Q-Q quantity plot and Box plot of data cleaning of brake disc A, respectively.

The Q-Q plot in Figure 8 (a) shows the relationship between the theoretical quantities and the actual quantities. The red line in Figure 8 (a) represents the theoretical quantities of Gaussian distribution that meets the training requirements after linear normalization. The slope of the red line represents the standard deviation, and the intercept of the red line represents the mean value. The blue points represent the measured data of the sound pressure level in dB(A) of brake disc A. Figure 8 (a) shows that, some experimental data are far away from the red straight line, indicating there are outliers of this feature in brake disc A (e.g. three outliers in the upper right corner of Figure 8 (a)), and these outliers need to be eliminated.

Figure 8 (b) is a Box plot. The biggest advantage of a Box plot is that it is not affected by outliers and can accurately depict the discrete distribution of data, and it is also conducive to data cleaning. It can be seen from Figure 8 (b) that there are a lot of data outside the upper and lower boundaries of the Box plot, indicating that these data are outliers and need to be deleted, so that the remaining data generally obeys the Gaussian distribution, and can then be used as machine learning data sets.



（a）Q-Q plot (b) Box plot

Figure 8. Illustration of the data cleaning of brake disc A. (a) Quantity plot; (b) Box plot

Table 3. Feature sets

Feature	Description	Unit	Feature	Description	Unit
f1	Brake speed	km/h	f7	Mean pressure	bar
f2	Release speed	km/h	f8	Maximal pressure	bar
f3	Braking period	s	f9	Mean friction coefficient	-
f4	Mean deceleration	m/s2	f10	Initial brake disc temperature	℃
f5	Mean torque	Nm	f11	Final brake disc temperature	℃
f6	Maximal torque	Nm	f12	Frictional braking noise	dB

5.1 MAIN Results

Prediction of COFs Based on the optimal combination of hyper parameters, the COFs of each disc in the last 30 braking operations in the SAE-J2521 standard are predicted by the LSTM algorithm. Figure 9 presents the predicted and measured COFs. It is observed from Figure 9 that, the LSTM prediction model tracks the change of the measured COFs very well. The overall observation of the predicted COFs of the five

discs verified that the overall prediction accuracy of the LSTM model is pretty good.

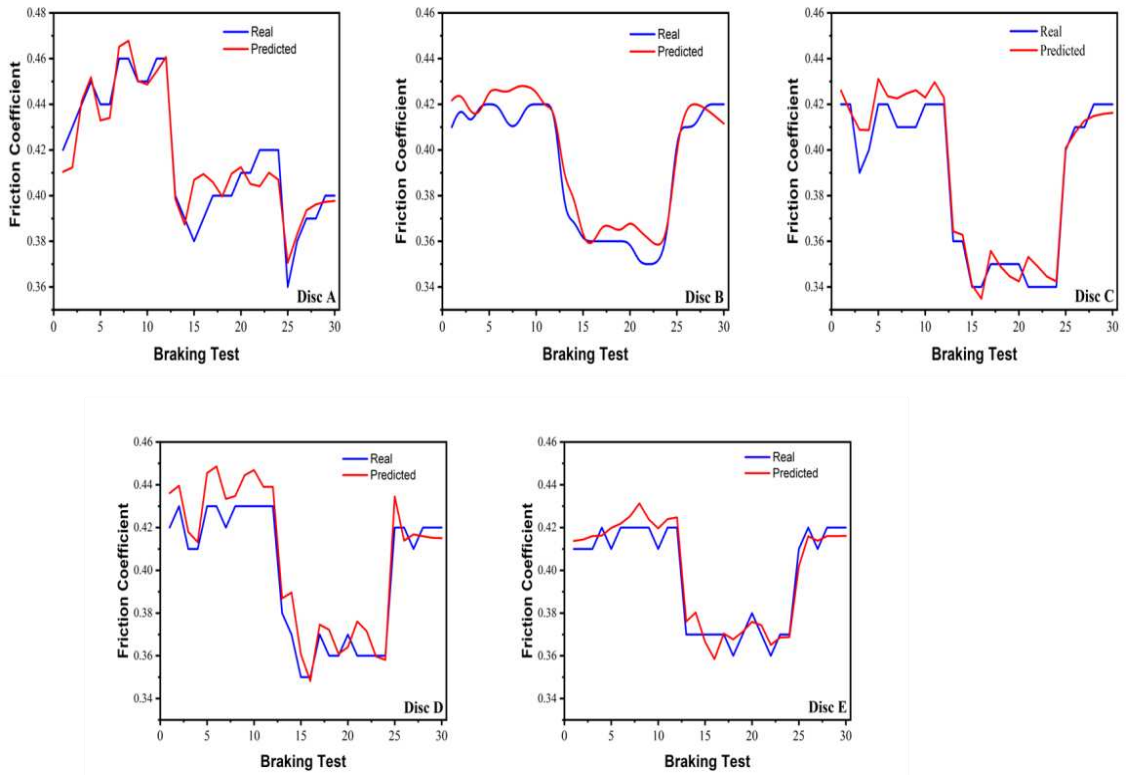


Figure 9. Predicted and measured coefficients of friction based on the LSTM

Table 4 illustrates the R^2, MAE, MSE and RMSE between the measured and predicted COFs by LSTM. It is observed from Table 4 that, in contrast to the R^2, the MAE, MSE and RMSE are very small, proving that based on the method proposed in this study, the LSTM model can be used to accurately and reliably predict the friction coefficients of each brake disc with different braking conditions.

Table 4. Performance of optimized LSTM model

Disc	Evaluation Metrics

	R^2	MSE (10-2)	RMSE (10-2)	MAE (10-4)
A	0.86	0.69	0.93	0.87
B	0.85	0.75	0.87	0.76
C	0.91	0.64	0.79	0.63
D	0.86	0.75	0.86	0.75
E	0.87	1.54	1.80	3.24

Prediction of frictional braking noise The R^2 of the five different algorithms are illustrated in Figure 10, from which it is observed that the XGBoost model has the best performance for the braking noise prediction. In addition, pretty good R^2 values are obtained from the KNN algorithm, but the KNN algorithm is lacking explanation, because it can not be used to definitely analyze the correlation between the operating parameters and the braking noise. Furthermore, as shown in Figure 11, the difference between the evaluated R^2 values of the XGBoost model of the validation dataset and the test dataset is the smallest among the five algorithms, indicating that this method has less tendency to be over fitted. In summary, the XGBoost algorithm is the best and is therefore selected for the frictional braking noise prediction in this study.

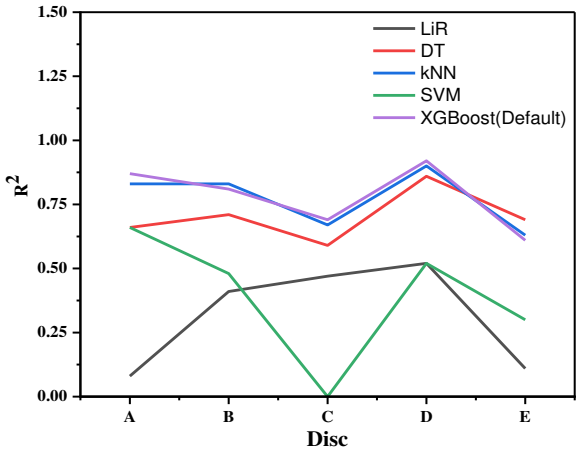


Figure 10. The R^2 of five discs in different methods

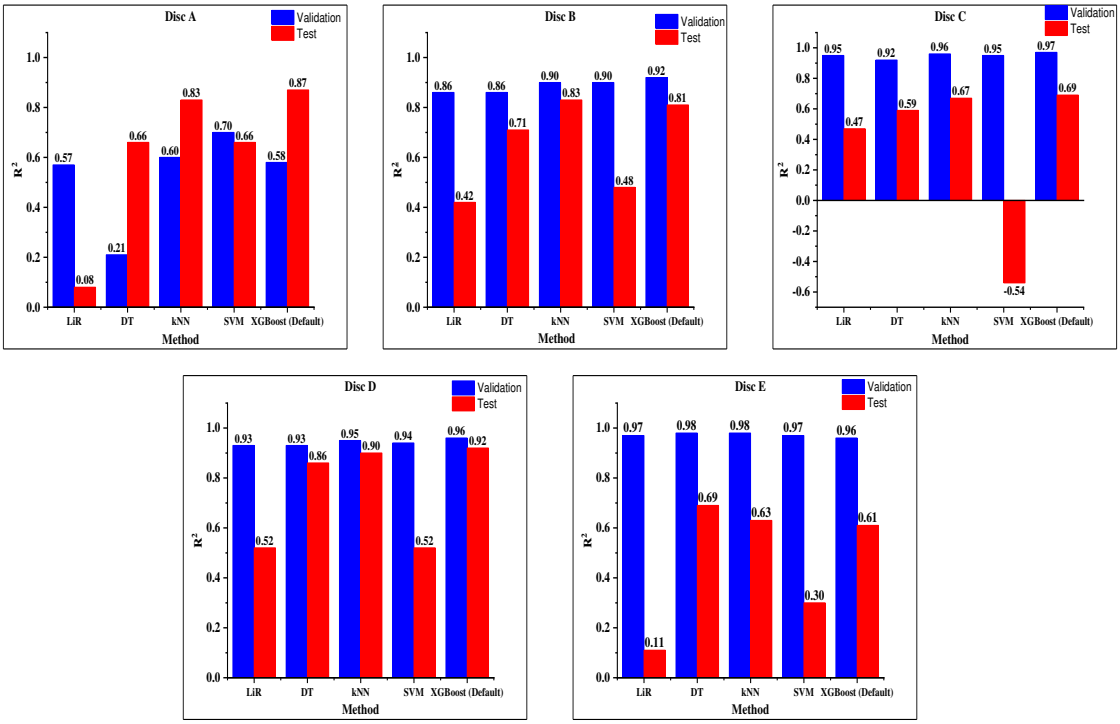


Figure 11. The validation and test R^2 values

The XGBoost model with grid searched hyper parameters enables the accurate prediction of samples in the test dataset. The predicted and measured noise values in dB

are shown in Figure 12. The horizontal coordinates are 30 braking conditions in the test datasets, and the vertical coordinates are the predicted and measured noise in dB. It can be seen from Figure 12 that the predicted noise of each brake disc based on the optimized XGBoost model can fit the measured noise quite well, which shows that the XGBoost model has good stability in the braking noise prediction. In addition, the MAE, MSE and RMSE values of braking noise prediction based on the XGBoost model for the five discs are listed in Table 5. The R^2 values of the five discs are reasonable large, and the three errors are reasonable small, proving that the XGBoost model optimized on the basis of the method presented herein can be used to predict the frictional braking noise with a reasonably good accuracy and reliability under different braking conditions.

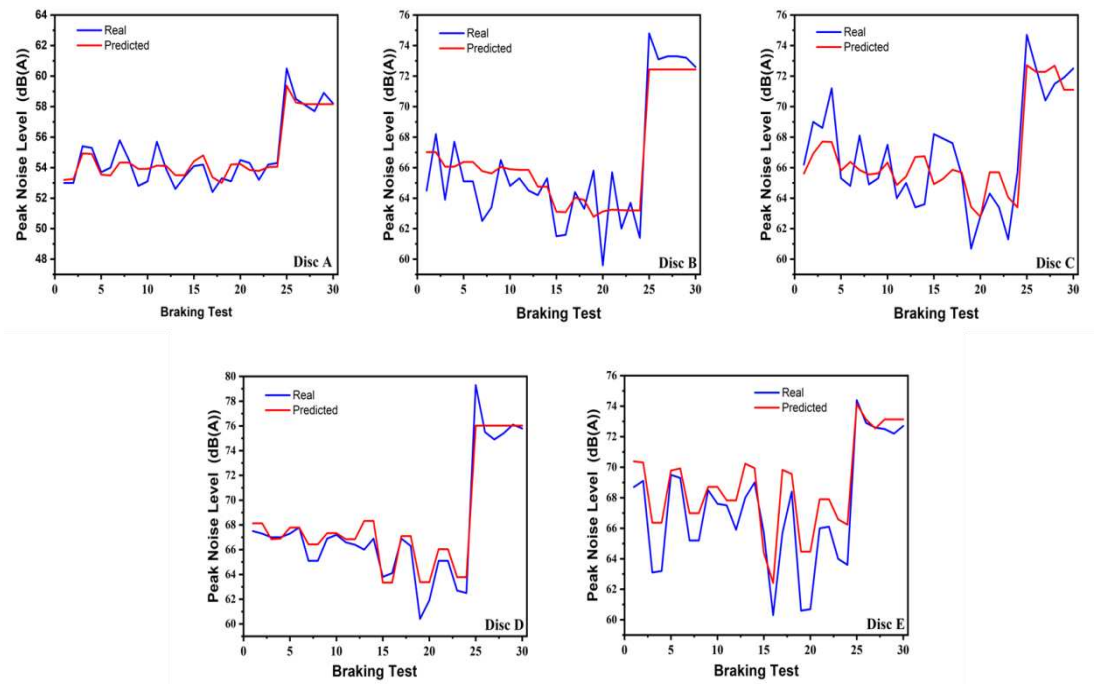


Figure 12. Measured and predicted noise based on the optimized XGBoost model

Table 5. Performance of optimized XGBoost model on different discs

Disc	Evaluation Metrics			
	R^2	MSE (10-2)	RMSE (10-2)	MAE (10-4)
A	0.90	0.46	0.68	0.53
B	0.83	2.80	1.67	1.41
C	0.71	3.63	1.91	1.60
D	0.93	1.41	1.19	0.89
E	0.72	3.95	1.99	1.61

Analysis of different features Based on the experimental data of the five brake discs, the XGBoost algorithm is used to sort the importance of the five features affecting the frictional braking noise. The importance of a feature is identified by averaging the sum of the results of a feature in all the ascension trees, the smaller the average importance ranking score of the feature, the more important the feature.

It is observed from Table 6 that the importance of the features influencing the braking noise is in the following order: the initial braking temperature, the mean COFs, the mean braking pressure, the initial braking speed, and the release braking speed of the brake discs; this order is in line with the general recognition of braking noise experts.

Table 6. Importance of different features

Discs	Initial Braking Speed	Release Braking Speed	Initial Braking Temperature	Friction Coefficient	Mean Braking Pressure	Weights
A	4	5	2	3	1	0.898
B	3	5	2	1	4	0.834
C	4	5	1	2	3	0.706
D	1	5	2	3	4	0.934
E	4	5	1	3	2	0.722
Averaged importance	2.548	4.094	1.352	1.981	2.306	——

ranking score

5.2 Hyper paramemters optimization

In order to find the optimal combination of the hyper parameters of the LSTM and XGBoost algorithms, the grid searching method is adopted in this work. As shown in Figure 13 (a), the R^2 values between the predicted and measured COFs of the five brake discs are all increased after optimization, which means that LSTM model has been improved after grid searching. Similarly, the XGBoost braking noise prediction model has also been optimized by the grid searching method. Figure 13 (b) shows that the braking noise prediction R^2 values of the five brake discs are all increased, and the average improvement is 5.58%. Therefore, it can be concluded that the grid searching method can improve the performance of the LSTM and XGBoost models.

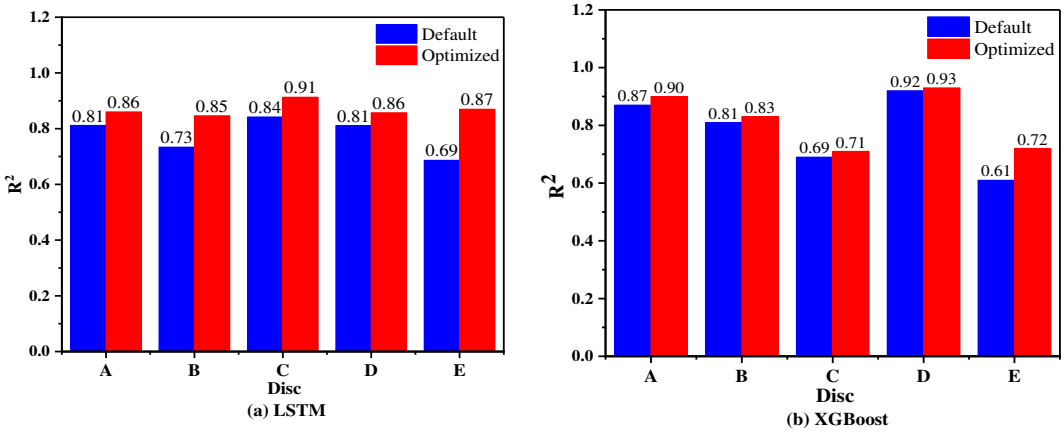


Figure 13. The R^2 values before and after grid searching method. (a) LSTM; (b) XGBoost

6 Conclusions and outlook

In the present work, in order to overcome the limitations of the current research in

the field of the frictional braking noise of automobiles, a frictional braking noise prediction model based on machine learning algorithms and experimental data from the braking dynamometer tests is developed for the first time in this field.

In this study, the braking COFs and braking noise are successfully predicted by the LSTM algorithm and XGBoost algorithm, respectively. In addition, the correlation and order of importance of braking features corresponding to the generation of braking noise are further analyzed and studied by the importance analysis of the XGBoost algorithms. In the two algorithms, the R^2 values are in most cases larger than 0.8.

Moreover, the other three metrics (MAE, MSE and RMSE) of the two prediction models for all the five brake discs reach a pretty good level of about 0.0395, 0.0199, and 0.0161, respectively. From the analysis of the importance of input features in the XGBoost model, the importance order of the features is identified as (from more important to less important): initial braking temperature, COF, braking pressure, initial braking speed, and the release braking speed.It can be concluded that the methods proposed in this study not only can be used to accurately and efficiently predict the COFs and frictional braking noise of a brake disc with specific structures, but also can be used as a novel way to analyze the braking noise effect factors by means of the XGBoost algorithm. Therefore, this approach is very worthy of continuous and comprehensive investigation in the future, in order to better understand the mechanism of braking noise generation. Specifically, the following further work is particularly

needed to be carried out.

First, more experimental data are needed to verify the accuracy and reliability of the prediction models; second, data for a large number of different brake disc materials and structures are required to construct and train a more comprehensive model with more features for the prediction of frictional braking noise of brake discs with more features; third, the state-of-the-art machine learning algorithms and strategies for data regression must be investigated to constantly improve the learning algorithms for the braking noise prediction.

CONFLECT OF INTEREST

On behalf of all the authors, the corresponding authors state that there is no conflict of interest.

ACKNOWLEDGEMENTS

This study was supported by the Science and Technology Committee of Shanghai Municipal Key Project (18060502400).

References

1. M. R. North, A mechanism of disc break squeal, 14th FISITA Congress, 1972, pp. 1-9.
2. D. A. Crolla and A. M. Lang. Paper VII (i) Brake Noise and Vibration - The State of the Art, Tribol. Ser., vol. 18, no. C, pp. 165–174, 1991, doi: 10.1016/S0167-8922(08)70132-9.
3. S. Yang and R. F. Gibson. Brake vibration and noise. Reviews, comments, and

proposals, Int. J. Mater. Prod. Technol., vol. 12, no. 4–6, pp. 496–513, 1997, doi. 10.1504/IJMPT.1997.036384.

4. N. M. Kinkaid, O. M. O Reilly and P. Papadopoulos. On the transient dynamics of a multi-degree-of-freedom friction oscillator. a new mechanism for disc brake noise, Journal of Sound and Vibration, vol.287, no. 4–5, 2005, pp. 901-917.
5. H. Ouyang, W. Nack, Y. Yuan, and F. Chen.Numerical analysis of automotive disc brake squeal. a review, Int. J. Veh. Noise Vib., vol. 1, no. 3–4, pp. 207–231, 2005, doi: 10.1504/ijvnv.2005.007524.
6. S. Oberst and J. C. S. Lai.Chaos in brake squeal noise, J. Sound Vib., vol. 330, no. 5, pp. 955–975, 2011, doi: 10.1016/j.jsv.2010.09.009.
7. M. R. North. Disc brake squeal – a theoretical model, The Motor Industry Research Association, 1972.
8. N. Millner,An analysis of disc brake squeal, SAE paper 780332.
9. H. Murakami, N. Tsunada, M. Kitamura.A study concerned with a mechanism of disc brake squeal, SAE paper 841233, 1984.
10. I. Ahme,Analysis of disc brake squeal using a ten-degree-of-freedom model, Int. J. Eng. Sci. Technol., vol. 3, no. 8, pp. 142–155, 2012, doi. 10.4314/ijest.v3i8.12.
11. M. Nishiwaki,Generalized theory of brake noise, Proc. Inst. Mech. Eng., Part D, vol. 207 no. 3, 1993, pp.195-202.
12. H. Ouyang, J. E. Mottershead, M. P. Cartmell, and M. I. Friswel. Friction-induced parametric resonances in discs, Effect of a negative friction-velocity relationship, J. Sound Vib., vol. 209, no. 2, pp. 251–264, 1998, doi. 10.1006/jsvi.1997.1261.
13. T. Butlin and J. Woodhouse. A systematic experimental study of squeal initiation, J. Sound Vib., vol. 330, pp. 5077–5095, 2011, doi: 10.1006/jsvi.2011.05.018.
14. G. P. Ostermeyer and M. Graf. Mode coupling instabilities induced by a periodic coefficient of friction, SAE International Journal of Passenger Cars - Mechanical Systems, vol. 3, no.2, Oct. 2010, doi: 10.4271/2010-01-1687.

15. G. Lou, T. W. Wu, and Z. Bai. Disk brake squeal prediction using the ABLE algorithm, J. Sound Vib., vol. 272, no. 3–5, pp. 731–748, 2004, doi: 10.1016/S0022-460X(03)00416-4.
16. P. Gao, Y. Du, J. Ruan, and P. Yan. Temperature-dependent noise tendency prediction of the disc braking system, Mech. Syst. Signal Process., vol. 149, p. 107189, 2021, doi: 10.1016/j.ymssp.2020.107189.
17. J. D. Fieldhouse and T. P. Newcomb.Double pulsed holography used to investigate noisy brakes, Opt. Lasers Eng., vol. 25, no. 6 SPEC. ISS., pp. 455–494, 1996, doi: 10.1016/0143-8166(95)00094-1.
18. G. D. Liles.Analysis of disc brake squeal using finite element methods, SAE Noise & Vibration Conference & Exposition, 1989.
19. Y. S. Lee, P. C. Brooks, D. C. Barton, and D. A. Crolla, A predictive tool to evaluate disc brake squeal propensity part 1. The model philosophy and the contact problem, Int. J. Veh. Des., vol. 31, no. 3, pp. 289–308, 2003, doi: 10.1504/IJVD.2003.003360.
20. Y. S. Lee, P. C. Brooks, D. C. Barton, and D. A. Crolla.A predictive tool to evaluate disc brake squeal propensity part 2. System linearisation and modal analysis, Int. J. Veh. Des., vol. 31, no. 3, pp. 309–329, 2003, doi: 10.1504/IJVD.2003.003362.
21. H. Ouyang, J. E. Mottershead, and W. Li.A moving-load model for disc-brake stability analysis, J. Vib. Acoust. Trans. ASME, vol. 125, no. 1, pp. 53–58, 2003, doi: 10.1115/1.1521954.
22. E. Denimal, S. Nacivet, L. Nechak, and J. J. Sinou. On the influence of multiple contact conditions on brake squeal, Procedia Eng., vol. 199, pp. 3260–3265, 2017, doi: 10.1016/j.proeng.2017.09.355.
23. S. Oberst and J. C. S. Lai. Statistical analysis of brake squeal noise, J. Sound Vib., vol. 330, no. 12, pp. 2978–2994, 2011, doi: 10.1016/j.jsv.2010.12.021.
24. D. Aleksendrić and D. C. Barton.Neural network prediction of disc brake performance, Tribol. Int., vol. 42, no. 7, pp. 1074–1080, 2009, doi:

10.1016/j.triboint.2009.03.005.

25. L. Chen, X. Yang, L. Chao, Y. Liu, H. Guo, H. Gao, and D. Cao.Hybrid-learning-based classification and quantitative inference of driver braking intensity of an electrified vehicle, IEEE Transactions on Vehicular Technology, doi:10.1109/TVT.2018.2808359, 2018.
26. X. Yang and L. Chen.Dynamic state estimation for the advanced brake system of electric vehicles by using deep recurrent neural networks, IEEE Transactions on Industrial Electronics, doi: 0.1109/TIE.2019.2952807, 2019.
27. M. Stender, M. Tiedemann, D. Spieler, D. Schoepflin, N. Hoffmann, and S. Oberst.Deep learning for brake squeal. Brake noise detection, characterization and prediction, Mech. Syst. Signal Process., vol. 149, p. 107181, 2021, doi: 10.1016/j.ymssp.2020.107181.
28. S. Hochreiter and J. Schmidhuber.Long-short-term memoryJ, Neural Computation, vol. 8, pp. 1735-1780, Dec. 1997, doi: 10.1162/neco.1997.9.8.1735.
29. W. Zhang et al.LSTM-based analysis of industrial IoT equipment, IEEE Access, vol. 6, pp. 23551-23560, 2018,doi: 10.1109/ACCESS.2018.2825538.
30. W. Jiang, N. Zhang, X. Xue, Y. Xu, J. Zhou and X. Wang.Intelligent deep learning method for forecasting the health evolution trend of aero-engine with dispersion entropy-based multi-scale series aggregation and LSTM neural network, IEEE Access, vol. 8, pp. 34361, 2020, doi:10.1109/ACCESS.2020.2974190.
31. M. Qiao, S. Yan, X. Tang and C. Xu. Deep convolutional and LSTM recurrent neural networks for rolling bearing fault diagnosis under strong noise and variable loads, IEEE Access, vol. 8, pp. 66257-66269, 2020, doi: 10.1109/ACCESS.2020.2985617.
32. R. J. Williams and D. Zipser. Gradient-based learning algorithms for recurrent networks and their computational complexity, back-propagation. theory, architectures and applications, 2nd ed., Hillsdale, MI, USA. L. Erlbaum Associates

Inc., 1995, pp. 433–486.

33. F. A. Gers, J. Schmidhuber, F. Cummins. Learning to forget. continual prediction with LSTM, Neural Computation, vol. 10, No. 10, Oct. 2000, pp.2451-2471,doi: 10.1162/089976600300015015.
34. T. Chen and C. Guestrin, XGBoost. A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, CA, USA, 2016, pp. 785–794.
35. D. Zhang, L. Qian, B. Mao, C. Huang, B. Huang and Y. Si.A data-driven design for fault detection of wind turbines Using random forests and XGboost, IEEE Access, vol. 6, pp. 21020-21031,2018,doi.10.1109/ACCESS.2018.2818678.
36. M. Chen, Q. Liu, S. Chen, Y. Liu, C. Zhang and R. Liu.XGBoost-based algorithm interpretation and application on post-fault transient stability status prediction of power system, IEEE Access, vol. 7, pp. 13149-13158, 2019, doi: 10.1109/ACCESS.2019.2893448.
37. Y. Feng, L. Liu and J. Shu.A link quality prediction method for wireless sensor networks based on XGBoost, IEEE Access, vol. 7, pp. 155229-155241, 2019, doi: 10.1109/ACCESS.2019.2949612.
38. J. H. Friedman.Greedy function approximation. a gradient boosting machine, The Annals of Statistics, vol. 29, no. 5, 2001, pp. 1189–1232, doi:10.1214/aos/1013203451.
39. M. J. Rovine and Alexander von Eye.A 14th way to look at a correlation coefficient. correlation as the proportion of matches, Am. Stat., vol. 51, no. 1, Feb. 1997,pp.42~6,doi:abs/10.1080/00031305.1997.10473586.
40. D. N. Reshef, Y. A. Reshef, H. K. Finucane, S. R. Grossman, G. McVean, P. J. Turnbaugh, E. S. Lander, M. Mitzenmacher and P. C. Sabeti. Detecting novel associations in large data sets, Science, vol. 334, no. 6062, pp. 1518-1524, doi: 10.1126/science.1205438.a

41. D. P. Kingma and J. Ba.Adam. A method for stochastic optimization, ICLR, San Diego, CA, USA, 2015.
42. X. Glorot and Y. Bengio.Understanding of the training deep feed forward neural networks, Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, Sardinia, Italy, 2010, pp. 9:249-256.
43. S. Wang, W. Guo, K. Zeng and G. Zhang. Characterization of automotive brake discs with laser-machined surfaces, Automot. Innov. 2, no. 3, pp.190–200 , 2019, doi: 10.1007/s42154-019-00068-y.