



## Improving efficiency of diesel engine fault detection based on multi-source data

Shengyu Piao<sup>1,2</sup> · Jae-Jung Hur<sup>3</sup> · Jong Hun Im<sup>3</sup> · Min-Kyoon Kang<sup>4</sup> · Won-Ju Lee<sup>2,3,†</sup>

(Received March 11, 2025 ; Revised April 3, 2025 ; Accepted May 12, 2025)

**Abstract:** As the core component of a ship's engine, the timely and effective maintenance of the cylinder is crucial for the vessel's stable navigation. Data-driven fault detection methods, which rely solely on historical data from the ship's systems, are widely used for fault detection. These methods focus on modeling and prediction based exclusively on target data. While previous studies have often employed more complex algorithms and composite models to improve prediction performance, this study introduces a novel approach by integrating a variety of engine room data, breaking through the limitations of traditional methods that rely solely on target data. By fusing the target cylinder data with additional engine room data, a multi-source data collaborative modeling framework was constructed. Building on this, we designed a hybrid model combining deep learning with an Exponentially Weighted Moving Average (EWMA) for real-time monitoring and fault warning of the diesel engine cylinder's operating status. The experiment compared three time series deep learning algorithms, and the results demonstrated that the collaborative modeling of multi-source data significantly improved prediction accuracy and effectively reduced the average run length (ARL). This innovative method improves the efficiency of fault detection.

**Keywords:** Fault Detection, Deep Learning, EWMA, Cylinder, Prediction

### 1. Introduction

80% of the world's total trade is transported by ships, making stable and continuous navigation essential for cargo transportation [1]. Timely detection of potential internal faults in ship machinery, followed by proactive maintenance before issues escalate, are key strategies for enhancing ship safety and reliability [2].

A research report by the marine insurance company, The Swedish Club, indicates that the proportion of claims related to ship machinery rose from 35% in 2010–2014 to 48% in 2015–2017. In this period, damage of main engine emerged as one of the costliest types of ship failure, accounting for 34% of all machinery-related claim amounts [3]. Therefore, timely and effective maintenance of marine diesel engines is crucial to minimizing the risk of failure and reducing downtime losses.

The marine diesel engine is the primary power unit of a vessel, and its operational status is directly related to the safety of navigation.

As a critical load-bearing component, the cylinder not only affects the power output performance of the diesel engine but also reflects the overall operating status of the entire diesel engine system. During long-term operation, the cylinder is continuously exposed to a high-temperature and high-pressure environment, where the energy generated by combustion drives reciprocating motion of the piston. Under the combined effects of high temperature, high pressure, and prolonged wear, the cylinder and its related components may develop faults, potentially impacting the overall performance of the diesel engine [4]. Therefore, monitoring cylinder data (including temperature, pressure etc.) can not only detect abnormalities in the cylinder itself but also help identify potential issues within the diesel engine system at an early stage, enabling predictive maintenance and early fault prevention. Condition-based maintenance can extend the maintenance cycle of mechanical equipment by 50% while reducing operating costs by 25% to 45% [5]–[7]. Furthermore, condition-based predictive maintenance is more accurate and timely than

† Corresponding Author (ORCID: <http://orcid.org/0000-0001-8380-8969>): Professor, Division of Marine System Engineering, Korea Maritime & Ocean University, 727, Taejong-ro, Yeongdo-gu, Busan 49112, Korea, E-mail: skywonju@g.kmou.ac.kr, Tel: +82-51-410-4262

1 Division of Marine Engineering, National Korea Maritime and Ocean University, E-mail: fillish1530386704@outlook.com

2 Interdisciplinary Major of Maritime AI Convergence, National Korea Maritime and Ocean University

3 Division of Marine System Engineering, National Korea Maritime and Ocean University

4 Division of Navigation Convergence Studies, National Korea Maritime and Ocean University

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

scheduled maintenance, ensuring not only the safe operation of ship systems but also reducing material waste and unnecessary labor costs.

The Exponentially Weighted Moving Average (EWMA) is a widely used method for monitoring time series data and process control. It smooths data fluctuations and enhances sequence stability [8][9]. This method is widely applied in various fields for fault detection by setting upper and lower limits for target data, enabling real-time anomaly detection [10]-[14]. Compared to traditional statistical control charts, EWMA typically offers higher accuracy and faster response speed [15][16].

Since engine-room data is continuously collected and belongs to typical time series data, it is well-suited for processing using Recurrent Neural Network (RNN) and its variants in neural network models [17]. In previous studies, most researchers employed composite algorithms to predict target data to improve prediction accuracy and integrated the prediction model into the EWMA control chart for fault detection. However, these algorithms are computationally expensive and do not clearly explain the relationship between prediction accuracy and the EWMA control chart.

This study aims to predict cylinder data by incorporating other engine-room parameters and operational features to develop models using deep learning algorithms. Various features are leveraged to enhance the prediction accuracy of the target variable. Finally, a numerical model is used to verify the relationship between prediction accuracy and the EWMA control chart. To achieve this, a large amount of engine room data has been collected, and models have been developed using RNN, Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU).

## 2. Methodology

### 2.1 Data

To comprehensively record the actual operating conditions of various equipment in the engine room, sensors were installed on a 9,196-ton training ship, with its specifications detailed in **Table 1**. The data was collected through the ship's VIMS system (**Table 1**) during a full day of training voyage on June 25, 2022, resulting in 8,420 data points.

The collected data features are categorized into three types: operational features, which include speed, main engine propeller pitch feedback, and main engine RPM, directly influenced by navigational operations and reflecting the vessel's running state—these parameters help distinguish between normal operational variations and abnormal mechanical failures in fault detection;

**Table 1:** Main specifications of the training ship

Subject	Value	Subject	Value
IMO No.	9807279	Breadth extreme (m)	19
Name	Hannara	Engine type	MAN 6S40MEB9.5
Vessel type	Training ship	Power	6618KW at 146rpm
Gross tonnage (t)	9196	Propeller blades	4
Summer deadweight (t)	3671	Propeller diameter (m)	4
Length overall (m)	133		

global features, which describe the overall performance of the main engine, including fuel supply, cooling system, and scavenging system—these features represent the overall health of the main engine rather than a specific component; and target features, which are the core indicators primarily used for fault detection, playing a crucial role in identifying anomalies and potential failures.

The complex operating environment inside the engine room may affect the reliability of sensors, leading to data loss or noise interference [3]. Therefore, effective data cleaning is essential to ensure the accuracy and reliability of sensor data. This study first classifies the voyage states to distinguish the characteristics of data under different operating conditions. When the ship's acceleration exceeds 0.2 knots per minute, it is defined as an unstable navigation state. Subsequently, data cleaning is performed separately for each voyage state [18]. The  $\mu \pm 3\sigma$  rule is applied to determine the upper and lower limits of each feature value, and outliers beyond this range are removed, where  $\mu$  represents the mean and  $\sigma$  denotes the standard deviation. For the removed outliers and originally missing data, the mean value of the adjacent data points is used for imputation. If too many consecutive data points are missing, a smoothing interpolation method is applied based on the number of missing values to reduce the impact of data fluctuations on the analysis.

Due to the different scales of each feature, and with the aim to reduce the impact of scale inconsistency on the model and decrease computational complexity, this study applies the Min-Max Normalization method to standardize the data, mapping the values to the range of 0 to 1. The algorithm is demonstrated in **Equation (1)**,

$$X'_t = \frac{X_t - X_{t_{min}}}{X_{t_{max}} - X_{t_{min}}} \quad (1)$$

where  $X'_t$  is the new value after normalization,  $X_t$  is the original value to be normalized,  $X_{t_{max}}$  is the maximum value of the data sample, and  $X_{t_{min}}$  is the minimum value of the data sample.

## 2.2 Models

### 2.2.1 RNN

RNN is a type of neural network designed for processing sequential data. Its core characteristic lies in the use of a recurrent structure (**Figure 1**), allowing computations at the current time step to depend on information from previous time steps. The RNN maintains a hidden state to store information from the previous time step (**Equation (2)**) and passes it to the next time step, thereby capturing long-term dependencies. The output at the current time step is determined by both the hidden state from the previous time step and the current input (**Equation (3)**).

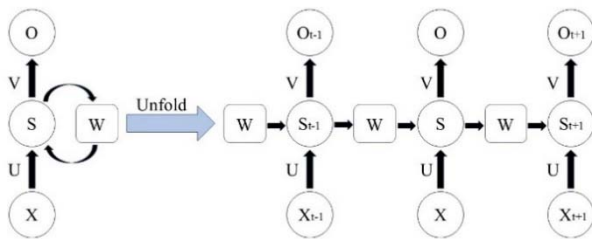
$$S_t = f(UX_t + WS_{t-1} + b) \quad (2)$$

$$O_t = g(VS_t) \quad (3)$$

Here,  $t$  represents the current time step,  $S$  denotes the hidden state, and  $X$  refers to the input information.  $U$  is the weight matrix for the input, while  $W$  is the weight matrix for the hidden layer.  $f$  represents the activation function,  $V$  is the weight matrix from the hidden layer to the output layer, and  $g$  is the output activation function. Additionally,  $b$  denotes the bias term.

### 2.2.2 LSTM

In an RNN structure, as information propagates through the network, the influence of early inputs gradually diminishes, potentially leading to the loss of long-term dependencies. To address this issue, the LSTM network was introduced. As a specialized type of recurrent neural network, LSTM is designed to capture long-term dependencies. Compared to traditional RNN,



**Figure 1:** Structure of RNN

LSTM incorporates a gating mechanism that effectively mitigates the vanishing gradient problem, enhancing stability and accuracy when processing long sequences (**Figure 2**).

The forget gate determines how much of the current unit's memory state should be discarded (**Equation (4)**).

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (4)$$

Where  $f_t$  is the output of the forget gate,  $h_{t-1}$  is the hidden state from the previous time step,  $x_t$  is the input at the current time step,  $W_f$  and  $b_f$  are the trainable weights and biases, and  $\sigma$  is the Sigmoid activation function.

The input gate determines which parts of the current input will be used to update the cell state (**Equation (5)**), and the candidate memory decides how the cell state is updated at the current time step (**Equation (6)**).

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (5)$$

Where  $i_t$  is the output of the input gate, which functions similarly to the forget gate,  $W_i$  and  $b_i$  are the weights and biases.

$$\tilde{C}_t = \tan h(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (6)$$

Where  $\tilde{C}_t$  is the candidate cell state, and  $\tan h$  is the hyperbolic tangent activation function.  $W_c$  and  $b_c$  are the weights and biases.

The current cell state is then computed by updating the memory from the previous time step's cell state, incorporating the memory portion controlled by the forget and input gates (**Equation (7)**).

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (7)$$

Where  $C_t$  is the current cell state.

The output gate determines the content of the hidden state  $h_t$  at the current time step, and the final hidden state  $h_t$  is computed by combining the current cell state  $C_t$  and the output gate  $O_t$  (**Equation (8), (9)**).

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (8)$$

$$h_t = O_t \cdot \tan h(C_t) \quad (9)$$

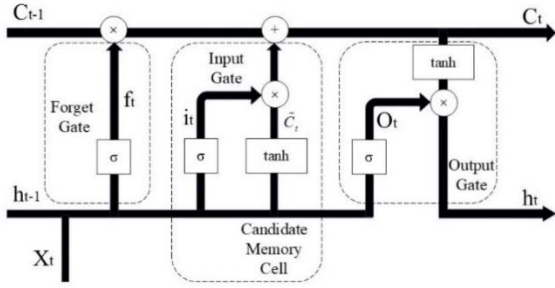


Figure 2: Structure of LSTM

### 2.2.3 GRU

GRU is a simplified version of the RNN, designed to address the vanishing gradient problem. Compared to LSTM, the GRU has a more streamlined structure, offering higher computational efficiency in some applications (Figure 3). It combines the input and forget gates of LSTM into a single update gate and uses a reset gate to control the focus on past information, effectively capturing long-term dependencies in sequential data.

The reset gate determines how much information should be copied from the hidden state of the previous time step. As calculated in Eq. 10,

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \quad (10)$$

where  $r_t$  is the reset gate vector,  $W_r$  is the weight matrix for  $x_t$ ,  $h_{t-1}$  is the hidden state from the previous time step,  $x_t$  is the input,  $\sigma$  is the Sigmoid activation function, and  $b_r$  is the bias term.

The update gate determines how much information from the previous hidden state should be incorporated into the current hidden state (Equation (11)). Then, the candidate hidden state is computed (Equation (12)), and the new hidden state is calculated based on the combination of the candidate hidden state and the update gate (Equation (13)).

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \quad (11)$$

$$\tilde{h}_t = \tanh(W_h \cdot [r_t \cdot h_{t-1}, x_t] + b_h) \quad (12)$$

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t \quad (13)$$

In Eq.11,  $z_t$  is the update gate vector,  $\tilde{h}_t$  in Equation (12) is the candidate hidden state,  $W_z$  in Equation (11) and  $W_h$  in Equation (12) are the weight matrices for  $x_t$ ,  $b_z$  in Equation (11) and  $b_h$  in Equation (12) are the bias terms, and  $\tanh$  in Equation (12) is the activation function. In Equation (13),  $1 - z_t$  controls the proportion of new information introduced.

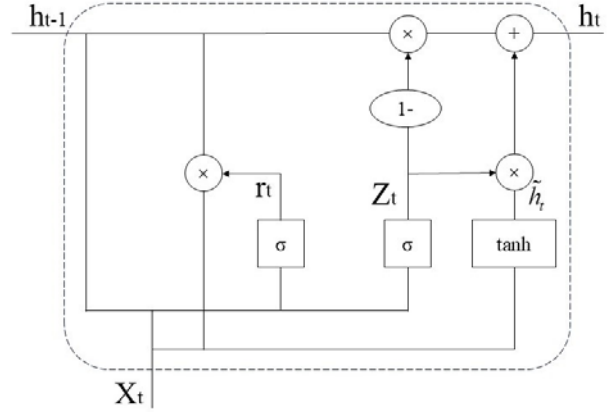


Figure 3: Structure of GRU

### 2.2.4 Hyperparameter Optimization

Bayesian optimization is particularly suitable for evaluating functions with high cost or computational complexity. Unlike traditional optimization methods such as gradient descent, Bayesian optimization builds a probabilistic model of the objective function and uses acquired data to identify regions most likely to yield the best results, optimizing progressively. This approach can effectively reduce the model training time. This paper uses expected improvement as the acquisition function for Bayesian optimization, which is the mostly commonly used and takes into account the expected improvement between the current best solution and all candidate points, thus balancing exploration and exploitation.

### 2.2.5 EWMA

The EWMA control chart is a process control mechanism commonly used to monitor changes in time series and achieve sequence stability. It does not require the assumption of normality in the data, making it robust for anomaly detection in time series. This method employs adaptive upper and lower boundaries as thresholds, assigns different weights to the data, incorporates a smoothing factor to calculate the moving average at the current moment, and classifies data points exceeding these boundaries as anomalies.

The EWMA control chart is normally built on data excluding those containing faults. In this paper, the EWMA control chart is constructed using residuals (Equation (14)) relying only on normal data. Then the EWMA value at the current time step is calculated (Equation (15)). The upper and lower control limits of the chart are calculated. By controlling parameters, the EWMA values are ensured to fluctuate within the limits (Equation (16), (17)).

$$r_k = \hat{y}_k - y_k \quad (14)$$

$$z_k = \lambda r_k + (1 - \lambda r_k) z_{k-1} \quad (15)$$

$$UCL = \mu_0 + L\sigma \sqrt{\frac{\lambda}{2-\lambda} [1 - (1-\lambda)^{2k}]} \quad (16)$$

$$LCL = \mu_0 - L\sigma \sqrt{\frac{\lambda}{2-\lambda} [1 - (1-\lambda)^{2k}]} \quad (17)$$

In **Equation (14)**,  $r_k$  is the residual of the k-th data point, which is the difference between the actual value and the predicted value,  $\hat{y}_k$  is the predicted value for the k-th data point,  $y_k$  is the actual value. In **Equation (15)**,  $z_k$  is the EWMA value for the k-th data point,  $\lambda$  is the smoothing factor. UCL and LCL in **Equation (16)** and **Equation (17)** are respectively the upper and lower control limits of the EWMA,  $\mu_0$  is the target initial value, and L can adjust the control limit width.

### 2.3 Model Evaluation

This paper uses  $R^2$ , Mean Absolute Error (MAE), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE) to evaluate the model, aiming to comprehensively measure the model's predictive ability and error characteristics.

$R^2$  is a standard metric for evaluating the fit of a regression model, representing the proportion of the variance in the target features that the model can explain (**Equation (17)**). The value of  $R^2$  ranges from 0 to 1, with values closer to 1 indicating better model fit.

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (18)$$

where  $SS_{res}$  is the sum of squared residuals,  $SS_{tot}$  is the total sum of squares,  $y_i$  is the observed value,  $\hat{y}_i$  is the predicted value, and  $\bar{y}$  is the mean of the observed values.

MAE measures the average absolute difference between the model's predicted values and the actual values, offering intuitive interpretability (**Equation (18)**). A smaller MAE indicates higher model accuracy.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (19)$$

MSE evaluates the model's predictive ability by calculating the square of the prediction errors (**Equation (19)**). MSE imposes a stronger penalty for larger errors, making it more

sensitive to the model's performance when there are significant deviations. A smaller MSE indicates higher model accuracy.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (20)$$

MAPE is a relative error metric that measures the percentage of prediction error relative to the actual value (**Equation (20)**). This metric is particularly useful for comparing data of different scales, providing an intuitive perception of the model's relative accuracy. A smaller MAPE indicates higher model accuracy.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (21)$$

## 3. Results and Discussion

### 3.1 Development of a Predictive Model

This study explores whether combining parameter manipulation with global engine characteristics for joint modeling can improve the prediction accuracy of target features. To achieve this, we constructed models using three deep learning algorithms—RNN, LSTM, and GRU—and evaluated their performance using four assessment metrics. The evaluation was conducted exclusively on the target features.

According to the results in **Tables 2** and **3**, the highest  $R^2$  was observed in the LSTM model built solely on target features, while the lowest MAE and MAPE were achieved by the GRU model using only target features. Meanwhile, the lowest MSE was obtained from the RNN and LSTM models built using all features.

Comparing the average values of the two modeling approaches (**Table 4**), the results indicate that, overall, modeling with all features yields better performance, which contradicts the sources of some individual optimal metrics. We believe this discrepancy is primarily due to the RNN model built solely on target features, whose evaluation metrics exhibited significant deviations, thereby affecting the overall average results. Consequently, the current analysis still supports modeling using only target features.

This result contradicts our initial expectations. Further analysis of the fitting performance of each feature (**Figure 4**) revealed that some global features exhibited poor fit, potentially failing to enhance the predictive capability of the target features. Therefore, in the following section, these features were removed during modeling, and the model was reconstructed to optimize prediction performance.

**Table 2:** Evaluation metrics for modeling based on all features

Models	R <sup>2</sup>	MAE	MSE	MAPE
RNN	0.8441	0.4795	0.686	0.66%
LSTM	0.8441	0.4795	0.686	0.66%
GRU	0.83	0.5421	0.8748	0.68%

**Table 3:** Evaluation metrics for modeling based on target features

Models	R <sup>2</sup>	MAE	MSE	MAPE
RNN	0.5235	0.8176	1.8542	0.81%
LSTM	0.8475	0.4897	0.6879	0.61%
GRU	0.8392	0.4623	0.7074	0.58%

**Table 4:** Mean score of evaluation metrics for each modeling approach

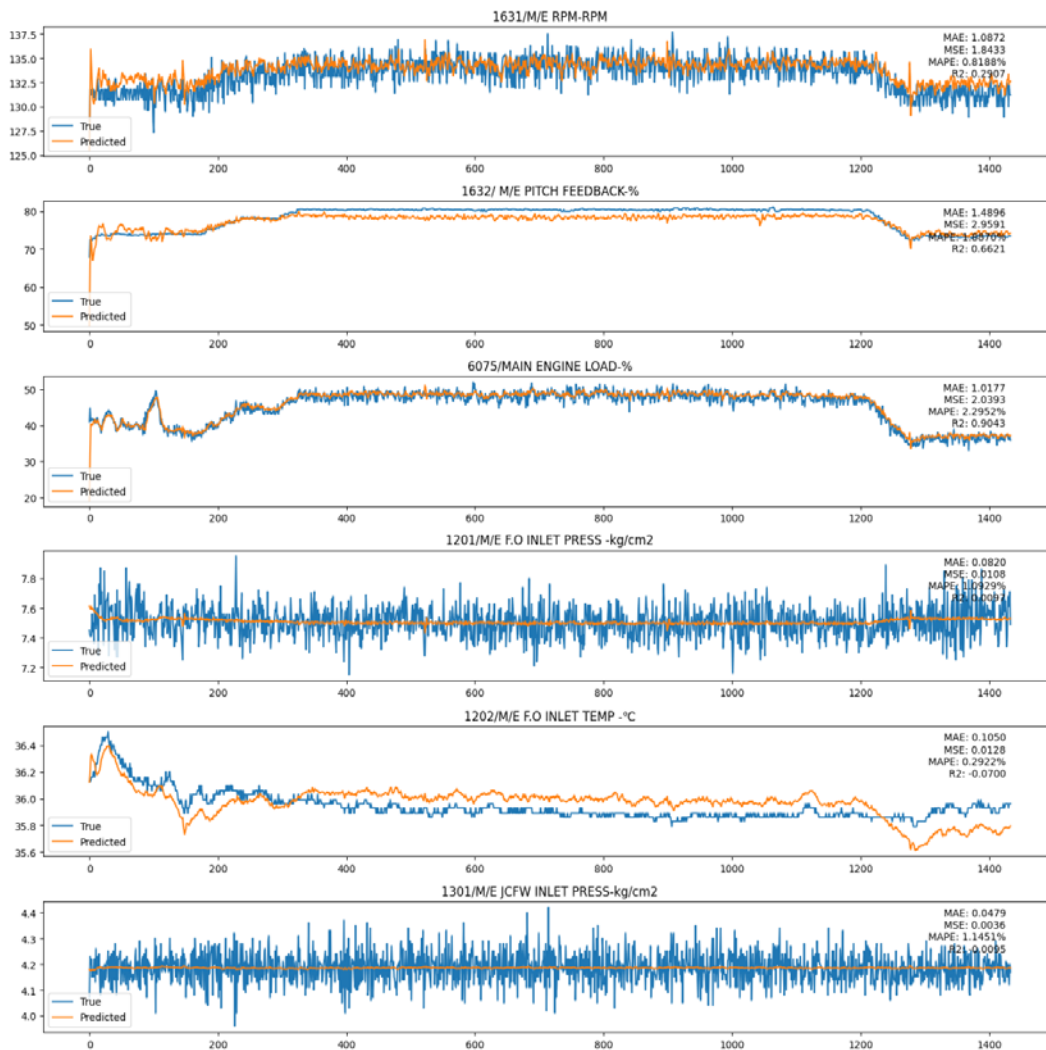
Features	R <sup>2</sup>	MAE	MSE	MAPE
All features	0.8394	0.5004	0.749	0.67%
Only target features	0.7367	0.5899	1.0832	0.67%

### 3.2 Modeling with All Features Excluding Unfitted Ones

As in **Figure 4**, the M/E F.O inlet pressure and the M/E JCFW inlet pressure cannot be fitted in the model. The R<sup>2</sup> of the main engine speed and the F.O inlet temperature are also too low. Likewise, as shown in this figure, the errors of other features are also relatively large. Such global features cannot bring help to the prediction of the whole model. Therefore, these six features were deleted and excluded from the training data. The modeling is renewed using the three algorithms.

The evaluation metrics for the three models are shown in **Table 5**. Compared to modeling with all features or using only target features, the RNN model achieved the highest R<sup>2</sup> and performed best in terms of MAE and MSE. Meanwhile, the GRU model obtained the lowest MAPE, and its performance was similar to that of the RNN model.

A comparison of the average evaluation metrics across the three



**Figure 4:** Cases of partially unfit features

**Table 5:** Evaluation metrics for models excluding unfitted features

Models	R <sup>2</sup>	MAE	MSE	MAPE
RNN	0.8763	0.4336	0.582	0.57%
LSTM	0.8494	0.5512	0.8266	0.75%
GRU	0.8453	0.4454	0.6097	0.56%

**Table 6:** Mean score of three modeling approaches to three evaluating indicators

Features	R <sup>2</sup>	MAE	MSE	MAPE
All features	0.8394	0.5004	0.749	0.0067
Only target features	0.7367	0.5899	1.0832	0.0067
All without unfitted features	0.857	0.4768	0.6728	0.0063

modeling approaches (**Table 6**) revealed that all optimal metrics were obtained from the modeling approach that excluded the poorly fitted features. Ultimately, we confirmed that incorporating selected data reflecting the overall operational state of the engine room, along with control data, can effectively improve the prediction accuracy of target features.

### 3.3 Modeling with All Features Excluding Unfitted Ones

The maritime industry is generally reluctant to share performance and status datasets, particularly when the data may contain errors. As a result, researchers typically use simulated faults to test models, employing linear adjustment and expansion methods to generate out-of-limit fault data [19].

When constructing the EWMA control chart, it is necessary to continuously predict the target value using residuals. Due to the model’s construction approach, predictions require a combined dataset of engine room data, operational data, and target data. However, simply using linear adjustments and expansions of target features does not yield the corresponding engine room and manipulation data, thus preventing the use of this method for model validation.

A literature review reveals that no study has explicitly explored the relationship between prediction accuracy and EWMA control charts. Therefore, this study developed a pure mathematical model and analyzed the relationship between prediction accuracy and EWMA through simulations of normal and fault data. For this, 100 arrays with an average value of 10 were created, and random small fluctuations were added to represent the true values. The predicted values were set with varying fluctuations depending on the accuracy and stability (**Table 7**).

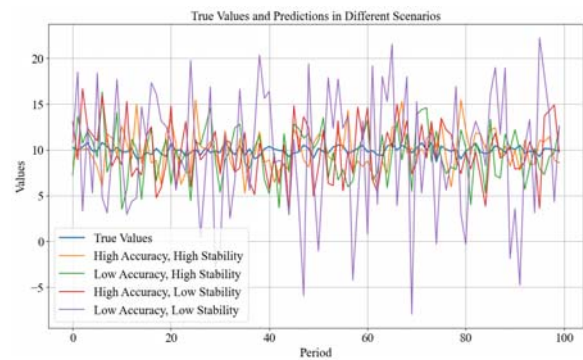
As revealed in **Table 7**, the values for low accuracy and high

stability were similar to those for high accuracy and high stability. This is because when accuracy is low, stability is necessarily poor, and when stability is poor, accuracy is also low. Therefore, numerical simulation can only ensure a relative ranking of both accuracy and stability.

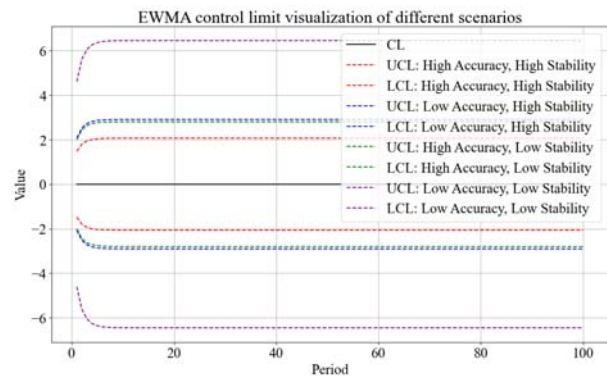
**Figure 5** shows the generated arrays, and the EWMA control lines based on these data are shown in **Figure 6**. It can be observed that higher accuracy and stability result in smaller EWMA control limits, making it easier to detect faults. It indicated that when the predictive model shows accurate and stable performance, the residuals tend to become smaller and accordingly the EWMA values and control limits are also smaller. When a fault occurs, the predicted value will deviate from the expected values and the residuals will change significantly, making them easier to be recognized by the EWMA.

**Table 7:** Evaluation of mathematical simulations of predicted and real values

	High accuracy		Low accuracy	
	MAE	MSE	MAE	MSE
High stability	1.7562	4.6928	2.3451	7.9731
Low stability	2.3217	8.1298	5.5162	45.7797



**Figure 5:** True Values and Predictions in Different Scenarios



**Figure 6:** EWMA control limit in four scenarios

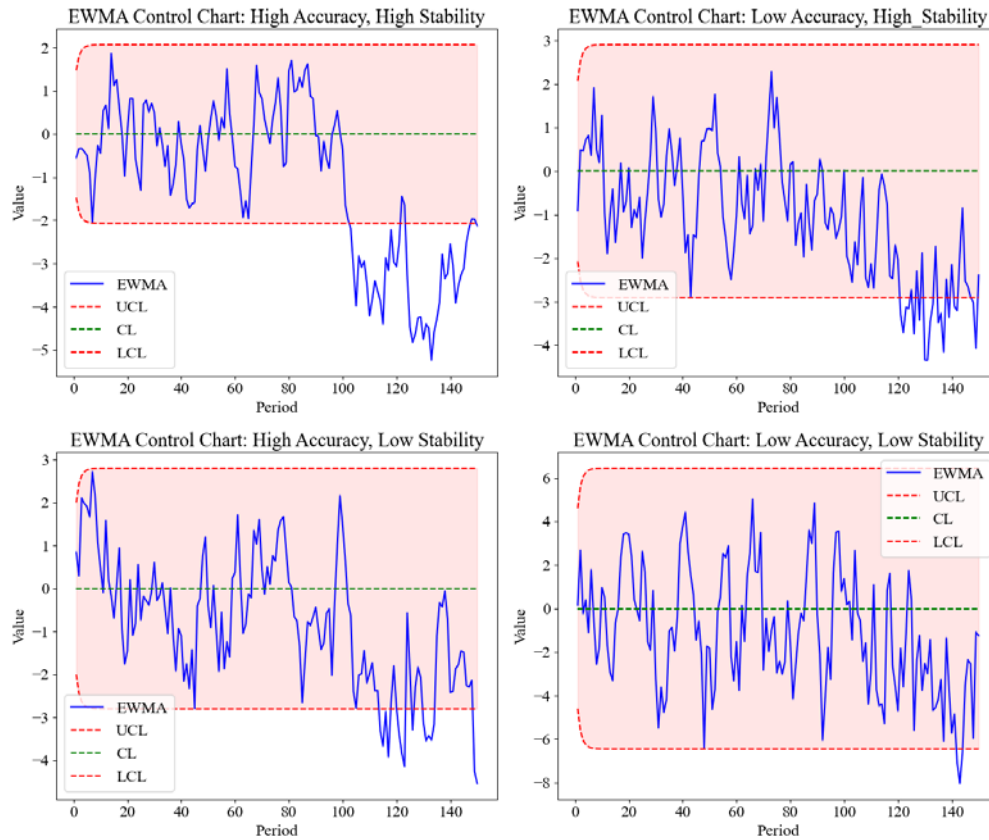


Figure 7: EWMA control charts

Additionally, the control limits for high accuracy and low stability are slightly smaller than those for low accuracy and high stability. This further indicates that when accuracy and stability are similar, higher accuracy facilitates fault detection.

To test the fault detection capability of the model, we added 50 sets of data with a mean value of 13 and introduced random small fluctuations. Meanwhile, the mean of the predicted values remained at 10, indicating that the model failed to correctly predict the true values. As shown in Fig. 7, the EWMA control chart with high accuracy and high stability detected faults the earliest. In contrast, as for the EWMA control chart with low accuracy and low stability, although from position 142 some points of EWMA value exceeded the control limits, the EWMA values were observed to quickly return within the control limits. Additionally, the data with high accuracy but low stability exceeded the control limits earlier than the data with low accuracy but high stability, but both sets of data continued to fluctuate both inside and outside the control limits.

The final conclusion is that under conditions of high accuracy and high stability, the EWMA control limits are smaller, making fault detection easier. Furthermore, during the early stages of a

fault, the data may fluctuate repeatedly inside and outside the control limits. Therefore, a threshold should be set to determine whether the duration of continuous exceedance of the control limits is sufficient to confirm the occurrence of a fault.

## 4. Conclusion

In conclusion, this study investigates whether combining operational features with global engine features for joint modeling can improve the prediction accuracy of target features. The results indicate that adding data reflecting the overall operating status of the engine room (global features) and the operational data to the target features to develop models can significantly enhance the FOC prediction accuracy. And among all the models, the RNN model performed the best.

Through the verification with mathematical models and the EWMA control chart, we find that higher accuracy and stability result in smaller EWMA control limits. Furthermore, when the accuracy and stability of the two predictive models are similar, their EWMA control limits also show similarity. But meanwhile it can be noticed that the model with relatively higher accuracy leads to a smaller control limit, compared with that with lower

accuracy but higher stability. Additionally, it is recommended to set a threshold to determine whether the EWMA values consistently exceed the control limits, thus defining the occurrence of faults.

Based on this study, it can be inferred that the modeling approach of incorporating operational data and data reflecting the overall operating status of the engine room has the potential to improve prediction accuracy and narrow the range of EWMA control limits, thereby achieving more efficient fault detection.

The limitation of this study lies in the lack of real fault data, which prevented us from verifying the actual performance of the EWMA control chart in practical applications. Although verification with simulated data provides some reference value, differences between simulated data and real fault data may exist, limiting the authenticity of the model and control chart evaluation.

### Author Contributions

Conceptualization, S. Piao and W. J. Lee; Methodology, S. Piao; Writing – Original Draft & Editing, S. Piao; Data Collection & Curation, J. J. Hur, J. H. Im, and M. K. Kang; Supervision, W. J. Lee; Project Administration, W. J. Lee; Writing – Review & Editing, S. Piao and W. J. Lee..

### Acknowledgement

This research was supported by Korea Basic Science Institute (National research Facilities and Equipment Center) grant funded by Ministry of Education (grant No. 2022R1A6C101B738); and was a part of the project titled 'Fostering Talent in Advanced Ship Blue Tech (RS-2025-02221147)', funded by the Ministry of Oceans and Fisheries, Korea.

### References

- [1] C. Dere and C. Deniz, "Load optimization of central cooling system pumps of a container ship for the slow steaming conditions to enhance the energy efficiency," *Journal of Cleaner Production*, vol. 222, pp. 206-217, 2019. Available: <https://doi.org/10.1016/j.jclepro.2019.03.030>.
- [2] I. Ančić, G. Theotokatos, and N. Vladimir, "Towards improving energy efficiency regulations of bulk carriers," *Ocean Engineering*, vol. 148, pp. 193-201, 2018. Available: <https://doi.org/10.1016/j.oceaneng.2017.11.014>.
- [3] J. Sun, H. Ren, Y. Duan, X. Yang, D. Wang, and H. Tang, "Fusion of multi-layer attention mechanisms and CNN-LSTM for fault prediction in marine diesel engines," *Journal of Marine Science and Engineering*, vol. 12, no. 6, 990, 2024. Available: <https://doi.org/10.3390/jmse12060990>.
- [4] L. A. Malm, J. Enstrom, L. M. Hager, and P. Stalberg, *Main Engine Damage Study*; The Swedish Club: Hong Kong, China, 2020.
- [5] R. Ahmad and S. Kamaruddin, "An overview of time-based and condition-based maintenance in industrial application," *Computers & Industrial Engineering*, vol. 63, no. 1, pp. 135-149, 2012. Available: <https://doi.org/10.1016/j.cie.2012.02.002>.
- [6] K. H. Park, "Effect of cylinder wall temperature on marine engine combustion," *Journal of Advanced Marine Engineering and Technology*, vol. 47, no. 6, pp. 309-316, 2023. Available: <https://doi.org/10.5916/jamet.2023.47.6.309>.
- [7] J. Im, B. Rho, and S. Lee, "Empirical case study of black-out incident caused by incomplete combustion and blow-by in ship generator engines," *Journal of Advanced Marine Engineering and Technology*, vol. 48, no. 4, pp. 186-197, 2024. Available: <https://doi.org/10.5916/jamet.2024.48.4.186>.
- [8] R. Isermann, *Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance*: Springer Science & Business Media, 2006.
- [9] E. Garoudja, F. Harrou, Y. Sun, K. Kara, A. Chouder, and S. Silvestre, "Statistical fault detection in photovoltaic systems," *Solar Energy*, vol. 150, pp. 485-499, 2017. Available: <https://doi.org/10.1016/j.solener.2017.04.043>.
- [10] M. Mansouri, A. Al-Khazraji, M. Hajji, M. F. Harkat, H. Nounou, and M. Nounou, "Wavelet optimized EWMA for fault detection and application to photovoltaic systems," *Solar Energy*, vol. 167, pp. 125-136, 2018. Available: <https://doi.org/10.1016/j.solener.2018.03.073>.
- [11] F. Harrou, M. N. Nounou, H. N. Nounou, and M. Madakyaru, "PLS-based EWMA fault detection strategy for process monitoring," *Journal of Loss Prevention in the Process Industries*, vol. 36, pp. 108-119, 2015. Available: <https://doi.org/10.1016/j.jlp.2015.05.017>.
- [12] M. I. Awad, M. AlHamaydeh, A. Faris, "Fault detection via nonlinear profile monitoring using artificial neural networks," *Quality and Reliability Engineering International*, vol. 34, pp. 1195-1210, 2018. Available: <https://doi.org/10.1002/qre.2318>.
- [13] N. A. Adegoke, S. A. Abbasi, A. B. A. Dawod, and M. D. M. Pawley, "Enhancing the performance of the EWMA

control chart for monitoring the process mean using auxiliary information,” *Quality and Reliability Engineering International*, vol. 35, pp. 920-933, 2019. Available: <https://doi.org/10.1002/qre.2436>.

- [14] P. Bangalore and M. Patriksson, “Analysis of SCADA data for early fault detection, with application to the maintenance management of wind turbines,” *Renewable Energy*, vol. 115, pp. 521-532, 2018. Available: <https://doi.org/10.1016/j.renene.2017.08.073>.
- [15] A. Mukherjee, Z. L. Chong, and M. B. C. Khoo, “Comparisons of some distribution-free CUSUM and EWMA schemes and their applications in monitoring impurity in mining process flotation,” *Computers & Industrial Engineering*, vol. 137, 106059, 2019. Available: <https://doi.org/10.1016/j.cie.2019.106059>.
- [16] M. Shamsuzzaman, S. Haridy, A. Maged, and I. Alsayouf, “Design and application of dual-EWMA scheme for anomaly detection in injection moulding process,” *Computers & Industrial Engineering*, vol. 138, 106132, 2019. Available: <https://doi.org/10.1016/j.cie.2019.106132>.
- [17] Z. C. Lipton, J. Berkowitz, and C. Elkan, A Critical Review of Recurrent Neural Networks for Sequence Learning, arXiv Preprint, CoRR, abs/1506.00019. Available: <https://doi.org/10.48550/arXiv.1506.00019>.
- [18] S. Piao, M. H. Park, S. Yeo, K. W. Chun, J. -H. Jee, and W. J. Lee, “Expanding the range of ship fuel consumption prediction: A multi-algorithm feature selection approach,” *Ocean Engineering*, vol. 316, 119944, 2025. Available: <https://doi.org/10.1016/j.oceaneng.2024.119944>.
- [19] M. Cheliotis, I. Lazakis, and A. Cheliotis, “Bayesian and machine learning-based fault detection and diagnostics for marine applications,” *Ships and Offshore Structures*, vol. 17, no. 12, pp. 2686-2698, 2022. Available: <https://doi.org/10.1080/17445302.2021.2012015>.