This is the author's accepted manuscript. The final version is available at 10.1016/j.measurement.2025.117252

Thermodynamic Simulation-assisted Random Forest: Towards explainable fault diagnosis of combustion chamber components of marine diesel engines

Congcong Luo[a], Minghang Zhao[a,*], Xuyun Fu[a], Shisheng Zhong[a], Song Fu[b], Kai Zhang[c], Xiaoxia Yu[d]

[a]*Department of Mechanical Engineering, Harbin Institute of Technology, Weihai 264209, China*
[b]*School of Mechatronics Engineering, Harbin Institute of Technology, Harbin 150001, China*
[c]*School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, China*
[d]*College of Mechanical Engineering, Chongqing University of Technology, Chongqing 400054, China*

Abstract

Aiming at the challenges that traditional intelligent fault diagnosis methods of marine diesel engines often suffer low generalizability due to the lack of fault training samples, as well as poor explainability due to the insufficient incorporation of domain knowledge on fault mechanism, this paper develops a Thermodynamic Simulation-assisted Random Forest (TSRF), which reveals fault characteristics through thermodynamic simulations and incorporates them as prior knowledge when designing the intelligent fault diagnosis model. Firstly, five thermodynamic fault models are developed by fine-tuning the essential system parameters to correspond with the distinct attributes of different faults. Then, potential thermodynamic indicators of combustion chamber component degradation are identified through numerical simulation results. By calculating SHapley Additive exPlanations (SHAP) values, a parameter selection process is conducted to retain only those variables demonstrating significant correlations with fault states. Finally, the selected parameters are leveraged to assess the condition of the combustion chamber and input into the fault diagnosis model. The proposed TSRF achieved exceptional classification performance, illustrating a mean accuracy of 99.07% on the fault dataset constructed in this paper. The estimation results of the model are interpreted from the local and global perspectives based on SHAP values. As a result, turbocharger exhaust temperature, blow-by heat flow, and cylinder liner heat flow are found to contribute significant to fault predictions.

Keywords: marine diesel engine, thermodynamic model, fault identification, Random Forest, explainability.

---

1. Introduction

The diesel engine is a critical component of the ship's propulsion system, with complicated internal structures and numerous interrelated parts [1]. In particular, the combustion chamber is crucial to engine functionality, and it comprises of the cylinder head, cylinder liner, piston, and other components [2]. Given its harsh operational environment, the combustion chamber is one of the most vulnerable sections in diesel engines [3–5]. Combustion chamber faults may adversely damage engine performance and safety, making fault detection vital for engine management. By monitoring thermodynamic parameters, diagnostic models could identify and locate defects, giving crucial information for maintenance scheduling, repair plan optimization, and cost control

---

[*] Corresponding author.
E-mail address: zhaomh@hit.edu.cn (M. Zhao).

[6–9]. However, diagnosing faults in practical ship operations remains challenging due to factors such as limited fault data and nonquantitative prior knowledge.

Currently, fault diagnosis methods for marine diesel engines are primarily categorized into three types, as presented in Table 1: : model-based, data-driven, and hybrid approaches [9–11]. In the model-based methods, a mathematical model of the diesel engine is constructed by merging physical laws, engineering knowledge, and operational principles. Then, the simulation software is applied to evaluate the model's performance under normal and faulty conditions. For example, He et al. [12] constructed a turbocharger efficiency model based on a thermodynamic mechanism and applied the T^2 test statistic for fault detection in diesel engines. Ricardo et al. [13] proposed thermodynamic fault models for the intake system, injection system, compression ratio, and fuel injection, applying the Levenberg-Marquardt method for parameter estimation to diagnose faults. Antonio et al. [14] developed a fault simulator based on a one-dimensional thermodynamic model of a diesel engine to accurately identify the symptoms in various fault states. Tang et al. [15] proposed a two-stroke marine diesel engine model to predict in-cylinder pressure in real-time and simulated the behavior under a single-cylinder fault condition. Moussa et al. [16] investigated a marine diesel engine model based on physical, semi-physical, mathematical, and thermodynamic equations, achieving the simulation of various faults. Lamaris et al. [17] developed a thermodynamic-based multi-area combustion model and a fault detection approach that leverages changes in model constants, which were applied to the fault diagnosis of the main propulsion and auxiliary diesel engine units of ships. However, constructing an accurate physical-mathematical model for marine diesel engines is exceedingly problematic under current conditions. Such models are often sophisticated and involve substantial subject expertise, encountering issues dealing with nonlinearities and interactions between various variables and environmental factors.

Data-driven methods bypass the need for detailed physical modeling of diesel engines. Instead, pre-processed data is input into a machine learning model, enabling it to learn fault characteristics and identify potential faults autonomously. For example, Hu et al. [18] processed acoustic emission signals through a one-dimensional CNN to achieve fault diagnosis of diesel engine exhaust valves with different degrees of blow-by. Yuankai et al. [19] proposed a Recurrence Plot and CNN-based approach (RP-CNN) for identifying the wear mechanisms of main bearings in marine diesel engines. Guo et al. [20] proposed a multi-scale and multi-view domain adversarial network (MMDAN) for identifying various fault types in diesel engines. Wang et al. [21] designed a Random Convolutional Neural Network (RCNN), a structure capable of automatically extracting discriminative features from vibration signals to identify the health states of diesel engines accurately. Wang et al. [22] proposed a novel hybrid fault monitoring method that integrates flow learning with isolation forests, enabling the model to achieve more accurate diagnostic results even with minor sample conditions. However, data-driven models rely on learning features from data rather than explicit physical laws, which makes it challenging to explain their decision-making rationale—especially in complicated models like deep learning. The intricate non-linear relationships and high-dimensional feature interactions often render these models opaque, turning them into black boxes where the reasoning behind decisions is difficult to interpret.

The hybrid approaches generally begin with a physical-mathematical fault model, then its calibration using experimental data. Diagnosis is then conducted based on simulation output or experimental and simulated data. For example, Yu et al. [23] constructed a digital physical model based on virtual prototyping and multi-body dynamics. By applying a data-driven approach, they extracted and selected features from actual and simulated signals to detect planetary gearbox faults via feature matching. Gryllias et al. [24] proposed a hybrid two-stage one-against-all support vector machine method based on physical model training to diagnose rolling bearing faults automatically. Li et al. [25] developed a transfer learning approach based on deep learning, which trains models using simulated data and applies them to fault diagnosis in real-world physical processes. Yan [26] proposed a generative adversarial network model incorporating a physical model for fault detection and diagnosis in air handling units. However, these approaches lack a detailed analysis of the decision-making process of models when integrated with the physical model. Instead, the physical models primarily serve as tools for fault simulation and sources of complementary data.

Table 1 Summary of previous studies

Method Type	Advantages	Limitations
Model-Based Methods	1. Clear physical interpretation. 2. Suitable for systems with well-understood mechanism.	1. Lack of flexibility and adaptability. 2. Struggle with complex nonlinear systems.
Data-Driven Methods	1. High adaptability. 2. Suitable for big data scenarios.	1. Highly dependent on data quality and quantity. 2. Poor explainability.
Hybrid Methods	1. Relatively higher accuracy. 2. Better interpretability.	1. High model complexity. 2. High computational cost.

Aiming at the above problems, this paper develops a fault diagnosis method for combustion chamber components, named Thermodynamic Simulation-assisted Random Forest (TSRF). Firstly, a one-dimensional thermodynamic model of a marine diesel engine is established. The model parameters are then calibrated with testing data, facilitating the simulation of five faults in combustion chamber components. To minimize computational complexity, SHAP values are calculated for each thermodynamic parameter, and those with high SHAP values are considered as key parameters. Finally, the selected parameters are leveraged for classification using Random Forest (RF), and the results are compared to different machine learning as well as feature selection methods. The evaluation framework incorporates both local and global interpretation perspectives based on SHAP values. Local interpretation identifies critical parameters contributing to individual fault sample through detailed waterfall plot analysis, while global significance assessment employs beeswarm plots to evaluate parameter contributions across the entire fault state. Further, parameters interaction plots and dependency plots are presented based on SHAP interaction values. The main innovations of this paper are summarized as follows:

(1) This paper develops a novel parameter fine-tuning method for modeling marine diesel engine combustion chamber faults, including head cracking (F1), piston ablation (F2), liner wear (F3), ring wear (F4), and ring sticking (F5). Traditional methods typically emphasize microscopic material properties under fault conditions, which are often associated with extended simulation durations, intricate processes, and a limited scope of thermodynamic parameter analysis. In contrast, the proposed method reproduces the fault characteristics through parameter fine-tuning to obtain a rapid system response to combustion chamber faults.

(2) This study evaluates the suitability of the SHAP method for parameter selection in combustion chamber fault diagnosis, comparing its performance with other feature selection methods, including the Chi-Square Test, Recursive Feature Elimination (RFE), and the Gini index. Through the computation of SHAP values, a quantitative assessment of parameter importance is conducted, enabling the meticulous selection of high-contribution parameters. Furthermore, the research incorporates advanced visualization techniques to elucidate the model's decision-making process, providing transparent insights into the specific mechanisms through which critical parameters influence fault classification outcomes.

(3) This paper presents a novel dual-perspective interpretation through data-driven method assisted by thermodynamic model in the decision-making process of RF for combustion chamber fault diagnosis. By leveraging the mechanistic foundation and physical interpretability provided by thermodynamic models, this method provides a novel insight into interpretable fault diagnosis for marine diesel engine combustion chamber faults.

The rest of the paper is arranged as follows: In section 2, the proposed TSRF is introduced, including thermodynamic modeling, parameter selection and interpretability methods. In section 3, the simulation results are theoretically analyzed. Results and discussions are presented in section 4.

2. Methodology

2.1. Overview

The proposed TSRF method for marine diesel engines is illustrated in Fig. 1. Firstly, a one-dimensional thermodynamic model is constructed, and the model is calibrated with data acquired from the Data Collecting Module (DCM). Secondly, five representative combustion chamber fault scenarios are simulated through fine-tuning of critical system parameters, enabling the extraction of potentially diagnostic-relevant thermodynamic indicators. Then, RF is leveraged for preliminary fault identification, followed by SHAP value computation to facilitate the selection of optimal parameter subsets for diagnostic analysis. The final stage incorporates multi-perspective model interpretation, combining SHAP value analysis with thermodynamic model insights to provide comprehensive diagnostic understanding.

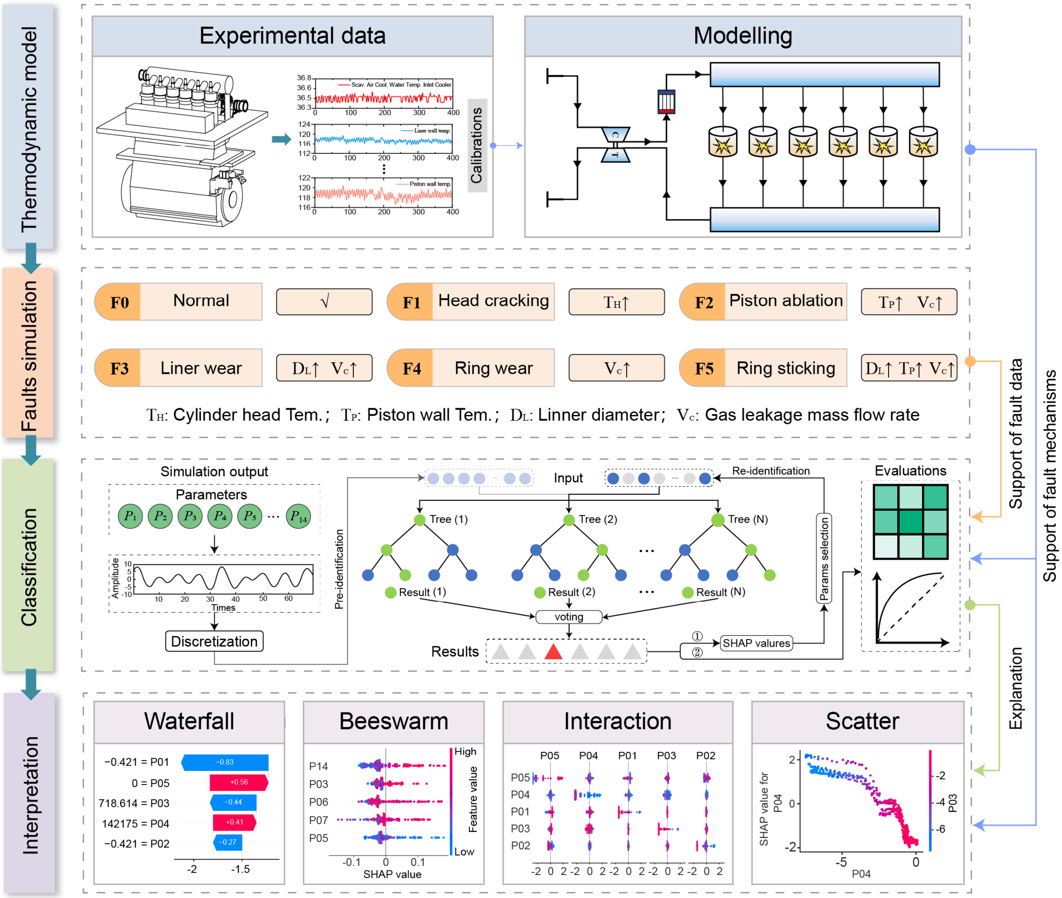


Fig. 1. Structure of TSRF

2.2. One-dimensional thermodynamic model of marine diesel engine

2.2.1. Model description

The structure of the one-dimensional thermodynamic model of the diesel engine is shown in Fig. 2, including the system boundary (SB1, SB2), intake and exhaust manifolds (PL1, PL2), turbocharger (TC1), intercooler (CO1), and six cylinders (C1 to C6). In addition, six monitoring points (MP1 to MP6) are established for monitoring critical parameters. For piping, pipe 1 indicates the compressor intake pipe, pipe 2 refers to the intercooler intake pipe, and pipe 3 is the intercooler exhaust pipe. Pipes 4 to 9 correspond to the intake pipes of the cylinders, while pipes 10 to 15 are the exhaust pipes of the cylinders. Pipe 16 is connected to the turbocharger's inlet pipe, and pipe 17 is the exhaust pipe of the turbocharger. The main specifications of the diesel engine are shown in Table 2.

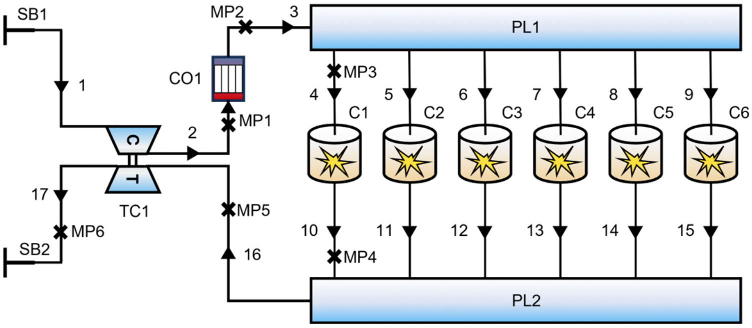


Fig. 2. One-dimensional thermodynamic model of the diesel engine

Table 2. Engine specifications

Engine Characteristic	Specifications	Engine Characteristic	Specifications
Cylinder bore	620 mm	Number of cylinders	6
Piston stroke	2658 mm	Arrangement of cylinders	Inline
Speed range	77 to 103 rpm	Firing Order	1-5-3-6-2-4
Mean piston speed	9.13 m/s	Cycle Type	2-stroke
Power	2900 kW	Stroke / bore ratio	4.29

2.3. Thermodynamic fault modeling

This study categorizes the health states of combustion chamber components into six distinct conditions, designated as F0 through F5: normal state (F0), head cracking (F1), piston ablation (F2), liner wear (F3), ring wear (F4), and ring sticking (F5). This subsection systematically elucidates the parameter fine-tuning method for the five fault states (F1-F5), followed by a comprehensive analysis of their respective failure mechanisms and underlying physical principles.

2.3.1. Cylinder head cracking (F1)

The cylinder head cracking is numerically investigated by fine-tuning the surface temperature of the cylinder head. Under operational conditions, the cylinder head experiences significant thermomechanical loading, predominantly in the combustion chamber-adjacent regions [27]. The initiation and propagation of cracks in

This is the author's accepted manuscript. The final version is available at 10.1016/j.measurement.2025.117252

these critical zones substantially degrade the structural integrity, thereby exacerbating localized stress concentration phenomena. Moreover, crack formation detrimentally affects the thermal dissipation, inducing localized thermal runaway through compromised heat dissipation efficiency.

2.3.2. Piston ablation (F2)

The piston ablation is numerically simulated through fine-tuning two critical parameters: surface temperature distribution and blow-by mass flow rate. The progressive degradation of piston surface material induces substantial thermal inhomogeneity, resulting in localized temperature elevation [28]. Furthermore, the material loss associated with thermal ablation compromises the piston-cylinder interface integrity, thereby exacerbating blow-by phenomena through increased gas leakage pathways.

2.3.3. Liner wear (F3)

The liner wear simulation is implemented through fine-tuning the bore diameter and corresponding blow-by mass flow characteristics. The primary wear mechanism originates from the infiltration of particulate contaminants, typically resulting from air filtration system failure or sealing interface degradation. These abrasive particles induce progressive material removal from the liner surface, manifesting as measurable roundness deviations and localized bore diameter enlargement. Such geometric alterations significantly degrade the sealing efficacy of the piston-liner interface, consequently amplifying blow-by phenomena through enhanced gas leakage pathways [29].

2.3.4. Ring wear (F4)

The piston ring wear simulation is conducted through controlled modulation of blow-by mass flow parameters. The predominant wear mechanism stems from the presence of abrasive contaminants within the cylinder assembly, which induce progressive material degradation of the ring surface. This wear process leads to measurable geometric deviations, particularly out-of-roundness deformation of the ring end face. Such geometric imperfections significantly compromise the ring-cylinder sealing integrity, thereby establishing a positive feedback loop that exacerbates blow-by phenomena through increased blow-by [30].

2.3.5. Ring sticking (F5)

The ring sticking is simulated through fine-tuning three critical parameters: bore diameter, liner surface temperature, and blow-by mass flow. This fault mode primarily originates from three contributing factors: excessive carbonaceous deposits, insufficient lubrication film formation, and sludge accumulation, which impede normal ring mobility and degrade sealing performance [31]. Under normal operating conditions, piston rings serve as essential thermal conduits, facilitating efficient heat transfer from the piston crown to the cylinder walls. However, ring sticking induces significant thermal resistance at the ring-liner interface, leading to substantial temperature elevation at the piston surface. Furthermore, this condition creates abnormal frictional interactions between the ring and liner surfaces, resulting in accelerated wear mechanisms and potential surface scoring of the cylinder liner.

2.4. Selection of thermodynamic parameters via Tree SHAP analysis

This paper presents a process for thermodynamic parameter selection leveraging SHAP values, as illustrated in Fig. 3. Firstly, potentially diagnostic-relevant thermodynamic parameters are filtered from the fault simulation output to comprise the preliminary dataset. Then, the dataset is input into RF for pre-identification. Building upon initial results, the Tree SHAP method is employed to quantitatively evaluate the contribution weights of individual parameters through SHAP value computation. The final stage involves the selection of

This is the author's accepted manuscript. The final version is available at 10.1016/j.measurement.2025.117252

optimal feature subsets based on parameter importance rankings, generating a refined dataset that is subsequently re-analyzed through the RF model for enhanced diagnostic accuracy.

Table 3 Comparison of feature selection methods

Method	Type	Computational Cost	Model Dependency	Interaction Effects	Nonlinear Relationships	Interpretability	Fairness
Chi-Square Test	Filter	Low	None	×	×	Statistical Significance	Low
RFE	Wrapper	High	Logistic Regression	√	×	Model Weights	Medium
Gini index	Embedded	Medium	Random Forest	√	√	Split Gain	Medium
SHAP	Explainability	High	Model-Agnostic	√	√	Global + Local Contributions	High

Compared to conventional feature selection methods, including the Chi-Square Test, RFE and Gini index, SHAP offers a comprehensive evaluation framework for thermodynamic parameter significance. This advanced method not only quantifies parameter importance but also elucidates directional influences (positive/negative effects), reveals parameter interaction dynamics, and characterizes sample distribution patterns. Such multifaceted analysis provides robust empirical support for investigating underlying physical mechanisms through thermodynamic modeling and uncovering potential parameter correlations. The comparative analysis between SHAP-based feature selection and traditional methods is presented in Fig. 3.

In this subsection, the calculation of SHAP values is presented in detail, concentrating on describing the core concept and computational process of the improved method for tree structures, known as Tree SHAP.

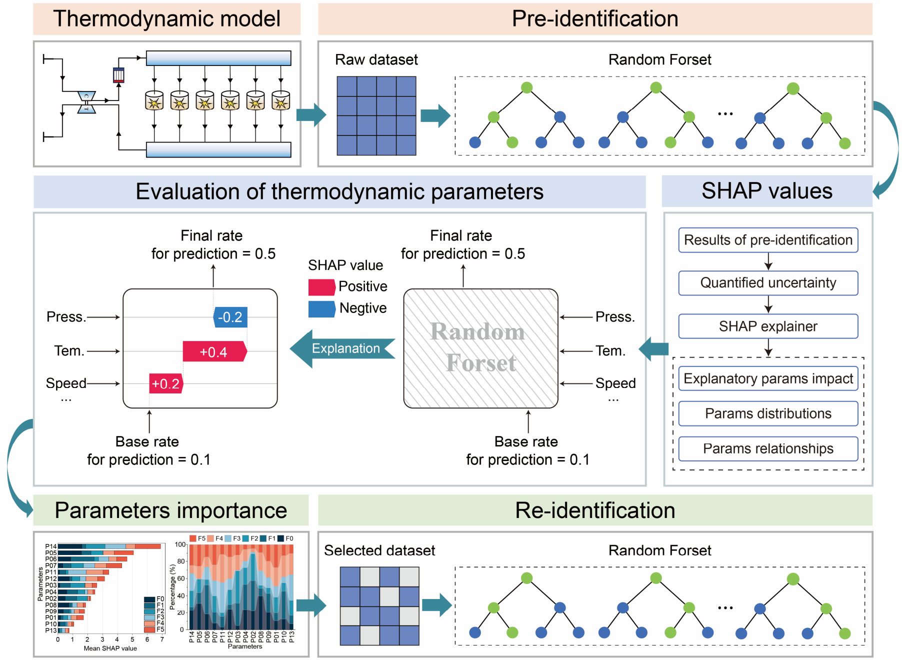


Fig. 3. SHAP-based parameter selection process

This is the author's accepted manuscript. The final version is available at 10.1016/j.measurement.2025.117252

2.4.1. SHAP value

SHAP has been extensively employed in machine learning, including feature selection and enhanced model interpretability, etc [32–38]. Within the SHAP value computation process, the concept of contribution value serves as a quantitative metric to characterize the influence of individual thermodynamic parameters on model predictions. The marginal contribution $\Delta_{val}(i,S)$ of a thermodynamic parameter i, when introduced to a specific parameter combination (as illustrated in Fig. 4a), is mathematically expressed through the following equation [39]:

$$\Delta_{val}(i,S)=f(S\cup\{i\})-f(S) \tag{1}$$

where S represents the set of parameters excluding parameter i; $f(S)$ is the contribution generated by the set S; $f(S\cup\{i\})$ is the contribution when parameter i is added.

After evaluating all possible parameter combinations, the average incremental contribution of parameter i, known as the SHAP value φ_i, can be calculated as follows [40]:

$$\varphi_i=\sum_{S\subseteq N\setminus\{i\}}\frac{|S|!(p-|S|-1)!}{p!}\cdot\Delta_{val}(i,S) \tag{2}$$

where N is the set of all parameters; S is a subset of the parameters used in the model; $N\setminus\{i\}$ encompasses all feasible parameter combinations excluding the target parameter i; p denotes the number of parameters in the subset S.

During the evaluation of a specific parameter combination S, the aggregate assessment is obtained through the linear superposition of individual parameter SHAP values, as illustrated in Fig. 4(b). Consequently, the composite SHAP value for parameter combination S is mathematically determined by the following formulation:

$$f(S)=\varphi_0+\sum_{i=1}^{M}\varphi_i \tag{3}$$

where: φ_0 is the baseline value, which represents the average output of the model without any parameter; $\sum_{i=1}^{M}\varphi_i$ is the sum of SHAP values predicted by each parameter for the model.

2.4.2. Tree SHAP

Tree SHAP represents an optimized computational approach specifically developed for tree-based models, including Random Forest and other tree-structured algorithms [41–43]. This method capitalizes on the inherent hierarchical architecture of tree models to achieve substantial computational efficiency gains compared to conventional SHAP value computation methods. As shown in Fig. 4(c), traditional SHAP value calculation requires exhaustive traversal of all possible parameter combinations, leading to significant computational overhead and exponential complexity growth with increasing parameter dimensionality.

The Tree SHAP algorithm fundamentally differs by computing marginal contributions exclusively along decision paths that satisfy specific input conditions (illustrated by the $i_1>5$ condition in Fig. 4d). This strategic optimization eliminates the necessity for complete parameter combination enumeration, thereby achieving remarkable computational efficiency improvements. For a given parameter i within a decision tree structure, the Tree SHAP-based value is computed through the following mathematical formulation [44]:

$$\varphi_i = \sum_{j=1}^{T} \sum_{S \subseteq P_j \setminus \{i\}} \frac{|S|!(p-|S|-1)!}{p!} \cdot \Delta_{val}(P_j, i, S) \tag{4}$$

where T represents the total number of decision trees in the ensemble model; P_j denotes the complete set of decision paths within the j-th tree structure; $P_j \setminus \{i\}$ encompasses all feasible parameter combinations excluding the target parameter i; S corresponds to the selected feature subset utilized in the model; and p indicates the cardinality of the parameter subset S.

2.5. Classification model

This paper employs RF as the base model for both parameter selection and fault classification tasks. This ensemble learning method shows superior capability in handling multi-class classification problems with limited sample sizes, while maintaining low preprocessing requirements. These characteristics render RF particularly effective for the complex multi-parameter fault diagnosis scenarios addressed in this research. Due to the continuity of thermodynamic parameters, a discretization process is implemented to transform these variables into appropriate formats compatible with the decision tree's input specifications.

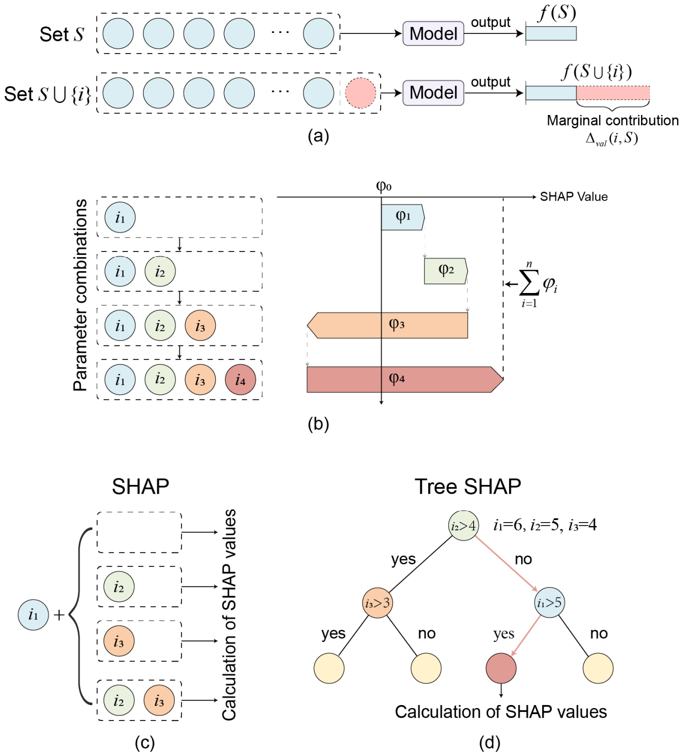


Fig. 4. Details of the SHAP algorithm: (a) Marginal contribution; (b) Additivity of SHAP values; (c) Calculation of SHAP; (d) Calculation of Tree SHAP.

2.5.1. Parameters discretization

The decision tree, serving as the fundamental component of RF, operates through recursive partitioning of the feature space to generate predictive outcomes. For a given sample set D, where p_k represents the proportion of samples belonging to the k-th class ($k = 1,2,\ldots,N$), the information entropy of D is defined as [45]:

$$\mathrm{Ent}(D) = -\sum_{k=1}^{N} p_k \log_2 p_k \tag{5}$$

When partitioning sample set D using a discrete attribute $a = \{a^1, a^2, \ldots, a^N\}$, which generates V distinct branch nodes, the subset of samples in the v-th branch node that satisfy the condition $a = a^v$ is denoted as D^v. The information entropy of D^v can be computed using Equation (5). To account for varying sample sizes across different branch nodes, a weighting factor $|D^v|/|D|$ is applied, enabling the calculation of information gain achieved through attribute-based partitioning [46]:

$$\mathrm{Gain}(D,a) = \mathrm{Ent}(D) - \sum_{v=1}^{V} \frac{|D^v|}{|D|}\mathrm{Ent}(D^v) \tag{6}$$

The magnitude of information gain serves as a quantitative indicator of an attribute's effectiveness in data category discrimination and uncertainty reduction. Attributes demonstrating higher information gain values are consequently prioritized for selection as splitting nodes within the decision tree structure.

Regarding thermodynamic parameters (e.g., temperature, pressure, and rotational speed), a discretization process is essential due to their continuous nature. For a given fault set D and continuous parameter P, where P assumes N distinct ordered values $\{P^1, P^2, \ldots, P^N\}$ in descending order, an optimal division point t is established. This division point segregates the samples into two subsets: D_t^- (samples with parameter values $\leq t$) and D_t^+ (samples with parameter values $> t$). Notably, any division point t within the interval $[P^i, P^{i+1}]$ yields identical partitioning results for adjacent values P_i and P_{i+1}. Therefore, for parameter P, the candidate division points can be systematically determined by computing the median value $(P^i + P^{i+1})/2$ for each interval $[P^i, P^{i+1}]$, resulting in the complete set of candidate division points T_P:

$$T_P = \left\{ \frac{P^i + P^{i+1}}{2} \mid 1 \leq i \leq N-1 \right\} \tag{7}$$

Consequently, the information gain achieved through partitioning the fault dataset D using parameter P can be expressed as [47]:

$$\begin{aligned} \mathrm{Gain}(D,P) &= \max_{t\in T_P} \mathrm{Gain}(D,P,t) \\ &= \max_{t\in T_P} \mathrm{Ent}(D) - \sum_{\lambda\in\{-,+\}} \frac{|D_t^\lambda|}{|D|}\mathrm{Ent}(D_t^\lambda) \end{aligned} \tag{8}$$

When the computed Gain (D, P) is sufficiently high, parameter P demonstrates superior discriminative capability and is consequently prioritized for selection as a splitting node in the decision tree architecture.

This is the author's accepted manuscript. The final version is available at 10.1016/j.measurement.2025.117252

2.5.2. Random forest

RF is an ensemble learning method that has demonstrated significant efficacy in fault diagnosis applications [48–51]. This algorithm enhances predictive performance through the strategic aggregation of multiple independent decision trees, effectively mitigating individual tree biases and variances. For a given sample characterized by thermodynamic parameter values $x_1,\ x_2\ \dots\ x_N$ the collective prediction $T_N(x)$ generated by an ensemble of n independently trained decision trees can be expressed as [52]:

$$T(x) = \frac{1}{N}\sum_{b=1}^{N} T_N(x) \tag{9}$$

Within the classification framework of random forests, two predominant metrics are conventionally employed for feature selection: the Gini index and entropy-based information gain. This research adopts the Gini index as the primary feature selection criterion. The Gini index, quantified within the interval [0, 1], serves as an inverse indicator of feature importance, where 0 signifies maximal information gain (denoting the most significant parameter) and 1 represents minimal information gain (indicating the least significant parameter). The mathematical formulation of the Gini index is expressed as follows [53]:

$$Gini(D) = 1 - \sum_{i=1}^{C} p_i^2 \tag{10}$$

where p_i is the proportion of category i in the dataset, and C is the number of categories.

3. Simulation of thermodynamic model

3.1. Data collection

3.1.1. Testing dataset

The testing dataset employed in this study is acquired from the mainframe sensor system of a prominent shipbuilding equipment manufacturer in China. The raw data collection is facilitated through an independently developed Data Collection Module (DCM) integrated with the marine propulsion system, as shown in Fig. 5. The DCM system implements hierarchical acquisition architecture, enabling synchronous data collection from multiple subsystems, including fast signal unit, tail gas treatment unit, and monitoring system.

The system records raw data at 10-second intervals, storing the information in an onboard database. To mitigate noise interference, a moving average is computed every 60 seconds. Following data encryption, the system transmits compressed CSV files containing average data from 720 sampling points to a shore-based server via maritime satellite every 12 hours, ultimately forming a six-month time-series dataset of marine diesel engine operations. The raw data undergoes rigorous preprocessing, including elimination of blank and duplicate records, interpolation of missing values, and anomaly detection and correction through a sliding window algorithm.

Following data preprocessing, an initial screening is conducted by referencing the vessel's operational manual, during which thermodynamic parameters exhibiting low relevance to the marine diesel engine are eliminated. Subsequently, the dataset undergoes further refinement by aligning with the specific data types that can be effectively monitored by the simulation model. This two-stage screening process establishes a robust foundation for subsequent model validation and performance evaluation.

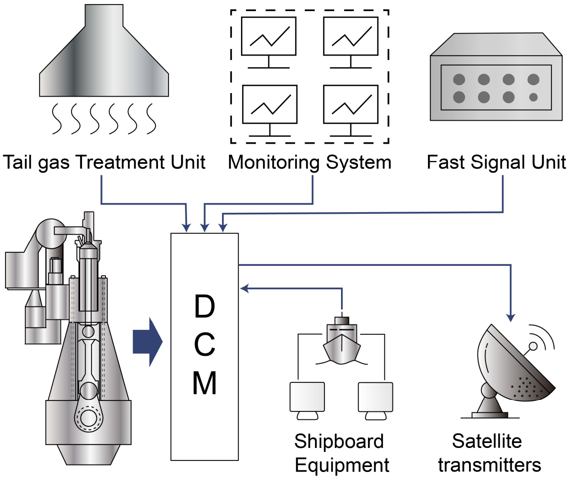


Fig. 5. Structure of the Data Collecting Module

3.1.2. Experimental dataset

The experimental samples are extracted from simulation results encompassing a complete operational cycle of the diesel engine model, with each health state comprising 120 samples. Given the substantial variation in value ranges and measurement scales across different thermal parameters, direct utilization of raw data would introduce bias in the training process, particularly in weight assignment. To address this issue, min-max normalization is implemented to standardize the parameter value ranges, ensuring equitable contribution from all features during model training. Following normalization, samples are randomly divided into a training dataset (70%) and a testing dataset (30%) proportionally.

3.2. Model validation

A representative data window is selected from the testing dataset during stable speed of the engine, and the values of key thermal parameters are averaged to validate the model. As shown in Table 4, the thermodynamic parameters from simulation results exhibit less than 5% deviation from experimental measurements under nominal operating conditions. This level of agreement confirms the model's reliability, enabling its effective utilization for both fault simulation scenarios and comprehensive diesel engine performance analysis.

Table 4. Validation of model validity

Item	Experimental value	Simulation value	Error
Power (kW)	2900	2895	-0.17%
Engine speed (rpm)	86.011	90	+4.64%
Temperature of exhaust gas before turbocharger (°C)	378.953	370.557	-2.22%
Temperature of exhaust gas after turbocharger (°C)	240.464	230.254	-4.25%
Inlet cooling water temperature of air cooler (°C)	36.101	37.521	+3.93%
Outlet cooling water temperature of air cooler (°C)	45.661	44.385	-2.79%
Temperature of cylinder liner (°C)	122.299	125.512	+2.63%
Temperature of exhaust gas (°C)	306.897	301.525	-1.75%

3.3. Fault modeling

This paper develops a systematic parameter fine-tuning method to emulate specific combustion chamber fault states, including head cracking (F1), piston ablation (F2), liner wear (F3), ring wear (F4), and ring sticking (F5). However, the established parameter ranges require further optimization to address potential discrepancies between simulation outcomes and actual operational conditions. Future research directions will focus on conducting detailed investigations into varying severity levels within individual fault modes, thereby enhancing the diagnostic precision and practical applicability of the model. Detailed information regarding the parameter calibration for each fault type can be found in Table 5.

Table 5. Details of fault parameters

Label	Fault type	Parameters			
		Blow-by mass flow (Kg/s)	Bore (mm)	Piston surface temp. (°C)	Cylinder head surface temp. (°C)
F0	Normal	0	620	276.85	246.85
F1	Head cracking	0	620	276.85	**346.85**
F2	Piston ablation	**0.01**	620	**376.85**	246.85
F3	Liner wear	**0.03**	**620.1**	276.85	246.85
F4	Ring wear	**0.02**	620	276.85	246.85
F5	Ring sticking	**0.02**	**620.1**	**326.85**	246.85

3.4. Results of fault simulation

In this paper, 14 thermodynamic parameters that exhibit significant diagnostic potential for combustion chamber condition assessment are identified. Details about these parameters are presented in Table 6. The variations of these parameters under various fault conditions are depicted in Fig. 6(a)-(n), with the crank angle represented on the horizontal axis and the amplitude of the parameters displayed on the vertical axis.

Table 6. Thermodynamic parameters of the simulation output

Label	Parameters	Unit
P01	Cylinder Pressure	Pa
P02	Cylinder Temperature	K
P03	Piston Wall Heat Flow	J/deg
P04	Head Wall Heat Flow	J/deg
P05	Liner Wall Heat Flow	J/deg
P06	Blow-By Heat Flow	J/deg
P07	Blow-By Mass Flow	Kg/s
P08	Turbine Power	J/s
P09	Pressure of Exhaust	Pa
P10	Temperature of Exhaust	K
P11	Pressure of Exhaust Gas before Turbocharger	Pa
P12	Temperature of Exhaust Gas before Turbocharger	K
P13	Pressure of Exhaust Gas after Turbocharger	Pa
P14	Temperature of Exhaust Gas after Turbocharger	K

4. Results and discussion

4.1. Parameters selection

Given the extensive array of thermodynamic parameters in diesel engines, comprehensive monitoring of all parameters for diagnostic purposes proves impractical. Consequently, a parameter selection process is implemented to identify the most diagnostic relevant parameters. This approach achieves an optimal balance between minimizing the number of monitored parameters and maintaining robust fault differentiation capability.

The SHAP values of the 14 parameters in six health states are shown in Fig. 7(a) and summarized in Fig. 7(b), with the horizontal axis denoting the mean SHAP values, and the vertical axis listing the 14 thermodynamic parameters ordered by importance from highest to lowest. Fig. 7(c) further illustrates the distribution of the percentage of importance scores for each parameter. As a result, eight core parameters with the highest cumulative contribution rate are selected for fault identification, including: P14 (Temperature of Exhaust Gas after Turbocharger), P05 (Liner Wall Heat Flow), P06 (Blow-By Heat Flow), P07 (Blow-By Mass Flow), P11 (Pressure of Exhaust Gas before Turbocharger), P12 (Temperature of Exhaust Gas before Turbocharger), P03 (Piston Wall Heat Flow), and P04 (Head Wall Heat Flow).

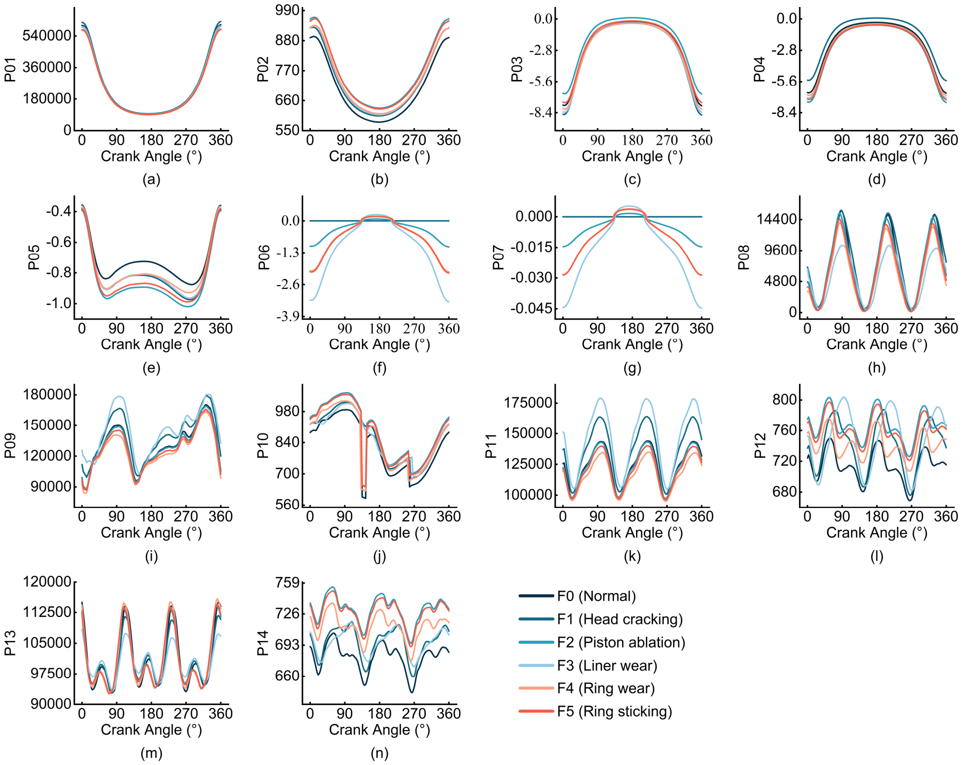


Fig. 6. Results of fault simulation: (a) Cylinder pressure; (b) Cylinder temperature; (c) Piston wall heat flow; (d) Head wall heat flow; (e) Liner wall heat flow; (f) Blow-by heat flow; (g) Blow-by mass flow; (h) Turbine power; (i) Pressure of exhaust gas before turbocharger; (j) Temperature of exhaust gas before turbocharger; (k) Pressure of exhaust gas after turbocharger; (l) Temperature of exhaust gas after turbocharger; (m) Pressure of exhaust; (n) Temperature of exhaust.

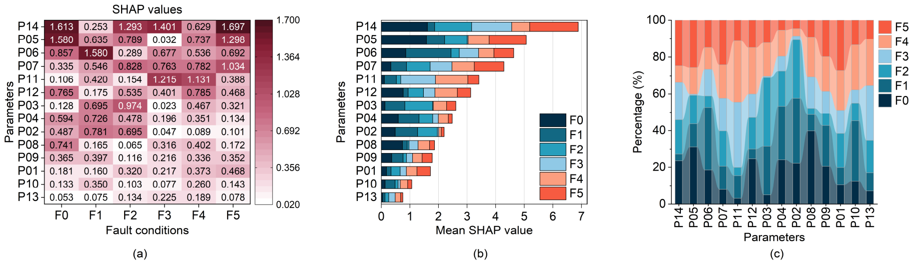


Fig. 7. Importance of thermodynamic parameters based on SHAP values: (a) Heat map of SHAP values; (b) Bar stacking chart, (c) Percentage chart.

4.2. Model evaluation

Given the limited sample size, this study implements an optimized hyper-parameter tuning strategy combining grid search method with five-fold cross-validation for all machine learning models under comparison. The optimal parameter configurations for K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and RF are presented in Table 7. The evaluation metrics across six health states for both the original dataset and the optimized parameter subset, including precision, recall, F1 score, and accuracy, are presented in Table 8. The confusion matrices and Precision-Recall curves for each model's performance on these datasets are illustrated in Fig. 8 and Fig. 9, respectively.

The confusion matrices reveal distinct performance of models: KNN exhibits relatively low accuracy in recognizing F3-F5 fault conditions on the original dataset, as shown in Fig. 8(a). However, its diagnostic capability for these fault conditions shows noticeable improvement when utilizing the optimal parameter subset, as depicted in Fig. 8(d). SVM exhibits suboptimal performance in identifying F1 and F4 faults on the original dataset (see Fig. 8b), with its recognition accuracy for these faults improving on the optimized parameter subset, albeit at the cost of reduced performance for F3 detection, as shown in Fig. 8(e). In comparison, RF initially underperforms in F5 fault recognition on the original dataset (see Fig. 8c) but achieves superior overall performance across all six health states when employing the optimal feature subset, as shown in Fig. 8(f). While its ability to differentiate between F4 and F5 faults is slightly diminished, it is noteworthy that both fault types are piston ring-related and exhibit similar system impact patterns. The evaluation shows that RF outperforms both SVM and KNN on the experimentally constructed fault dataset.

Table 7 Selection of hyperparameters for models

Model	Core hyperparameters for training		
KNN	n_neighbors: 3	metric: manhattan	weights: distance
SVM	C: 1.0	kernel: linear	gamma: scale
RF	n_estimators: 20	max_depth: None	max_features: sqrt

Table 8. Comparison of the performance of various models.

Model	Performance evaluation (on the original dataset, %)				Performance evaluation (on the optimal subset, %)			
	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score	Accuracy
KNN	90.94±3.72	89.81±3.43	90.03±3.72	**89.81±3.73**	95.14±3.10	94.44±3.23	94.54±3.25	**94.44±3.24**
SVM	92.91±2.33	92.13 ±2.42	91.90±2.02	**92.13±2.72**	95.55±2.13	94.44±2.73	94.41±1.52	**94.44±1.85**
RF	94.66±0.73	94.44 ±0.52	94.41±0.33	**94.44±0.46**	99.12±0.45	99.07±0.66	99.07±0.51	**99.07±0.46**

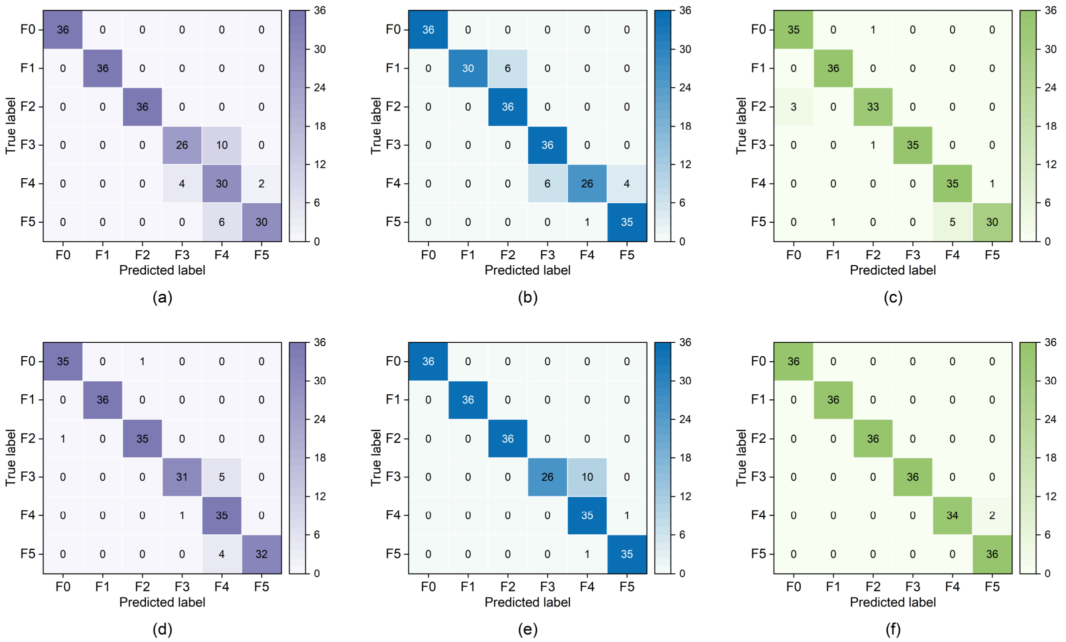


Fig. 8 Confusion matrix: (a)-(c) present the confusion matrices of KNN, SVM and RF on the original dataset, respectively; (d)-(f) present the confusion matrices of KNN, SVM and RF on the optimal subset, respectively.

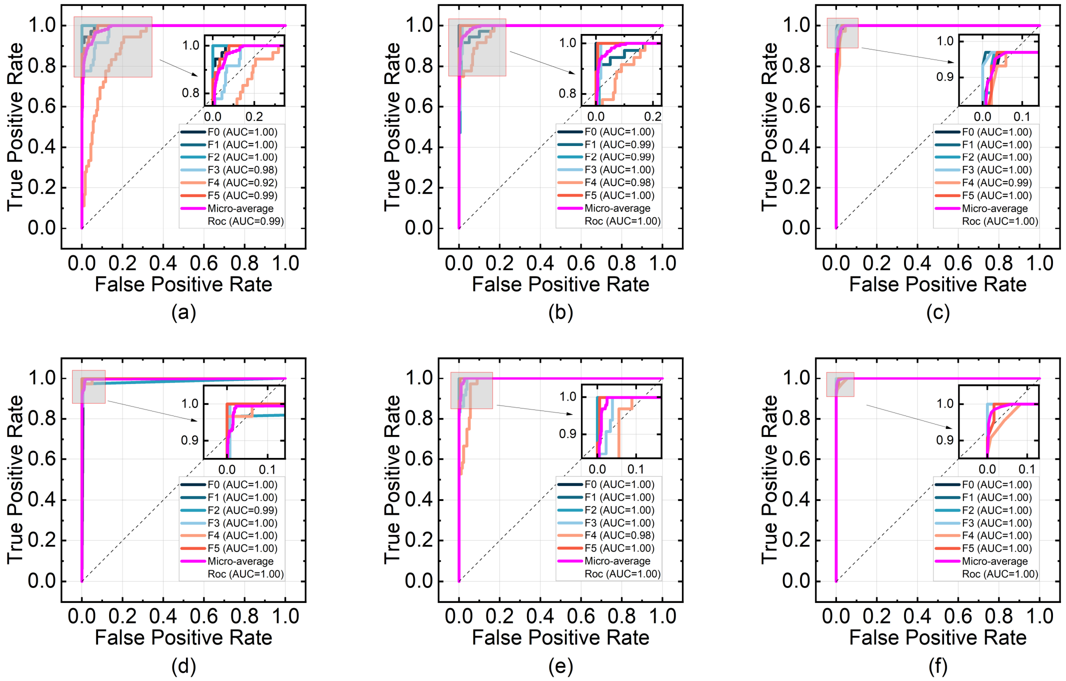


Fig. 9 Precision-Recall curves: (a)-(c) present the Precision-Recall curves of KNN, SVM and RF on the original dataset, respectively; (d)-(f) present the Precision-Recall curves of KNN, SVM and RF on the optimal subset, respectively.

To establish a comprehensive comparative analysis, multiple feature selection methods are implemented, with the resulting optimal parameter subsets from each method visually presented in Fig. 10. Utilizing RF as the classification model, each optimized parameter subset is evaluated for diagnostic performance, with average accuracy summarized in Table 9. The results reveal that the parameter subset obtained through SHAP exhibits superior diagnostic capability compared to subsets derived from alternative feature selection methods.

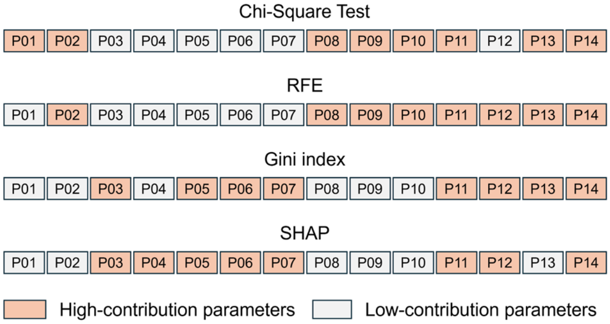


Fig. 10 Optimal subset of different feature selection methods

Table 9 Accuracy of RF using different feature selection methods

Method	Average accuracy (%)
Chi-Square Test	78.51± 0.53
RFE	81.62± 0.22
Gini index	88.95± 0.45
SHAP	**99.07±0.21**

4.3. Interpretation assisted by thermodynamic model

To demonstrate the enhanced interpretability assisted by thermodynamic modeling, this study employs piston ring failure (F4) as a representative case study. The analysis adopts a dual-scale approach, initially examining the model's decision-making mechanism at the individual sample perspective, followed by a global perspective of the fault characteristics.

4.3.1. Individual sample analysis

The waterfall plot is leveraged as an effective visualization tool for analyzing individual sample significance assessment, as illustrated in Fig. 9(a). The information in the waterfall plot is interpreted as follows:

(1) The waterfall plot begins at the bottom with the expected value $E[f(x)]$ of the model outputs.

(2) Each horizontal bar quantitatively represents a parameter's contribution magnitude and direction, with red bars indicating positive contributions and blue bars denoting negative influences.

(3) The terminal value $f(x)$ is computed through the algebraic summation of all individual parameter contributions, representing the model's final prediction for the specific sample.

As shown in Fig. 11(a), certain parameter values directly influence the predictive direction of the model. However, the relationship between these values and the standard reference values is not immediately apparent. By integrating the output from the thermodynamic model in Fig. 6, it becomes clear that at specific moments, such as P06=1.641, P07=-0.024 and P12=728.568, the values of P14 and P04 are significantly low, which negatively affects the model’s predictions. Given that piston ring wear exacerbates the blow-by, this phenomenon is incorporated into the fault model as prior knowledge. The presence of blow-by increases the values of P14 and P04, which directly contradicts the low values observed at the sample points, ultimately reducing the model's prediction accuracy.

4.3.2. Global interpretation

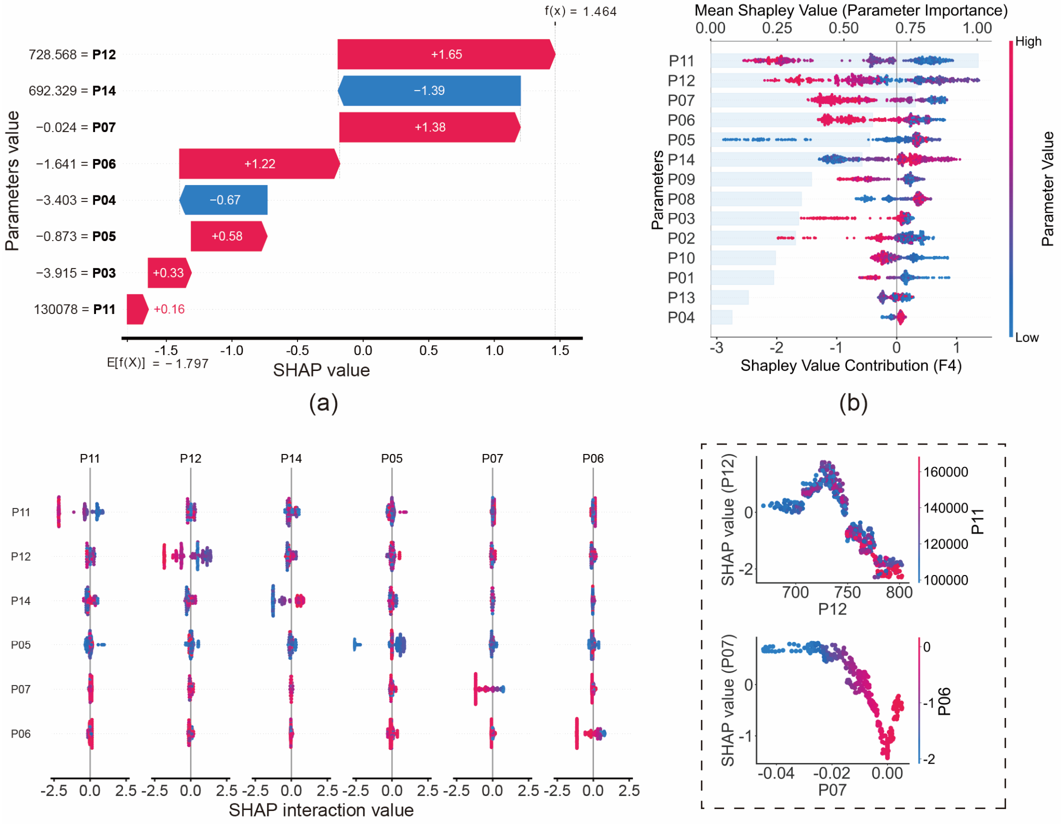
In this paper, the evaluation results of all samples of F4 fault state are summarized, and the distribution of the samples, induced by variations in thermodynamic parameters, are presented through beeswarm plots (see Fig. 11b). Further, the average SHAP value for each parameter is combined with beeswarm plots, facilitating a more intuitive representation of the importance distribution of each parameter. The information in the beeswarm plots is interpreted as follows:

(1) The y-axis displays the parameters arranged by their level of importance, with each bar representing the mean Shapley value for that parameter.

(2) Each sample is illustrated by a dot, where the x-axis position of the dot corresponds to the SHAP value, and the dots are stacked within each row to depict their density.

(3) Colors are utilized to represent the raw values of the parameters.

(4) The horizontal distribution of points reflects the magnitude of impact; a wider distribution indicates a greater impact.



(a)

(b)

(c)

(d)

Fig. 11 Fault analysis of piston ring wear (F4) based on SHAP values: (a) Waterfall plot; (b) Beewarm plot; (c) Interaction plot; (d) Dependence plot.

Fig. 11(b) reveals that depressed values of P11, P12, P07, and P06 significantly influence the model's predictive tendency toward identifying fault type F4. To thoroughly investigate the interdependencies among thermodynamic parameters, this study further calculates the SHAP interaction values. The six most influential parameter interactions are presented in Fig. 11(c), with detailed visualizations of the specific interactions between P11-P12 and P06-P07 provided in Fig. 11(d). The information in the dependence plots is interpreted as follows:

(1) Each point corresponds to the prediction for an individual sample.
(2) The x-axis represents the actual value of the parameter.
(3) The y-axis displays the SHAP value of the parameter.
(4) The variations in color reflect the value of the parameter with which it interacts.

The results reveal significant positive correlations between P11-P12 and P06-P07, consistent with fundamental thermodynamic principles. When integrated with thermodynamic modeling results (see Fig. 6k), the value of P11 is the lowest of all the health states. Due to the positive correlation between P11 and P12, this results in correspondingly reduced P12 values. Furthermore, piston ring wear is predominantly characterized by blow-by, where P06 and P07 exhibit negative correlations with blow-by severity. Consequently, the depressed values of P06 and P07 serve as effective indicators for identifying piston ring wear conditions.

5. Conclusion

The integration of data-driven methodologies with model-based approaches is a significant task in enhancing both simulation reliability and model interpretability. This paper introduces an innovative explainable fault diagnosis framework for marine diesel engine combustion chambers, termed the Thermodynamic Simulation-assisted Random Forest (TSRF) method. The simulation component incorporates priori fault knowledge to establish five thermodynamic fault models for key combustion chamber components. Within the diagnostic framework, SHAP values are synergistically combined with thermodynamic modeling to facilitate enhanced diagnostic capability. The proposed method aims to provide reliable support for fault diagnosis and improve the interpretability of combustion chamber diagnosis results.

The validity of the constructed one-dimensional thermodynamic model is confirmed through testing data collected from DCM. Additionally, the efficacy of the developed TSRF in improving diagnostic accuracy has been validated through experimental comparisons with conventional machine learning methods. The developed TSRF outperformed the KNN and SVM on the fault dataset constructed in this paper by achieving mean accuracy of 99.07%.

Further, the contribution and interactions of thermodynamic parameters for a specific fault case (F4) are investigated and interpretive analyses of model decision results are provided. According to the findings from TSRF model estimation and SHAP analysis, turbocharger exhaust temperature, blow-by heat flow and cylinder liner heat flow are the key thermodynamic parameters affecting the fault diagnosis results.

In future work, varying marine conditions will be investigated to further evaluate and upgrade the proposed method. Additionally, fault parameter settings will be optimized and integrated with varying fault severities to achieve more precise classification and higher diagnostic accuracy. Further, our research will focus on the synergistic integration of advanced models (e.g., deep learning) with various thermodynamic modeling approaches, including multi-zone models or models incorporating detailed chemistry. This strategy aims to leverage the complementary strengths of data-driven learning and physics-based modeling, enabling both enhanced predictive accuracy and robust physical interpretability in marine diesel engine fault diagnosis.

References

[1] J. Castresana, G. Gabiña, L. Martin, A. Basterretxea, Z. Uriondo, Marine diesel engine ANN modelling with multiple output for complete engine performance map, Fuel. 319 (2022) 123873.
[2] C. Luo, M. Zhao, K. Zhang, X. Yu, J. Li, Z. Cui, Thermal-fluid-structure coupling simulation of piston in large marine diesel engine, J. Phys.: Conf. Ser. 2827 (2024) 012020.
[3] Y. Li, J. Liu, G. Zhong, W. Huang, R. Zou, Analysis of a diesel engine cylinder head failure caused by casting porosity defects, Engineering Failure Analysis. 127 (2021) 105498.
[4] M.D.C. Rubiales Mena, A. Muñoz, M.Á. Sanz-Bobi, D. Gonzalez-Calvo, T. Álvarez-Tejedor, Application of ensemble machine learning techniques to the diagnosis of the combustion in a gas turbine, Applied Thermal Engineering. 249 (2024) 123447.
[5] Z. Long, M. Bai, M. Ren, J. Liu, D. Yu, Fault detection and isolation of aeroengine combustion chamber based on unscented Kalman filter method fusing artificial neural network, Energy. 272 (2023) 127068.
[6] B. Cai, Fault detection and diagnostic method of diesel engine by combining rule-based algorithm and BNs/BPNNs, Journal of Manufacturing Systems. 57 (2020) 148–157.
[7] B. Liu, H. Gan, D. Chen, Z. Shu, Research on Fault Early Warning of Marine Diesel Engine Based on CNN-BiGRU, JMSE. 11 (2022) 56.
[8] S. Zhong, S. Fu, L. Lin, A novel gas turbine fault diagnosis method based on transfer learning with CNN, Measurement. 137 (2019) 435–453.
[9] H. Gharib, G. Kovács, A Review of Prognostic and Health Management (PHM) Methods and Limitations for Marine Diesel Engines: New Research Directions, Machines. 11 (2023) 695.
[10] Y. Lv, X. Yang, Y. Li, J. Liu, S. Li, Fault detection and diagnosis of marine diesel engines: A systematic review, Ocean Engineering. 294 (2024) 116798.
[11] Z. Gao, C. Cecati, S.X. Ding, A Survey of Fault Diagnosis and Fault-Tolerant Techniques—Part I: Fault Diagnosis With Model-Based and Signal-Based Approaches, IEEE Trans. Ind. Electron. 62 (2015) 3757–3767.
[12] Z. He, Y. Yang, H. Han, J. Wang, Y. Zhang, H. Li, A key performance indicator-based fault detection scheme for marine diesel turbocharging systems, Journal of the Franklin Institute. 358 (2021) 9346–9363.
[13] R.H.R. Gutiérrez, C.R.P. Belchior, L.A. Vaz, U.A. Monteiro, Diagnostic methodology in four-stroke marine diesel engine by identifying operational parameters, J Braz. Soc. Mech. Sci. Eng. 40 (2018) 500.
[14] J.A. Pagán Rubio, F. Vera-García, J. Hernandez Grau, J. Muñoz Cámara, D. Albaladejo Hernandez, Marine diesel engine failure simulator based on thermodynamic model, Applied Thermal Engineering. 144 (2018) 982–995.
[15] Y. Tang, J. Zhang, H. Gan, B. Jia, Y. Xia, Development of a real-time two-stroke marine diesel engine model with in-cylinder pressure prediction capability, Applied Energy. 194 (2017) 55–70.
[16] H.M. Nahim, R. Younes, C. Nohra, M. Ouladsine, Complete modeling for systems of a marine diesel engine, J. Marine. Sci. Appl. 14 (2015) 93–104.
[17] V.T. Lamaris, D.T. Hountalas, A general purpose diagnostic technique for marine diesel engines – Application on the main propulsion and auxiliary diesel units of a marine vessel, Energy Conversion and Management. 51 (2010) 740–753.
[18] J. Hu, Y. Yu, J. Yang, H. Jia, Research on the generalisation method of diesel engine exhaust valve leakage fault diagnosis based on acoustic emission, Measurement. 210 (2023) 112560.
[19] Y. Zhou, Z. Wang, X. Zuo, H. Zhao, Identification of wear mechanisms of main bearings of marine diesel engine using recurrence plot based on CNN model, Wear. 520–521 (2023) 204656.
[20] Y. Guo, J. Zhang, Fault Diagnosis of Marine Diesel Engines under Partial Set and Cross Working Conditions Based on Transfer Learning, JMSE. 11 (2023) 1527.
[21] R. Wang, H. Chen, C. Guan, Random convolutional neural network structure: An intelligent health

monitoring scheme for diesel engines, Measurement. 171 (2021) 108786.
[22] R. Wang, H. Chen, C. Guan, W. Gong, Z. Zhang, Research on the fault monitoring method of marine diesel engines based on the manifold learning and isolation forest, Applied Ocean Research. 112 (2021) 102681.
[23] J. Yu, S. Wang, L. Wang, Y. Sun, Gearbox fault diagnosis based on a fusion model of virtual physical model and data-driven method, Mechanical Systems and Signal Processing. 188 (2023) 109980.
[24] K.C. Gryllias, I.A. Antoniadis, A Support Vector Machine approach based on physical model training for rolling element bearing fault detection in industrial environments, Engineering Applications of Artificial Intelligence. 25 (2012) 326–344.
[25] W. Li, S. Gu, X. Zhang, T. Chen, Transfer learning for process fault diagnosis: Knowledge transfer from simulation to physical processes, Computers & Chemical Engineering. 139 (2020) 106904.
[26] K. Yan, X. Chen, X. Zhou, Z. Yan, J. Ma, Physical Model Informed Fault Detection and Diagnosis of Air Handling Units Based on Transformer Generative Adversarial Network, IEEE Trans. Ind. Inf. 19 (2023) 2192–2199.
[27] W. Yang, J. Pang, L. Wang, X. Kang, S. Zhou, C. Zou, S. Li, Z. Zhang, Thermo-mechanical fatigue life prediction based on the simulated component of cylinder head, Engineering Failure Analysis. 135 (2022) 106105.
[28] K. Xie, W. Long, Y. Wang, H. Tian, Study on the forming process of the cylinder wall fuel film and the piston top land ablation mechanism of a medium speed marine diesel engine under cold start condition, Applied Thermal Engineering. 243 (2024) 122634.
[29] B. Zhang, X. Ma, L. Liu, Y. Wang, H. Yu, A. Morina, X. Lu, Reciprocating sliding friction behavior and wear state transition mechanism of cylinder liner and piston ring, Wear. 546–547 (2024) 205293.
[30] B. Zhang, X. Ma, L. Liu, H. Yu, A. Morina, X. Lu, Study on the sliding wear map of cylinder liner − piston ring based on various operating parameters, Tribology International. 186 (2023) 108632.
[31] R. Ferreira, J. Martins, Ó. Carvalho, L. Sobral, S. Carvalho, F. Silva, Tribological solutions for engine piston ring surfaces: an overview on the materials and manufacturing, Materials and Manufacturing Processes. 35 (2020) 498–520.
[32] S. Li, Z. Zhao, T. Miao, X. Li, S. Yu, Y. Zhang, J. Wei, K. Zou, Discrimination of Pb-Zn deposit types using the trace element data of galena based on deep learning, Ore Geology Reviews. 170 (2024) 106133.
[33] L. Longo, M. Brcic, F. Cabitza, J. Choi, R. Confalonieri, J.D. Ser, R. Guidotti, Y. Hayashi, F. Herrera, A. Holzinger, R. Jiang, H. Khosravi, F. Lecue, G. Malgieri, A. Páez, W. Samek, J. Schneider, T. Speith, S. Stumpf, Explainable Artificial Intelligence (XAI) 2.0: A manifesto of open challenges and interdisciplinary research directions, Information Fusion. 106 (2024) 102301.
[34] K. Sadeghi R., D. Ojha, P. Kaur, R.V. Mahto, A. Dhir, Explainable artificial intelligence and agile decision-making in supply chain cyber resilience, Decision Support Systems. 180 (2024) 114194.
[35] J.W. Goodell, S. Ben Jabeur, F. Saâdaoui, M.A. Nasir, Explainable artificial intelligence modeling to forecast bitcoin prices, International Review of Financial Analysis. 88 (2023) 102702.
[36] R. Zuo, Q. Cheng, Y. Xu, F. Yang, Y. Xiong, Z. Wang, O.P. Kreuzer, Explainable artificial intelligence models for mineral prospectivity mapping, Sci. China Earth Sci. 67 (2024) 2864–2875.
[37] Z. Li, Extracting spatial effects from machine learning model using local interpretation method: An example of SHAP and XGBoost, Computers, Environment and Urban Systems. 96 (2022) 101845.
[38] F. Klauschen, J. Dippel, P. Keyl, P. Jurmeister, M. Bockmayr, A. Mock, O. Buchstab, M. Alber, L. Ruff, G. Montavon, K.-R. Müller, Toward Explainable Artificial Intelligence for Precision Pathology, Annu. Rev. Pathol. Mech. Dis. 19 (2024) 541–570.
[39] Z. Li, Extracting spatial effects from machine learning model using local interpretation method: An example of SHAP and XGBoost, Computers, Environment and Urban Systems. 96 (2022) 101845.
[40] Y. Kim, Y. Kim, Explainable heat-related mortality with random forest and SHapley Additive

exPlanations (SHAP) models, Sustainable Cities and Society. 79 (2022) 103677.
[41] S.M. Lundberg, G. Erion, H. Chen, A. DeGrave, J.M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, S.-I. Lee, From local explanations to global understanding with explainable AI for trees, Nature Machine Intelligence. 2 (2020) 56–67.
[42] D. Wang, S. Thunéll, U. Lindberg, L. Jiang, J. Trygg, M. Tysklind, Towards better process management in wastewater treatment plants: Process analytics based on SHAP values for tree-based machine learning methods, Journal of Environmental Management. 301 (2022) 113941.
[43] O.O. Bifarin, Interpretable machine learning with tree-based shapley additive explanations: Application to metabolomics datasets for binary classification, Plos One. 18 (2023) e0284315.
[44] R. Mitchell, E. Frank, G. Holmes, GPUTreeShap: massively parallel exact calculation of SHAP scores for tree ensembles, PeerJ Computer Science. 8 (2022) e880.
[45] Q. Hu, X. Che, L. Zhang, D. Zhang, M. Guo, D. Yu, Rank Entropy-Based Decision Trees for Monotonic Classification, IEEE Trans. Knowl. Data Eng. 24 (2012) 2052–2064.
[46] Y. Wang, S.-T. Xia, J. Wu, A less-greedy two-term Tsallis Entropy Information Metric approach for decision tree classification, Knowledge-Based Systems. 120 (2017) 34–42.
[47] X. Han, X. Zhu, W. Pedrycz, Z. Li, A three-way classification with fuzzy decision trees, Applied Soft Computing. 132 (2023) 109788.
[48] Y. Wu, Y. Bai, S. Yang, C. Li, Extracting random forest features with improved adaptive particle swarm optimization for industrial robot fault diagnosis, Measurement. 229 (2024) 114451.
[49] A.F. Amiri, H. Oudira, A. Chouder, S. Kichou, Faults detection and diagnosis of PV systems based on machine learning approach using random forest classifier, Energy Conversion and Management. 301 (2024) 118076.
[50] S. Tao, R. Ma, Y. Chen, Z. Liang, H. Ji, Z. Han, G. Wei, X. Zhang, G. Zhou, Rapid and sustainable battery health diagnosis for recycling pretreatment using fast pulse test and random forest machine learning, Journal of Power Sources. 597 (2024) 234156.
[51] L.C. Brito, G.A. Susto, J.N. Brito, M.A.V. Duarte, An explainable artificial intelligence approach for unsupervised fault detection and diagnosis in rotating machinery, Mechanical Systems and Signal Processing. 163 (2022) 108105.
[52] S. Tangirala, Evaluating the Impact of GINI Index and Information Gain on Classification using Decision Tree Classifier Algorithm*, IJACSA. 11 (2020) 612–619.
[53] V.R.E. Christo, H.K. Nehemiah, J. Brighty, A. Kannan, Feature Selection and Instance Selection from Clinical Datasets Using Co-operative Co-evolution and Classification Using Random Forest, IETE Journal of Research. 68 (2022) 2508–2521.