

Hybrid deep learning method for predicting vertical bending moments at midship

Jun Ding^{1,2} , Peiqiao Zhu^{1,2}, Zhang Zhu^{1,2}, and Yiming Qiang^{1,2}

¹China Ship Scientific Research Center, Wuxi 214082, China

²Taihu Laboratory of Deepsea Technological Science, Wuxi 214082, China



Cite This: *Ocean*, 2026, 2, 9470022



Read Online

ABSTRACT: Rapid and accurate prediction of the vertical bending moment at midship is critical for ensuring the safety of ships in severe sea conditions. However, conventional prediction methods are not effective for achieving highly precise and reliable uncertainty quantification. This is primarily attributed to the strong nonlinearity and chaotic nature of marine load signals, which is further complicated by high-frequency disturbances such as wave-induced vibration and slamming events. To overcome these limitations, this study proposes a hybrid deep learning prediction method that integrates variational mode decomposition, a bidirectional gated recurrent unit, and a quantile regression neural network. The three key stages of this method are as follows: First, the original bending-moment signal is decomposed into a set of low-complexity subsequences to mitigate interference induced by the inherent nonlinearity of the signal. Thereafter, the Bidirectional Gated Recurrent Unit (BiGRU) network is used to capture the temporal dependencies of the subsequences for deterministic prediction. Finally, the prediction uncertainty is quantified based on the conjunction and confidence intervals from the Quantile Regression Neural Network (QRNN). The hybrid model was validated in evaluations of a ship model with wave-induced vibration and slamming events. Satisfactory performance was achieved, with an average coefficient of determination of 0.9533 and a coverage rate of 94.22% at the 95% confidence level. These results confirm that the model effectively enables accurate deterministic predictions and effective uncertainty quantification for the vertical bending moment at midship, offering practical support for ship design and real-time operational decisions.

KEYWORDS: midship vertical bending moment, hybrid model, experimental validation, intelligent prediction

1 Introduction

The safety of hull structures is a fundamental concern throughout the life-cycle of ships and offshore structures [1]. A critical factor governing this safety is the wave-induced vertical bending moment at midship, which directly determines the global dynamic load response of the hull. Consequently, accurate and real-time determination of this parameter essential for rational structural design and informed decision-making related to operation [2]. In modern ship design, classification societies have mandated strict verification of the structural integrity of the hull girder when subjected to specified vertical bending moments. Such data are derived from extensive theoretical and experimental research. These design parameters incorporate significant safety margins but cannot fully account for the complex and stochastic nature of real sea conditions. This issue is becoming more prominent as ships become larger and faster, which intensifies hydroelastic effects such as wave-induced vibration and slamming [3]. Consequently, the ability to rapidly and accurately predict the actual vertical bending moment at midship is of great practical importance in optimizing hull designs and providing essential data for improving navigational

safety and ensuring structural integrity during ship operation.

Recently, the development of deep learning techniques has provided novel approaches for rapidly predicting the performance of ships and offshore structures. Deep learning has been used in predicting the seakeeping characteristics of ships [4,5], assessing maneuverability [6,7], forecasting navigation trajectories [8,9], and predicting flow-fields [10,11]. Numerous studies have focused on predicting the motion performance of ships, whereas research on the rapid and intelligent forecasting of global hull girder loads, such as the midship bending moment, remains relatively limited. This disparity is mainly attributable to the strongly nonlinear and chaotic characteristics inherent in the time-series of such loads.

In early studies, Xu and Haddara [12] pioneered the use of an artificial neural network for predicting ship loads and estimating the wave-induced vertical bending moment in real-time using the heave and pitch motions collected by a hull response monitoring system. Moreira and Soares [13] further developed a technique for time-domain mapping between wave-induced loads and ship motions. Zhou et al. [14] used a deep feedforward neural network for mapping 20 strain measurement points on the hull and the ice loads on the ship, enabling real-time prediction of ice loads. However, these feedforward neural networks do not consider the temporal characteristics of load sequences; therefore, their prediction accuracy is significantly constrained. Some scholars have employed recurrent neural network (RNN) models, specifically designed for time-series analysis, for addressing ship dynamic

Received: December 26, 2025; **Revised:** January 20, 2026

Accepted: February 2, 2026

 Address correspondence to Jun Ding, dingjun@cssrc.com.cn

response loads. Wang et al. [15] introduced a method combining an long short-term memory (LSTM) network with an error-correction mechanism and employed two serial neural networks for achieving more accurate prediction of the bending moment at midship. Liu et al. [16] proposed a method for predicting the vertical bending moment at midship based on an LSTM model and demonstrated its predictive advantage in head-sea conditions.

Although RNN models, such as LSTM, can effectively capture temporal information regarding bending-moment loads at midship, they cannot fully model the high-frequency load components, such as those induced by wave springing and slamming whipping, which are significant for large vessels. To address this challenge, scholars have adopted a “decomposition-superposition” strategy. Liu et al. [17] proposed a method integrating a generalized modal superposition approach with LSTM. This method establishes mappings between ship motions and individual load modes for predicting vertical bending moments and accelerations, successfully capturing the slamming-induced signals in the structural response. Other scholars [18,19] employed modal decomposition methods, such as empirical mode decomposition (EMD) and empirical wavelet transform (EWT), to preprocess ship-motion sequences, achieving a certain degree of improvement in the prediction accuracy.

However, the existing methods still have two key limitations. First, most models rely on ship-motion data as input, which means any error in the motion prediction is directly propagated to the predicted load. Second, existing methods primarily focus on deterministic forecasting and lack the ability to quantify the uncertainty in load predictions, making it difficult to meet the needs of risk assessment.

To address these challenges, this study proposes an autoregressive hybrid model comprising variational mode decomposition–bidirectional gated recurrent unit–quantile regression neural network (VMD-BiGRU-QRNN). The main innovations are as follows: (1) This approach departs from the conventional dependence on ship motion inputs. Instead, autoregressive forecasting is performed directly on the time-series of the bending-moment load at midship, which prevents propagation of errors from motion prediction. (2) The model uses variational mode decomposition instead of empirical mode decomposition for signal decomposition. It solves the inherent issue of variable mode numbers in EMD while effectively reducing the signal complexity. (3) The model integrates the bidirectional temporal modeling of BiGRU with the quantile regression capability of QRNN. This hybrid model allows for both deterministic load prediction and probabilistic interval forecasting. The following sections detail the theoretical basis and framework design of the model. The hybrid model is validated through tests using a ship model to demonstrate its effectiveness and engineering applicability.

2 Theoretical basis and model framework

2.1 Theoretical basis

2.1.1 Variational mode decomposition (VMD)

Variational mode decomposition (VMD) [20] is an adaptive and fully nonrecursive method for modal variation and signal processing. The underlying concept involves decomposing the

original signal into several sparse modal components. A constrained variational model is constructed, which seeks a solution that minimizes the sum of the bandwidths of all modes, effectively separating different frequency components in the signal. Implementation of this approach can be divided into three steps.

First, a constrained variational model is constructed. The Hilbert transform is applied to each modal function and the frequency of each mode is modulated to the corresponding baseband by taking the center frequency of the mode as the basis. The L^2 norm of the gradient of each signal is calculated for estimating the sum of the bandwidths of all modes. The constrained model is expressed as follows:

$$\min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right)^* u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \quad \text{s.t.} \quad \sum_{k=1}^K u_k = f \quad (1)$$

where K is the number of modal components to be decomposed, u_k and ω_k are the k th modal component and the corresponding center frequency after the decomposition, respectively. $\delta(t)$ is the Dirac delta function.

A quadratic penalty term and Lagrange multipliers are introduced to transform the constrained problem into an unconstrained augmented Lagrangian form, which is expressed as follows:

$$L(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right)^* u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle \quad (2)$$

where α is the punishment coefficient and λ represents the Lagrange multiplier.

Finally, the alternating direction method of multipliers is adopted for iterative solutions. The modal components u_k , center frequencies ω_k , and Lagrange multipliers λ_k are updated sequentially until convergence is achieved to find the saddle point of the augmented Lagrangian function. Finally, inverse Fourier transform is applied to the solved modal components $\hat{u}_k^{n+1}(\omega)$, generating all the modal components decomposed by VMD.

2.1.2 Bidirectional gated recurrent unit (BiGRU)

The gated recurrent unit (GRU) is a special type of recurrent neural network proposed by Chung et al. [21] in 2014. It improves the LSTM network by simplifying the three traditional gating mechanisms (forget, input, and output gates) into two core gates, namely, a “reset gate” and an “update gate”. The reset gate controls how much hidden-state information is preserved from the previous step, whereas the update gate dynamically balances the contributions of the prior hidden state and the current candidate state to the current hidden state.

This simplification of GRU reduces the computational complexity while maintaining the ability to capture long-term dependencies, making it more suitable for real-time processing of ship-motion signals under complex sea conditions.

The formulas for calculating each neuron unit are given in Eqs. 3–6:

$$R_t = \sigma(X_t \mathbf{W}_{xr} + h H_{t-1} \mathbf{W}_{hr} + b_r) \quad (3)$$

$$Z_t = \sigma(X_t \mathbf{W}_{xz} + h_{t-1} \mathbf{W}_{hz} + b_z) \quad (4)$$

$$\tilde{h}_t = \tanh[X_t \mathbf{W}_{xh} + (Z_t \odot h_{t-1}) \mathbf{W}_{zh} + b_h] \quad (5)$$

$$h_t = Z_t \odot h_{t-1} + (1 - Z_t) \odot \tilde{h}_t \quad (6)$$

where h_t is the hidden state of the unit at time t , X_t is the input of the unit at time t , $\tilde{h}_t$ is the candidate hidden state of the unit. $\mathbf{W}_{xz}$ and $\mathbf{W}_{hz}$ are the two weight matrices of the reset gate, $\mathbf{W}_{xh}$ and $\mathbf{W}_{zh}$ are the weight matrices of the update gate, and $\mathbf{W}_{xh}$ and $\mathbf{W}_{zh}$ are the weight matrices for calculating the candidate hidden states. The symbol $\odot$ represents the matrix multiplication.

However, GRU cannot capture the long-term dependencies of the time-series from both forward and backward directions. To enhance modeling of the time-series, a BiGRU neural network is constructed by integrating forward and backward GRUs, which recursively feed back the information of both past and future hidden states.

2.1.3 Quantile regression neural network (QRNN)

QRNN [22] is a reliable method for solving quantile regression, which integrates the concept of quantile regression into the loss function design of neural networks. By defining specific loss terms for each quantile, the network can precisely fit across different distribution positions. This unique design allows QRNN to capture the complex nonlinear relationship between the predictor and response variables and fully characterize the conditional distribution of the response variables. The specific quantile loss function is given by Eq. 7:

$$\text{QuantileLoss} = \frac{1}{N} \sum_{i=1}^N \mathbb{I}_{Y_i \geq y_i} (1 - \gamma) |y_i - Y_i| + \mathbb{I}_{Y_i < y_i} \gamma |y_i - Y_i| \quad (7)$$

where y_i is the predicted value of the ship's dynamic response, Y_i is the true value of the ship's dynamic response, and α is the quantile.

2.2 Model framework and structure

The VMD-BiGRU-QRNN hybrid model proposed in this study consists of three main parts: data processing, deterministic prediction, and probabilistic prediction. To construct the dataset, time-domain data for the vertical bending moment at midship under different working conditions are obtained through tests using ship models. For each working condition, the collected data for the vertical bending moment at midship are further divided into three subsets, namely, the training, validation, and test sets.

First, data processing is performed. The data collected from the tests employing the ship model are converted into full-scale data according to the scale ratio, followed by data filtering and normalization. The normalization formula is given in Eq. 8:

$$x'_i = \frac{x_i - \mu}{\sigma} \quad (8)$$

where μ is the mean value, σ is the standard deviation, and x'_i denotes the normalized data.

The vertical bending moment at midship is characterized by randomness and time-variability, making the direct development of an accurate prediction model based on the raw data difficult. Therefore, in this study, the normalized vertical bending moment at midship $X(t)$ is decomposed into multiple subsequences $[\text{Sub}_1, \text{Sub}_2, \dots, \text{Sub}_n]$ using VMD, as expressed in Eq. 9:

$$[\text{Sub}_1, \text{Sub}_2, \dots, \text{Sub}_n] = \text{VMD}[X(t)] \quad (9)$$

The data for each operating condition in this study are derived from time-domain data collected through five separate tests of the model. Accordingly, when performing VMD, the decomposition is carried out individually for each test of the model. Thus, a total of five VMD operations is conducted per operating condition.

Second, deterministic prediction is implemented using the BiGRU network. For each decomposed subsequence, the data are partitioned into training, validation, and test sets in a fixed ratio of 3:1:1. For each subsequence, the BiGRU model is trained independently, where the validation set is used to detect potential overfitting or underfitting tendencies. The prediction results for the test set of each subsequence are summed to generate the deterministic prediction of the vertical bending moment at midship.

Finally, probabilistic prediction is achieved using the QRNN. The raw data are input into the VMD-BiGRU framework to generate the deterministic predictions of the vertical bending moment at midship, denoted as $y = [y_1, y_2, \dots, y_n]$; the true value is denoted as $Y = [Y_1, Y_2, \dots, Y_n]$. The training dataset for the QRNN model is constructed as $[y_i, Y_i]$. By specifying different quantiles, the upper and lower bounds for predicting the vertical bending moment at midship can be derived using the equations given below:

$$y_i^u = \text{QRNN}(y_i, \alpha^u) \quad (10)$$

$$y_i^l = \text{QRNN}(y_i, \alpha^l) \quad (11)$$

where α^u and α^l are the upper and lower quantile, respectively; y_i^u and y_i^l represent the upper and lower prediction bounds, respectively. The region covered by the lower and upper bounds of the probabilistic prediction is denoted as $CL = [y_i^l, y_i^u]$, which is the probabilistic prediction interval at the given confidence level. The detailed flowchart is illustrated in Fig. 1.

3 Dataset construction and evaluation indicators

3.1 Dataset construction

The training effectiveness and generalization ability of deep learning models are highly dependent on the quality and scale of the training data. A common approach for constructing the dataset for the vertical bending moment at midship involves integrating the numerical simulation results with the test data from the model. However, the simulation results based on traditional potential flow theory often fail to capture the critical high-frequency components of the vertical bending-moment loads at midship. These components are induced by slamming and wave-induced vibration and their absence directly compromises the efficacy of model training. To address this limitation, this study constructs the dataset based on the test results for a model of a 205,000 DWT bulk carrier in a towing tank, with a scale ratio of 1:50. The principal particulars of the full-scale ship and the model are provided in Table 1.

Table 2 presents the model-scale parameters for the three sets of irregular wave conditions, designed to excite wave-induced vibration and slamming responses. To ensure statistical reliability, each irregular wave condition was evaluated 5 times. At the model scale, the total sample duration exceeded 500 s, which is equivalent to approximately 59 min for the full-scale ship. This dataset provides a sufficient number of subrecord samples for

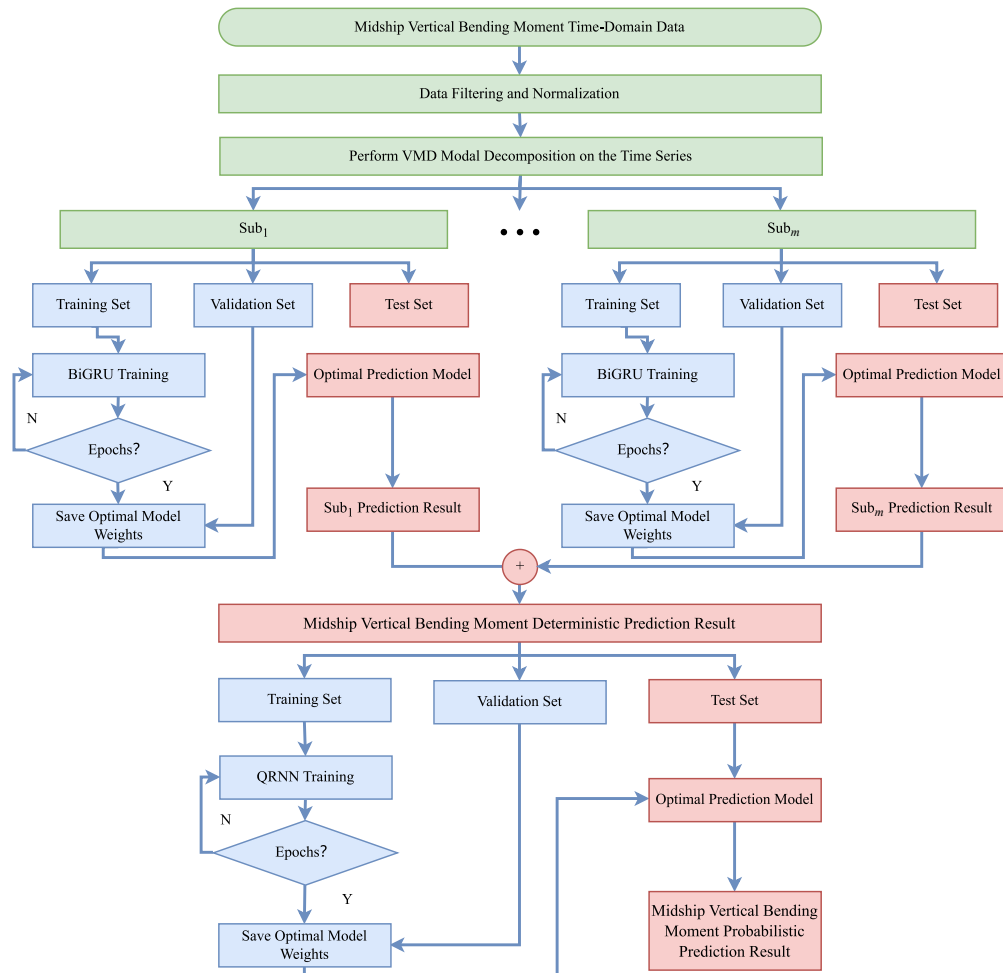


Figure 1 Flowchart of VMD-BiGRU-QRNN.

Table 1 Main dimensions of the full-scale ship and ship model.

Physical quantity or status	Full-scale ship	Ship model
Scale ratio λ	1	50
Overall length L_{OA}/m	300.0	6.0
Length between perpendiculars L_{BP}/m	295.0	5.9
Waterline width B/m	50.0	1.0
Molded depth H/m	24.7	0.494
Forward draft T_f/m	17.71	0.354
Aft draft T_a/m	18.73	0.375
Displacement Δ	235610.1	1.8389
Block coefficient C_b	0.8449	0.8449
Height of the center of gravity Z_g/m	8.03	0.161
Longitudinal position of the center of gravity X_g/m	5.23	0.105

Table 2 Specific parameters of irregular wave conditions.

Condition	Significant wave height $h_{1/3}/mm$	Period T_2/s	Ship speed/(m/s)	Wave direction angle
Series #1	90	1.061	1.077	180°
Series #2	90	1.189	1.077	180°
Series #3	160	1.485	1.077	180°

comprehensive statistical analysis of the response-time histories.

Figure 2 shows the typical time-domain data collected under three irregular wave conditions. The spectrum obtained via Fourier-transform analysis is shown in Fig. 3. The results indicate that the energy spectral density of the high-frequency loads in Condition Series #1 and Condition Series #3 is comparable to that of the low-frequency wave-induced load components. Compared with the traditional dataset containing only low-frequency wave excitation, the current dataset offers considerable value for training deep learning models to accurately predict bending-moment loads at midship under complex sea conditions. Additionally, note that Condition Series #1 corresponds to the wave-induced vibration occurring under mild-sea states, whereas Condition Series #3 is specifically designed to capture the slamming loads in high-sea states.

3.2 Evaluation indicators

To accurately evaluate the effectiveness of the prediction model, this study employs three indicators to systematically evaluate the modal decomposition, deterministic forecasting, and probabilistic forecasting. The three indicators are discussed in the following sections.

3.2.1 Indicators of sequence complexity

The sample entropy (SampEn) [23] can be used as an indicator for quantifying the complexity of a time-series. Generally, a larger value of SampEn indicates weaker autocorrelation and higher sequence complexity. For a given time-series $\mathbf{X} = [x_1, x_2, x_3, \dots, x_n]$, the formula for calculating SampEn is as follows:

$$\text{SampEn} = -\ln \frac{E^{\text{ed}+1}(k)}{E^{\text{ed}}(k)} \quad (12)$$

where $E^{\text{ed}+1}(k)$ and $E^{\text{ed}}(k)$ represent the statistical averages of the time-series under the embedding dimensions $\text{ed}+1$ and ed , respectively.

3.2.2 Indicators of deterministic prediction

Three common indicators, namely, the mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R^2), are selected for evaluating the performance of the deterministic prediction model. These indicators are calculated as follows:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - Y_i| \quad (13)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - Y_i)^2} \quad (14)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - Y_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (15)$$

where n is the number of data points and $\bar{Y}$ is the average of the true values of the ship's dynamic response.

3.2.3 Indicators of probabilistic prediction

To precisely evaluate the results predicted using the probabilistic prediction model, the prediction interval coverage probability (PICP), prediction interval normalized average width (PINAW), and coverage width-based criterion (CWC) were selected as indicators of the probabilistic prediction. The PICP value represents the probability that the predicted target falls within the upper and

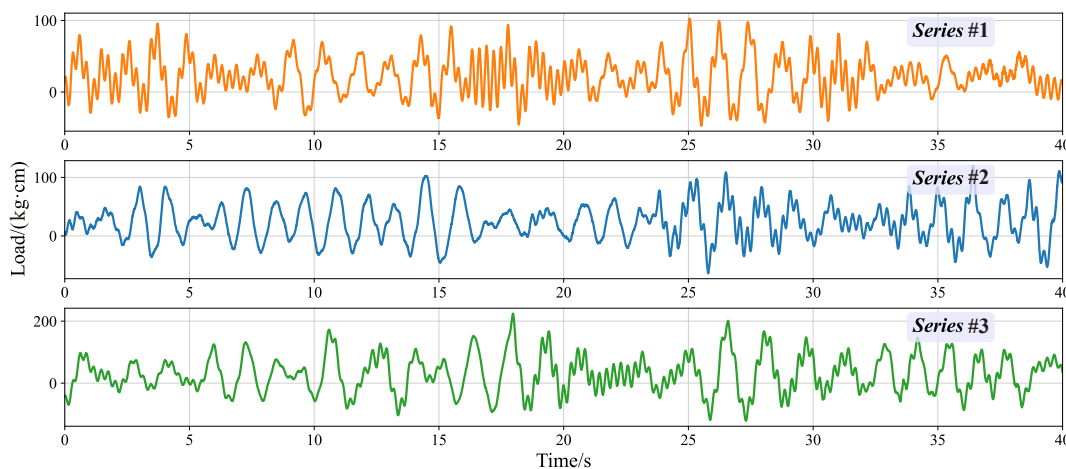


Figure 2 Typical time-domain data for three irregular operating conditions.

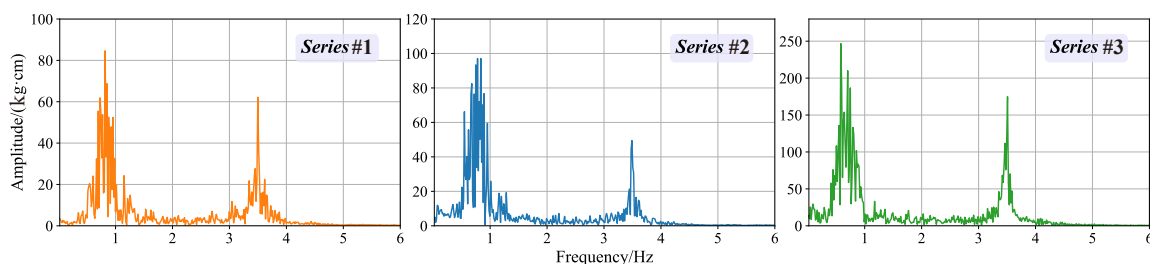


Figure 3 Spectral analysis of operation under irregular wave conditions.

lower limits of the prediction interval. The PINAW is defined as a measure of the width of the prediction interval. The above probabilistic prediction evaluation indicators are calculated as follows:

$$\text{PICP} = \frac{1}{N^p} \sum_{n=1}^{N^p} c_n \times 100\%, \text{ and } \begin{cases} c_n = 1, x_n \in [Q_n^L, Q_n^U] \\ c_n = 0, \text{ otherwise} \end{cases} \quad (16)$$

$$\text{PINAW} = \frac{1}{N^p (\text{Max}_x - \text{Min}_x)} \sum_{n=1}^{N^p} (Q_n^U - Q_n^L) \quad (17)$$

$$\text{CWC} = \text{PINAW} [1 + \text{I}(\text{PICP}) e^{-\rho(\text{PICP}-\alpha)}],$$

$$\text{and } \begin{cases} (\text{PICP}) = 1, \text{PICP} < \alpha \\ (\text{PICP}) = 0, \text{ otherwise} \end{cases} \quad (18)$$

where N_p is the length of the lower bound of the probabilistic prediction, Q_n^U and Q_n^L represent the lower and upper bounds of the probabilistic prediction, Max_x and Min_x are the maximum and minimum values of the actual results respectively, ρ is the constant in [50, 100].

4 Results of analysis

4.1 Model setup

The hybrid VMD-BiGRU-QRNN model was built based on PyTorch and the experimental environment was configured on a regular desktop computer (Memory: 32G; CPU: Intel i5-13500HX; GPU: RTX 4060). The hyperparameters of the prediction model have a significant impact on the prediction results. Thus, a rational parametric analysis is conducted to determine the optimal hyperparameters. Taking Condition *Series #1* as an example, this section analyzes the effects of the window length, VMD level, BiGRU hidden layers, and number of neurons.

4.1.1 Window length

A sliding window of length d is defined for predicting the amplitude of the vertical bending moment at midship for future moments. The window length d is set to 50, 100, 150, and 200, and the results are presented in Table 3. For these tests, the number of BiGRU layers is fixed at 1, the number of hidden layer units is 64, and the Adam optimizer is used. The advance prediction step is set to 40, which corresponds to a real-scale time of approximately 2.8 s.

Theoretically, a longer input window can provide more information for the model, facilitating prediction of the vertical bending moment at midship for future time-steps. However, an excessively long window may introduce excessive invalid information, which is not conducive to model training. Table 3 shows the influence of different input-window lengths on the

prediction accuracy of the model. The results indicate that the model achieves the highest accuracy with an input window length of 100 steps (MAE = 26.4797, RMSE = 34.2121, R^2 = 0.9514). In contrast, a shorter window of 50 steps fails to capture sufficient historical information, leading to inadequate modeling of the long-term load dependencies. An excessively long window (150 or 200 steps) may introduce outdated or weakly correlated data, which introduces noise and degrades the predictive performance of the model, resulting in an R^2 value below 0.942.

4.1.2 VMD modal decomposition

Figure 4 presents a comparison of SampEn for the original sequence and the decomposed intrinsic mode function (IMF) subsequences obtained via VMD. For the three independent conditions (*Series #1–3*), the original signal exhibits a relatively high SampEn value (approximately 0.4), whereas the SampEn of the decomposed IMF subsequences is significantly reduced, ranging from 30% to 80% relative to that of the original sequence. This result indicates that VMD can effectively reduce the complexity and uncertainty of the signal. By decomposing the nonstationary signal into a series of more regular subcomponents, VMD lays a solid foundation for subsequent feature extraction.

Figure 5 presents the results of time- and frequency-domain VMD. In the time domain, the original signal (red curve) exhibits complex nonstationary fluctuations. In contrast, the 6 decomposed IMFs (green curves) all show clear amplitude modulation characteristics, which effectively captured the local time-varying features of the signal. The spectrum from the original signal in the frequency-domain analysis has a wide-band distribution, whereas the spectrum for each IMF is concentrated around a single dominant frequency, with significantly improved energy concentration. From a physical perspective, the frequency distributions of IMF1, IMF4, IMF5, and IMF6 fall within the low-frequency range of the original signal, corresponding to the low-frequency, wave-induced bending moment of the ship. In contrast, IMF2 and IMF3 represent the high-frequency modal components of the vertical bending moment at midship, which under the conditions of *Series #1* are caused by wave-induced vibration. This decomposition effectively solves the coupling issues of the multiple frequency components in the original signal, simplifying the complex spectrum into sparse, narrow-band subsignals. Consequently, the decomposition successfully reduces the dimensions of the nonlinear feature extraction, improves the signal interpretability, and provides a more robust feature for deep learning models. These improvements can improve the accuracy and generalization ability of prediction models.

The decomposition level K of VMD is an important factor affecting the prediction accuracy. A low K cannot fully decouple the nonlinearity and high-frequency noise, thereby degrading the model performance. Conversely, an excessively high K leads to component redundancy and model overfitting. Table 4 presents the model prediction results for different K values.

Table 3 Influence of input window on prediction accuracy.

Window length	MAE	RMSE	R^2
50	28.6029	36.5907	0.9442
100	26.4797	34.2121	0.9514
150	27.6469	37.5349	0.9420
200	29.3682	37.4573	0.9419

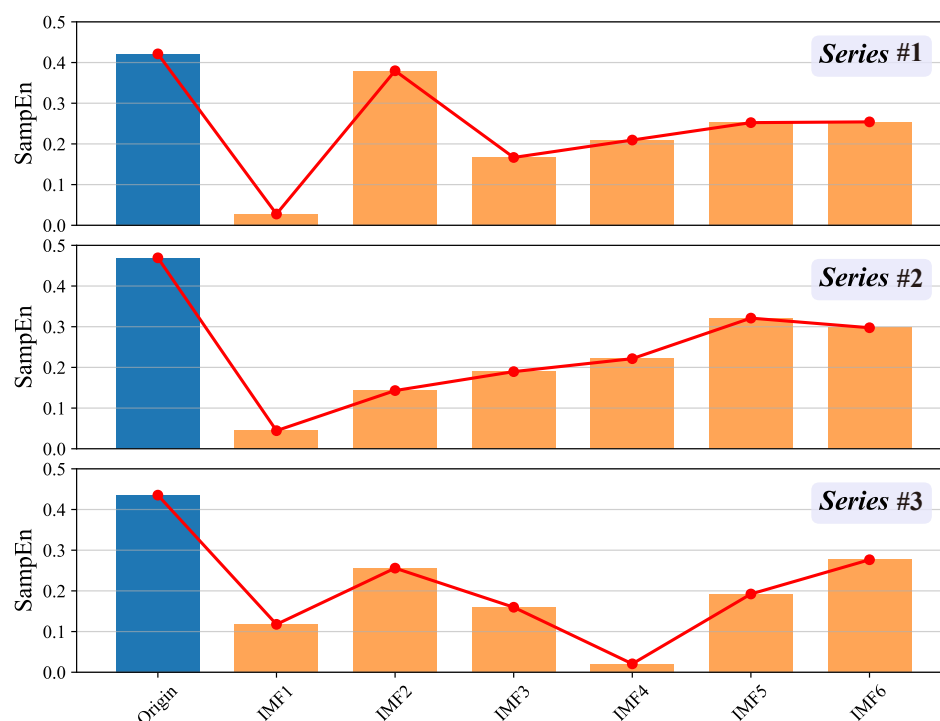


Figure 4 SampEn of the original sequence and the sequences after VMD.

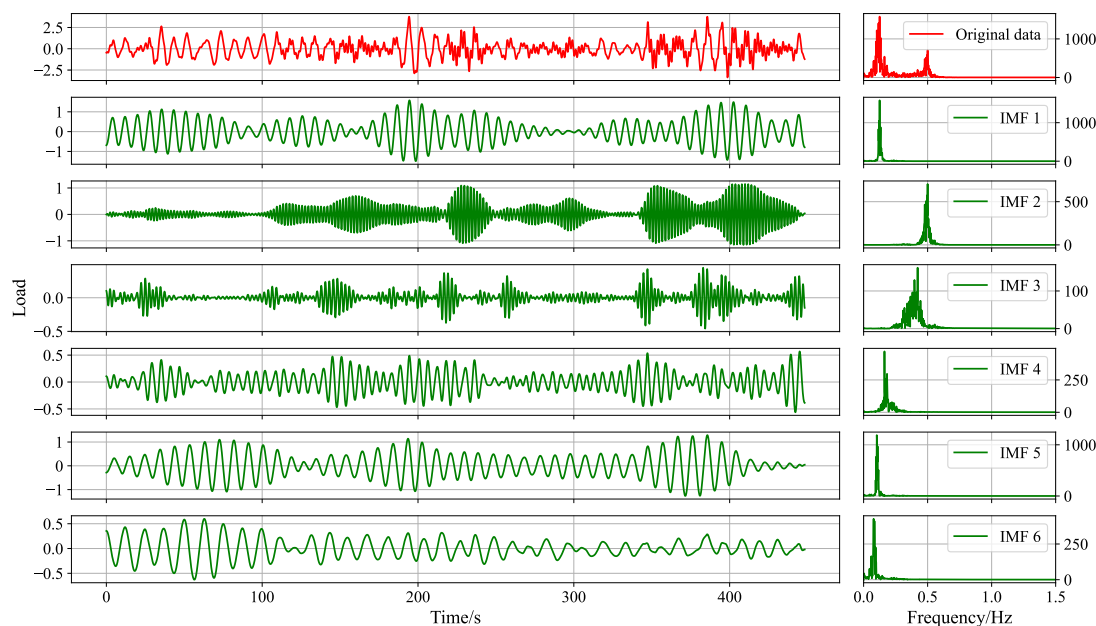


Figure 5 Results of VMD (Condition Series #1).

The empirical results indicate that the model performance is optimal when $K = 6$. At this level, the model reaches its peak accuracy ($\text{MAE} = 26.4797$, $\text{RMSE} = 34.2121$, $R^2 = 0.9514$). Compared with $K = 5$, the errors are reduced by 31.1% (MAE) and 27.4% (RMSE), whereas the coefficient of determination increases by 4.4% (ΔR^2). The superiority of $K = 6$ stems from its optimal balance between modal decomposition and data overfitting: for lower K values ($K < 6$), insufficient decoupling of the signal for the vertical bending moment at midship prevents the BiGRU model from fully learning the data trends. In contrast, a higher K ($K > 6$) causes data overfitting, increasing the MAE by 2.4% and reducing R^2 to 0.9466. Therefore, $K = 6$ is validated theoretically and

empirically as the optimal choice for maximizing the generalization capability of the VMD-BiGRU framework.

4.1.3 BiGRU hidden layers and number of neurons

The number of layers and units are important parameters that affect the prediction accuracy of a neural network. Table 5 presents the results obtained with different layers and units. The single-layer BiGRU model exhibits significant advantages in predicting the vertical bending moment at midship, where the 64-unit configuration affords the optimal performance ($\text{MAE} = 24.9396$, $R^2 = 0.9588$). Compared with the shallower model having 16 units in the same layer (MAE increased by 11.7%, R^2 decreased by 1.18%)

Table 4 Model prediction results for different VMD levels.

Decomposition level	MAE	RMSE	R^2
3	48.3998	59.6458	0.8502
4	42.0926	51.1843	0.8904
5	38.4144	47.1174	0.9077
6	26.4797	34.2121	0.9514
7	27.1045	35.8925	0.9466

Table 5 Influence of model hyperparameters on prediction accuracy.

Layers	Number of neurons	MAE	RMSE	R^2
1	16	27.8622	35.5771	0.9475
	32	25.6304	32.3251	0.9567
	64	24.9396	31.4949	0.9588
2	16	27.3279	34.6140	0.9503
	32	26.9984	33.9173	0.9523
	64	26.4797	34.2121	0.9514
3	16	28.5787	35.8003	0.9468
	32	25.7135	32.8876	0.9551
	64	27.5563	35.8183	0.9468

and the deeper model with 64 units across 3 layers (MAE increased by 9.5%, R^2 decreased by 1.27%), the proposed model exhibits significantly improved accuracy.

This phenomenon reveals two key findings: (1) The nonlinear characteristics of the wave-induced bending-moment loads at midship can be fully characterized by a single-layer structure with a high number of units, eliminating the need for complex deep stacking; (2) multi-layer BiGRU may lead to decreased generalization arising from gradient attenuation and high-frequency noise amplification. The results show that the computationally efficient, lightweight structure (single layer with 64 units) has engineering applicability and robustness for predicting the vertical bending-moment at midship.

4.2 Comparison of deterministic prediction results

To validate the effectiveness of the prediction models, the BiGRU, EWT-BiGRU, VMD-GRU, and VMD-BiGRU models were comparatively analyzed. The optimal hyperparameters for each model are determined by adjusting the respective hyperparameters, as detailed in Table 6. In conjunction with Figs. 6 and 7, Table 7 summarizes the prediction results of Condition Series #1–3 using the different models. The most accurate solutions are represented by bold text in Table 7.

The following key conclusions can be drawn based on the comparative analysis:

(1) Models integrated with modal decomposition (VMD-BiGRU, VMD-GRU, and EWT-BiGRU) achieve significantly higher accuracy than the single BiGRU model. For example, in Condition Series #3, the MAE of VMD-BiGRU is 54% lower than that of BiGRU, validating the ability of modal decomposition to capture high-frequency loads.

(2) VMD-BiGRU affords optimal performance across all conditions. In terms of error minimization and prediction accuracy, it achieves the lowest MAE and RMSE while maintaining the highest R^2 value, demonstrating its capability to capture patterns

under complex working conditions.

(3) VMD exhibits better stability in modal decomposition than EWT. Although VMD-BiGRU and EWT-BiGRU have similar accuracies in Condition Series #2 and #3, VMD-BiGRU achieves significantly better prediction results in Condition Series #1, with a relative decrease of 47.5% in the RMSE.

4.3 Analysis of probabilistic prediction results

In this study, the QRNN model is used to learn the mapping between deterministic prediction results and true values, thereby generating valid and reliable intervals. This enables probabilistic prediction of the vertical bending moment at midship and quantifies the uncertainty associated with the deterministic predictions. Taking Condition Series #1 as an example, the probabilistic prediction results using the proposed VMD-BiGRU-QRNN model at the 90% and 95% confidence levels are presented in Table 8 and Fig. 8, respectively. The results indicate that:

(1) At the 95% confidence level, the PICP of all three conditions is close to the theoretical value (0.9422). Specifically, the PICP of Condition Series #3 reaches 0.9460, indicating that the model can effectively quantify the load uncertainty under slamming conditions.

(2) The PINAW increases in moving from the 90% to the 95% confidence level (24.8%–39.39%). This increase is higher than the theoretical increase in the normal distribution (19.15%), indicating that the model effectively learns the true statistical distribution of hull girder loads.

(3) At the 90% confidence level, the CWC of Condition Series #2 (0.9167) is significantly higher than that of the other conditions, reflecting the unsatisfactory trade-off between the inherent coverage range and width. Moreover, the analysis of the CWC indicator shows that a higher confidence level comes at the cost of a wider PINAW to maintain the target PICP. This trade-off is essential for reliable interval predictions.

Table 6 Selection of main parameters for related methods.

Method	Parameter	Value
VMD	Number of decompositions	6
	Penalty factor	3000
EWT	Number of decompositions	5
BiGRU	Number of hidden layer units	64
	Number of hidden layers	2
	Batch size	64
	Learning rate	0.001
GRU	Number of hidden layer units	64
	Number of hidden layers	2
	Batch size	64
	Learning rate	0.001

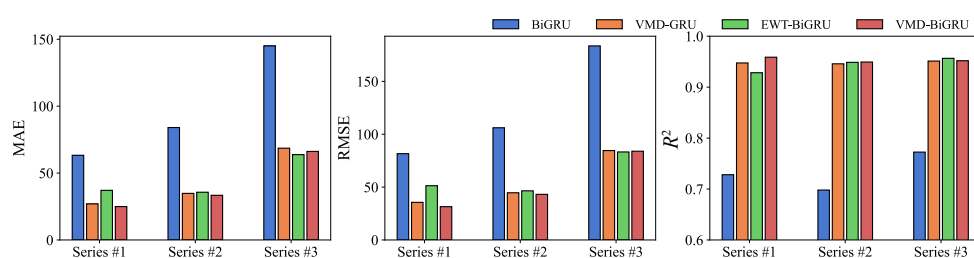
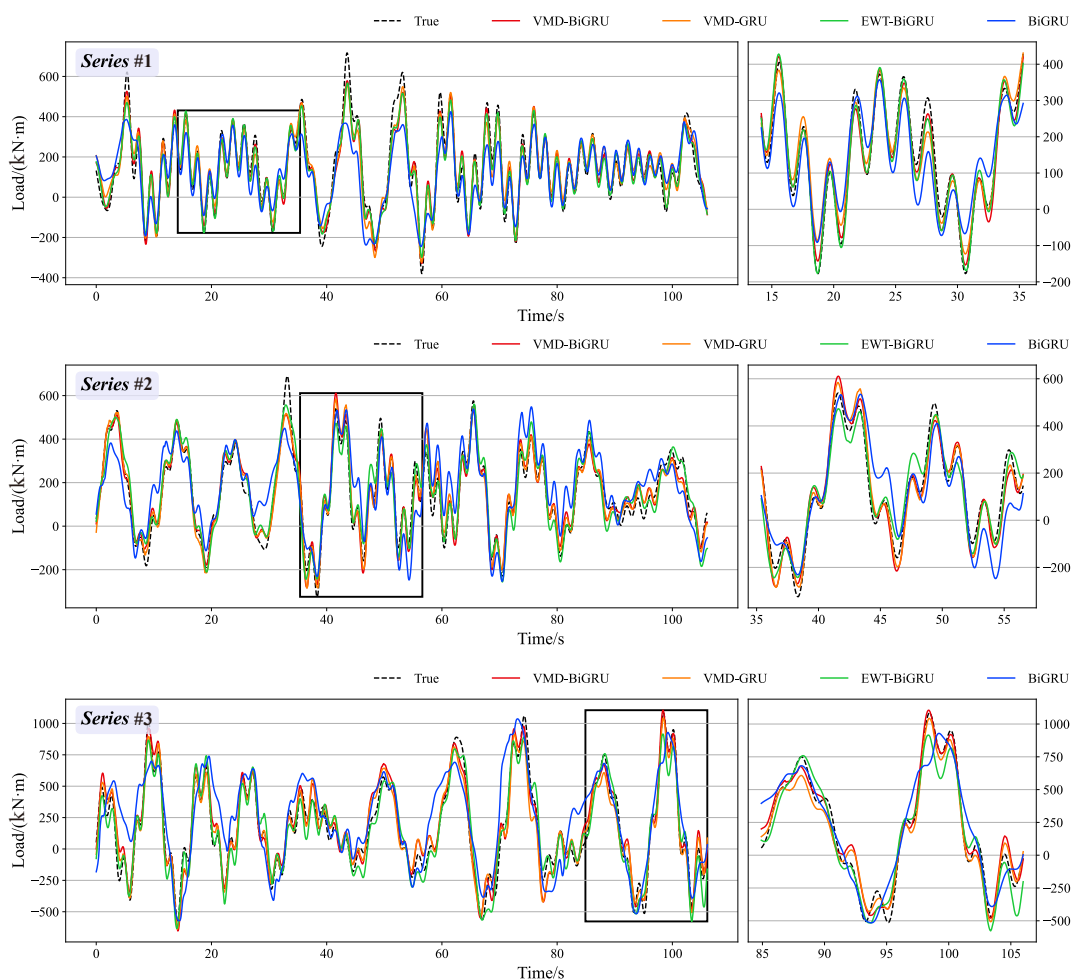
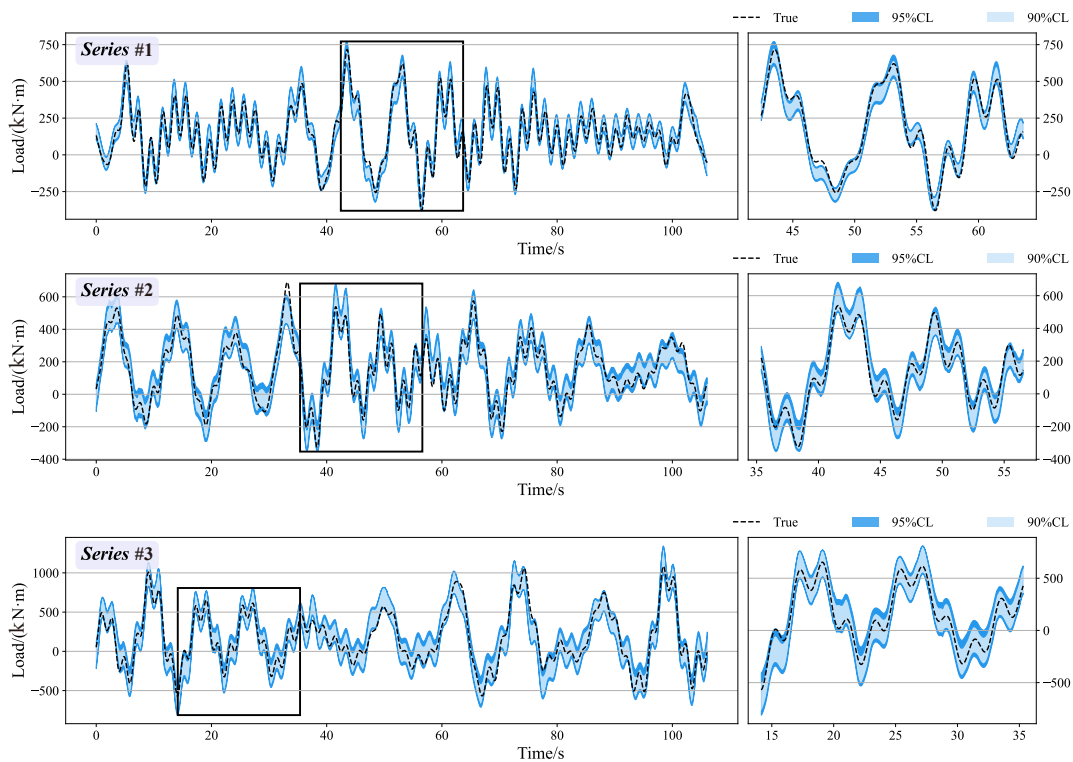
**Figure 6** Prediction indicators for Condition Series #1-3.**Figure 7** Time-domain deterministic prediction results for Condition Series #1-3.

Table 7 Model prediction indicators for Condition Series #1–3.

Condition	Model	MAE	RMSE	R ²
Series #1	VMD-BiGRU	24.9396	31.4949	0.9588
	VMD-GRU	26.9563	35.5777	0.9475
	EWT-BiGRU	37.0900	51.3198	0.9283
	BiGRU	63.2956	81.6247	0.7281
Series #2	VMD-BiGRU	33.3706	43.1571	0.9493
	VMD-GRU	34.8118	44.6307	0.9458
	EWT-BiGRU	35.6702	46.4661	0.9487
	BiGRU	84.0209	106.1332	0.6980
Series #3	VMD-BiGRU	66.1839	83.9554	0.9519
	VMD-GRU	68.5831	84.5549	0.9512
	EWT-BiGRU	67.7764	84.1457	0.9514
	BiGRU	145.0187	183.5525	0.7725

Table 8 Probabilistic prediction indicators for Condition Series #1–3.

Condition	Confidence level	PICP	PINAW	CWC
Series #1	90%	0.8813	0.0835	0.2966
	95%	0.9426	0.1042	0.2553
Series #2	90%	0.8584	0.1018	0.9167
	95%	0.9379	0.1424	0.4036
Series #3	90%	0.8863	0.1185	0.3534
	95%	0.9460	0.1486	0.3296

**Figure 8** Time-domain probabilistic prediction results for Condition Series #1–3.

4.4 Cross-validation

To validate the stability of the proposed prediction model and reduce the potential errors from random sampling, a 5-fold cross-validation is applied to the VMD-BiGRU model. For each working

condition, the data from the five ship model tests are used as the test set, whereas the remaining data serve as the training and validation set. The average of the results from all five folds is calculated to quantify the stability of the model. The detailed results are presented in Table 9.

Table 9 5-Fold cross-validation results for Condition Series #1–3.

5-Fold Cross-Validation	Series #1			Series #2			Series #3		
	MAE	RMSE	R^2	MAE	RMSE	R^2	MAE	RMSE	R^2
1	43.423	55.137	0.870	42.092	65.757	0.889	67.100	89.700	0.933
2	19.670	29.795	0.963	34.369	44.818	0.951	69.262	87.804	0.941
3	20.856	26.514	0.970	34.535	43.379	0.950	57.932	74.159	0.955
4	19.997	25.060	0.973	30.072	39.523	0.959	56.324	72.053	0.961
5	24.940	31.495	0.959	33.371	43.157	0.949	66.184	83.955	0.952
Average	25.777	33.600	0.947	34.888	47.327	0.940	63.360	81.534	0.949

The average R^2 values for the three conditions are 0.947, 0.940, and 0.949, respectively, demonstrating the good generalization capability of the model. For certain conditions, the accuracy observed in the first test is slightly lower. This may be attributed to the discrepancies between the test conditions used in the initial and subsequent ship-model trials, which affected the variation patterns present in the collected data. In the future, the model can be further optimized by increasing the number of tests.

5 Conclusion

This study proposes a rapid and intelligent method for predicting the vertical bending moment at midship using the VMD-BiGRU-QRNN hybrid model. Using the test data for a model of a large bulk carrier, the intelligent hybrid model was compared with existing models and cross-validated, with emphasis on examining the influence of signal processing and the model parameters on the prediction accuracy. The main conclusions are as follows:

(1) The integration of VMD into the hybrid model effectively improves the ability to extract the temporal features of the vertical bending moment at midship. Comparative analysis across three different working conditions indicates that the signal processing methods involving VMD afford better performance than the single model without decomposition. This validates the effectiveness of VMD in dealing with the nonstationary and nonlinear characteristics of midship bending-moment sequences. Particularly in complex scenarios involving wave-induced vibrations and high-frequency slamming, VMD-assisted models retain high accuracy (coefficient of determination $R^2 > 0.94$), highlighting the essential contribution of VMD in improving the prediction robustness.

(2) As a core prediction module, BiGRU effectively captures the forward and backward temporal dependencies of bending moment data after fusing VMD-decomposed signals. Validation by comparison with experiments revealed that BiGRU outperforms unidirectional structures in forecasting complex wave loads. It also demonstrates strong fitting capability and generalization performance across varying sea states, thereby establishing a crucial foundation for high-accuracy deterministic prediction.

(3) The QRNN-based probabilistic prediction model demonstrates reliable interval estimation at the 90% and 95% confidence levels, with PICP values closely approaching the corresponding nominal probabilities. Notably, under the working conditions involving significant wave-induced vibration (Series 1) and slamming events (Series #3), the model achieves better performance in indicators such as PINAW and CWC. This phenomenon indicates that QRNN can effectively characterize the strong, nonlinear, dynamic response characteristics caused by wave-induced vibration and slamming, thereby providing valuable

support for risk assessment in naval architecture and ocean engineering.

The proposed method was developed and validated using data from a 205,000 DWT bulk carrier under three representative irregular wave conditions (Series #1–3). Although the current dataset covers key nonlinear phenomena such as springing and slamming, the model still exhibits limitations in terms of the diversity of ship types and sea states. In principle, the proposed framework is also applicable to other ship types with similar hydroelastic characteristics and is not limited to bulk carriers. However, if applied to vessels with significantly different hydrodynamic properties, such as container ships, the model hyperparameters should be adjusted accordingly and the model should be retrained using a new dataset specific to the target ship type. In future work, we will collect data from more ship types and a wider range of sea conditions to systematically validate the generalization ability of the model.

Acknowledgments

None.

Author contribution statement

Jun Ding proposed the preliminary concept of the forecasting scheme and provided the experimental data, overseeing the overall progress of the paper and its critical revisions. Peiqiao Zhu constructed the forecasting model and drafted the manuscript. Zhang Zhu and Yiming Qiang were responsible for the design of the experimental scheme and reviewed and revised the manuscript. All the authors have approved the final manuscript.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declaration of competing interest

All the contributing authors report no conflicts of interest in this work.

Funding

None.

Use of AI statement

None.

References

- [1] Temarel, P., Bai, W., Bruns, A., Derbanne, Q., Dessi, D., Dhavalikar, S., Fonseca, N., Fukasawa, T., Gu, X., Nestegård, A., et al. (2016). Prediction of wave-induced loads on ships: Progress and challenges. *Ocean Eng.* 119, 274–308.
- [2] Sun, H., Jiao, J. L., Xu, W. H., Ma, B. W., Chang, X. (2026). Intelligent prediction of ship wave loads and whipping responses from input motions by decomposition-superposition method. *Ocean Eng.* 343, 123267.
- [3] Hirdaris, S., Parunov, J., Qui, W., Iijima, K., Wang, X. L., Wang, S., Brizzolara, S., Soares, C. G. (2023). Review of the uncertainties associated to hull girder hydroelastic response and wave load predictions. *Mar. Struct.* 89, 103383.
- [4] Zhang, B., Wang, S., Deng, L. W., Jia, M. Q., Xu, J. Z. (2023). Ship motion attitude prediction model based on IWOA-TCN-attention. *Ocean Eng.* 272, 113911.
- [5] Hou, L. X., Wang, Q. Y., Liu, D. Y. (2025). Real time prediction of ship motion based on iTransformer with EWT-SampEn data processing. *Ocean Eng.* 342, 122987.
- [6] Zhou, X., Zou, L., He, H. W., Wu, Z. X., Zou, Z. J. (2024). Real-time prediction of full-scale ship maneuvering motions at sea under random rudder actions based on BiLSTM-SAT hybrid method. *Ocean Eng.* 314, 119664.
- [7] Liu, D. Y., Gao, X. P., Huo, C., Su, W. T. (2025). Research on maneuvering motion prediction for intelligent ships based on LSTM-multi-head attention model. *JMSE* 13, 503.
- [8] Wang, W. H., Yi, Z., Zhao, L. C., Jia, P., Kuang, H. B. (2025). Application of switching-input LSTM network for vessel trajectory prediction. *Appl. Intell.* 55, 289.
- [9] Deng, Z. R., Cui, J. H., Zhu, Z. X., Dan, Y. W., Xie, H. W., Hu, Y. J. (2025). Adaptive trajectory tracking control of underactuated ships using parameter prediction based neural network. *Ocean Eng.* 329, 121188.
- [10] Yang, Z. S., Dong, Y. D., Deng, X. G., Zhang, L. P. (2022). AMGNET: Multi-scale graph neural networks for flow field prediction. *Connect. Sci.* 34, 2500–2519.
- [11] Xu, Q., Zhuang, Z. J., Pan, Y. C., Wen, B. H. (2023). Super-resolution reconstruction of turbulent flows with a transformer-based deep learning framework. *Phys. Fluids* 35, 055130.
- [12] Xu, J. S., Haddara, M. R. (2001). Estimation of wave-induced ship hull bending moment from ship motion measurements. *Mar. Struct.* 14, 593–610.
- [13] Moreira, L., Soares, C. G. (2020). Neural network model for estimation of hull bending moment and shear force of ships in waves. *Ocean Eng.* 206, 107347.
- [14] Zhou, C. Y., Chen, L., Liu, W. T., Yin, Q., Zhang, J. N. (2025). Research on the inversion method of ice loads on ships in broken ice channels using deep neural networks. *Results Eng.* 27, 107017.
- [15] Wang, Q., Yu, P. Y., Lv, M. D., Wu, X. C., Li, C. F., Chang, X., Wu, L. H. (2024). Real-time prediction of wave-induced hull girder loads for a large container ship based on the recurrent neural network model and error correction strategy. *Int. J. Naval Archit. Ocean Eng.* 16, 100587.
- [16] Liu, R. X., Li, H., Zou, J., Ong, M. C. (2023). Reconstruction and prediction of global whipping responses on a large cruise ship based on LSTM neural networks. *Ocean Eng.* 285, 115393.
- [17] Liu, R. X., Li, H., Ong, M. C., Zou, J., Lu, L. (2025). Prediction of global structural responses for an ultra-large container ship under long-crested irregular waves using generalized modes and artificial neural networks. *Ocean Eng.* 318, 120152.
- [18] Duan, W. Y., Huang, L. M., Han, Y., Zhang, Y. H., Huang, S. (2015). A hybrid AR-EMD-SVR model for the short-term prediction of nonlinear and non-stationary ship motion. *J. Zhejiang Univ. -Sci. A* 16, 562–576.
- [19] Wei, Y. Y., Chen, Z. Z., Zhao, C., Chen, X., Yang, R., He, J. H., Zhang, C. Y., Wu, S. T. (2022). Deterministic and probabilistic ship pitch prediction using a multi-predictor integration model based on hybrid data preprocessing, reinforcement learning and improved QRNN. *Adv. Eng. Inf.* 54, 101806.
- [20] Dragomiretskiy, K., Zosso, D. (2014). Variational mode decomposition. *IEEE Trans. Signal Process.* 62, 531–544.
- [21] Chung, J., Gulcehre, C., Cho, K., Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint*, arXiv: 1412.3555.
- [22] Taylor, J. W. (2000). A quantile regression neural network approach to estimating the conditional density of multiperiod returns. *J. Forecast.* 19, 299–311.
- [23] Richman, J. S., Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* 278, H2039–H2049.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits reusers to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the original author(s) and the source, a link to the license is provided, and any changes made are indicated. See <https://creativecommons.org/licenses/by/4.0/>

© The author(s) 2026. Published by Tsinghua University Press.