scientific reports

OPEN

Modeling and prediction for diesel performance based on deep neural network combined with virtual sample

Hainan Zheng[1], Honggen Zhou[1], Chao Kang[1]✉, Zan Liu[1], Zhenhuan Dou[2], Jinfeng Liu[1], Bingqiang Li[1] & Yu Chen[1]

The performance models are the critical step for condition monitoring and fault diagnosis of diesel engines, and are an important bridge to describe the link between input parameters and targets. Large-scale experimental methods with higher economic costs are often adopted to construct accurate performance models. To ensure the accuracy of the model and reduce the cost of the test, a novel method for modeling the performances of marine diesel engine is proposed based on deep neural network method coupled with virtual sample generation technology. Firstly, according to the practical experience, the four parameters including speed, power, lubricating oil temperature and pressure are selected as the input factors for establishing the performance models. Besides, brake specific fuel consumption, vibration and noise are adopted to assess the status of marine diesel engine. Secondly, small sample experiments for diesel engine are performed under multiple working conditions. Moreover, the experimental sample data are diffused for obtaining valid extended data based on virtual sample generation technology. Then, the performance models are established using the deep neural network method, in which the diffusion data set is adopted to reduce the cost of testing. Finally, the accuracy of the developed model is verified through experiment, and the parametric effects on performances are discussed. The results indicate that the overall prediction accuracy is more than 93%. Moreover, power is the key factor affecting brake specific fuel consumption with a weighting of 30% of the four input factors. While speed is the key factor affecting vibration and noise with a weighting of 30% and 30.5%, respectively.

Diesel engines have been widely used in ships and automobiles due to their large power range, high thermal efficiency, and low fuel consumption rate[1-3]. The status of the diesel engine is intrinsically linked to the controlled parameters during the running process of the diesel engine. To understand the relationships between diesel engine performance and control parameters, many studies have focused on modeling methods based on large sample experiments, theory, and simulation. Recently developments have attracted more scholars to concentrate their studies on this field[4-6].

Nowadays, numerical simulation method has been applied to deal with complex nonlinear problems of diesel engine. As for the performance prediction of diesel engine, Nahim et al.[7] presented a numerical simulation model of a marine diesel engine based on physical, semi-physical, mathematical and thermodynamic equations. This model can accurately predict the pressure, temperature, and emissions of diesel engines. Gosala et al.[8] used a phase-angle approach to build a vibration prediction model. It is demonstrated that the weighted phase-angle approach can accurately and quickly predict the frequencies and relative amplitudes of the vibration. The proposed model can be further used for online, real-time selection of ignition modes under steady-state operating conditions. In terms of parametric optimization of diesel engine, Muse et al.[9] simplified the in-cylinder combustion process of a diesel engine and developed a 1-dimensional large two-stroke low-speed diesel engine model. Diesel engine characteristics were analyzed by changing the model injection parameters, and optimization of injection characteristics and exhaust timing was achieved, resulting in a 9% reduction in NO_x emissions. Besides, based on the GT-SUITE, the simulation models of diesel engines were also established to predict and optimize fuel consumption, noise and emissions[10,11]. The optimization process was carried out through numerical

[1]School of Mechanical Engineering, Jiangsu University of Science and Technology, Zhenjiang, China. [2]Shaanxi Diesel Heavy Industry, LTD, Xi'an, China. ✉email: kaochao_017@163.com

No.	Innovation points
1	In order to reducing the cost of experimental modeling, the DNN coupled with VSG technology is proposed to establish the accurate performance models of the diesel. The models achieved in this paper could reflect the relationship of parameters using small experiment samples
2	The influence law of diesel engine parameters on the performance is obtained based on the proposed models. Moreover, the parametric sensitivity is determined based on impact factors analysis. These would provide a guideline for engineer to optimize and assess performance in practice

Table 1. Innovation points.

simulation with suitable objective functions aimed at minimizing BSFC while not exceeding the target NO_x emission levels. Recently, the numerical simulation method has also been used for fault diagnosis of diesel engine. Rubio et al.[12] built a one-dimensional thermodynamic model using AVLBoost software. This model was able to effectively simulate 15 typical thermodynamic faults such as turbine failure, exhaust manifold leakage and intake valve seat failure, which provided information for the development of a diesel engine failure simulator. Overall, the numerical simulation approach reduces the amount of expensive experimental trials required to evaluate the performance of the engines. However, as diesel engine systems become increasingly complex, the time for model calculation and optimization increases. This approach becomes gradually unsuitable for large marine diesel engine modeling.

With the development of algorithm theory, intelligent algorithms have been applied to diesel engine performance studies. Aghbashlo et al.[13] designed a novel method to estimate the exergetic performance of a DI diesel engine based on extreme learning machine with a wavelet transform algorithm. The method allows estimating, optimizing, and controlling the operating performance of a diesel engine in real time under known engine operating conditions and fuel characteristics. Similarly, an adaptive neuro-fuzzy inference system tuned by particle swarm algorithm was used to predict diesel engine performance and emissions by Atarod et al.[14]. The effective prediction of CO_2, CO, and NO_x emissions was achieved after the continuation of the modeling system based on experimental data, which provides a reference for the next optimization of diesel engines. The intelligent algorithm could predict the diesel engine operating parameters very well. Moreover, it also provides ideas and directions to optimize the diesel engine operating parameters and fuel composition[15]. Shin et al.[16] combined the deep neural network and Bayesian method to optimizing the diesel engine parameters and predicting the NOx transient emission, which enhanced the model stability and accuracy significantly. Dinesha et al.[17] incorporated cerium oxide (CeO_2) nanoparticles into diesel for effectively reducing diesel emissions and improving diesel engine performance. Besides, intelligent algorithms are also applied for the fault diagnosis. Zhang et al.[18] proposed a convolutional neural network-based (CNN) method for diesel engine misfire fault diagnosis. The results showed that the CNN method can accurately detect a complete misfire in one or two cylinders when the diesel engine is operating under steady-state conditions.

From the various literature, it can be observed that many modeling approaches were proposed to discuss diesel engine performance prediction and fault diagnosis. Much research had focused on software simulations or intelligent algorithms based on large samples. However, models obtained by simulation methods often lack the necessary experimental information, resulting in models that fail to fully reflect the effects introduced by interfering factors. Besides, extensive experimental information is necessary to obtained accurate performance models using intelligent algorithms. Therefore, it is necessary to propose a modeling approach that includes a small amount of experimental information to construct an accurate model.

In this paper, a novel method is proposed to establish the performance models of marine diesel engines based on the Deep Neural Network (DNN) method coupled with Virtual Sample Generation (VSG). First, the predicted performances are analyzed, and an experiment is designed. Then, the diffusion trend of experimental data is mined and analyzed based on VSG technology. Finally, the marine diesel engine performance model is established by deep neural network for guaranteeing the accuracy of prediction. Here, the innovation points of this paper are summarized in Table 1 below. In addition, in order to better understand the article, Table 2 of abbreviations and symbol nations of terms has been added.

Experimental set-up and design

Experimental equipment.   In this paper, the experimental tests are conducted using a 20-cylinder, four-stroke, water-cooled diesel engine, which is provided by Shaanxi Diesel Heavy Industry, LTD. The detailed parameters of experimental equipment are shown in Table 3. During the experimental process, the hydro dynamometer is mounted on the diesel engine to provide different load conditions. Meanwhile, the diesel engine speed is adjusted to achieve different test conditions by controlling the throttle position lever.

Experimental design.   It is well known that the performances of diesel engines are sensitive to power and speed, which could be used for reflecting the operating conditions[19]. Moreover, the oil pressure and temperature are also important to analyze the combustion of the diesel engine[20]. Therefore, power, speed, oil temperature, and pressure are selected as input factors to study parametric effects on the diesel engine in this paper. Meanwhile, three critical parameters including brake specific fuel consumption, vibration, and noise are adopted as the outputs. In the design of experiment, eleven working conditions based on the propulsion characteristic line of the diesel engine are designed to completely investigate the status of the diesel engine during the entire operation. According to the practice experience, the diesel engine speed varies from 600 to 1500 rpm, the lubricating

Abbreviation	
DNN	Deep neural network
VSG	Virtual sample generation
TSA	Trend similarities among attributes
IQR	Interquartile range
MIV	Mean influence value
BSFC	Brake specific fuel consumption (g/kWh)
Symbol notations	
a_w	Weighted acceleration
T	Measurement time
a_{total}	Total vibration acceleration
L, U	Lower limit and upper limit
MF	Sample distribution value
$g(i)_{p, q}$	Sample trend value
$S_{p,q}$	Strength of correlation
v	Evaluation symbol
$\theta_{p, q}$	Offset
v_{Xp}^-, v_{Xp}^+	Possible virtual value
Ri_{max}, Ri_{min}	Predicted results after expanding samples

Table 2. Abbreviation and symbol notations.

| Parameter	Value
Engine type	4 cycle-20 cylinder
Bore and stroke	230 mm × 230 mm
Type of ignition	Compression ignition
Total working volume	191.2 L
Compression ratio	12:1
Rated power	3750 kW
Rated speed	1455 rpm
Cooling system	Water cooling

Table 3. Marine diesel engine specifications.

No.	Speed (rpm)	Power (kW)	Lubricating oil temperature (°C)	Lubricating oil pressure (MPa)
1	600	260	77	0.41
2	750	500	79	0.48
3	920	940	84	0.56
4	1150	1875	87	0.55
5	1320	2810	88	0.70
6	1455	3750	92	0.69
7	1455	3750	92	0.69
8	1455	3750	92	0.68
9	1455	3750	92	0.67
10	1455	3990	92	0.66
11	1500	4380	94	0.65

Table 4. Experimental design.

oil temperature range is 77–95 °C, the power range covers 260–4380 kW and lubricating oil temperature range covers 0.3–0.7 MPa. The designed tests are shown in Table 4.

Experimental measurement.   As shown in Fig. 1, the 1A307E vibration sensor is attached to the cylinder head of the diesel engine for measuring the vibration signal. The noise signal is obtained by a BSWA 308 sound
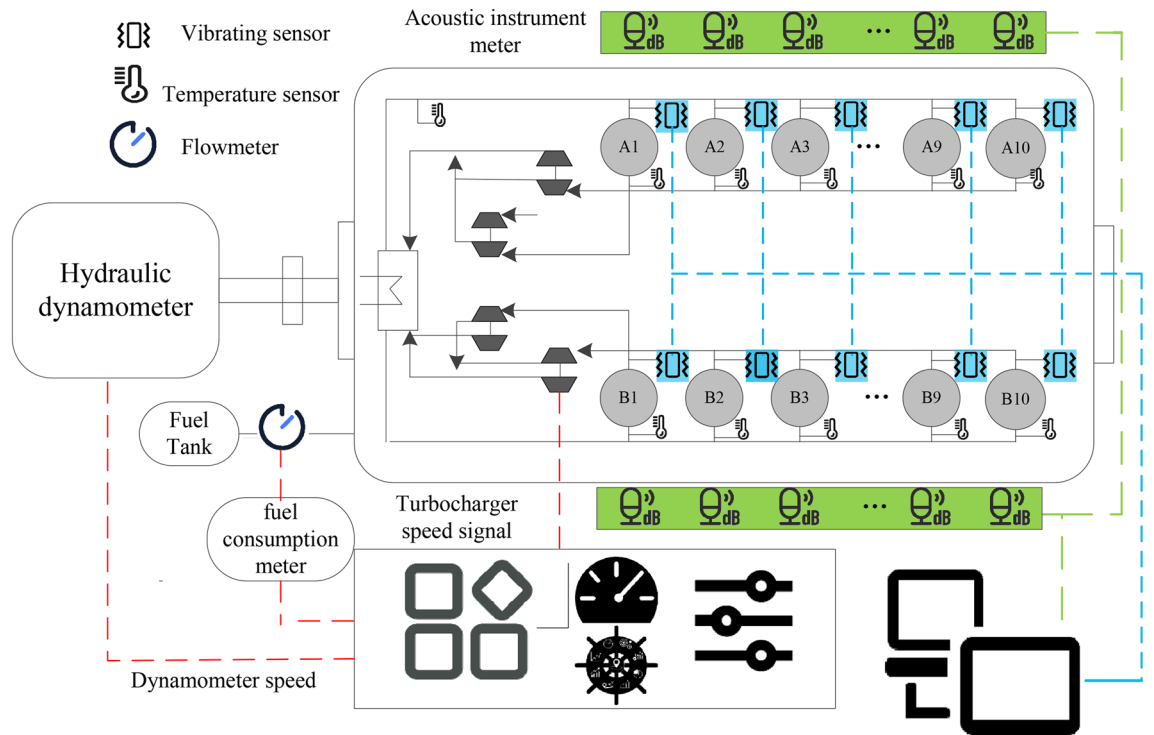
Figure 1. The detailed schematic of the experimental set up.

level meter at 1 m from the cylinder head. The fuel consumption rate is collected by the fuel consumption meter. During the experiment, the speed and load are controlled to achieve the expected working conditions. Then the speed, power, lubricating oil temperature and pressure are measured by the speed sensor, hydraulic dynamometer, temperature sensor and pressure sensor, respectively. The experimental details of the diesel engine are shown in Fig. 2.

Figure 3 shows the noise and vibration results measured in the experiment process. These results are continuous and vary too fast so that it is hard to express the continuous values through a certain model. Thus, the root mean square value is proposed to characterize the state of the diesel engine. The root mean square can be calculated as follows[21],

$$a_w = \sqrt{\frac{1}{T} \int_0^T a_w^2(t)dt}, \tag{1}$$

where a_w represents the weighted acceleration and T represents measurement time.

Due to the presence of vibration in three directions, the total vibration acceleration is used to evaluate the vibration state of the diesel engine. The acceleration equation is presented in Eq. (2),

$$a_{total} = \sqrt{a_x^2 + a_y^2 + a_z^2}, \tag{2}$$

where total vibration acceleration (a_{total}) is the value to show the combined acceleration of vertical (a_x), lateral (a_y) and longitudinal (a_z).

Uncertainty analysis. In laboratory experiments, uncertainty analysis deals with evaluating the uncertainty in any measurements. It allows the estimation of the numerical value of a physical variable and how it is affected by errors due to instrumentation. In the present work, the uncertainty of a dependent variable is calculated using errors involved in measuring independent parameters such as power, speed, and lubricating oil temperature. The uncertainty value can be derived by Eq. (3)[17,22].

$$W_R = \left(\left[\frac{\partial R}{\partial x_1} w_1 \right]^2 + \left[\frac{\partial R}{\partial x_2} w_2 \right]^2 + \cdots + \left[\frac{\partial R}{\partial x_1} w_n \right]^2 \right)^{1/2}. \tag{3}$$

The uncertainty values for various parameters are listed in Table 5
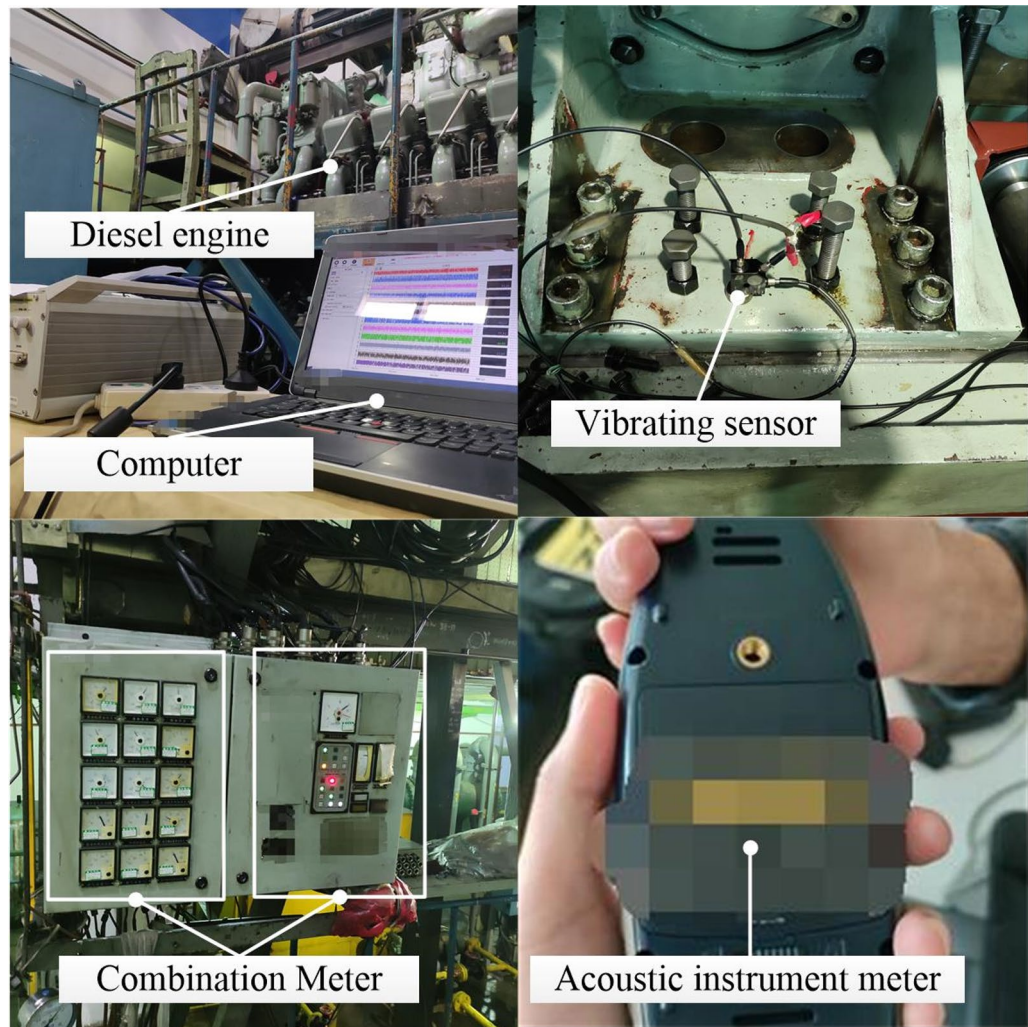
Figure 2. The performances measurement for the multicylinder engine in experiment.

Modeling method

Compared with other algorithms, the DNN method has a strong advantage for modeling nonlinear complex systems. However, the DNN algorithm requires a large amount of sample data to ensure accuracy of model[23]. To reduce the cost of the diesel experiment, virtual sample generation technology is adopted to diffuse the experimental samples. The virtual sample generation technique is also a current sample diffusion method with high accuracy, and has been used in conjunction with various intelligent algorithms with good results. To clearly illustrate the proposed approach, the flowchart of DNN coupled with VSG is displayed in Fig. 4. Firstly, the target data is achieved through conducting the experiments. Then the sample distribution is determined by dealing with experimental samples. By comparing the magnitude of the values, calculate the trend similarities among attributes (TSA), and the virtual sample data is generated further. Finally, the performance model is obtained by training and validating the sample based on DNN method.

Virtual sample generation. *Build sample distribution.* In order to achieve the virtual sample, the sample distribution should be firstly determined according to the experimental data. The procedure is given as follows,

Step 1: The test data are collected and divided into input parameters and output parameters of the diesel engine. As for the diesel engine system, the input parameters include speed, power, lubricating oil temperature and pressure. The output parameters are brake specific fuel consumption, total vibration acceleration and noise.

Step 2: Based on the small data sets obtained from the above experiments, seven sample domain boundaries are estimated. For obtaining the boundaries, the range should be achieved firstly. Here the interquartile range (IQR) is used to describe the range, which can be derived as follows,

$$IQR = Q_3 - Q_1, \tag{4}$$

where Q_1 is the first quartile of each sample set, and Q_3 is the third quartile of each sample set.
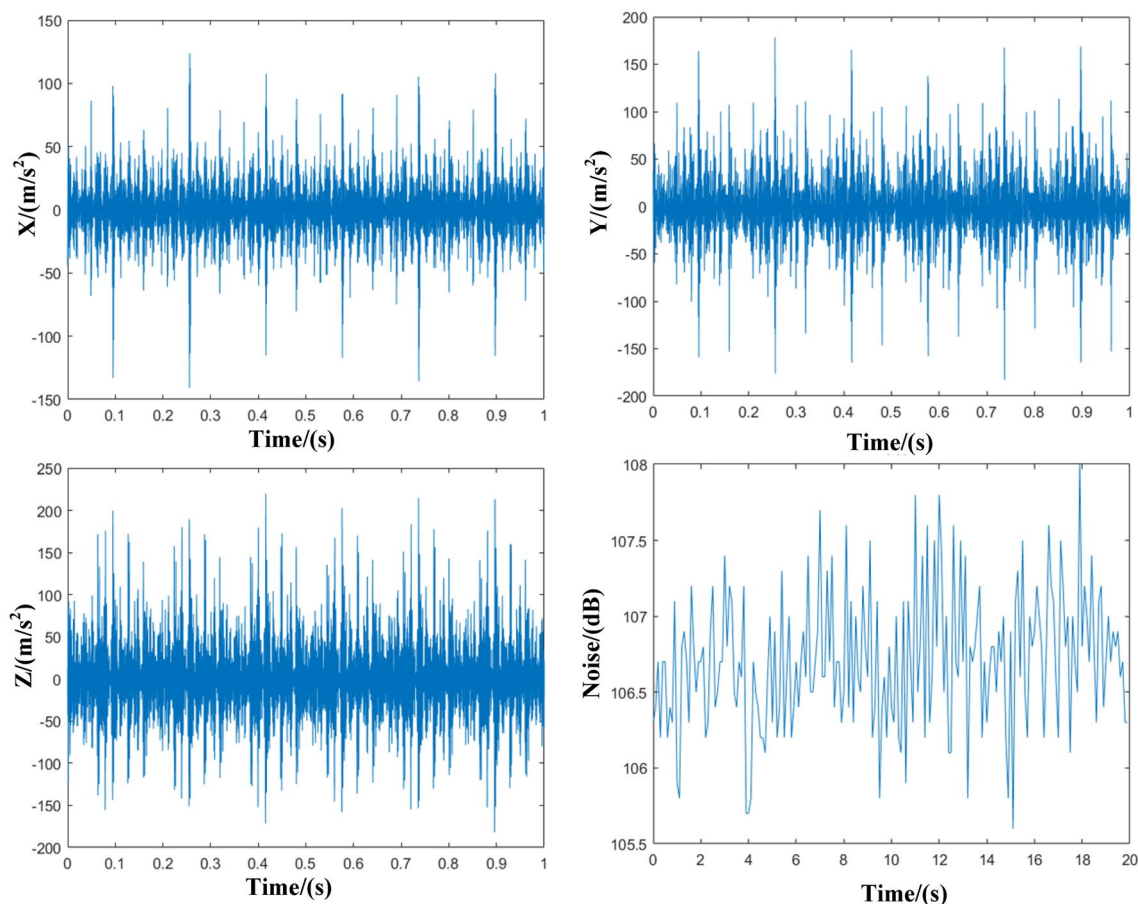
Figure 3. Measurement results of vibration and noise.

| Measured quantity	Percentage uncertainty
Speed	0.64
Power	1.3
Lubricating oil temperature	0.11
Lubricating oil pressure	0.36
BSFC	0.21
Vibration	0.37
Noise	0.17

Table 5. Uncertainty analysis of the measured parameters.

Step 3: The lower limit (L) and upper limit (U) of the sample domain boundary should be obtained. The reasonable bounds [L, U] can be calculated by the following Eqs. (5) and (6),

$$L = \begin{cases} Q_1 - 1.5 \times IQR, & L \leq \min \\ \min, & Q_1 - 1.5 \times IQR > \min, \end{cases} \tag{5}$$

$$U = \begin{cases} Q_3 + 1.5 \times IQR, & U \leq \max \\ \max, & Q_3 + 1.5 \times IQR < \max, \end{cases} \tag{6}$$

where min and max are the minimum and maximum values of the observations, respectively.

Step 4: Determine the sample distribution MF. When the domain bounds of observations are determined, a triangular MF based on L, Q_2 (Me, taken as the location center of sample range as depicted in Fig. 5), and U could represent the estimated sample distribution. MF is formulated as follows[24,25],
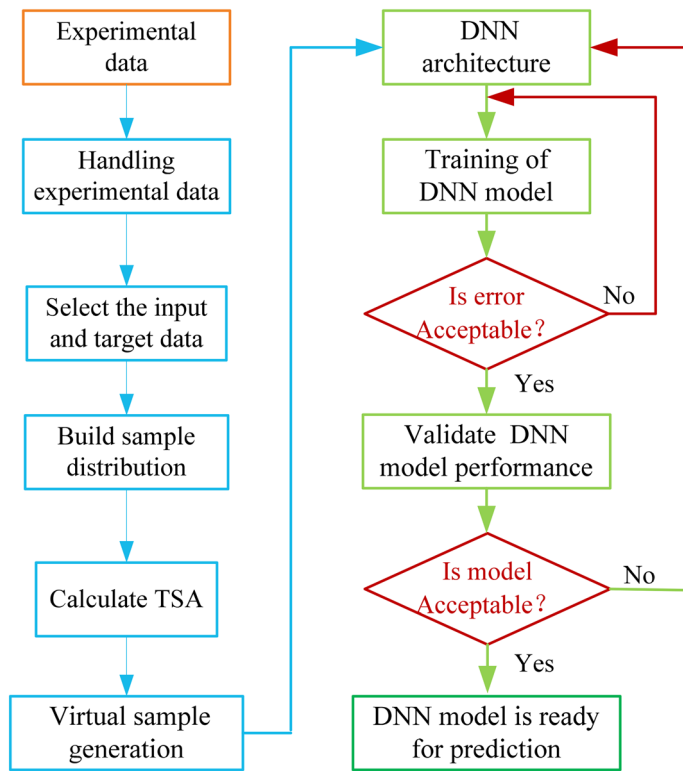
Figure 4. The flowchart of DNN method coupled with VSG.
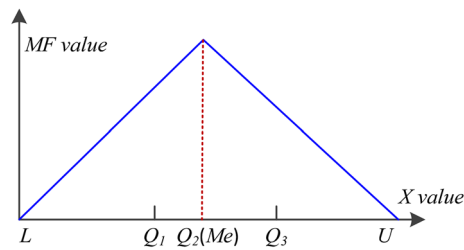
Figure 5. The shapes of attribute sample distribution.

$$MF(x) = \begin{cases} (x - L)/(Me - L), & L \le x < Me \\ (U - x)/(U - Me), & Me \le x < L \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

Based on the above formula and the measured small data set, the shape of the parameter distribution is plotted and shown in Fig. 5.

Calculate trend similarities among attributes. In this section, the trend similarities between attributes are measured according to a non-parametric process[24,25]. First, two parameters Xp and Xq are selected in two related attribute domains. Here Xp is the rotational speed, and Xq is an arbitrary parameter. Then the trend assessment function $g(i)_{p,q}$ of the i-th observation between X_p and X_q can be formulated as,

$$g(i)_{p,q} = \begin{cases} 1, & if\ (x_{p,i} - Me_{X_P})(x_{q,i} - Me_{X_q}) > 0 \\ 0, & if\ (x_{p,i} - Me_{X_P})(x_{q,i} - Me_{X_q}) = 0 \\ -1, & if\ (x_{p,i} - Me_{X_P})(x_{q,i} - Me_{X_q}) < 0. \end{cases} \tag{8}$$

The strength of the trend similarity between X_p and X_q is derived as the average of all available observations, as shown in Eq. (9).

7

Relation strength	Absolute value of $S(|S|)$
Low	[0,0.3)
Medium	[0.3,0.7)
High	[0.7,1]

Table 6. The relation strengths of trend similarities among attributes.

$$S_{p,q} = \frac{1}{n} \sum_{i=1}^{n} g(i)_{p,q}, \quad p,q \in \{1,2,\ldots,m\}, p \neq q. \tag{9}$$

Based on the experimental data, the values of $|S_{p,q}|$ are obtained as shown in Table 6. Here, the larger this value is, the stronger the correlation between Xp and Xq.

Virtual sample generation. Based on the value $g(i)_{p,q}$ obtained above, a value interval is projected to produce a suitable virtual value. The specific steps are given as follows[24,25]:

Step 1: Determine the value of v_{X_p}, and then produce v_{X_q}. Randomly select one temporary value (tv) from U (L_{X_p}, U_{X_p}), and then calculate the $MF_{X_p}(tv)$. Choose a random seed (rs) from $U(0,1)$ to assess whether tv can be kept as a suitable virtual value v_{X_p}.

Step 2: The cumulative distribution function value $F(rs)$ of the uniform distribution represents the cumulative probability of rs. The probability that rs is lower than $MF_{X_p}(tv)$ is the possibility of the value of $MF_{X_p}(tv)$ itself occurring. When rs is lower than $MF_{X_p}(tv)$, tv will thus be kept as v_{X_p}, otherwise, tv will be discarded. Therefore, if $MF_{X_p}(tv)$ is larger, tv will have a higher probability of being v_{X_p}. The evaluation criteria are Eq. (10):

$$v = tv, \quad if\ rs \leq MF_{X_p}(tv). \tag{10}$$

Select the v_{X_p} value that meets the criteria, and further calculate the offset as given in Eq. (11)

$$\theta_{p,q} = -0.8 \times |S_{p,q}| + 0.9. \tag{11}$$

The interval bounds $\left[v_{X_p}^-, v_{X_p}^+ \right]$ are given in Eqs. (12) and (13)

$$v_{X_p}^- = \begin{cases} v_{X_p} - \theta_{p,q}\left(U_{X_p} - L_{X_p}\right), & v_{X_p}^- \geq L_{X_p} \\ L_{X_p}, & v_{X_p} - \theta_{p,q}\left(U_{X_p} - L_{X_p}\right) < L_{X_p}, \end{cases} \tag{12}$$

$$v_{X_p}^+ = \begin{cases} v_{X_p} + \theta_{p,q}\left(U_{X_p} - L_{X_p}\right), & v_{X_p}^+ \leq U_{X_p} \\ U_{X_p}, & v_{X_p} + \theta_{p,q}\left(U_{X_p} - L_{X_p}\right) > U_{X_p}. \end{cases} \tag{13}$$

Then substitute the value $v_{X_p}^-$ and $v_{X_p}^+$ into the following equation. The calculation of $MF(x)$ is given in Eq. (14):

$$v_{X_q} = \begin{cases} L_{X_q} + MF\left(v_{X_p}\right)\left(Me_{X_q} - L_{X_q}\right), & L_{X_q} \leq v_{X_q} \leq Me_{X_q} \\ U_{X_q} - MF\left(v_{X_p}\right)\left(U_{X_q} - Me_{X_q}\right), & Me_{X_q} \leq v_{X_q} \leq U_{X_q} \\ 0, & otherwise. \end{cases} \tag{14}$$

Step 3: Sample generates. Based on the above formula, the interval $\left[v_{X_p}^-, v_{X_p}^+ \right]$ is determined, and the random value v_{X_q} from the interval is selected. The schematic diagram is shown in Fig. 6. According to the above method, all dummy samples are generated in turn.

Deep neural network model. The deep neural network model is widely used to predict various application results with high accuracy. The basic structure of the deep neural network model comprises one input layer, one output layer, and more than two hidden layers. The DNN algorithm trains internal parameters such as weights and matrix biases according to relationships between input features and output results. The training procedure results in a higher accuracy comparing with the equation-based modeling[26,27]. In this paper, the deep neural network prediction model is built by the following steps.

Step 1: Input/output parameters. The input parameters of the network are the operating parameters of the diesel engine, which includes diesel engine speed, power, lubricating oil temperature and pressure. The output parameters of the network are brake specific fuel consumption, vibration, and noise in this paper.

Step 2: Normalization process. Since the neural network is a parallel processing system, the network weights are parallel in order of magnitude during its training and prediction. If the difference in the order of magnitude of the input/output parameters of the network is too large, the influence of a smaller order of magnitude parameter on the network weight may be masked by a larger order of magnitude parameter. This would cause a degrada-
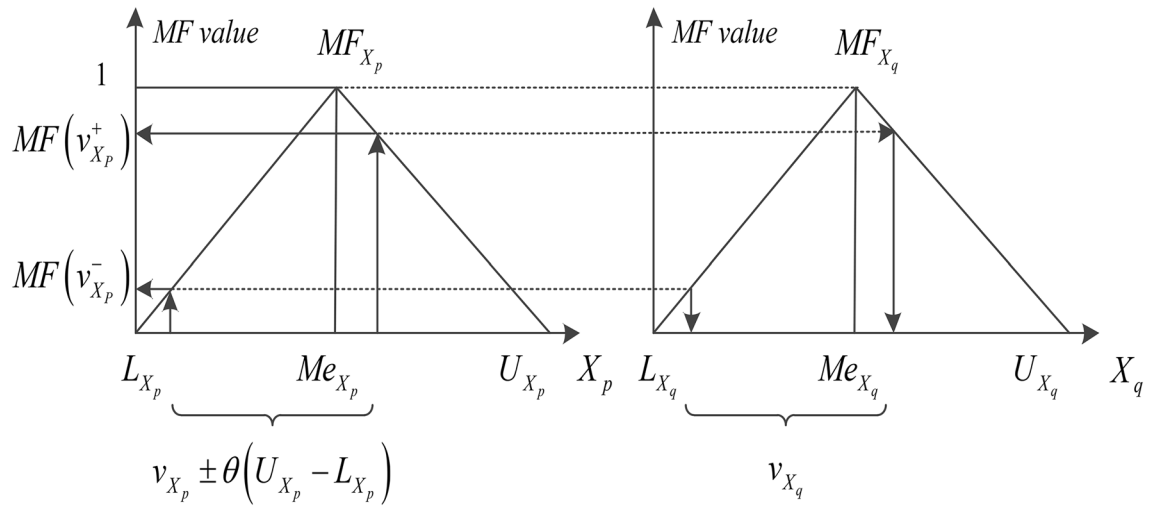
Figure 6. Estimating the possible attribute ranges based on TSA.

tion of the network prediction performance. Therefore, it is necessary to normalize the input/output parameters before training the network. The selected normalization function is given as,

$$\bar{x} = -1 + \frac{2(x - x_{\min})}{x_{\max} - x_{\min}}, \tag{15}$$

where x is the vector to be normalized, x_{max} is the maximum value of the sample, x_{min} is the minimum value of the sample, and $\bar{x}$ is the normalized vector.

Step 3: Network structure and network training. Compared with traditional feed-forward neural networks, DNNs have multiple implicit layer structures. Each hidden layer requires the input vector of the previous layer, and performs a nonlinear transformation using the activation function of the hidden layer. Then, the obtained vectors are passed from inputs to the next layer of neurons. Finally, the output is passed to the network through the iterate method. To determine the best network structure, the predicted performance under different numbers of neurons in the hidden layer is first compared. Then, the best network structure is selected for further optimization. In this paper, the training and testing are set to 3:1. The hidden layer is determined to be 4 layers, the number of nodes is 10, and the weight matrix W is expressed as,

$$W = \begin{bmatrix} W^1 & W^2 & W^3 & W^4 & W^5 \end{bmatrix}. \tag{16}$$

The connection weight matrix W^1 between the input layer and the hidden layer neurons, and the connection weight matrix W^5 between hidden layer and output layer neurons are as follows,

$$W^1 = \begin{bmatrix} W^1_{1,1} & W^1_{1,2} & W^1_{1,3} & W^1_{1,4} \\ W^1_{2,1} & W^1_{2,2} & W^1_{2,3} & W^1_{2,4} \\ \vdots & \vdots & \vdots & \vdots \\ W^1_{10,1} & W^1_{10,2} & W^1_{10,3} & W^1_{10,4} \end{bmatrix} \tag{17}$$

$$W^5 = \begin{bmatrix} W^5_{1,1} & W^5_{1,2} & \cdots & W^5_{1,10} \\ W^5_{2,1} & W^5_{2,2} & \cdots & W^5_{2,10} \\ W^5_{3,1} & W^5_{3,2} & \cdots & W^5_{3,10} \end{bmatrix}. \tag{18}$$

Result and discussion

Model accuracy analysis. *Correlation analysis.* Based on the large amount of data generated by VSG, the deep neural network model was trained and tested for prediction accuracy. The residual plots of the brake specific fuel consumption, vibration and noise prediction models are shown from Figs. 7, 8, 9. From the plots, it can be found that the coefficients of determination (R^2) between the predicted and sample values of fuel consumption rate, vibration and noise are 0.90, 0.94 and 0.91, respectively. That indicates a high correlation between predicted model and experiment data. Besides, the observed and predicted values of the three responses are concentrated around the zero-error line, which fully indicates that the DNN model has a good correlation.
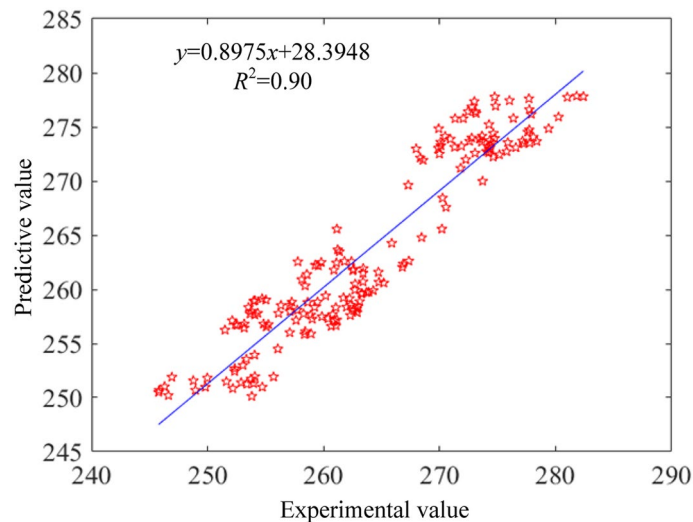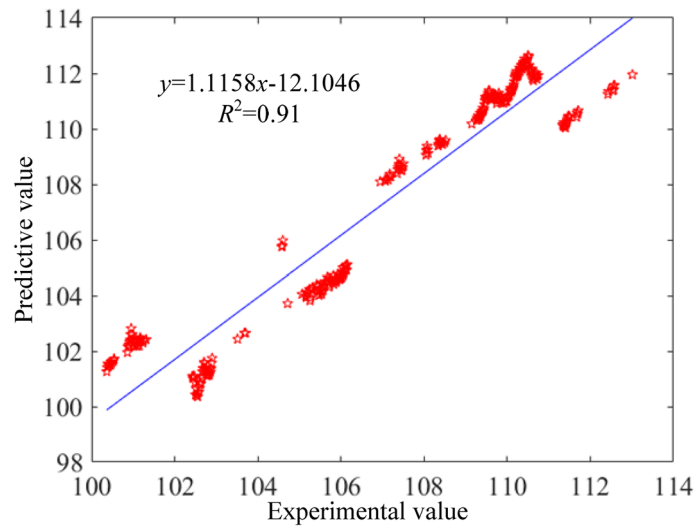
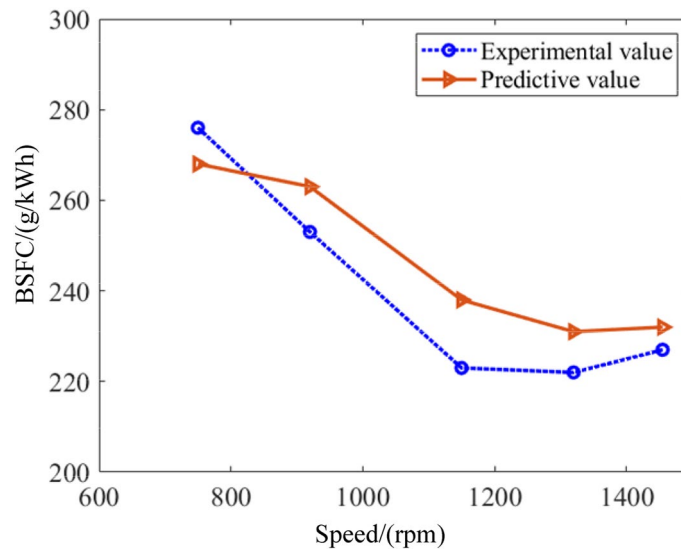**Figure 7.** Residual comparison of brake specific fuel consumption between predictive and experimental values.
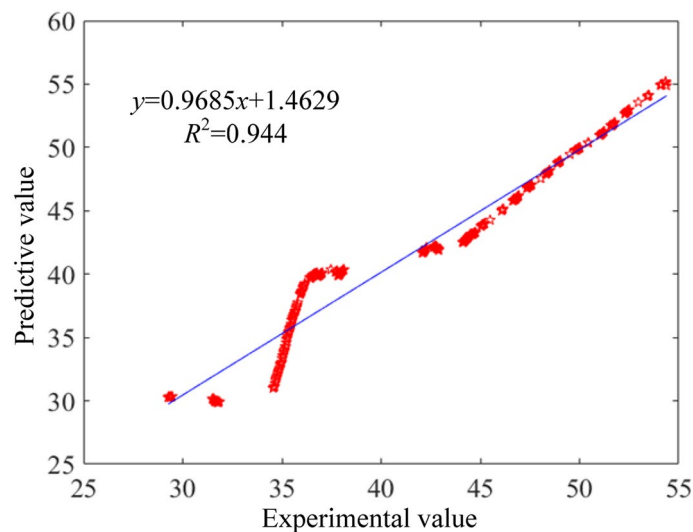
Figure 8. Residual comparison of vibration between predictive and experimental values.

Experimental verification. In order to further experimentally verify the reliability of the performance model proposed in this paper, the brake specific fuel consumption, total vibration acceleration, and noise are measured under five working conditions including the speed at 750 rpm, 920 rpm, 1150 rpm, 1320 rpm and 1455 rpm. Figure 10 describes the comparison of the brake specific fuel consumption between experimental and predictive values. As can been seen, the maximum error occurs at 750 rpm, and the error between the predicted and experimental values is close to 7%. When the rotational speed rises to 1300 rpm, the error between the predicted and experimental values is close to 3.9%. Moreover, the trend of experimental and predicted values is generally consistent. Figure 11 depicts the results of vibration. It is found that the trends of the experimental and predicted values of vibration are similar. When the speed increases to 1150 rpm, the error between the predicted and experimental values of the total vibration acceleration is the largest, close to 7%. While the speed reaches 1450 rpm, the error is the smallest, close to 1%. The noise of marine diesel is demonstrated as shown in Fig. 12. It is observed that the predicted results roughly overlap with the experimental values. It means the prediction model is accurate. The maximum error occurs at a speed of 750 rpm, which is close to 2%. The minimum error occurs at a speed of 1150 rpm, which is close to 1%. This result shows the model is very accurate and fully compliant with engineering requirements.

Effect of input parameters. For the marine diesel engines, the main reason of vibration and noise are piston movement and parts wear. Rotational speed and power are closely related to piston motion. Higher speed and power lead to faster piston motion and increased inertia forces such as crankshaft speed. This leads to high fuel consumption, vibration and noise. As for the oil pressure and temperature, they are related to the wear of
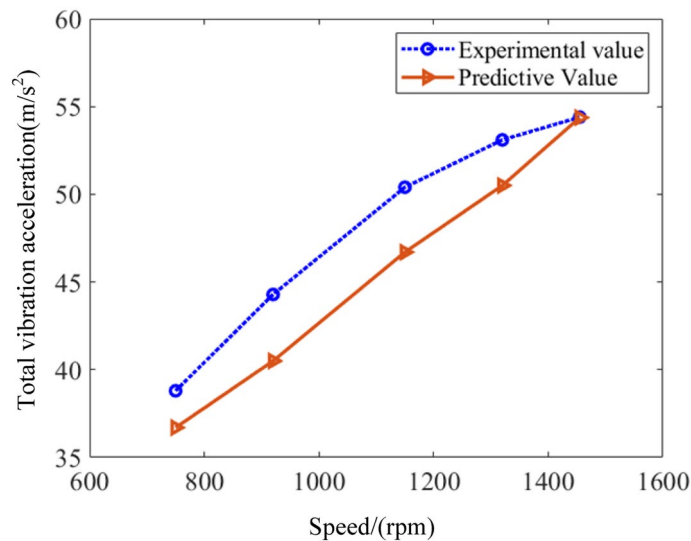
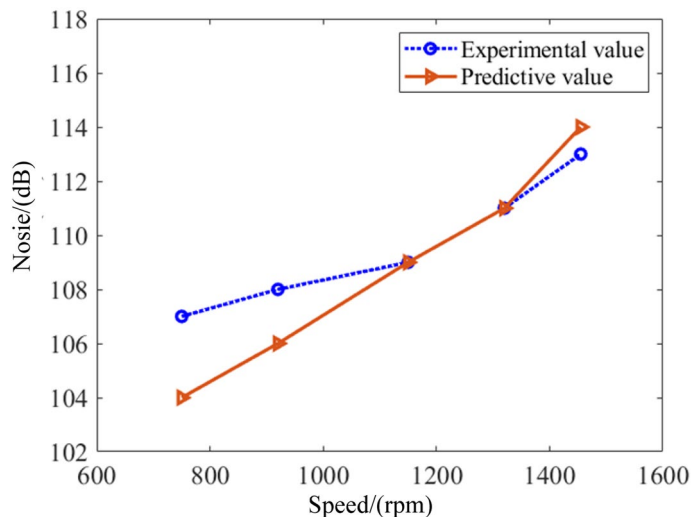**Figure 9.** Residual comparison of noise between predictive and experimental values.

Figure 10. Comparison of brake specific fuel consumption between predictive and experimental values.

parts. When the lubricating oil pressure and temperature are too low, the lubricating oil will not flow smoothly and the diesel engine will not work properly. High oil pressure and temperature will make the oil pump parts overload, the friction surface is not easy to form oil film, resulting in unreliable lubrication, increase parts wear, oil consumption increases. This causes an increase of vibration, noise, and fuel consumption. Therefore, in order to reveal the influence of diesel engine operating parameters on its performance, the diesel engine mechanism is discussed below in conjunction with the prediction model.

Brake specific fuel consumption. Brake specific fuel consumption is one of the important economic index of diesel engines. In order to improve the power and economy of diesel engines, the effect of parameters on brake specific fuel consumption needs to be studied. Figure 13 shows the effect of input parameters on brake specific fuel consumption. It is found that the brake specific fuel consumption decreases moderately for all parameters as the input parameters increase. That is due to the losses of the low mechanical efficiency and increased leakage at the start of the diesel engine, which results in higher brake specific fuel consumption values. The trend in brake specific fuel consumption is generally consistent with that in the previous literature[28]. However, the brake specific fuel consumption drops slowly after 1000 rpm. It can be explained that the brake specific fuel consumption is relatively low as the diesel engine gradually reach near rated operating conditions. When the diesel engine speed reaches 1350 rpm, fuel consumption is minimal. Subsequently, the fuel consumption increases and shows an upward trend.

**Figure 11.** Comparison of vibration between predictive and experimental values.

Figure 12. Comparison of noise between predictive and experimental values.

Vibration of the engine. Figure 14 shows the effect of all input parameters on the vibration based on the model data. From the Fig. 14, it can be observed that the total vibration is smaller during the starting phase of the diesel engine, with a minimum value of 29 m/s². As the diesel engine runs to the rated condition, the vibration increases to the maximum value of 54.8 m/s². Besides, the output parameters are overall positively correlated with the variation of vibration. This is attributed to the gradual increase in lubricant temperature and pressure as the speed of the diesel engine increases. Moreover, the torque become progressively larger with the rise of the inertia force. This situation makes the piston motion more violent currently, which is the main vibration source of the diesel engine. The trend in vibration was also found in the Ref.[29].

Sound pressure level of the engine. The variation of noise with the input parameters is shown in Fig. 15. It can be seen from the graph that the sound pressure level of diesel engine rises with increasing speed, power, and lubricating oil temperature. The sound pressure level is basically the same as that of vibration. This trend in brake specific fuel consumption is also validated in Ref.[21]. Meanwhile, it is found that the minimum noise is 100 dB which presents at 460 rpm of speed. The maximum value is 113 dB at the maximum speed. This is attributed to the fact that the decay of sound pressure level may be related to the vibration decay of the diesel engine.

Impact factors analysis. Although the parametric effect on the performance of diesel engines has been discussed, the rank of input parameters on performances is still ambiguous. In this paper, the mean influence value
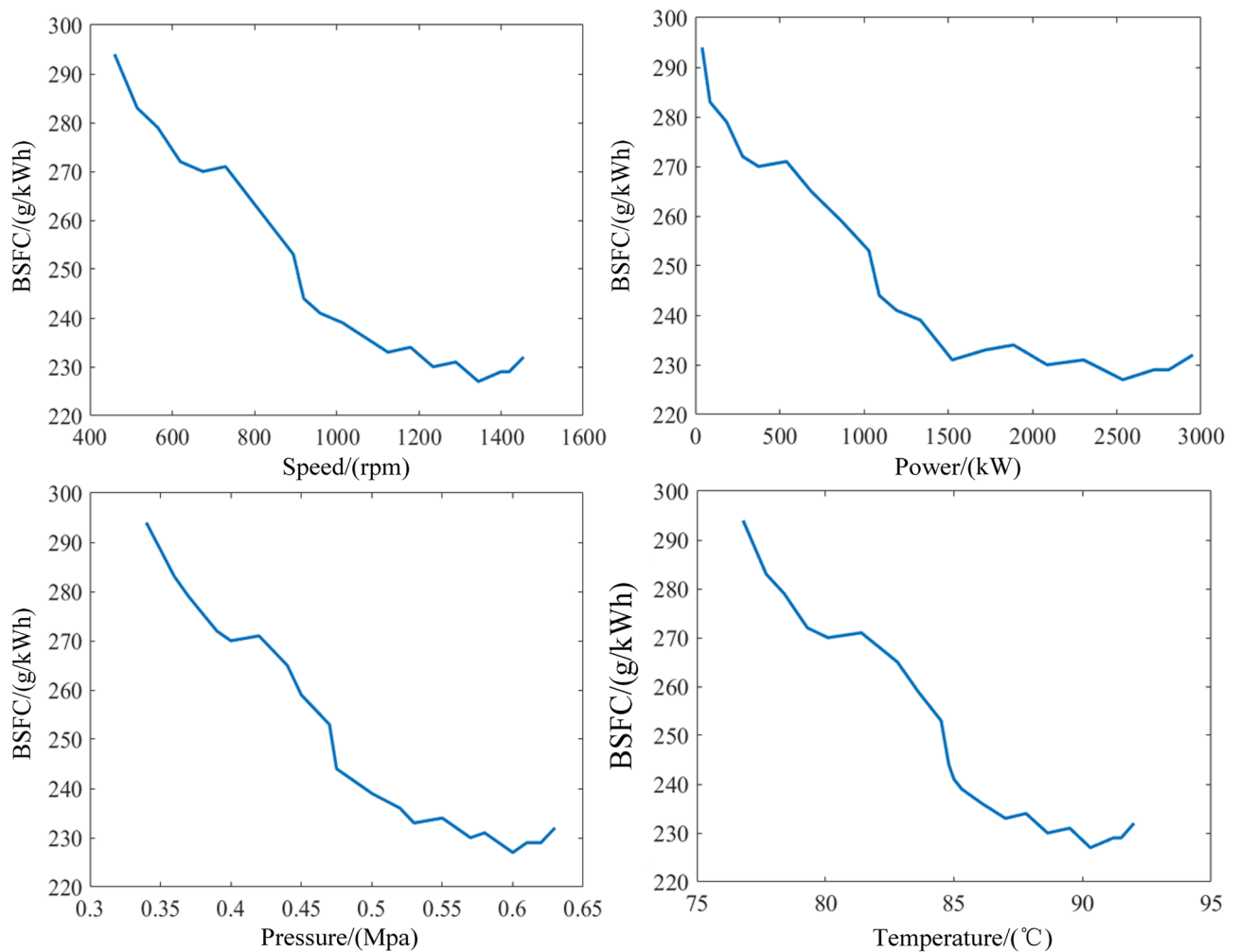
Figure 13. Influence of input parameters on brake specific fuel consumption.

(MIV) is used to determine the magnitude of the influence of input neurons on output neurons. The MIV_i can be derived as following[30,31],

$$MIV_i = abs(mean(Ri_{\max} - Ri_{\min})), \tag{19}$$

where Ri_{max} and Ri_{min} are new samples by adding and subtracted by KI on the basis of the original samples. Here KI represents the increment of input parameters.

In this study, four adjustment rates of MIV are set, which include K1 = 5%, K2 = 10%, K3 = 15%, and K4 = 20%, respectively. For each adjustment rate, multiple test trials are conducted to achieve the mean value. Finally, the |MIV| of each variable is calculated. The detailed results are listed in Table 7.

Figure 16 shows the |MIV| weight ratio of input parameters on performances of diesel engine. It can be seen from Fig. 16 that the weight ratios of speed on vibration and noise are relatively larger, followed by power. While the weight ratios of lubricating oil temperature and pressure on vibration and noise are relatively small. As for the brake specific fuel consumption, the power takes on a slightly significant weighting, followed by the speed. While the parameters that have less influence on brake specific fuel consumption are the lubricating oil pressure and temperature. Besides, the |MIV| standard deviation of the four input variables is high that indicates the performance indexes are more sensitive to all of input parameters. This also illustrates that brake specific fuel consumption, vibration, and noise are neither affected by a single factor, nor by the combination of multi-parameter.

Conclusions and future directions

In this study, a hybrid approach combining DNN and VSG is used to achieve an accurate performance model for marine diesel engines. The effects of diesel engine parameters on performances are analyzed, which could provide a guide for engineers in performance assessment and fault diagnosis. The main conclusions of this study can be summarized as follows.

(1) The sample data diffused by the VSG method keeps a high accuracy due to an error less than 7%. This indicates that the VSG method is capable of better diffusion of the experimental data. Further, the per-
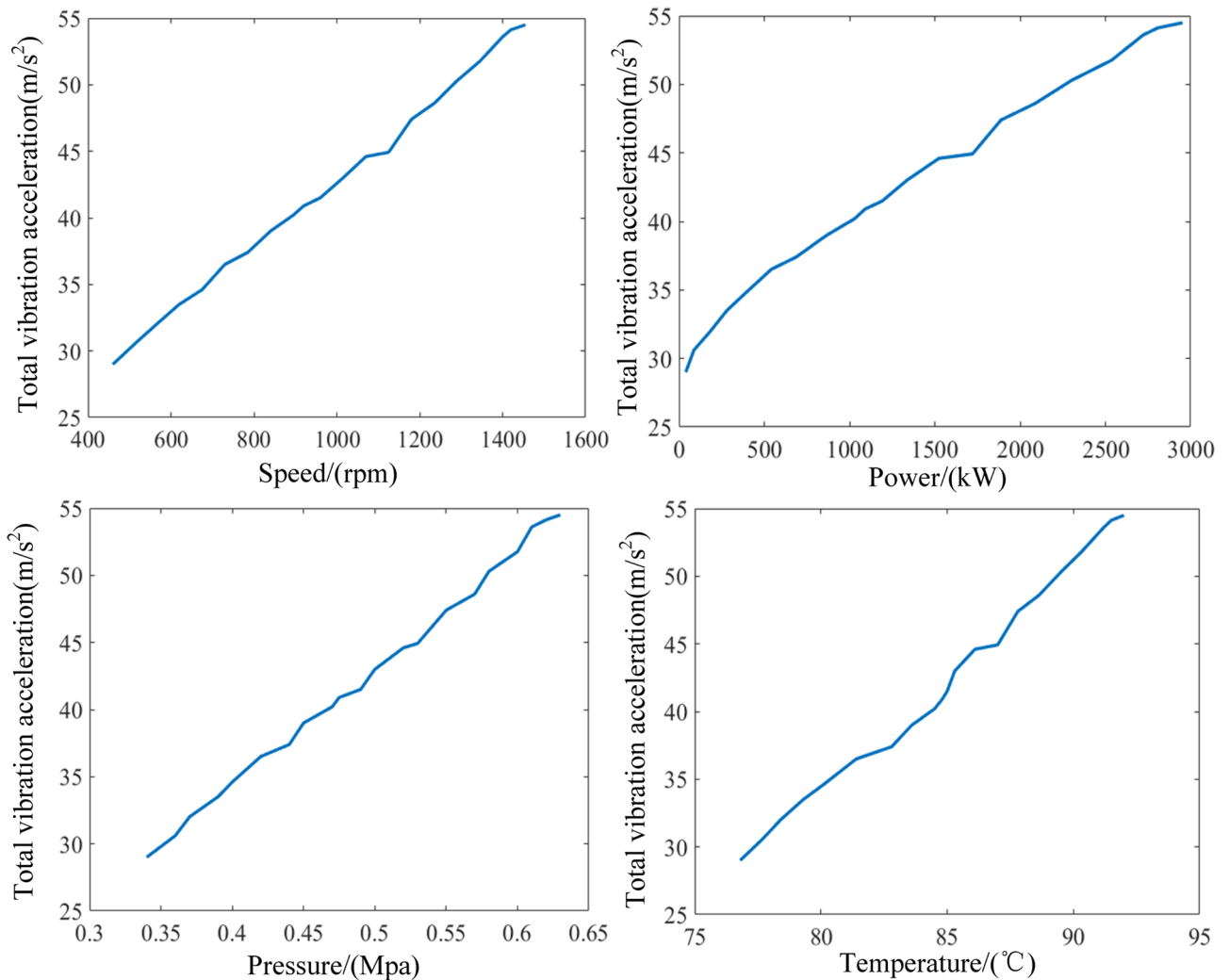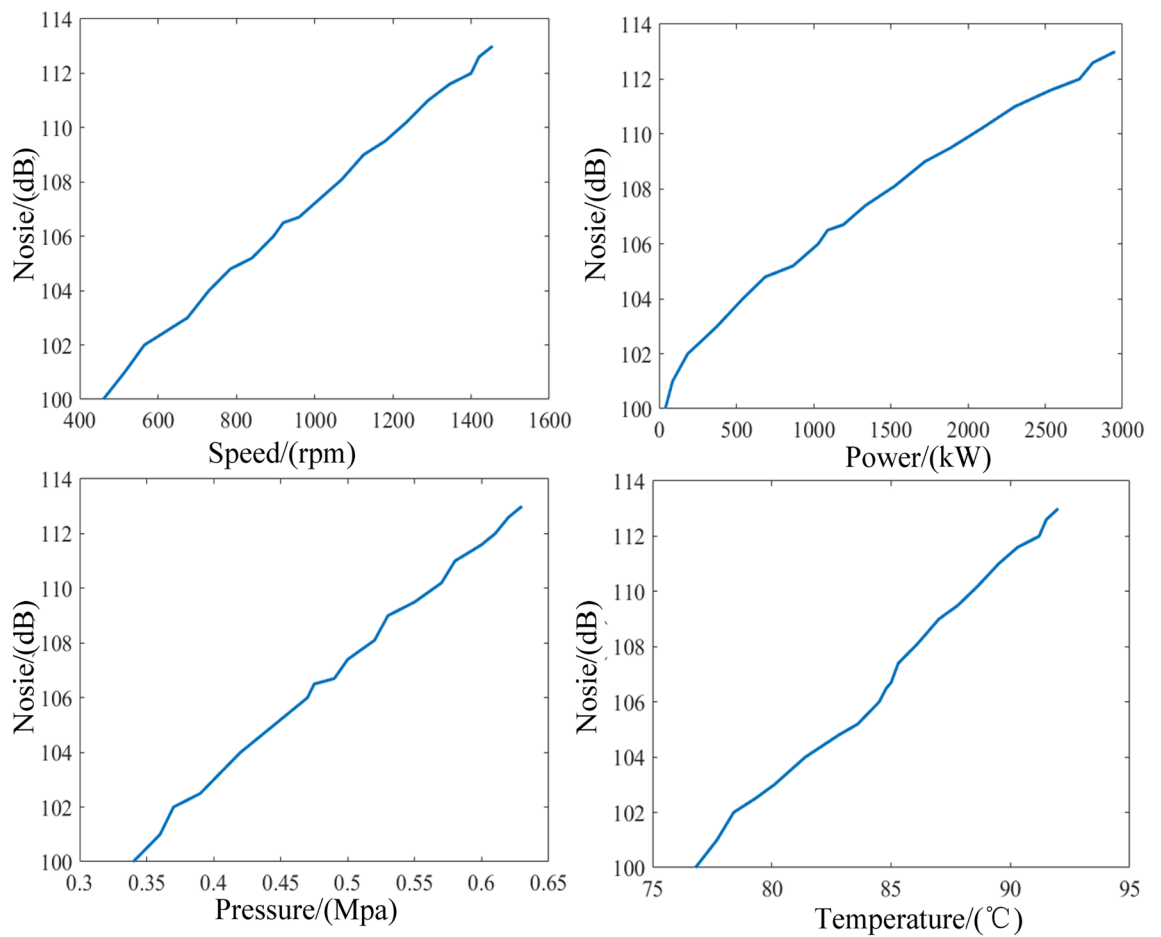
Figure 14. Influence of input parameters on vibration.

formance models of marine diesel engine are established based on the proposed hybrid method of DNN coupled with VSG. The coefficients of determination (R^2) between the predicted and experimental values of fuel consumption rate, vibration, and noise are 0.90, 0.94 and 0.91, respectively. The overall prediction accuracy is more than 93%. The results indicate the proposed model can be effectively applied for predicting and assessing performances of marine diesel engine.

(2)   Based on the DNN models of performances, the effect of diesel engine parameters on performance is discussed. With the increase of speed, power, lubricating oil temperature and pressure, the fuel consumption rate reduces moderately. Moreover, the change of vibration is in general positively related to the diesel engine parameters. It is also found that the noise attenuation trend is parallel to the changing trend of cylinder head vibration.

(3)   The MIV algorithm has been used to quantify the weighting of the influence of each input parameter on the diesel engine performance parameters. The results show that speed has the greatest effect on vibration and noise with a weighting of 30% and 30.5% of the four input factors, respectively. While brake specific fuel consumption is sensitive to power due to a weighting of 30%. Moreover, the MIV standard deviation indicates the all of performance indexes are more sensitive to the input parameters.

This paper achieves the prediction of diesel engine performance and emissions, which effectively reduces time consumption and test costs. However, due to the limited experimental budget, only four input variables are designed in this paper, and the effects of multiple input parameters such as compression ratio, combustion starting point, and injection timing, are not considered. Moreover, due to the application of VSG technology, the four input variables selected in this paper are coupled each other, which makes it difficult to carry out the optimal design. However, the method proposed in this paper can be applied for optimizing, design, and analysis of independent input variables as well. Therefore, the research in this paper provides a foundation for the next work. Firstly, the method proposed in this paper provides a modeling basis for the optimization of complex diesel engine system parameters in the future work. Secondly, the results of this paper can provide a basis for performance evaluation and fault diagnosis of the subsequent diesel engine in the next work.

Figure 15. Influence of input parameters on noise.

Output	Input			
	Speed	Power	Lubricating oil temperature	Lubricating oil temperature
Brake specific fuel consumption	0.2225	0.2510	0.1750	0.1895
Total vibration	0.25	0.2	0.1975	0.1972
Noise	0.24	0.188	0.1731	0.1875

Table 7. Value of |MIV| for input parameters.
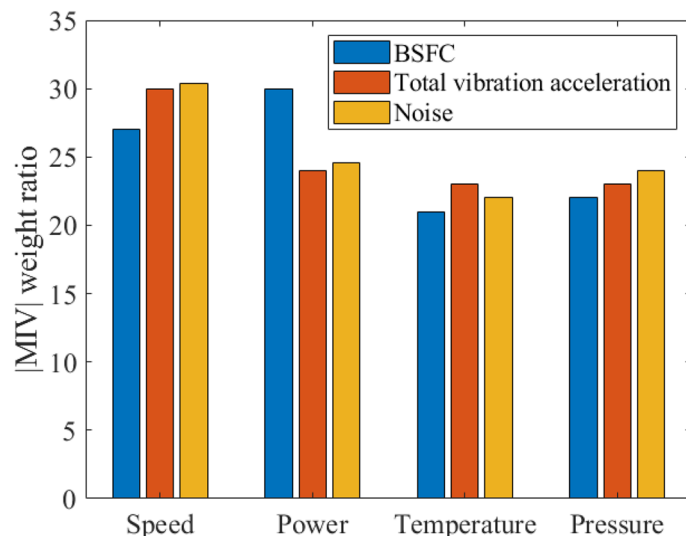
Figure 16. MIV weighting ratio of the input parameters.

References

1. Emaish, H., Abualnaja, K. M., Kandil, E. E. & Abdelsalam, N. R. Evaluation of the performance and gas emissions of a tractor diesel engine using blended fuel diesel and biodiesel to determine the best loading stages. *Sci. Rep.* https://doi.org/10.1038/s41598-021-89287-0 (2021).
2. Yessian, S. & Varthanan, P. A. Optimization of performance and emission characteristics of catalytic coated IC engine with biodiesel using Grey-Taguchi Method. *Sci. Rep.* https://doi.org/10.1038/s41598-019-57129-9 (2020).
3. Ni, P., Wang, X. & Li, H. A review on regulations, current status, effects and reduction strategies of emissions for marine diesel engines. *Fuel* https://doi.org/10.1016/j.fuel.2020.118477 (2020).
4. Shi, Y. & Reitz, R. D. Assessment of multiobjective genetic algorithms with different niching strategies and regression methods for engine optimization and design. *J. Eng. Gas Turbines Power Trans. Asme* https://doi.org/10.1115/1.4000144 (2010).
5. Wang, G., Yu, W., Yang, R., Li, X. & Su, Y. Experimental and numerical study on the combustion limiting characteristics and design margin of a marine diesel engine. *Energy Sour. Part A Recovery Util. Environ. Effects* **41**(18), 2266–2280. https://doi.org/10.1080/15567036.2019.1602234 (2019).
6. Zhang, P. *et al.* Real-time angular velocity-based misfire detection using artificial neural networks. *J. Eng. Gas Turbines Power Trans. Asme.* https://doi.org/10.1115/1.4041962 (2019).
7. Nahim, H. M., Younes, R., Nohra, C. & Ouladsine, M. Complete modeling for systems of a marine diesel engine. *J. Mar. Sci. Appl.* **14**(1), 93–104. https://doi.org/10.1007/s11804-015-1285-y (2015).
8. Gosala, D. B. *et al.* Model-based design of dynamic firing patterns for supervisory control of diesel engine vibration. *Control. Eng. Pract.* https://doi.org/10.1016/j.conengprac.2020.104681 (2021).
9. Muse, A., Juric, Z., Racic, N. & Radica, G. Modelling, performance improvement and emission reduction of large two-stroke diesel engine using multi-zone combustion model. *J. Therm. Anal. Calorim.* **141**(1), 337–350. https://doi.org/10.1007/s10973-020-09321-7 (2020).
10. Zhao, M., Wei, M., Song, P., Liu, Z. & Tian, G. Performance evaluation of a diesel engine integrated with ORC system. *Appl. Therm. Eng.* **115**, 221–228. https://doi.org/10.1016/j.applthermaleng.2016.12.065 (2017).
11. Sapio, F., Piano, A., Millo, F. & Pesce, F. C. Digital shaping and optimization of fuel injection pattern for a common rail automotive diesel engine through numerical simulation. SAE Technical Paper. 0148-7191 (2017).
12. Pagan Rubio, J. A., Vera-Garcia, F., Hernandez Grau, J., Munoz Camara, J. & Albaladejo Hernandez, D. Marine diesel engine failure simulator based on thermodynamic model. *Appl. Therm. Eng.* **144**, 982–995. https://doi.org/10.1016/j.applthermaleng.2018.08.096 (2018).
13. Aghbashlo, M., Shamshirband, S., Tabatabaei, M., Yee, P. L. & Larimi, Y. N. The use of ELM-WT (extreme learning machine with wavelet transform algorithm) to predict exergetic performance of a DI diesel engine running on diesel/biodiesel blends containing polymer waste. *Energy* **94**, 443–456. https://doi.org/10.1016/j.energy.2015.11.008 (2016).
14. Atarod, P. *et al.* Soft computing-based modeling and emission control/reduction of a diesel engine fueled with carbon nanoparticle-dosed water/diesel emulsion fuel. *J. Hazard. Mater.* https://doi.org/10.1016/j.jhazmat.2020.124369 (2021).
15. Aghbashlo, M. *et al.* Machine learning technology in biodiesel research: A review. *Prog. Energy Combust. Sci.* https://doi.org/10.1016/j.pecs.2021.100904 (2021).
16. Shin, S. *et al.* Deep neural network model with Bayesian hyperparameter optimization for prediction of NOx at transient conditions in a diesel engine. *Eng. Appl. Artif. Intell.* https://doi.org/10.1016/j.engappai.2020.103761 (2020).
17. Dinesha, P., Kumar, S. & Rosen, M. Effects of particle size of cerium oxide nanoparticles on the combustion behavior and exhaust emissions of a diesel engine powered by biodiesel/diesel blend. *Biofuel Res. J.* **8**(2), 1374–1383 (2021).
18. Zhang, P., Gao, W., Li, Y. & Wang, Y. Misfire detection of diesel engine based on convolutional neural networks. *Proc Inst. Mech. Eng. Part D J. Automob. Eng.* https://doi.org/10.1177/0954407020987077 (2021).
19. Niu, X., Yang, C., Wang, H. & Wang, Y. Investigation of ANN and SVM based on limited samples for performance and emissions prediction of a CRDI-assisted marine diesel engine. *Appl. Therm. Eng.* **111**, 1353–1364. https://doi.org/10.1016/j.applthermaleng.2016.10.042 (2017).

20. Tauzia, X., Maiboom, A., Karaky, H. & Chesse, P. Experimental analysis of the influence of coolant and oil temperature on combustion and emissions in an automotive diesel engine. *Int. J. Engine Res.* **20**(2), 247–260. https://doi.org/10.1177/1468087417749391 (2019).
21. Uludamar, E., Tosun, E. & Aydin, K. Experimental and regression analysis of noise and vibration of a compression ignition engine fuelled with various biodiesels. *Fuel* **177**, 326–333. https://doi.org/10.1016/j.fuel.2016.03.028 (2016).
22. Kumar, S., Dinesha, P., Ajay, C. M. & Kabbur, P. Combined effect of oxygenated liquid and metal oxide nanoparticle fuel additives on the combustion characteristics of a biodiesel engine operated with higher blend percentages. *Energy* https://doi.org/10.1016/j.energy.2020.117194 (2020).
23. Lee, S. *et al.* EGR prediction of diesel engines in steady-state conditions using deep learning method. *Int. J. Automot. Technol.* **21**(3), 571–578. https://doi.org/10.1007/s12239-020-0054-3 (2020).
24. Wanigasekara, C., Swain, A., Nguang, S. K. & Prusty, B. G. Improved learning from small data sets through effective combination of machine learning tools with VSG techniques. In *2018 International Joint Conference on Neural Networks (IJCNN)* 1–6 (IEEE, 2018) https://doi.org/10.1109/IJCNN.2018.8489759.
25. Li, D.-C., Lin, W.-K., Lin, L.-S., Chen, C.-C. & Huang, W.-T. The attribute-trend-similarity method to improve learning performance for small datasets. *Int. J. Prod. Res.* **55**(7), 1898–1913. https://doi.org/10.1080/00207543.2016.1213447 (2017).
26. Yu, Y. *et al.* A novel deep learning approach to predict the instantaneous NO_x emissions from diesel engine. *IEEE Access* **9**, 11002–11013. https://doi.org/10.1109/ACCESS.2021.3050165 (2021).
27. Pravin, M. *et al.* Predicting the emissive characteristics of an IC engine using DNN. In *IOP Conference Series: Materials Science and Engineering*, vol. 995, no. 1 012010 (IOP Publishing, 2020) https://doi.org/10.1088/1757-899X/995/1/012010.
28. Noor, C. M., Mamat, R., Najafi, G., Nik, W. W., & Fadhil, M. Application of artificial neural network for prediction of marine diesel engine performance. In *IOP Conference Series: Materials Science and Engineering*, vol. 100, no. 1 012023 (IOP Publishing, 2015).
29. Zhang, J. H. & Han, J. CAE process to simulate and optimise engine noise and vibration. *Mech. Syst. Signal Process.* **20**(6), 1400–1409. https://doi.org/10.1016/j.ymssp.2004.11.005 (2006).
30. Zhang, Z. & Jin, X. Prediction of peak velocity of blasting vibration based on artificial neural network optimized by dimensionality reduction of FA-MIV. *Math. Probl. Eng.* https://doi.org/10.1155/2018/8473547 (2018).
31. Zhang, X. *et al.* Determination of zinc oxide content of mineral medicine calamine using near-infrared spectroscopy based on MIV and BP-ANN algorithm. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **193**, 133–140. https://doi.org/10.1016/j.saa.2017.12.019 (2018).

Acknowledgements

Author contributions

H.Z., H.Z. and C.K. wrote the main manuscript textH.Z., Z.L., Z.D. and J.L. prepared figuresAll authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to C.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.