journal homepage: www.brodogradnja.fsb.hr

Brodogradnja

An International Journal of Naval Architecture and Ocean Engineering for Research and Development

Deep learning-based research on fault warning for marine dual fuel engines

Lingkai Meng, Huibing Gan, Haisheng Liu, Daoyi Lu*

College of Marine Engineering, Dalian Maritime University

ARTICLE INFO

Keywords:
Marine dual fuel engine
CNN
BiLSTM
KAN
Fault warning

ABSTRACT

Dual fuel engines are crucial for ensuring the safe navigation of ships. Predicting the working status of these engines can provide advanced knowledge of their condition and thereby guarantee safe navigation. In this study, a novel deep learning model, the CNN-BiLSTM-KAN, was designed to forecast exhaust gas temperature (EGT) in dual fuel engines operating in gas mode. The model integrated convolutional neural networks (CNN), bidirectional long short-term memory (BiLSTM) networks, and Kolmogorov-Arnold networks (KAN) to perform feature extraction from multi-dimensional time series data, autonomously identify temporal patterns within the data, and directly learn parameterized nonlinear activation functions, respectively. The results reveal that the model obtained a mean square error (MSE) of 0.000051, a root mean square error (RMSE) of 0.007135, a mean absolute error (MAE) of 0.003185, and a mean absolute percentage error (MAPE) of 0.000386. The proposed model demonstrated higher accuracy compared to other forecasting models. Additionally, residual value distribution curves and statistical process control methods were employed to set alarm thresholds for residuals. A sliding window approach was used to establish the alarm threshold for residual standard deviation, with an upper boundary of the residual threshold set at 0.15 and a lower boundary at -0.1. The upper boundary of the residual standard deviation was set at 0.343. Furthermore, the model was validated through a fault dataset. The findings suggest that this approach effectively achieved fault warnings for marine dual-fuel engines. This research provides new references for studies on fault prediction and health management of dual-fuel engines for ships.

1. Introduction

The engine is crucial to the power plant of a ship, and its safety can directly affect the operation of the ship, [1]. Networked construction is becoming increasingly advanced in the field of ship maintenance and support attributed to the rapid progress and widespread application of computer technology and artificial intelligence technology. It provides valuable reference and solid technical support for the development of ship intelligence while reinforcing the importance of ship maintenance and supporting system management in modern economic development and safety assurance [2]. Compared with marine diesel engines, dual fuel engines with the highest economy and low emissions have become one of the vital technical measures to

* Corresponding author.
E-mail address: ghbzq@dlmu.edu.cn

http://dx.doi.org/10.21278/brod76303
Submitted 15 December 2024; Received in revised form 03 March 2025; Accepted 07 April 2025
Available online 10 April 2025
ISSN 0007-215X; eISSN 1845-5859

comply with the strict economic and emission regulations of the International Maritime Organization in the future [3]. Therefore, there is a huge demand in the domestic and international markets. However, dual-fuel engines, which have complex structures and diverse operating modes, are influenced by sea temperature and humidity, resulting in a relatively humid operating environment and harsh working conditions. The existing maintenance methods primarily include two categories: post maintenance and preventive maintenance. Post maintenance and preventive maintenance have the phenomenon of "insufficient maintenance" and over maintenance of equipment, respectively [4-5]. Conventional monitoring and alarm technologies are incapable of providing early warning of faults or sending out alarm signals in advance before faults occur [6]. Additionally, current research on marine engines focuses on diesel engines, and there is little research on dual-fuel engines [7]. Many studies and innovations have been performed on dual-fuel engine fault prediction technology to guarantee that dual-fuel engines can provide safe, reliable, and continuous power, reduce the possibility of threatening the safety of ship personnel and property, and achieve an early understanding of faults before they occur [8-10]. Therefore, it is of great significance to investigate the warning technologies for marine dual-fuel engines.

At present, fault diagnosis methods typically include physics-based [11], expert system-based [12], and data-driven approaches [13]. The physics approach requires the construction of precise mathematical or physical models to describe the predicted object [14]. However, ship equipment is generally in a constantly changing situation, rendering it difficult to build accurate models. With the knowledge of domain experts, the expert system approach predicts failures through experience and logic. This approach commonly requires significant expert experience and knowledge accumulation. Improved methods such as combining expert systems with neural networks [15] and fusing expert systems with fault trees [16] have also been proposed. The data-driven approach can be employed to collect historical operational data of devices [17], and process and analyse the data with algorithms to establish a fault warning model, thereby achieving fault warning. As various intelligent algorithms achieving good results in various fields, fault warning based on data methods has gradually attracted many researchers' attention [18] and has been introduced into the fault warning of ship equipment.

Exhaust gas temperature (EGT) is one of the critical thermodynamic performance parameters of marine dual fuel engines for evaluating and optimizing the performance of marine dual fuel engines [19]. It is related to multiple aspects such as combustion efficiency, emission control, thermal management, and power output of the engine, involves. extensive status information on dual-fuel engines and reflects the combustion and dynamic characteristics of dual-fuel engines [20]. The surveillance and forecasting of EGT in dual-fuel engines can provide real-time insights into their health status owing to the characteristics of slow changes in EGT, minimal interference, and pronounced fault signals [21]. This allows for fast identification of potential faults and timely measurement to address them.

Ozsari et al. [22] employed artificial neural networks to forecast engine power and emissions for container ships, cargo ships, and oil tankers, experimenting with various configurations of hidden neurons to refine the model structure. The analysis yielded highly accurate results in terms of MSE and MAPE for both regression and error metrics. Yao et al. [23] introduced a deep transfer reinforcement learning framework with long short-term memory (LSTM) networks to manage the challenges of RUL prediction in scenarios with scarce fault data. They transferred a deep reinforcement learning model, initially trained on fault data, to a new RUL prediction instrument. This method demonstrates high accuracy and good adaptability. Sahu et al. [24] established a pioneering method to estimate the RUL of bearings through the integration of a health indicator with the LSTM deep learning model. The experimental results confirmed the superior performance of the method. Lin et al. [25] proposed a prediction mechanism combining the random forest method and LSTM for the predictive maintenance of wind turbines. The forecast model, with certain advantages in prediction accuracy and recall rate, can effectively anticipate wind turbine malfunctions. Batool et al. [26] applied deep learning techniques to software fault prediction and forecasted software faults using LSTM, bidirectional LSTM (BiLSTM), and radial basis function network (RBFN). The findings suggest that while LSTM and BiLSTM outperformed in accuracy, RBFN offered swift predictions, contributing to a more precise and efficient approach to software defect forecasting.

Bazai et al. [27] adopted two techniques to predict the behaviour of complex systems such as fluidized beds: computational fluid dynamics (CFD) and convolutional neural networks (CNN). They revealed that the combination of these two techniques enabled CNN to learn faster and better while lowering the amount of computation, making the entire process more efficient. Bai et al. [28] established a novel network called the time convolutional network (TCN) based on dilated convolution and residual connection design. Through multiple experiments, they verified that this method performed significantly better than general cyclic architectures in complex sequence modelling tasks. Haung et al. [29] proposed a novel periodic time series forecasting model grounded in DA-RNN to overcome the shortcomings of existing deep learning models in handling periodic and long-distance dependency sequences. The proposed model has advantages in capturing the periodicity and long-range dependency features of sequences. Wadie Bendali et al. [30] combined a Deep Echo State Network with a Binary Genetic Algorithm to develop predictions of energy consumption over different time ranges. The proposed model demonstrated a faster processing speed. In different time ranges of prediction, the error index is the best. Jiang et al. [31] designed an enhanced slime mold algorithm (SMA) to optimize the interpolation points of the trajectory modelling, managing the precise demands for trajectory tracking control of autonomous underwater vehicles. This improved SMA was compared with the artificial fish swarm algorithm, particle swarm optimization, and compact cuckoo search, unveiling that the refined SMA shortened the search duration, efficiently sidestepped local optima, and produced high-precision trajectory models more quickly. Zhang et al. [32] proposed a new method to simulate and monitor the dynamic performance of high-speed bearings. Combining short time Fourier transform (STFT) with Convolutional Neural Networks and incorporating Multi-Head Attention Mechanism provide an innovative perspective for fault diagnosis and performance evaluation of high-speed bearings in complex production environments.

In existing research, fault warning methods typically consist of state forecasting and state classification, [33]. The latter requires a lot of fault data, whereas there is very little early fault data for marine dual-fuel engines. Therefore, a CNN-BiLSTM-KAN-based state forecast model was introduced in this paper to forecast the EGT of marine dual-fuel engines in gas mode. Specifically, statistical methods were employed to examine the discrepancies between predicted and actual values, known as residual, and set a residual alarm limit. Attributed to the limitations of the model itself, significant differences between predicted and actual values appear even under normal conditions, bringing about erroneous alarms. With the purpose of avoiding this situation, the sliding window algorithm was adopted to determine the standard deviation of residuals, and the alarm threshold was set accordingly. The system only triggered an alert when both the residual and its standard deviation surpassed the predefined limits. Finally, the effectiveness of this method was validated through experiments.

The rest of this paper is organized as follows. Section 2 introduces the deep learning theory used, including CNN, BiLSTM, and KAN. Section 3 details the source of predicted sample data and the procedures for data pre-processing. Particularly, the construction of the forecast model involved the optimal selection of hyperparameters and verified the advantages of the established marine dual fuel engine fault warning model through comparative experiments. In Section 4 of the fault warning, a method for determining the alarm threshold is provided and validated through experimentation. Finally, conclusions and outlook are drawn in Section 5.

2. Deep learning model theory

2.1 Convolutional neural network (CNN)

CNN primarily comprises input layer, convolutional layer, pooling layer, fully connected layer, output layer, and activation function [34], as depicted in Figure 1. CNN can automatically learn features and patterns suitable for specific tasks without manually designing features while minimizing the number of parameters by utilizing shared convolutional kernels, hereby lessening the model's complexity and computational demands, [35]. Moreover, it can automatically extract feature information, effectively avoiding manual feature extraction, [36].

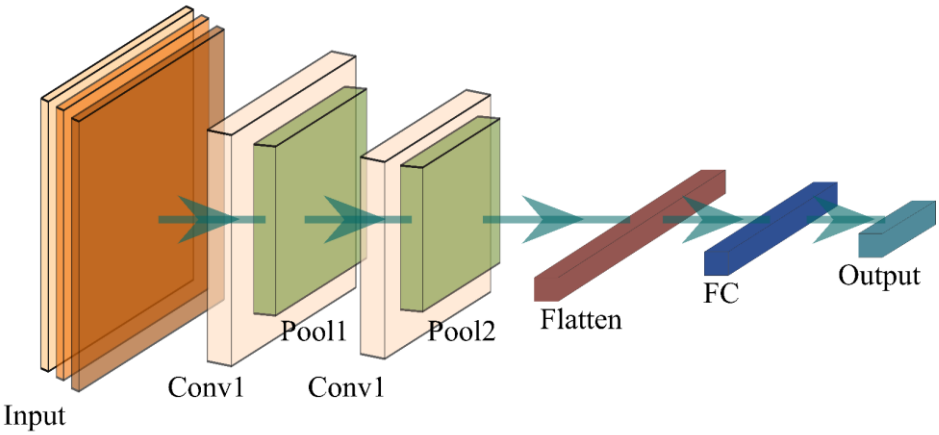

Fig. 1 Structure diagram of Convolutional Neural Network

The convolutional layer and pooling layer constitute the feature extractor of CNN [37]. In convolutional layers, each neuron only connects adjacent neurons, and each convolutional layer contains several feature maps composed of rectangular units [38]. In the same feature map, each node shares a weight, which is the convolution kernel. The initialization of it applies a random decimal matrix, which is trained to obtain appropriate weights [39]. Its core advantage lies in that it can effectively curtail the connections between different levels and minimize the risk of overfitting. The pooling layer consists of both average and maximum pooling. Pooling is essentially a special convolution operation that significantly diminishes the complexity of the model and the number of parameters it contains. The fully connected layer synthesizes the features gleaned from the convolutional and pooling layers for classification, recognition, or forecast.

Through convolution, pooling operations, and weight sharing, CNN can better capture local features in feature data, revealing the advantages in feature extraction of data. One-dimensional convolutional neural networks (1DCNNs) are particularly adopted to analyse time series data, a domain where traditional two-dimensional CNNs (2DCNNs) are not well-suited due to dimensionality mismatch. These 1DCNNs excel at extracting local features from time series and thus reinforce the precision and reliability of forecast models, [40].

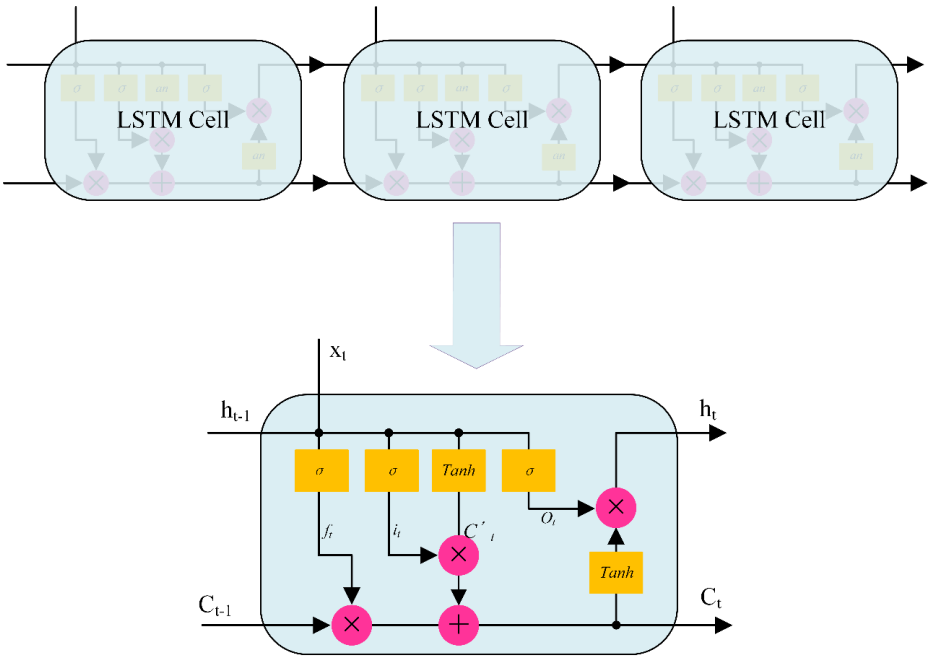

Fig. 2 The schematic diagram of the LSTM structure

2.2 Long short-term memory network (LSTM)

LSTM, based on Recurrent Neural Networks (RNN), solves the short-term memory deficiency that exists when processing sequential data [41]. The structure of LSTM is illustrated in Figure 2. LSTM controls the storage, updating, and discarding of memory states through three gate structures: input gate, output gate, and forget gate. Additionally, there is a storage long-term memory control unit C. The advantage of LSTM is that it selectively remembers both long-term and short-term information from multiple time series through gate structures. Since the storage long-term memory control unit C in LSTM, similar to a conveyor belt, only performs a small amount of information exchange, the information can remain relatively stable during transmission and will not undergo significant changes [42].

In the figure, x_t denotes the input value at the current time; h_{t-1} indicates the output value of the previous moment; C_{t-1} represents the state of the memory control unit from the previous moment; h_t embodies the output value at the current time; C_t refers to the current state of the memory control unit. Besides, f_t, i_t, $C_t^{'}$, and O_t form the gate computing structure of the LSTM network, describing the forget gate, the memory gate, the temporary cell state, and the output gate, respectively. σ expresses the activation function sigmoid and $\tanh$ is the hyperbolic tangent function operation. The operation process of LSTM is to forget the information of neurons and retain new information, transmit valuable information for subsequent operations, remove unnecessary information, and output the hidden state h_t at each step. The forget gate f_t, memory gate i_t, and output gate O_t in an LSTM architecture are responsible for managing the forgetting, memory retention, and output functions, respectively. These gates are derived from the interaction between the previous hidden layer state h_{t-1} and the current input x_t [43]. The calculation formula is:

$$\begin{cases} \sigma(x) = \dfrac{1}{1-e^{-x}} \\ \tanh(z) = \dfrac{e^{z}-e^{-z}}{e^{z}+e^{-z}} \\ f_t = \sigma(W_f^T \times h_{t-1} + U_f^T \times x_t + b_f) \\ i_t = \sigma(W_i^T \times h_{t-1} + U_i^T \times x_t + b_i) \\ C_t^{'} = \tanh(W_c^T \times h_{t-1} + U_c^T \times x_t + b_c) \\ C_t = f_t * C_{t-1} + i_t * C_t^{'} \\ O_t = \sigma(W_o^T \times h_{t-1} + U_o^T \times x_t + b_o) \\ h_t = O_t * \tanh(C_t) \end{cases} \tag{1}$$

where W and U represent the weight parameters corresponding to the gate structure, and b denotes bias.

Compared with RNN, LSTM is particularly suitable for processing sequence data with large time spans and can effectively address long-term dependency issues. It has excellent data-fitting ability, high robustness, and wide applicability [44]. Besides, the recurrent layers of LSTM adopt the same network parameters and do not change as the time series span increases.

2.3 Bidirectional long short-term memory neural network (BiLSTM)

Since the operating status of ship engines is influenced by both historical and forthcoming information, traditional LSTM can only apply historical information to predict future state outputs and cannot obtain previous information from later information. Given this issue, a bidirectional LSTM has been proposed [45]. BiLSTM integrates past and future information and comprehensively analyses time series data through forward and backward propagation mechanisms to more accurately predict the current operating status of ship engines [46] and capture bidirectional temporal dependencies [47]. Therefore, BiLSTM can provide more representative feature representations, stronger modelling capabilities, and more accurate prediction capabilities in marine engine fault prediction problems, making it an effective model structure. The BiLSTM structure diagram is illustrated in Figure 3.

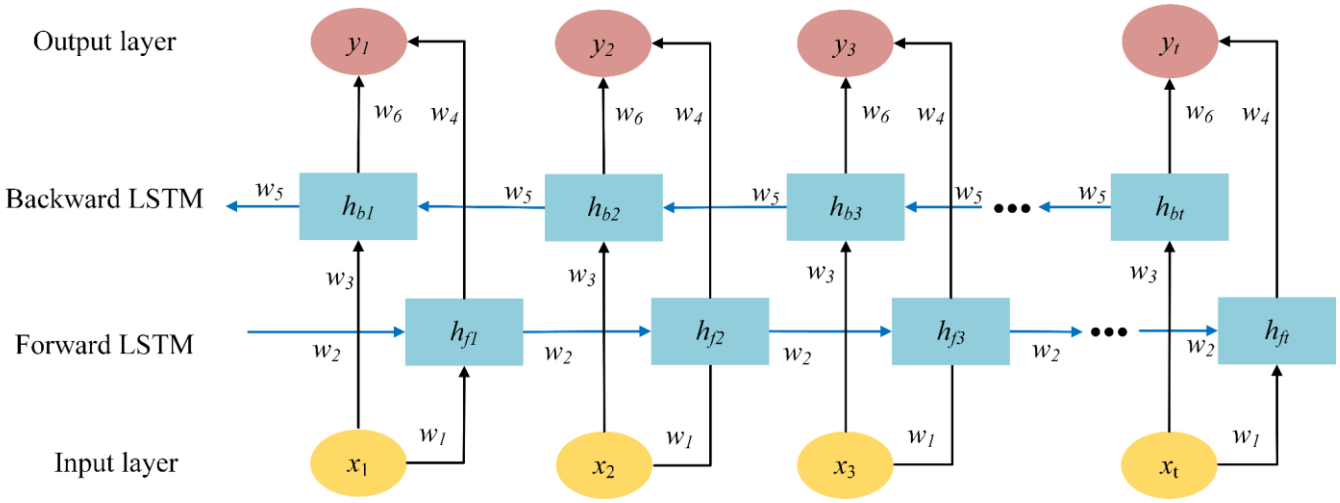

Fig. 3 Structure of a bidirectional long short-term memory neural network

As observed from Figure 3, the input sequence is fed into both the backward LSTM's hidden layer h_{bt} and the forward LSTM's hidden layer h_{ft}. Subsequently, these two layers are merged and linked to the output layer to generate predictions, culminating in the BiLSTM model's output h_t. The calculation formula is:

$$\begin{cases} h_{ft} = f(w_1 x_t + w_2 h_{ft-1} + b_{ft}) \\ h_{bt} = f(w_3 x_t + w_5 h_{bt-1} + b_{bt}) \\ h_t = g(w_4 h_{ft} + w_6 h_{bt} + b_{ot}) \end{cases} \tag{2}$$

where w_1 denotes the weight coefficient from the input layer to the preceding LSTM; w_2 signifies the weight coefficient among the forward LSTM unit layers; w_3 indicates the weight coefficient from the input layer to the backward LSTM; w_5 represents the weight coefficient between the layers of the backward LSTM unit; w_4 embodies the weight coefficient from the forward LSTM to the output layer ;w_6 stands for the weight coefficient from the backward LSTM to the output layer; b_{ft}, b_{bt}, and b_{ot} correspond to the bias matrices associated with their respective components.

2.4 Kolmogorov–Arnold networks (KAN)

KAN is based on the Kolmogorov Arnold representation theorem. It was proposed by two Russian mathematicians in 1957 to represent any continuous multivariate function using simple functions, expressed as [48]:

$$f(x) = \sum_{q=1}^{2n+1} \Phi_q \left(\sum_{p=1}^{n} \phi_{q,p}(x_p) \right) \tag{3}$$

where x refers to the input; $\phi_{q,p}(x_p)$ denotes basic unary functions, and the inner layer summation is to put them together; Φq represents the outer function, each accepting the sum of the inner layers as input. The outer summation $\sum$ indicates that the entire function $f(x)$ is the sum of the subfunctions Φq.

The principle of the entire KAN network architecture is presented in Figure 4, where many sine functions similar to three-quarters of a period can be combined to fit any shape of function.

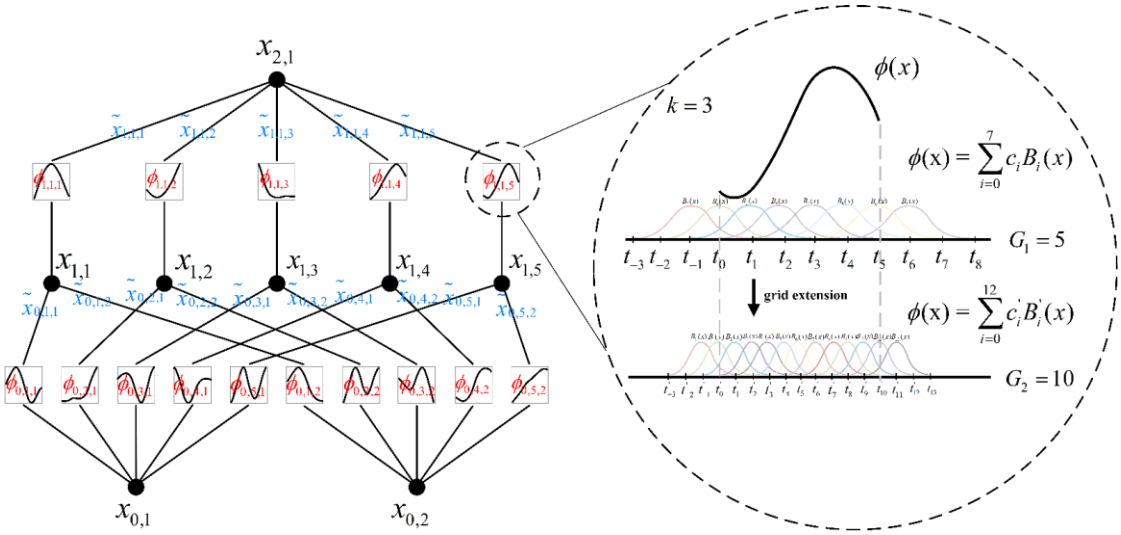

Fig. 4 Structure of Kolmogorov–Arnold Networks neural network

Figure 4 demonstrates a combination of two scales or resolutions: coarse-grained and fine-grained grids, which more accurately capture and adapt to changes in the function while maintaining computational efficiency.

The KAN layer is a one-dimensional function matrix, expressed as:

$$\Phi = \{\phi_{q,p}\}, \qquad p = 1, 2, \cdots, n_{in}, \qquad q = 1, 2, \cdots, n_{out} \tag{4}$$

Among them, function $\phi_{q,p}$ has trainable parameters, the internal function forms a KAN layer of $n_{in} = n$, $n_{out} = 2n + 1$, and the external function forms a KAN layer of $n_{in} = 2n + 1$, $n_{out} = 1$. Simply stacking them allows for deeper function. Simply put, it aims to find the transition matrix between the input and output of each layer, expressed as:

$$x_{l+1} = \underbrace{\begin{pmatrix} \phi_{l,1,1}(\cdot) & \phi_{l,1,2}(\cdot) & \cdots & \phi_{l,1,n_l}(\cdot) \\ \phi_{l,2,1}(\cdot) & \phi_{l,2,2}(\cdot) & \cdots & \phi_{l,2,n_l}(\cdot) \\ \vdots & \vdots & & \vdots \\ \phi_{l,n_{l+1},1}(\cdot) & \phi_{l,n_{l+1},2}(\cdot) & \cdots & \phi_{l,n_{l+1},n_l}(\cdot) \end{pmatrix}}_{\Phi_l} x_l \tag{5}$$

where l denotes the layer number, with the right side and left side as the input and output, respectively. The left image in Figure 4 illustrates the corresponding relationship. Because the input is 2, the second layer has $2 * 2 + 1 = 5$. Besides, $\phi_{l,i,j}$ denotes the activation function on each edge, which is a nonlinear transformation. It is equivalent to having 5 clones of each x and then combining them separately. Among them, i and j mark the nodes of the current layer and the next layer, respectively. The output of each node $x_{l,i}$ is processed by activation function $\phi_{l,i,j}$ and contributes to the computation of all the next layers $x_{l+1,j}$.The left figure in Figure 4 exhibits 2 nodes in the input layer and 5 nodes in the second layer, and thus the matrix is $5 * 2$. The first column of the matrix represents the 5 activation functions corresponding to $x_{0,1}$, the second column corresponds to $x_{0,2}$, and then they are combined pairwise. Notably, the number of nodes in the KAN network layer is determined by the number of input nodes n, which is $2n + 1$, and the required parameters or connections are $(2n + 1) * n$, which is significantly less than fully connected.

Overall, the original two-layer KAN network had a shape of $[n, 2n + 1, 1]$. Nonetheless, it has become a multi-layer cascade, further relaxing the structural constraint of $2n + 1$ and enabling the number of hidden nodes to be freely determined.

3. Prediction process

3.1 Data acquisition and processing

The complete model of the W6X72DF marine dual fuel engine built using GT-POWER software is illustrated in Figure 5. The relevant technologies and performance parameters applied in the process of building the model were based on the diesel engine bench test report issued by HSD Engine, which was certified by China Classification Society and possesses the credibility and authority required for modelling. It has imperative reference significance for modelling and simulation research.

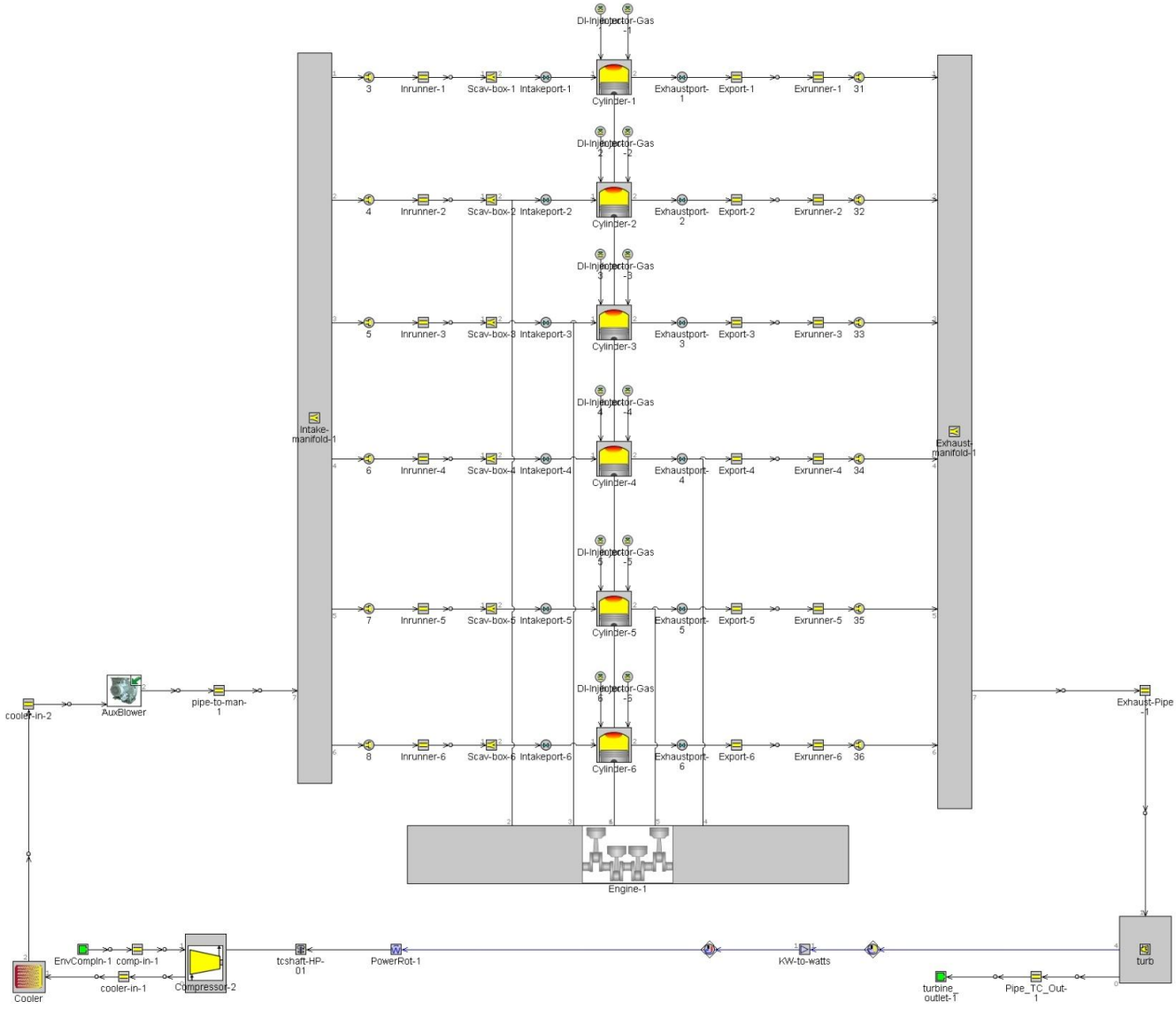

Fig. 5 Engine simulation model in GT-POWER

Following the engine bench test report issued by HSD Engine, the effectiveness of the engine simulation model was calibrated, verified, and validated in this study when the load is 25 %, 50 %, 75 %, and 100 % stable operation under two fuel modes. Notably, the fuel consumption of the engine when operating in gas mode is not calculated upon the consumption of diesel fuel under this operating condition. Instead, the total fuel consumption was calculated based on the minimum calorific value of the ignited diesel and natural gas consumed during operation under this operating condition. The low calorific value of diesel is 42.920 MJ/kg, and the low calorific value of natural gas is 47.555 MJ/kg. EGT refers to the EGT behind the cylinder.

(1) Diesel mode

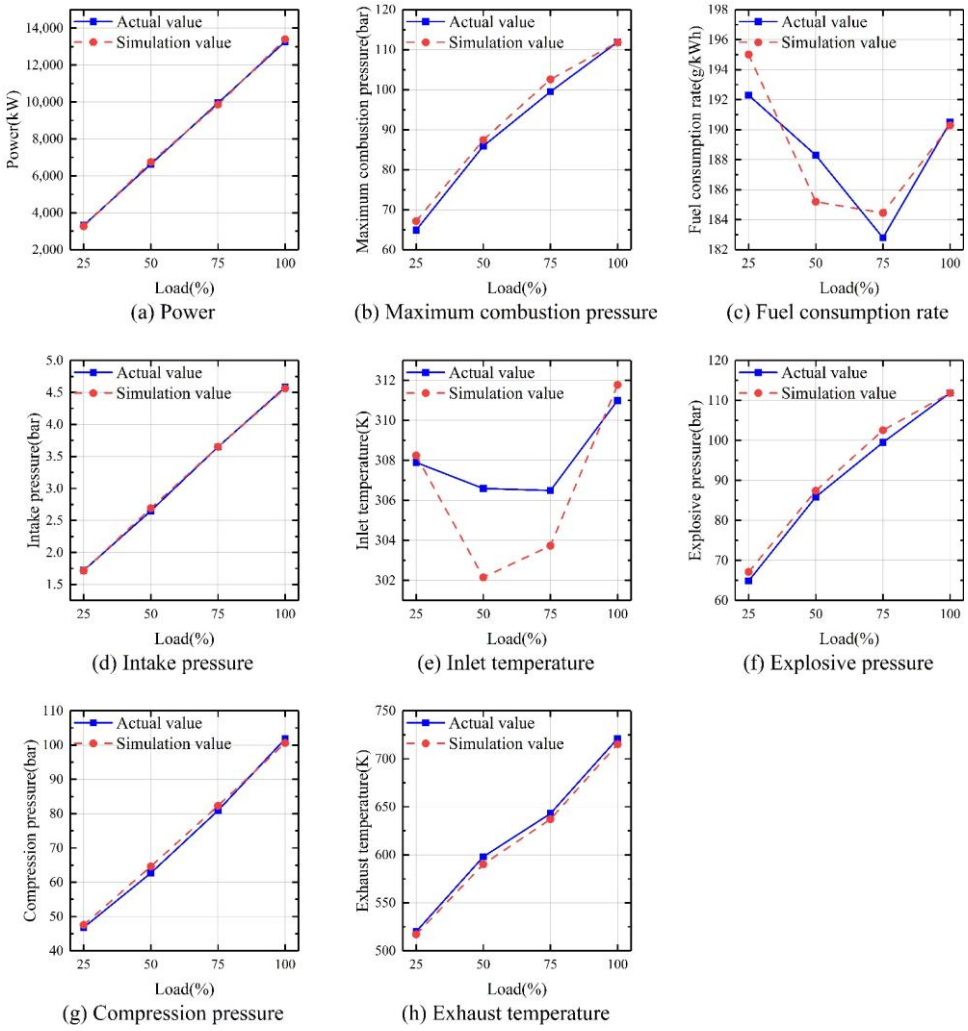

Fig. 6 Comparison between simulation value and experiment value in diesel mode

Table 1 Simulation error of the diesel mode

Load (%)	25	50	75	100
Error (%)				
Power (kW)	-1.44	1.67	-0.95	1.01
Maximum combustion pressure (bar)	3.47	1.79	3.09	-0.06
Fuel consumption rate (g/kWh)	1.41	-1.65	0.90	-0.11
Intake pressure (bar)	-0.30	1.54	0.02	-0.36
Inlet temperature (K)	0.11	-1.45	-0.90	0.25
Explosive pressure (bar)	3.45	1.77	3.07	-0.05
Compression pressure (bar)	1.61	3.06	1.74	-1.18
EGT (K)	-0.55	-1.31	-0.91	-0.81

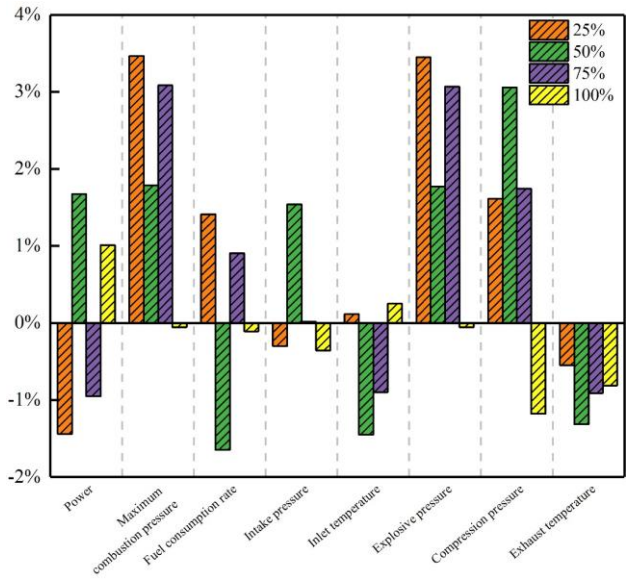

Fig. 7 The simulation error under different working conditions

Figures 6 and 7, and Table 1 reveal that the simulation outcomes for the dual fuel engine model, which is developed for the steady diesel model, closely match the bench test data. The parameters such as power output, peak combustion pressure, fuel consumption, intake pressure, intake temperature, ignition pressure, compression pressure, and EGT exhibit minimal discrepancies, with all relative errors falling below 5 %.

(2) Gas mode

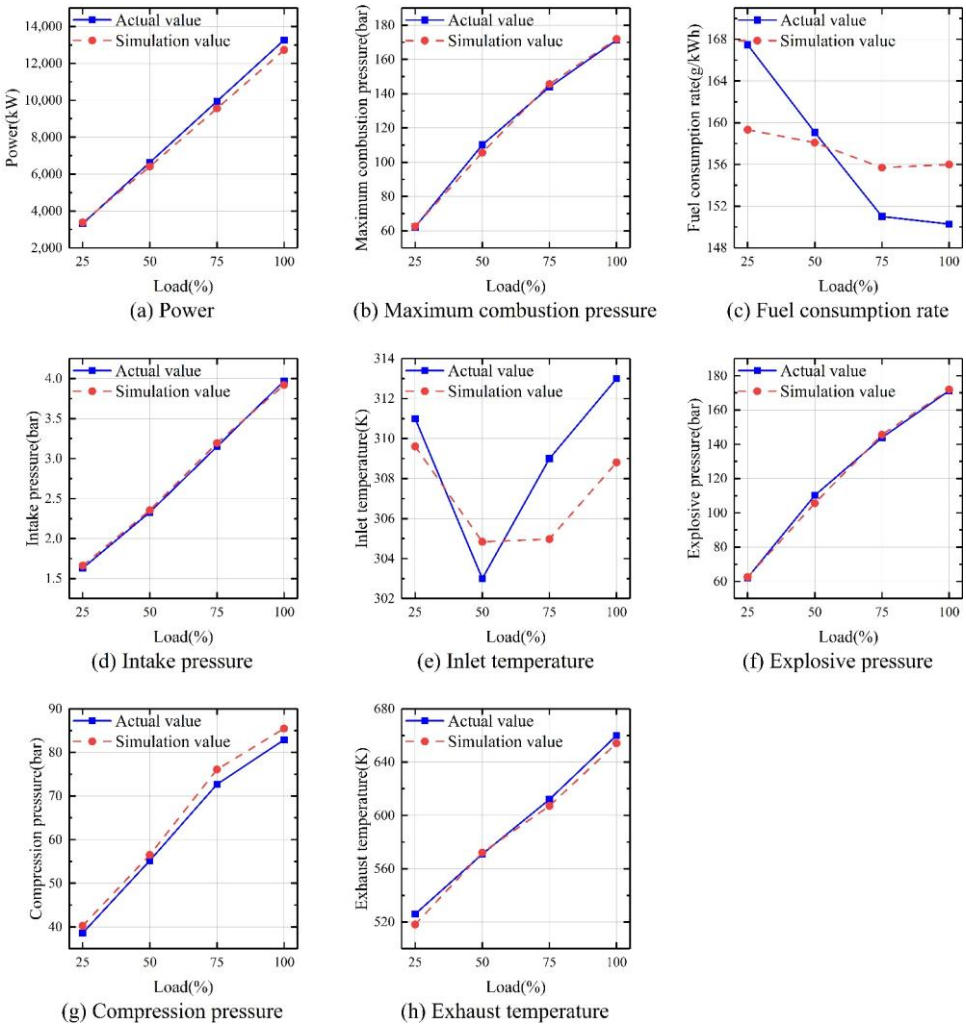

Fig. 8 Comparison between simulation value and experiment value in gas mode

Table 2 Simulation error of the gas mode

Load (%)	25	50	75	100
Error (%)				
Power (kW)	1.98	-3.35	-3.91	-4.02
Maximum combustion pressure (bar)	0.58	-4.13	1.16	0.38
Fuel consumption rate (g/kWh)	-4.86	-0.62	3.10	3.80
Intake pressure (bar)	1.93	1.04	1.39	-1.27
Inlet temperature (K)	-0.45	0.61	-1.30	-1.34
Explosive pressure (bar)	0.58	-4.13	1.16	0.38
Compression pressure (bar)	4.24	2.38	4.66	3.11
EGT (K)	-1.49	0.20	-0.79	-0.89

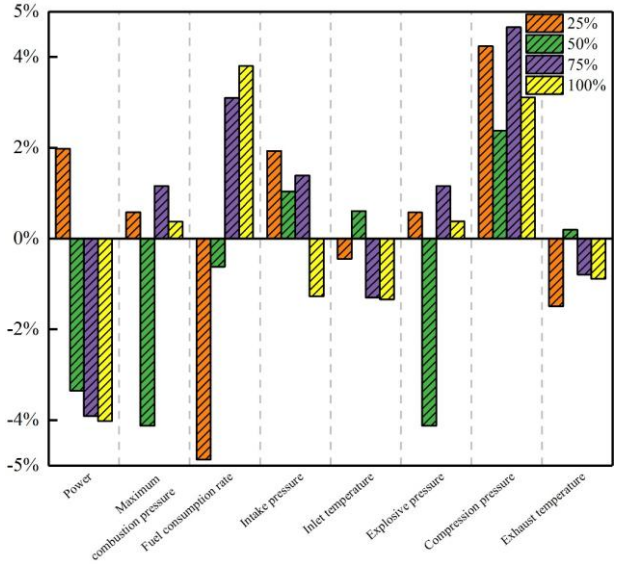

Fig. 9 The simulation error under different working conditions

Figures 8 and 9, and Table 2 demonstrate that the simulation outcomes for the dual fuel engine model, which is developed for the steady gas model, are consistent with the bench test data. The parameters such as power output, peak combustion pressure, fuel consumption, intake pressure, intake temperature, ignition pressure, compression pressure, and EGT exhibit minimal discrepancies, with all relative errors falling below 5 %.

Therefore, the main performance parameter errors of the dual fuel engine simulation model built in both diesel and gas modes are within 5 %, implying that the engine simulation model can simulate real situations and should be further studied.

Gas mode is the main operating mode distinguishing dual fuel engines from traditional diesel engines [49], Therefore, this paper focuses on modelling and simulation research of gas models. The comparison and error between simulation values and experimental values in fuel mode and gas mode suggest that the simulation model built has the best simulation effect at 100 % load. Hence, simulation analysis in gas mode at 100 % load was conducted. Additionally, a data point was recorded every 2 s for data simulation, with a total of 7200 data points selected, to decrease the number of calculations, improve processing speed, and simulate the real situation of dual-fuel engines as much as possible. This model of dual fuel engine has selected a total of 7 operational parameters, as detailed in Table 3. These thermodynamic parameters include extensive details regarding the engine's condition. EGT is indicative of combustion efficiency and overall engine performance. Excessively high EGT levels reflect either incomplete combustion or engine malfunction. The EGT after the turbine reveals the combustion efficiency. The concentration of NOx specifies the

combustion process and fuel quality. The outlet temperature of the cooler implies the circulation effect of the engine cooling water and the thermal load of the engine. Power indicates the operating status and fuel efficiency of the engine. The fuel consumption rate unveils the fuel economy and operating efficiency of the engine. The temperature at the compressor outlet demonstrates both the main engine's combustion efficiency and the operational condition of the compressor itself.

Table 3 Table of characteristic parameter types for dual fuel engines

Characteristic parameter	Symbol	Unit
EGT	T1	K
EGT after turbine	T2	K
NOx concentration	NOx	ppm
Cooler outlet temperature	T3	K
Power	P	kW
Fuel consumption rate	BSFC	g/kWh
Compressor outlet temperature	T4	K

A ship dual fuel engine EGT forecast dataset was established with the obtained data, with 70 % for training and 30 % for testing, to predict the ship dual fuel engine EGT and verify its effectiveness. Choosing input variables is crucial for constructing a prediction model. Considering that the EGT of dual-fuel engines is majorly predicted in this study, feature variables having a strong correlation with EGT should be selected when selecting input variables to minimize the model's input dimensionality and to avoid overfitting stemming from high input dimensionality.

Pearson correlation coefficients (PCCs) quantify the correlation among various variables, with the strength of the relationship indicated by the magnitude of these coefficients [50], expressed as:

$$r_{X,Y} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}} \tag{6}$$

where $r_{X,Y}$ represents the correlation size between variables X and Y, with a range of [-1,1]. The absolute value of $r_{X,Y}$ is directly proportional to the degree of correlation between variables. X_i denotes the i-th sample in variable X; Y_i indicates the i-th sample in variable Y; $\overline{X}$embodies the average value of all samples in variable X; $\overline{Y}$ refers to the average value of all samples in variable Y. The evaluation indicators for PCCs are listed in Table 4.

Table 4 The evaluation criteria for Pearson correlation

The range of values for $\|r\|$	Relevance level
$\|r\| \in [0.0,0.2)$	Very weakly correlated or unrelated
$\|r\| \in [0.2,0.4)$	Weak correlation
$\|r\| \in [0.4,0.6)$	Moderately related
$\|r\| \in [0.6,0.8)$	Strong correlation
$\|r\| \in [0.8,1.0]$	Very strongly correlated

Table 5 presents the correlation degree between various thermodynamic parameters collected through PCCs analysis and EGT. According to Table 5, the coefficient dimension of EGT and nitrogen oxides after the turbine of the dual-fuel engine has the highest correlation with EGT. Furthermore, the correlation coefficients for the outlet temperature, power, fuel consumption rate, and EGT of the cooler are all above 0.7, verifying a strong correlation. Consequently, these five thermodynamic parameters are considered influential

factors and employed together with EGT as inputs for the prediction model to forecast the future trend of EGT changes.

Table 5 Result of operating parameters correlation analysis

Parameter type	T1	T2	NOx	T3	P	BSFC	T4
T1	1	0.937	0.857	0.740	-0.738	0.727	0.149
T2	0.937	1	0.650	0.476	-0.755	0.747	-0.202
NOx	0.857	0.650	1	0.971	-0.715	0.704	0.585
T3	0.740	0.476	0.971	1	-0.591	0.581	0.744
P	-0.738	-0.755	-0.715	-0.591	1	-1.000	0.018
BSFC	0.727	0.747	0.704	0.581	-1.000	1	-0.026
T4	0.149	-0.202	0.585	0.744	0.018	-0.026	1

Given the disparate scales and substantial fluctuations among the parameters serving as inputs for the forecast model, all dataset variables should be standardized before their introduction into the model. The process of normalization is delineated by:

$$X_N = \frac{X - X_1}{X_2 - X_1} \tag{7}$$

where X_N represents normalized data; X denotes raw parameters; X_2 and X_1 stand for the upper and lower bounds of the parameters, respectively.

Concerning the different input formats of different neural networks, the data must be transformed into a format that meets the requirements of the neural network. In this study, sliding window technology was utilized to adjust data, as depicted in Figure 10. Among them, x_i represents the raw data of time i. Specifically, x_{ij} is the value of the j-th feature at time i. Figure 10 suggests that the size of the window is 8, with the first six data (x_1, x_2, x_3, x_4, x_5, x_6) used as input samples and the remaining data (x_7, x_8) in the window used as output labels. Then, the window one step is shifted to the right to obtain the next input sample and output label, contributing to the establishment of a complete dataset. Figure 11 displays the time serialization method, and the left image presents the original dataset, which contains 6 features. It is transformed into a matrix through a sliding window method to fulfil the input specifications of the prediction model. A s indicates that each input sample contains historical data from the past s time periods, and n represents the number of features contained in the input data. In this paper, the value of n is 6.

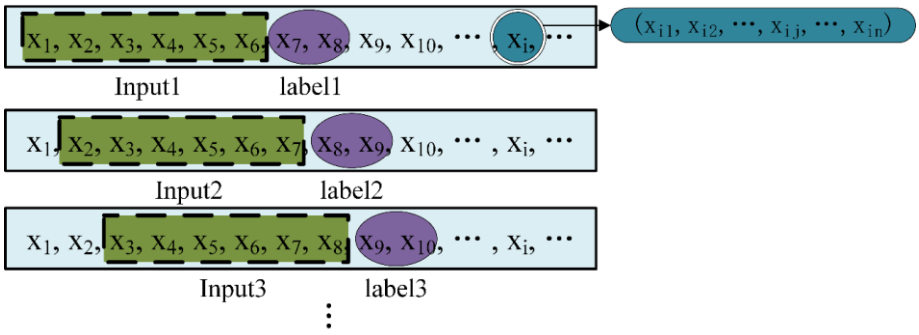

Fig. 10 The sliding window method

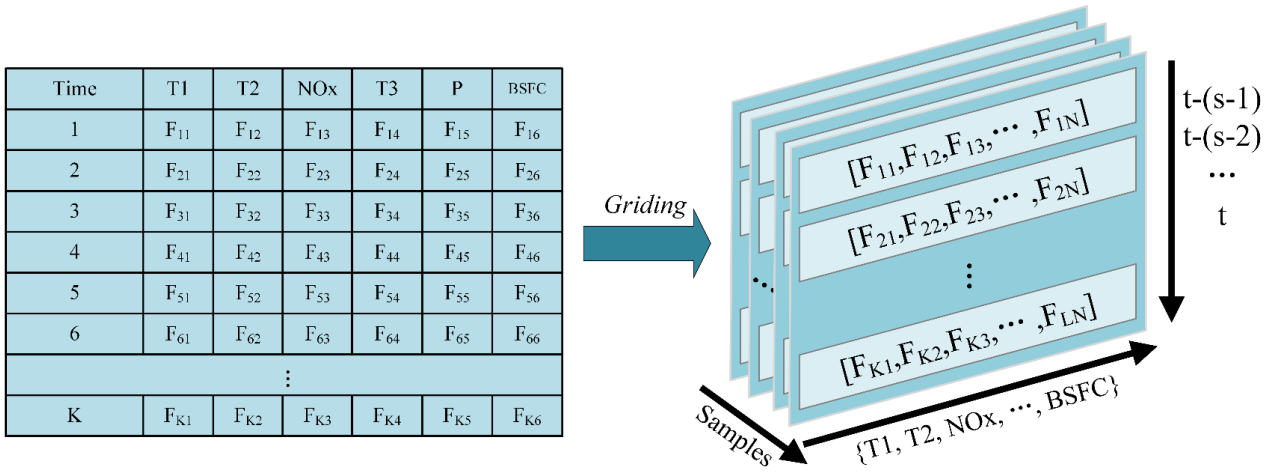

Fig. 11 Data time serialization method

Mahalanobis distance can be used to measure the gap between multivariate data and effectively tackle the issue of non-independent and non-identically distributed dimensions within high-dimensional linear datasets. In this experiment, this distance measure was utilized to identify and exclude outliers, which are data points markedly diverging from the norm. After correlation analysis, the Mahalanobis distance, which quantifies the separation between EGT and the remaining five features, is ascertained through the computation of this distance metric, expressed as:

$$D_M = \sqrt{(x_i - y_i)S^{-1}(x_i - y_i)} \tag{8}$$

where D_M indicates the calculated Mahalanobis distance; x_i and y_i refer to column vectors; S^{-1} describes the inverse matrix of covariance.

The Chi-square test was performed to identify outliers within the EGT data. A total of 5 degrees of freedom was considered for the dataset. The critical value for Mahalanobis distance was set at a significance level of 0.005. The calculation formula for the Chi-square test is:

$$x^2 = \sum_{i=1}^{k} \frac{(A_i - np_i)^2}{np_i} \tag{9}$$

where x^2 denotes the Chi-square degree of freedom, A_i represents the cell observation value, and p_i embodies the expected frequency of the cell.

Since the Chi-square test calculates a Chi-square degree of freedom of 16.74960, the points with a Mahalanobis distance greater than 16.74960 are considered outliers. Through this method, 24 outlier points were successfully eliminated from the sample data, contributing to tackling the headache of inaccurate simulated engine operation data caused by the model itself and enhancing the accuracy of EGT forecasting.

Missing values were fixed by the mean interpolation method to address the issue of intermittent temporal data after removing outliers. The general average interpolation method was employed to interpolate this attribute through the average value of all data. From a practical perspective, the exhaust data at a specific moment is closely related to its previous short-term data, while the original dataset is the data spanning a longer period of time. Using the average of the overall data for direct interpolation does not conform to natural laws. Therefore, an optimized window mean interpolation method was established in this study.

The window size S was manually set and adjusted according to the characteristics of the data. Starting from the missing values, $S/2$ sampling points were taken on each side. If the number of sampling points on one side is less than $S/2$, it will expand to the other side, keeping the number of sampling points within the window at S. As revealed in Figure 12, S is 6, and three different situations exist. The six data points near missing $null$ are x_a, x_b, x_c, x_d, x_e and x_f. Finally, the missing values were replaced with the calculated mean of the six data points.

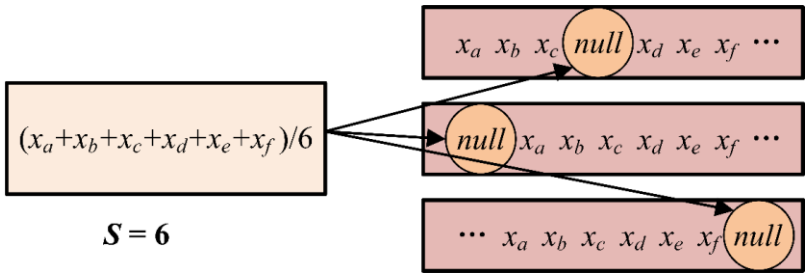

Fig. 12 Three cases of missing value

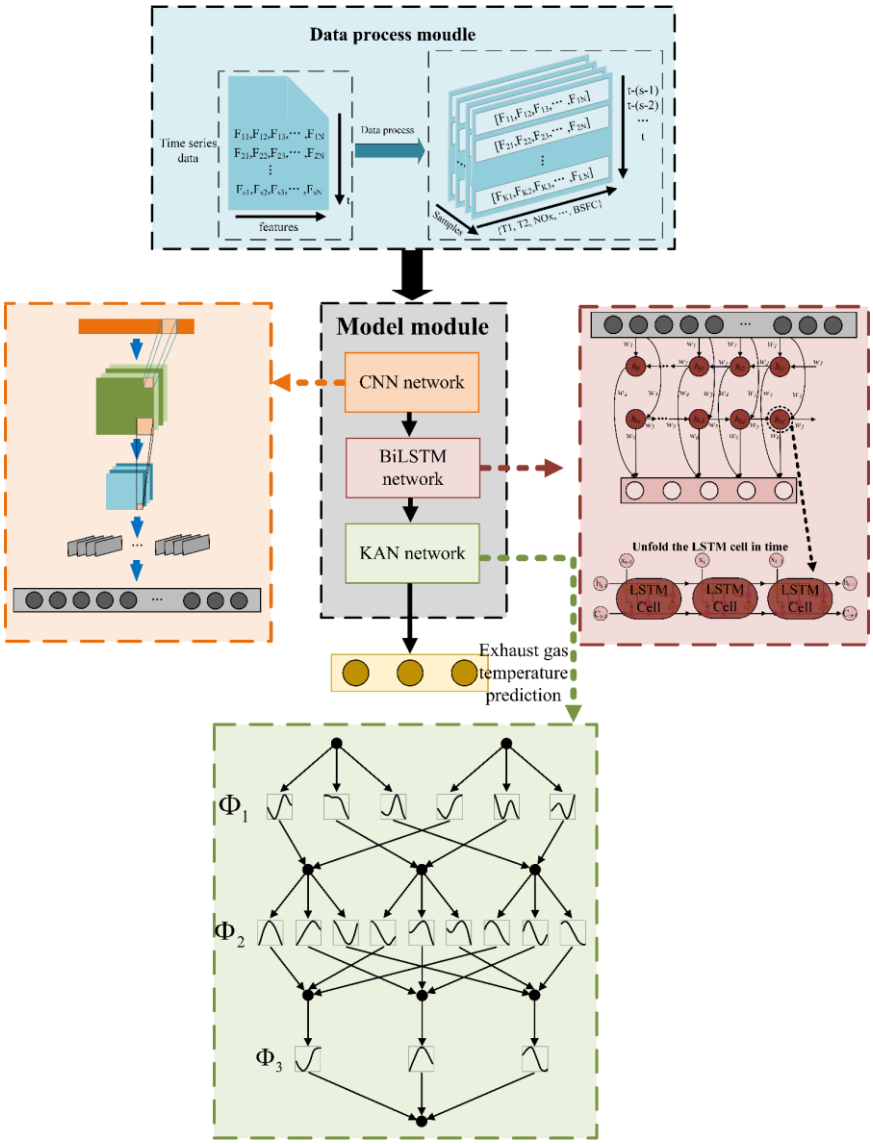

Fig. 13 Structure of the CNN-BiLSTM-KAN prediction model

3.2 Prediction model

The complexity requirements for fault warning of current marine dual fuel engines exceed the capabilities of traditional single neural networks. Therefore, a fault prediction method for marine dual-fuel engines was proposed in this paper based on the CNN-BiLSTM-KAN neural network. CNN can effectively capture local features in data, which are crucial for understanding the operational state of marine dual fuel engines based on their EGT. These local features lay a vital foundation for subsequent fault warnings. BiLSTM processes forward and backward temporal information simultaneously, enabling it to be well-suited for the time-series nature of EGT data. It can identify long-term dependencies within these sequences to enhance fault

prediction accuracy. The KAN layer directly learns parameterized nonlinear activation functions, further strengthening the expressiveness and generalization capabilities of the model. This enhancement also improves training efficiency and prediction accuracy. Overall, the CNN-BiLSTM-KAN model integrates the strengths of each component to effectively manage the complex EGT data from marine dual-fuel engines, to reinforce the accuracy and real-time performance of fault warnings.

The combustion process of a marine dual-fuel engine in gas mode is influenced by multiple factors, such as fuel properties, substitution rate, load, and engine speed. Variations in these factors trigger fluctuations in EGT, complicating the prediction process. Additionally, EGT data demonstrate clear time-series characteristics, with changes affected by not only current operating conditions but also historical states. These dynamic changes necessitate models that can capture long-term dependencies for accurate temperature forecasting. Abnormal variations in EGT indicate potential failures, highlighting the importance of real-time monitoring and precise prediction. Since the EGT patterns of dual-fuel engines vary under different operating conditions, the model must possess robust feature extraction capabilities. Moreover, the complex operating environment of ships demands high real-time and accuracy standards for fault warnings, requiring the model to quickly and accurately detect anomalies and issue timely alerts. Considering these unique challenges of EGT prediction, the CNN-BiLSTM-KAN model is chosen for this study. The structure diagram of the CNN-BiLSTM-KAN forecasting model is displayed in Figure 13.

3.3 Model evaluation indicators

The precision and efficacy of the model were evaluated, and various models were compared to determine which one performed better. Furthermore, these metrics can be used to gradually improve the performance of the model. Four commonly used prediction evaluation indicators were selected, including MSE, RMSE, MAE, and MAPE. A lower value of these metrics indicates higher predictive accuracy. MSE measures the fit of a model by calculating the mean of the sum of squares of the difference between predicted and true values. RMSE evaluates the precision of predictions by taking the square root of the mean of these squared discrepancies. MAE denotes the mean of the absolute value of the difference between forecasts and true values. MAPE reflects the average relative error as a percentage and measures the proportionate error between predictions and actuals, independent of the scale of absolute errors. The formula is:

$$\begin{cases} MSE = \dfrac{1}{N}\displaystyle\sum_{N=1}^{N}(\hat{Y}_i - Y_i)^2 \\ RMSE = \sqrt{\dfrac{1}{N}\displaystyle\sum_{N=1}^{N}(\hat{Y}_i - Y_i)^2} \\ MAE = \dfrac{1}{N}\displaystyle\sum_{N=1}^{N}\left|\hat{Y}_i - Y_i\right| \\ MAPE = \dfrac{1}{N}\displaystyle\sum_{N=1}^{N}\left|\dfrac{\hat{Y}_i - Y_i}{Y_i}\right| \end{cases} \tag{10}$$

where Y_i represents the actual value of EGT; $\hat{Y}_i$ denotes the forecast value of EGT; N refers to the number of samples in the test set.

3.4 Neural network hyperparameter selection

The neural network proposed in this paper has six input parameters, bringing about six neurons at the input level. Because the sole predicted output of the model is the EGT, there is a single neuron in the output layer. The architecture is composed of a CNN layer, a BiLSTM layer, and a KAN layer. Variables were controlled to refine the architecture of the BiLSTM model, to ascertain hyperparameters including the number of layers. When exploring the impact of model layers on performance, other parameters were kept constant and only the number of layers was adjusted to assess model efficacy. As suggested in the experimental results

in Table 6, all evaluation indicators reach their lowest when the BiLSTM model is set to 1 layer, demonstrating the best state. An increase in the number of layers leads to higher indicator values, suggesting overfitting. With the same approach, the optimal CNN-BiLSTM-KAN model was trained, with other parameter settings detailed in Table 7.

Table 6 Outcomes of predictions of BiLSTM with different numbers of network layers

BiLSTM network layers	MAE	RMSE	MSE	MAPE
1	0.001532	0.002034	0.000004	0.002843
2	0.001841	0.002449	0.000006	0.003437
3	0.002426	0.003025	0.000009	0.004526
4	0.002099	0.002759	0.000008	0.003919

Table 7 CNN-BiLSTM-KAN network structure parameters

Hyperparameter Name	Optimal parameter values
CNN layer	3
CNN Kernel	3
Activation function	ReLU
Pooling layer	Max Pool
LSTM layers	1
Input size	6
Hidden size	32
Sequence length	8
Batch size	100
Epochs	1000
Learning rate	0.005
Dropout	0.4
Loss	MSE loss
Optimizer	Adam

3.5 Prediction result analysis

In this study, the processed multidimensional data were input into the established forecast model. The predicted results are presented in Figure 14. As observed in Figure 14, the output EGT of the prediction model appears a small deviation from the actual EGT value, in line with the actual value. The evaluation indicators for the predicted results are provided in Table 8, specifying that the CNN-BiLSTM-KAN forecasting model has high accuracy and precision.

Table 8 Summary of the assessment of the predictive indicators of the combined model

Evaluating indicator	Optimal Model Evaluation Results
MAE	0.003185
RMSE	0.007135
MSE	0.00005091
MAPE	0.00038602

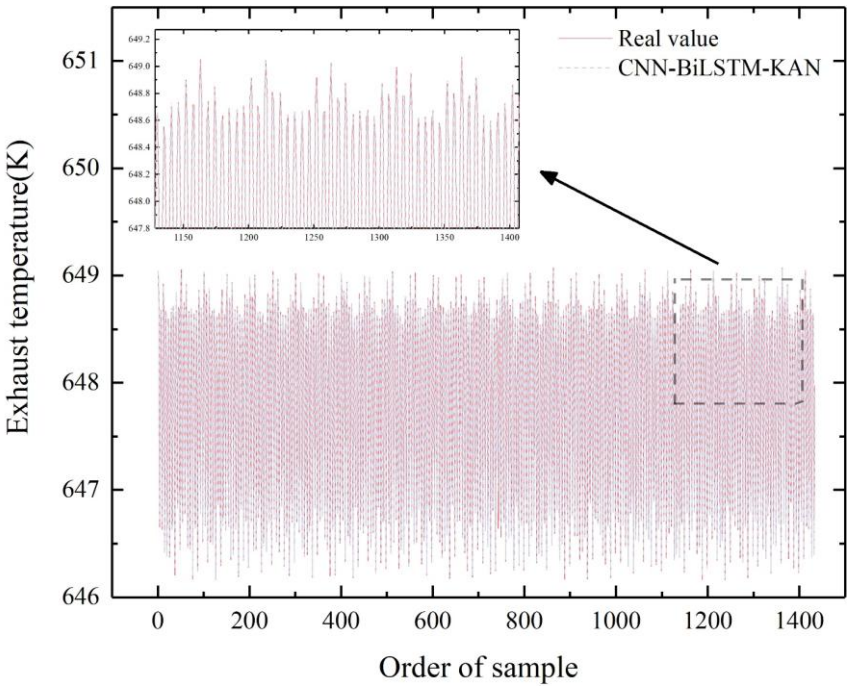

Fig. 14 CNN-BiLSTM-KAN prediction result graph

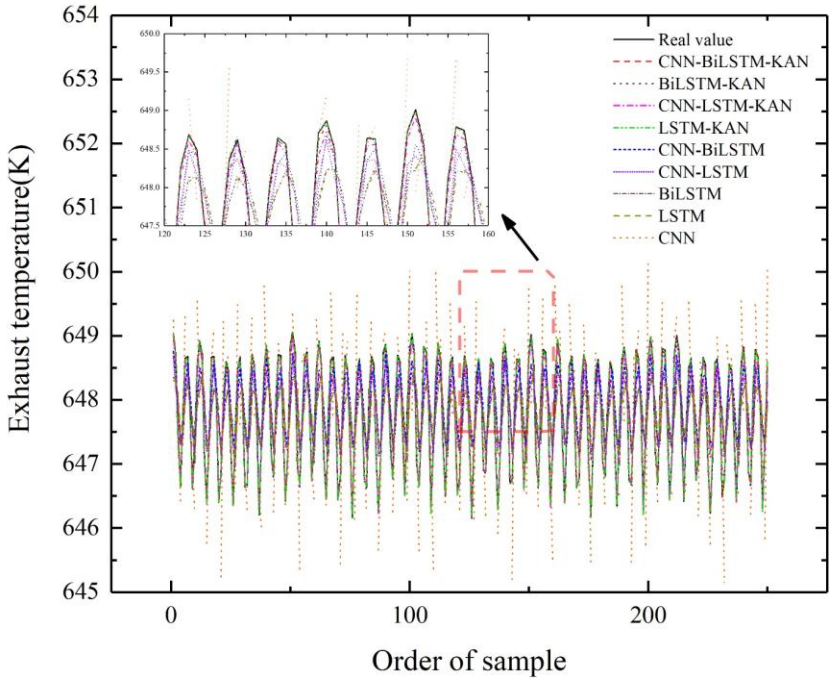

Fig. 15 Comparing the results of different models

The accuracy of the model was verified through experimental comparison to substantiate the efficacy and reliability of the combination forecast model introduced in this paper. The CNN-BiLSTM-KAN model was compared with eight alternative models: CNN, LSTM, BiLSTM, CNN-LSTM, CNN-BiLSTM, LSTM-KAN, BiLSTM-KAN, and CNN-LSTM-KAN. Before experimental comparisons, the same standard dataset after data processing, was taken as the input dataset for each model, and the hyperparameters for the other models were aligned with those of the CNN-BiLSTM-KAN forecast model presented in this paper. These hyperparameters were obtained in the proposed model by the control variable method. Moreover, the hyperparameters included in the other comparison models were consistent with those of the proposed model to ensure that the influence of other factors is avoided in the model comparison experiments. This allows for a more accurate assessment of the prediction accuracy and training efficiency of the various models. Figure 15 illustrates the comparison of the prediction outcomes of various models. Figure 16 exhibits a scatter

diagram of predicted and true values from different models, with RMSE, R2, and linear fitting functions calculated. Among them, the black dashed line is the diagonal line used as a reference, and the red line is the straight line obtained from the scatter plot.

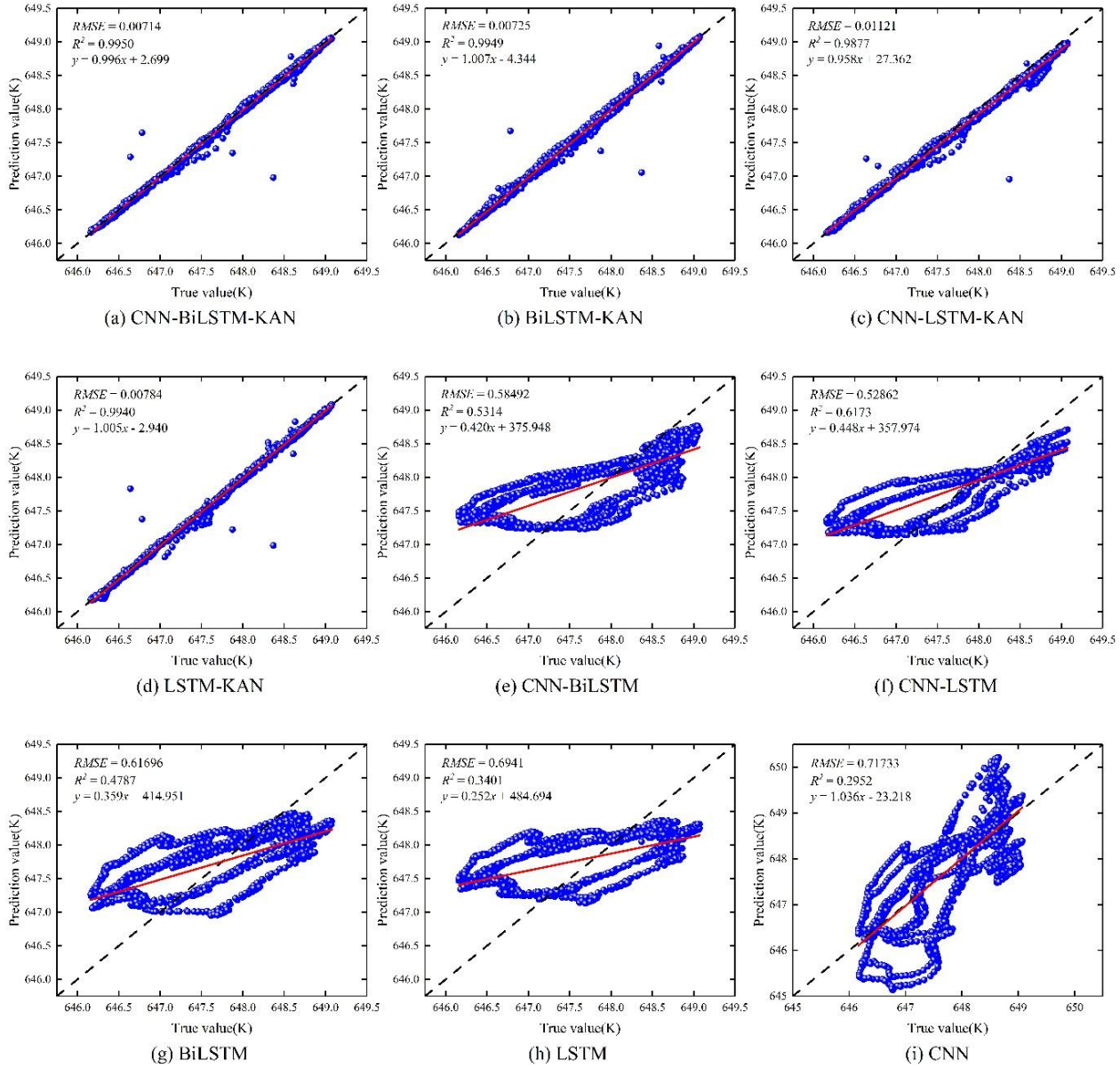

Fig. 16 Scatter diagram of predicted values and true values from different models

The evaluation indicators for different models are presented in Table 9.

Table 9 Overview of evaluation metrics for predictions across various models

Prediction model	MAE	RMSE	MSE	MAPE
CNN	0.584288	0.717325	0.514556	0.000902
LSTM	0.609975	0.694120	0.481803	0.000942
BiLSTM	0.543093	0.616961	0.380641	0.000838
CNN-LSTM	0.442076	0.528623	0.279442	0.000683
CNN-BiLSTM	0.478685	0.584917	0.342128	0.000739
LSTM-KAN	0.003254	0.007843	0.000615	0.000463
BiLSTM-KAN	0.003987	0.007251	0.000053	0.000616
CNN-LSTM-KAN	0.008530	0.011211	0.000126	0.000565
CNN-BiLSTM-KAN	0.003185	0.007135	0.000051	0.000386

Figure 15 suggests that the CNN-BiLSTM-KAN forecast model outperforms other prediction models. Figure 16 demonstrates that the CNN-BiLSTM-KAN forecasting model exhibits the highest correlation with actual values, the best-fitting line, and the smallest dispersion. As observed from Table 9, the CNN-BiLSTM-KAN model records the lowest error among the evaluation indices. In other words, the prediction results of the model have the smallest difference from the actual values and higher accuracy than other models. Furthermore, the residual box plot in Figure 17 reveals that the mean prediction error of all models in the experiment tends towards 0. Nonetheless, the CNN-BiLSTM-KAN model presents a markedly lower prediction error. In conclusion, CNN-BiLSTM-KAN has high predictive accuracy and significant advantages in time series data prediction, meeting the basic application of dual fuel engine EGT time series prediction. Particularly, the excellent feature extraction capability of the CNN layer, the good time series modeling capability of the BiLSTM layer, and the unique parametric nonlinear activation function of the KAN layer allow the proposed model to manage complex time series more comprehensively compared with other models. Moreover, the exhaust temperature of the marine dual-fuel engine is affected by various factors, and the model can automatically learn the relationship between these complex factors, making it more compatible with the experimental dataset.

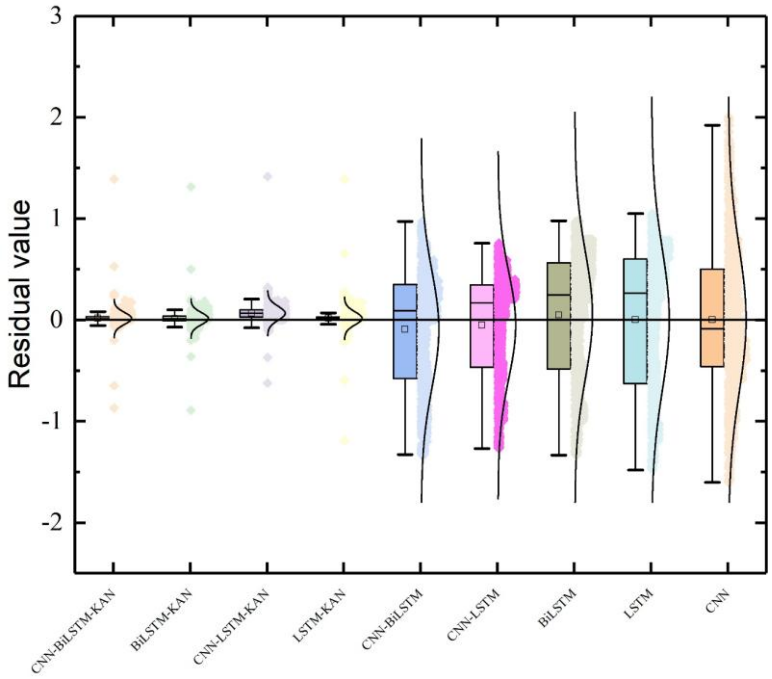

Fig. 17 Residual box line plots for different models

4. Research on fault warning

4.1 Warning indicator setting

The setting of warning indicators is a crucial aspect of ship fault warning research. Too high values of these indicators can impede timely alerts in the early stages of ship equipment failures. If the warning indicator has too low values, false alarms will occur when there are small fluctuations in equipment load.

Feature selection is performed in this paper to simulate the normal operation data of a dual-fuel ship engine simulation model built with simulation software. The five thermodynamic parameters exerting the highest impact on EGT are selected as inputs for the proposed model, and the predicted value of EGT for the dual-fuel engine is obtained. The proposed real-time monitoring method for EGT of dual fuel ship engines can achieve a rapid evaluation of the stability and accuracy of dual fuel ship engines, ensuring the realization of fault warnings for dual fuel ship engines. During normal operation, the EGT of a dual-fuel engine is generally maintained at a relatively constant level, enabling the forecast model to accurately predict the EGT and maintain a small prediction error. Nevertheless, once there is a potential malfunction in the engine, the severity of the malfunction will gradually increase with the extension of operating time, and the EGT will exceed the normal fluctuation range, leading to a decrease in the accuracy of the prediction model. Therefore, residual value distribution curves and statistical methods are adopted to establish the upper and lower bounds for residual values [51]. Since few residual values deviate significantly beyond the set threshold range, the sliding window method is utilized to obtain the standard deviation of each window residual distribution, and a threshold upper limit for the standard deviation is established. The warning index for EGT of dual fuel ship engines is set following two calculated thresholds. The standard deviation of a sliding window is calculated by:

$$S = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(e_i - \overline{X})^2} \tag{11}$$

where N denotes the sliding window's size; e_i signifies the residual value at position i of the window; $\overline{X}$ indicates the average of all residual values within the window.

Figure 18 displays the residual distribution curve derived from the prediction outcomes of the simulated normal operating data through the simulation model. As observed in Figure 18, the residuals generally vary between -0.1 and 0.15 when predictions are compared to actual values. The residual numbers of each region in Table 10 unveil that the cumulative percentage of residual numbers within the range of -0.1 to 0.15 is over 99 %. Thus, under normal conditions, there is a 99 % probability that the discrepancy between predicted and actual values falls within the range of -0.1 to 0.15.

Consequently, the upper threshold and lower threshold for the alert indicator are established at 0.15 and -0.1, respectively, as depicted in Figure 19. The figure demonstrates that most of the residual values are included within the established boundary range, whereas a small portion exceeds the range, and a very small number is far from the range. This difference may stem from the model itself, leading to false alarms. Standard deviation serves as a widely recognized metric for quantifying the dispersion within a dataset. Hence, sliding windows are employed to examine the residual distribution and compute the standard deviation of these residuals. Following practical knowledge, a window size of 30 is designated for calculating the standard deviation of residual values. Figure 20 illustrates the trend of the residual standard deviation from this processing. The highest value of standard deviation is 0.343 when the dual-fuel ship engine operates under stable conditions. Therefore, the residual alert range for dual fuel ship engines is -0.1~0.15, and the alert threshold for residual standard deviation is 0.343. The alarm will only be triggered when both alert values are surpassed simultaneously.

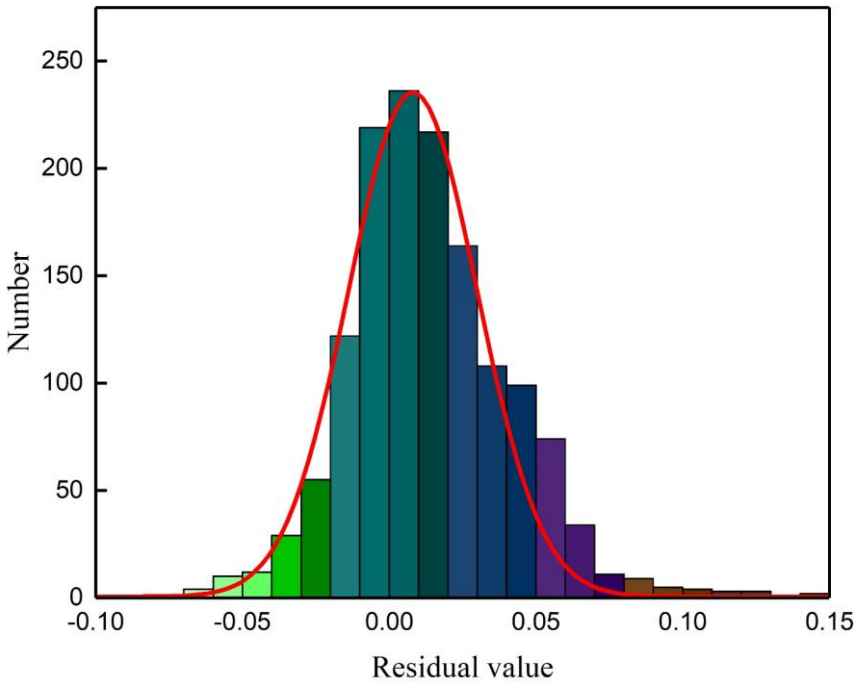

Fig. 18 Residual Distribution Chart

Table 10 Number and cumulative percentage of residual values in different intervals

Interval	Quantity	Cumulative Percentage	Interval	Quantity	Cumulative Percentage
$-0.3 \leq R < -0.25$	0	0	$0 \leq R < 0.05$	824	89.05923
$-0.25 \leq R < -0.2$	1	0.06969	$0.05 \leq R < 0.1$	133	98.32753
$-0.2 \leq R < -0.15$	1	0.13937	$0.1 \leq R < 0.15$	12	99.16376
$-0.15 \leq R < -0.1$	0	0.13937	$0.15 \leq R < 0.2$	4	99.44251
$-0.1 \leq R < 0.05$	15	1.18467	$0.2 \leq R < 0.25$	3	99.65157
$-0.05 \leq R < 0$	437	31.63763	$0.25 \leq R < 0.3$	1	99.72125

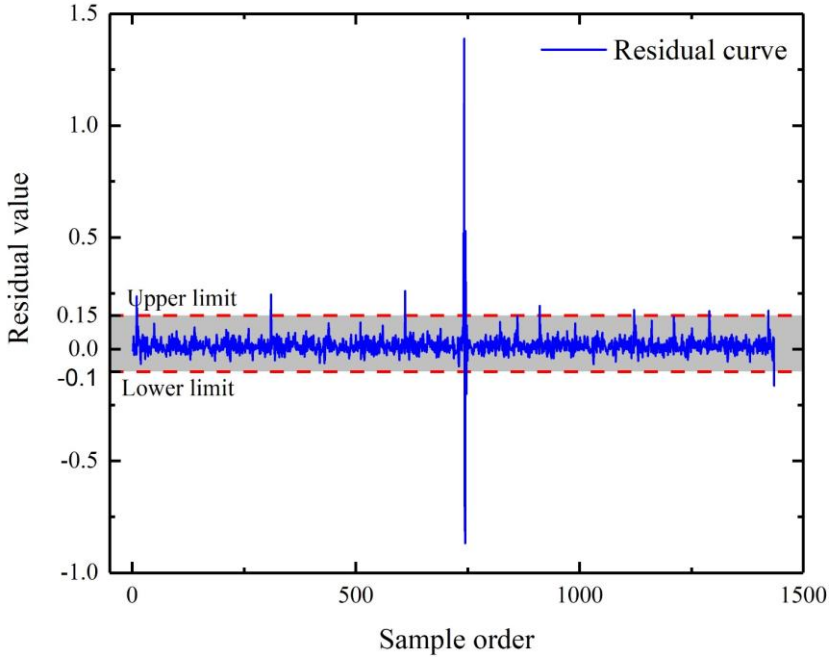

Fig. 19 Residual value threshold setting

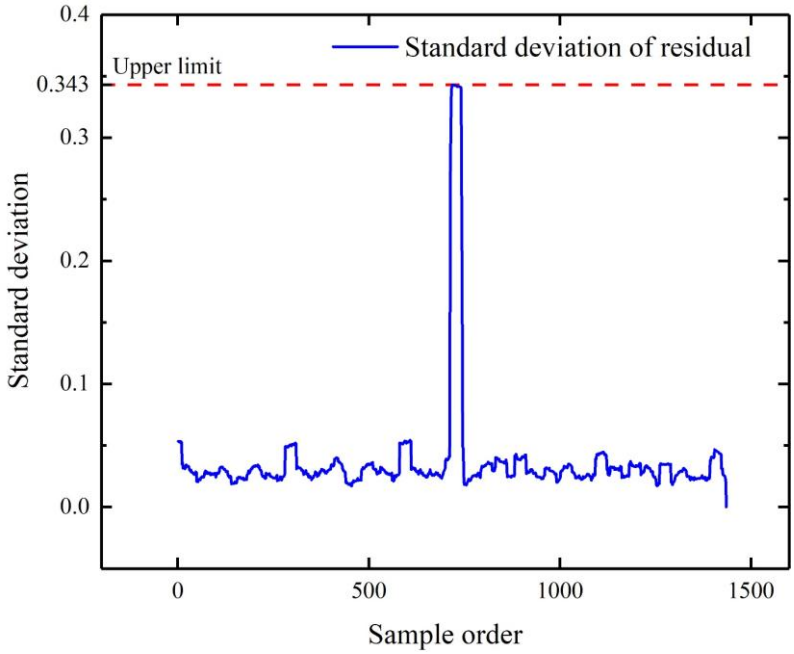

Fig. 20 Standard deviation curve of residuals

4.2 Verification of warning function

Fault data is required to substantiate the efficacy of the proposed fault alert methodology. The typical faults of dual-fuel engines primarily include fuel supply pump failure, gas supply delay failure, and gas filter blockage failure. In this paper, gas supply delay failure is selected to validate the effectiveness of the suggested approach. The fault data can be obtained by simulating the gas supply delay fault in the built dual fuel engine simulation model. The specific operation is performed to modify the parameters of the crankshaft angle Injection-timing-Gas corresponding to the gas supply for simulating the actual situation of gas supply delay.

Figure 21 demonstrates that all residual values are within the set range before sample point 1321. After the introduction of fault data at sample point 1321, the residual values begin to increase sharply, exceeding the set range. Figure 22 reveals that the residual value is within the set range, even though the standard deviation of the residuals has exceeded the set threshold at sample point 1261. Hence, the fault warning will not appear until sample point 1321. After a fault is set at sample point 1321, the residual value quickly exceeds the threshold, whereas the standard deviation has exceeded the set threshold at sample point 1261. This reflects that the dual fuel engine may have malfunctioned, resulting in a fault warning at sample point 1321. Therefore, the fault warning approach outlined in this paper can not only meet the real-time surveillance of EGT of the marine dual fuel engine but also issue alerts for anticipated EGT anomalies when the marine dual fuel engine malfunctions, satisfying the contemporary requirements for EGT fault warning in marine dual-fuel engines.

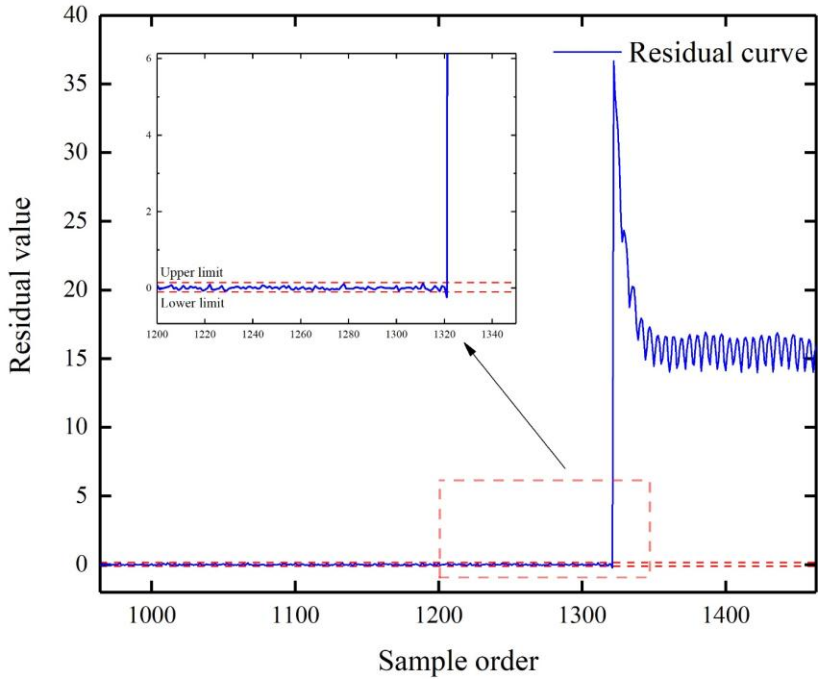

Fig. 21 Residual curve of abnormal working condition

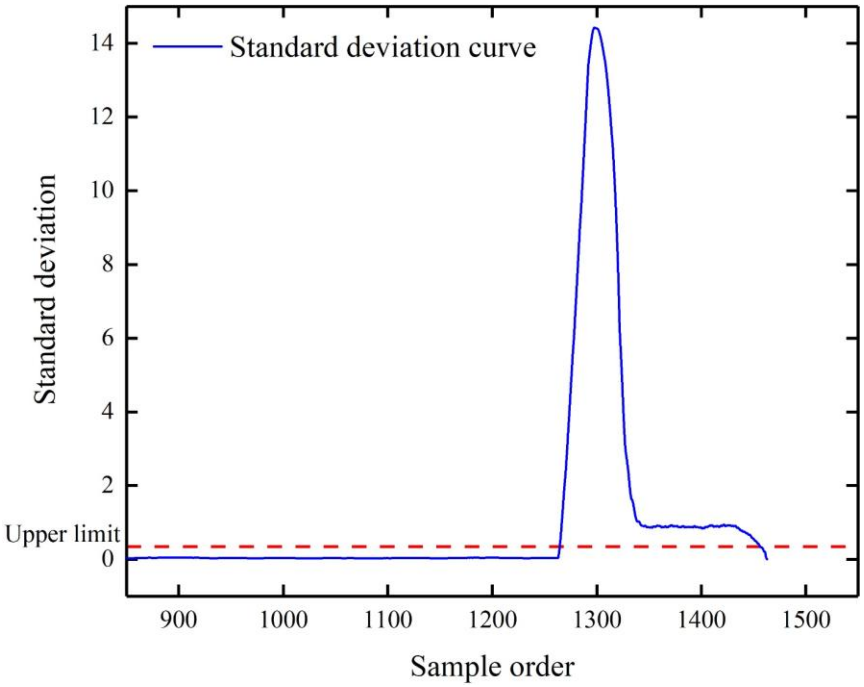

Fig. 22 Standard deviation curve of abnormal working condition

5. Conclusion

In this study, a CNN-BiLSTM-KAN prediction model was established to predict the EGT of marine dual-fuel engines. The crucial conclusions are summarized as follows.

With PCCs, thermodynamic parameters highly correlated with EGT were selected as inputs for the prediction model, and the input data were normalized to reduce the dimensionality of the input data, simplify the architecture of the model, decrease computation time, and improve model prediction accuracy. Simultaneously, Mahalanobis distance was adopted to screen outliers and mitigate their detrimental effects on the model's computational efficiency and accuracy.

The CNN-BiLSTM-KAN forecasting model proposed integrates CNN, BiLSTM, and KAN. The CNN is adept at extracting features from multidimensional time series, the BiLSTM automatically learns, and

extracts features from these series, and the KAN layer directly learns parameterized nonlinear activation functions. This method surpasses the constraints of previous single neural networks, and the inclusion of the KAN layer improves the predictive accuracy of the neural network. In this way, the spatiotemporal characteristics of EGT of marine dual fuel engines can be more comprehensively extracted, contributing to better prediction results.

With the purpose of demonstrating the superiority of the proposed model, a comparative experiment was conducted to compare the predictive model with eight prediction models: CNN, LSTM, BiLSTM, CNN-LSTM, CNN-BiLSTM, LSTM-KAN, BiLSTM-KAN, and CNN-LSTM-KAN. As unveiled from the evaluation metrics, the CNN-BiLSTM-KAN model constructed in this study demonstrates high accuracy, highlighting its strengths in time series forecasting.

By analyzing the predicted residual values, alarm thresholds are set for both the residuals and their standard deviation. Experimental validation and analysis confirm that the approach can timely identify anomalies in marine dual fuel engines while enlightening their state detection and health management.

To encapsulate, this study aimed to investigate the marine dual-fuel engine, analyze its exhaust temperature by deep learning techniques, and hence achieve fault prediction and early warning capabilities. Deep learning was introduced into the condition prediction of marine equipment, offering a novel approach for its condition monitoring, health management, and the development of intelligent ships. The innovative contributions of this paper are twofold. First, it offers a novel perspective in selecting the research object. Different from most prior studies that focus on marine diesel engines, the gas mode of marine dual fuel engines - a key distinction from other marine engines, was examined in this study. This choice holds significant value for the development of marine dual-fuel ships. Second, a new prediction model architecture was constructed based on existing research. The experimental results demonstrate that the proposed model can effectively predict the state and trigger fault alarms for marine dual-fuel engines, providing insights into building fault warning models for these engines. In the future, our research group will concentrate on researching and analyzing the multi parameters of marine dual fuel engines. Furthermore, the model functions and fault warning method will be optimized and improved to achieve fault prediction classification and provide support for the development of intelligent ships.

ACKNOWLEDGEMENTS

This research was supported by the National Key R&D Program of China (grant number 2022YFB4301400), and the High-technology Ship Research Program (grant number CBG3N21-3-3).

REFERENCES

[1] Zhang, H., Fei, T., Guan, W., Zhang, S., Jin, Z., Tang, W., 2019. Research on visual 3D assembly process design and simulation for marine diesel engine. *Cluster Computing-the Journal of Networks Software Tools and Applications*, 22, pp. S5505-S5519. https://doi.org/10.1007/s10586-017-1342-1

[2] Karatu, C., Arslanoglu, Y., 2022. Development of condition-based maintenance strategy for fault diagnosis for ship engine systems. *Ocean Engineering*, 256, 111515. https://doi.org/10.1016/j.oceaneng.2022.111515

[3] Tuswan, T., Sari, D. P., Muttaqie, T., Prabowo, A. R., Soetardjo, M., Murwantono, T. T. P., Utina, R., Yuniati, Y., 2023. Representative application of LNG-fuelled ships: a critical overview on potential GHG emission reductions and economic benefits. *Brodogradnja*, 74(1), pp. 63-83. https://doi.org/10.21278/brod74104

[4] Poór, P., Basl, J., Zenisek, D., 2019. Predictive Maintenance 4.0 as next evolution step in industrial maintenance development. *In 2019 International research conference on smart computing and systems engineering (SCSE)*, March 28, Colombo, Sri Lanka. https://doi.org/10.23919/SCSE.2019.8842659

[5] Jimenez, V. J., Bouhmala, N., Gausdal, A. H., 2020. Developing a predictive maintenance model for vessel machinery. *Journal of Ocean Engineering and Science*, 5(4), pp. 358-386. https://doi.org/10.1016/j.joes.2020.03.003

[6] Che, C. C., Wang, H. W., Fu, Q., Ni, X. M., 2021. Intelligent fault prediction of rolling bearing based on gate recurrent unit and hybrid autoencoder. *Proceedings of the Institution of Mechanical Engineers Part C-Journal of Mechanical Engineering Science*, 235(6), pp. 1106-1114. https://doi.org/10.1177/0954406220941037

[7] Zou, F. K., Zeng, H., Wang, H. Y., Wang, X. X., Xu, Z. X., 2021. Implementation and Parameter Analysis of the Knock Phenomenon of a Marine Dual-Fuel Engine Based on a Two-Zone Combustion Model. *Processes*, 9(4), 602. https://doi.org/10.3390/pr9040602

[8] Gan, J., Jin, H. B., Shang, Q. M., Sheng, C. X., 2024. A New Method of Intelligent Fault Diagnosis of Ship Dual-Fuel Engine Based on Instantaneous Rotational Speed. *Journal of Marine Science and Engineering*, 12(11), 2046. https://doi.org/10.3390/jmse12112046

[9] Xu, X. A., Lin, Y., Ye, C., 2024. Real-time prediction for temperature distribution on the cylinder head of dual-fuel engines via a novel deep learning framework. *Expert Systems with Applications*, 238, 122357. https://doi.org/10.1016/j.eswa.2023.122357

[10] Stoumpos, S., Theotokatos, G., 2022. A novel methodology for marine dual fuel engines sensors diagnostics and health management. *International Journal of Engine Research*, 23(6), pp. 974-994. https://doi.org/10.1177/1468087421998635

[11] Jiang, Y., Cukic, B., Ma, Y., 2008. Techniques for evaluating fault prediction models. *Empirical Software Engineering*, 13, pp. 561-595. https://doi.org/10.1007/s10664-008-9079-3

[12] Nan, C., Khan, F., Iqbal, M. T., 2008. Real-time fault diagnosis using knowledge-based expert system. *Process safety and environmental protection*, 86(1), pp. 55-71. https://doi.org/10.1016/j.psep.2007.10.014

[13] Xian, Z. Z., Jian, K., Hao, L., Li, W., 2010. Overview of fault prediction technology. *Fire Control Command Control*, 35(1), pp. 1-5.

[14] Tang, J. X., Liu, B. L., Fan, W. H., Zhong, D. W., Liu, L., 2024. Degradation Diagnosis and Control Strategy for a Diesel Hybrid Powertrain Considering State of Health. *Energies*, 17(21), 5413. https://doi.org/10.3390/en17215413

[15] Ding, Q. Q., Tao, M., Li, R. L., 2013. Study on the Method of Radar Servo System Faults Diagnosis Based on Expert System and Neural Network. *Advanced Materials Research*, 787, pp. 959-964. https://doi.org/10.4028/www.scientific.net/AMR.787.959

[16] Xiao, Y. J., Han, F. R., Ding, H., Liu, W. L., 2021. Research on fault diagnosis method of rapier loom based on the fusion of expert system and fault tree. *Journal of Intelligent & Fuzzy Systems*, 41(2), pp. 3429-3441. https://doi.org/10.3233/JIFS-210741

[17] Caliwag, A., Han, S. Y., Park, K. J., Lim, W., 2022. Deep-Learning-Based Fault Occurrence Prediction of Public Trains in South Korea. *Transportation Research Record*, 2676(4). pp. 710-718. https://doi.org/10.1177/03611981211064893

[18] Kusiak, A., Li, W. Y., 2011. The prediction and diagnosis of wind turbine faults. *Renewable Energy*, 36(1), pp. 16-23. https://doi.org/10.1016/j.renene.2010.05.014

[19] Abdelghany, E. S., Mohamed, E. S., Sarhan, H. H., 2023. Exhaust heat recovery performance analysis of a Bi-fuel engine utilizing a thermoelectric generation kit and fuel economy evaluation. *Case Studies in Thermal Engineering*, 49, 103288. https://doi.org/10.1016/j.csite.2023.103288

[20] Kumar, A., Srivastava, A., Goel, N., McMaster, J., 2015. Exhaust gas temperature data prediction by autoregressive models. *In IEEE 28th Canadian conference on electrical and computer engineering (CCECE)*, May 3-6, Halifax, Canada. https://doi.org/10.1109/CCECE.2015.7129408

[21] Zhang, Y. F., Liu, P. P., He, X., Jiang, Y. P., 2020. A prediction method for exhaust gas temperature of marine diesel engine based on LSTM. *In 2nd IEEE International conference on civil aviation safety and information technology (ICCASIT)*, October 14-16, Wuhan, China. https://doi.org/10.1109/ICCASIT50869.2020.9368826

[22] Ozsari, I., 2023. Predicting main engine power and emissions for container, cargo, and tanker ships with artificial neural network analysis. *Brodogradnja*, 74(2), pp. 77-94. https://doi.org/10.21278/brod74204

[23] Yao, J. C., Lu, B. C., Zhang, J. L., 2022. Tool remaining useful life prediction using deep transfer reinforcement learning based on long short-term memory networks. *International Journal of Advanced Manufacturing Technology*, 118(3-4), pp. 1077-1086. https://doi.org/10.1007/s00170-021-07950-2

[24] Sahu, P. K., Rai, R. N., 2024. LSTM-based deep learning approach for remaining useful life prediction of rolling bearing using proposed C-MMPE feature. *Journal of Mechanical Science and Technology*. 38, pp. 2197–2209. https://doi.org/10.1007/s12206-024-0402-8

[25] Lin, K. C., Hsu, J. Y., Wang, H. W., Chen, M. Y., 2024. Early fault prediction for wind turbines based on deep learning. *Sustainable Energy Technologies and Assessments*, 64, 103684. https://doi.org/10.1016/j.seta.2024.103684

[26] Batool, I., Khan, T. A., 2023. Software fault prediction using deep learning techniques. *Software Quality Journal*, 31(4), pp. 1241-1280. https://doi.org/10.1007/s11219-023-09642-4

[27] Bazai, H., Kargar, E., Mehrabi, M., 2021. Using an encoder-decoder convolutional neural network to predict the solid holdup patterns in a pseudo-2D fluidized bed. *Chemical Engineering Science*, 246, 116886. https://doi.org/10.1016/j.ces.2021.116886

[28] Bai, S. J., Kolter, J. Z., Koltun, V., 2018. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*, 10.

[29] Huang, B. Q., Zheng, H. N., Guo, X. B., Yang, Y., Liu, X. M., 2022. A Novel Model Based on DA-RNN Network and Skip Gated Recurrent Neural Network for Periodic Time Series Forecasting. *Sustainability*, 14(1), 326. https://doi.org/10.3390/su14010326

[30] Bendali, W., Saber, L., Boussetta, M., Bourachdi, B., Mourad, Y., 2022. Optimization of deep reservoir computing with binary genetic algorithm for multi-time horizon forecasting of power consumption. *Journal Europeen des Systemes Automatises*, 55(6), pp. 701. https://doi.org/10.18280/jesa.550602

[31] Jiang, C. M., Tang, Y. M., Wang, J. G., Zhang, W. C., Zhou, M., Niu, J. Y., Wan, L., Chen, G. F., Wu, G. X., Cheng, X. D., 2024. An optimized method for AUV trajectory model in benthonic hydrothermal area based on improved slime mold algorithm. *Brodogradnja*, 75(4). https://doi.org/10.21278/brod75401

[32] Zhang, L. J., Yang, S. J., Wang, S., Zeng, T. M., Hua, W. C., Li, G. L., 2023. High-speed bearing dynamics and applications in production lines. *International Journal of Simulation Modelling*, 22(4), pp. 701-711. https://doi.org/10.2507/IJSIMM22-4-CO17

[33] Dashti, R., Daisy, M., Mirshekali, H., Shaker, H. R., Aliabadi, M. H., 2021. A survey of fault prediction and location methods in electrical energy distribution networks. *Measurement*, 184, 109947. https://doi.org/10.1016/j.measurement.2021.109947

[34] Ren, L., Dong, J. B., Wang, X. K., Meng, Z. H., Zhao, L., Deen, M. J., 2021. A Data-Driven Auto-CNN-LSTM Prediction Model for Lithium-Ion Battery Remaining Useful Life. *IEEE Transactions on Industrial Informatics*, 17(5), pp. 3478-3487. https://doi.org/10.1109/TII.2020.3008223

[35] Chen, L., Yang, Y. Y., Wang, Z. H., Zhang, J., Zhou, S. W., Wu, L. H., 2023. Underwater Target Detection Lightweight Algorithm Based on Multi-Scale Feature Fusion. *Journal of Marine Science and Engineering*, 11(2), 320. https://doi.org/10.3390/jmse11020320

[36] Zhang, X. L., Liang, W., Bai, J. L., Luo, X., Wang, F, Z., 2024. Intelligent fault diagnosis of roller bearing based on dual-flow convolutional neural network. *Proceedings of the Institution of Mechanical Engineers Part C-Journal of Mechanical Engineering Science*, 238(17), pp. 8889-8902. https://doi.org/10.1177/09544062241242595

[37] Wazery, Y. M., Saleh, M. E., Ali, A. A., 2023. An optimized hybrid deep learning model based on word embeddings and statistical features for extractive summarization. *Journal of King Saud University-Computer and Information Sciences*, 35(7), 101614. https://doi.org/10.1016/j.jksuci.2023.101614

[38] Yin, X., Liu, Q. S., Huang, X., Pan, Y. C., 2021. Real-time prediction of rockburst intensity using an integrated CNN-Adam-BO algorithm based on microseismic data and its engineering application. *Tunnelling and Underground Space Technology*, 117, 104133. https://doi.org/10.1016/j.tust.2021.104133

[39] Xu, C. Y., Wang, H., 2022. Research on a Convolution Kernel Initialization Method for Speeding Up the Convergence of CNN. *Applied Sciences*, 12(2), 633. https://doi.org/10.3390/app12020633

[40] Chen, Z. S., Zhang, L. K., Hua, J. M., Kim, B., Li, K., Xue, X. Y., 2023. A framework of data-driven wind pressure predictions on bluff bodies using a hybrid deep learning approach. *Measurement & Control*, 56(1-2), pp. 237-256. https://doi.org/10.1177/00202940221099064

[41] Mishra, A., Tripathi, K., Gupta, L., Singh, K. P., 2017. Long Short-Term Memory Recurrent Neural Network Architectures for Melody Generation. In *Proceedings of the 7th International conference on soft computing for problem solving (SocProS)*, Indian Institutes of Technology, December 23-24, Bhubaneswar, India.

[42] Lu, X. Y., Lu, C., Yu, W. X., Qiao, L., Liang, S. Y., Lau, A. P. T., Chi, N., 2019. Memory-controlled deep LSTM neural network post-equalizer used in high-speed PAM VLC system. *Optics Express*, 27(5), pp. 7822-7833. https://doi.org/10.1364/OE.27.007822

[43] Wang, X. R., Wang, J. C., Cao, Y. Q., Ye, Z. Y., Zhang, W. L., 2022. Principle Analysis of Voiceprint Identification and Authentication System Based on LSTM. *Academic Journal of Computing & Information Science*, 5(7), pp. 85-89. https://doi.org/10.25236/AJCIS.2022.050714

[44] Yang, H. F., Hu, J. J., Cai, J. H., Wang, Y. P., Chen, X., Zhao, X. J., Wang, L. L., 2023. A New MC-LSTM Network Structure Designed for Regression Prediction of Time Series. *Neural Processing Letters*, 55(7), pp. 8957-8979. https://doi.org/10.1007/s11063-023-11187-3

[45] Shi, H. F., Miao, K., Ren, X. C., 2023. Short-term load forecasting based on CNN-BiLSTM with Bayesian optimization and attention mechanism. *Concurrency and Computation-Practice & Experience*, 35(17). https://doi.org/10.1002/cpe.6676

[46] De, L. M., Ballesteros, M., Nivre, J., Assoc Computat, L., 2019. Recursive Subtree Composition in LSTM-Based Dependency Parsing. *In Proceedings of the 2019 Annual conference of the North American chapter of the association for computational linguistics (NAACL-HLT)*, June 2-7, Minneapolis, Minnesota, USA.

[47] Li, X. T., Wang, H. L., Xiu, P. F., Zhou, X. Y., Meng, F. H., 2022. Resource Usage Prediction Based on BILSTM-GRU Combination Model. *In 13th IEEE International conference on joint cloud computing (JCC)*, August 15-18, Fremont, California, USA. https://doi.org/10.1109/JCC56315.2022.00009

[48] Liu, Z. M., Wang, Y. X., Sachin, V., Fabian, R., James, H., Marin, S., Thomas, Y. H., Max, T., 2024. Kan: Kolmogorov-arnold networks. *arXiv preprint arXiv*, 2404.19756.

[49] Mestemaker, B. T. W., Castro, M. B. G., van den Heuvel, H. N., Visser, K., 2020. Dynamic simulation of a vessel drive system with dual fuel engines and energy storage. *Energy*, 194, 116792. https://doi.org/10.1016/j.energy.2019.116792

[50] Li, Z. Q., Wang, L. S., Wang, D, D., 2021. Analysis of music similarity based on Pearson correlation coefficient. *Art and Performance Letters*, 2(5).

[51] Wang, F. F., Xiao, X. Q., Zhao, H. S., 2011. Wind turbine gearbox failure prediction based on time series analysis and statistical process control. *In Proceedings of the International conference on energy, environment and sustainable development (ICEESD 2011)*, March 29-31, Manila, Philippines.