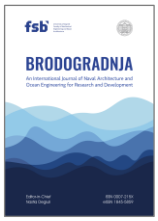
Brodogradnja Volume 77 Number 3 (2026) 77311

University of Zagreb
Faculty of Mechanical
Engineering and Naval
Architecture

journal homepage: www.brodogradnja.fsb.hr

Brodogradnja

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



Study on the prediction performance of ship motion in waves by LSTM under missing data

Jiaye Gong, Pengsheng Ni, Zheng Fu, Yongzhi Qin*

College of Ocean Science and Engineering, Shanghai Maritime University, Shanghai, China

ARTICLE INFO

Keywords:
Ship motion prediction
LSTM
Missing data
Data imputation

ABSTRACT

Ship motion prediction is essential in marine engineering, but missing data caused by sensor faults or signal interruptions often degrades the accuracy of long short-term memory (LSTM) models. This study investigates how different missing data rates and imputation methods affect LSTM prediction performance. A ship-motion dataset under various speeds and wave conditions was used to examine model feasibility and hyperparameter sensitivity. Traditional filling strategies, including zero and mean filling, were compared under missing data scenarios. Results show that data loss significantly reduces prediction accuracy. The mean-filling method generally performs better than zero-filling, though its effectiveness decreases with higher data diversity. Proper data clustering can effectively enhance its performance.

1. Introduction

With the rapid advancement of artificial intelligence (AI), data-driven modeling has become a widely adopted approach in engineering and scientific domains for tasks such as classification, prediction, and regression, yielding promising results. The awarding of the 2024 Nobel Prize to breakthroughs in AI underscores the significance of this field. As society enters the era of big data, the availability of high-quality and reliable data has become critical for effective data analysis and model training. However, in actual situations, due to equipment aging, sensor failure, external environment or human factors, etc., the acquired data will be affected by data missing, which is also a common and classic problem in neural networks, and will have a certain impact on prediction accuracy and system security. Deep neural networks may make high-confidence erroneous predictions for meaningless input data, revealing the high dependence of models on the integrity and continuity of input data, as well as potential generalization problems [1].

As early as 2006, artificial neural networks (ANN) were used in the field of medical treatment for computer-aided diagnosis, that is, to diagnose diseases through patient image data. However, since medical sensors and imaging data are often missing, medical neural network models are particularly sensitive to missing values in diagnosis or survival prediction, which has a certain impact on the accuracy of prediction [2]. Similarly, in the domain of meteorology, the work by Hua et al. [3] has increasingly focused on accurate air quality prediction due to rising environmental concerns. They emphasize that such predictions

* Corresponding author.
E-mail address: fuzheng@shmtu.edu.cn

https://doi.org/10.21278/brod77311
Received 15 September 2025; Received in revised form 20 November 2025; Accepted 26 November 2025
Available online 26 January 2026
ISSN 0007-215X; eISSN 1845-5859

are crucial for planning preventive measures and formulating policies to address potential health hazards and environmental issues caused by poor air quality. Since air quality data are typically time series in nature, researchers also note that missing values frequently occur during data preparation and aggregation stages for various reasons. Junninen et al. [4] argue that the inability to properly analyze and handle these missing values can significantly hinder the overall data analysis process. Moreover, traffic prediction has become a critical component of intelligent transportation systems and smart cities. Both travelers and urban planners depend on accurate traffic information for route optimization and traffic management. However, traffic data is also prone to missing values caused by both human-related and natural factors. Consequently, addressing the impact of missing data on traffic flow prediction has emerged as a key research topic, gaining increasing attention in both academic and practical contexts [5].

Among various engineering applications affected by missing data, ship motion prediction stands out as a critical task in the field of naval architecture and ocean engineering. Accurate prediction of ship motion plays a vital role in ensuring navigation safety, structural integrity, energy efficiency, and real-time decision-making, especially under complex sea conditions [6]. Recent studies published in journal Brodogradnja have advanced data-driven and hybrid approaches for ship motion prediction. Atasayan et al. [7] developed a grey-box identification approach combining physics-based knowledge with machine learning, achieving improved maneuvering performance but relying heavily on complete datasets. With the increasing development of intelligent marine systems and unmanned surface vessels (USVs), ship motion prediction has become one of the cornerstones of autonomous navigation and motion control. In practice, ship motion involves six degrees of freedom (6DOF), including surge, sway, heave, roll, pitch, and yaw, which are influenced by various environmental and operational factors such as wave parameters, hull form, and control inputs [8]. Due to this complex and highly nonlinear dynamic behavior, traditional physics-based prediction models often suffer from limitations in generalization and adaptability under diverse conditions. In addition, CFD-based studies of added resistance in short regular waves highlight that post-processing choices materially affect the reliability of extracted responses, underscoring the sensitivity of downstream predictive models to input quality [9].

Therefore, this study has practical significance in ship data processing. Due to the 6DOF in the ship motion and the complexity of the external environment, there are more parameters for predicting ship motion, and the prediction model is more complex [10]. Long short-term memory (LSTM) networks have been widely used in ship motion prediction due to their good performance in time series data [11]. The model can learn complex dynamic patterns directly from data without explicit physical assumptions. For instance, D'Agostino et al. [12] employed recurrent neural networks (RNN, LSTM, GRU) for short-term forecasting of ship motions in irregular waves, proving the feasibility of machine-learning-based nowcasting. Diez et al. [13] combined dynamic mode decomposition (DMD) with neural networks for equation-free prediction of ship motions, bridging data-driven modeling and system identification**,** yet their framework also assumed complete time histories during training. Moreover, explainable machine-learning work on operational ship data has shown that model accuracy and interpretability hinge on input reliability, reinforcing the importance of data completeness for maritime datasets [14].

Given the prevalence of missing data across various domains, substantial research has been conducted to investigate its impact on recurrent models, particularly LSTM networks. As LSTM models are widely used for time series prediction due to their ability to capture temporal dependencies, their performance can degrade significantly in the presence of missing values. Missing values were generated under missing completely at random (MCAR) mechanism at 10 %, 15 %, 25 %, and 35 % rates of missingness using complete data of 24-h ambulatory diastolic blood pressure readings. The performance of the mean, Kalman filtering, linear, spline, and Stineman interpolations, exponentially weighted moving average (EWMA), simple moving average (SMA), *k*-nearest neighborhood (KNN), and last-observation-carried forward (LOCF) imputation techniques on the time series structure and the prediction performance of the LSTM and autoregressive integrated moving average (ARIMA) models were compared on imputed and original data [15].

Recently the critical role of data completeness and continuity in the accuracy and reliability of LSTM models is highlighted for predicting algal blooms. Zhang et al. [16] demonstrates that missing chlorophyll

concentration (Chl) has a greater negative impact on the predictive performance of LSTM models compared to missing temperature, dissolved oxygen and pH data. Chl, being a primary indicator of algal biomass, directly influences the accuracy of LSTM forecasts, and its absence resulted in substantial increases in root mean square error and lower correlations. And in the ship design field, generative adversarial network (GAN)-based generative models have been used to reconstruct incomplete design parameters when the data is incomplete, illustrating a trend toward deep generative imputation [17].

To address the issue of missing data in time series, researchers have explored a range of imputation techniques. Traditional approaches include listwise deletion and simple statistical methods, such as replacing missing values with the mean, median, or mode of the observed data [18]. However, these methods may lead to information missing or bias, especially in time-dependent datasets. More advanced techniques such as regression imputation estimate missing values based on correlations with other variables. Another commonly used approach in time series is LOCF, which assumes temporal consistency by filling each missing value with the last available observation. Although LOCF is easy to implement and maintains the temporal structure, it may not reflect true dynamics when values fluctuate rapidly over time [19]. Machine learning-based techniques like k-nearest neighbors imputation (KNNI) and miss forest, based on random forests, improve imputation by capturing nonlinear patterns in the data [20]. Recent years, deep learning approaches have been increasingly adopted. GANs and generative adversarial imputation networks (GAIN) leverage adversarial training to reconstruct missing components [21]. Kalman filtering provides recursive estimation under known system dynamics [22]. Cutting-edge methods such as diffusion models [23] and transformer-based hybrid models [24] further enhance imputation performance by capturing long-range dependencies and adapting to real-time changes in missing data scenarios. In the market-forecasting domain, Ribeiro et al. [25] provided a review-type study that systematically compared multiple imputation strategies and analyzed their influence on predictive performance. As a survey, their work highlighted that different interpolation or reconstruction methods can yield substantially different forecasting outcomes, underscoring that missing data handling is itself a decisive factor in time-series modeling.

It is evident that missing data can adversely affect the predictive performance of LSTM models to varying degrees. In the maritime domain, this issue becomes even more complex due to the high dimensionality of ship motion data, the diversity of missing data types, and the variability in feature-missing combinations [26]. These factors raise important questions regarding how such missing patterns influence the accuracy of LSTM-based ship motion prediction under different sea states, wave lengths, and vessel configurations, and what strategies can be employed to mitigate their impact. Understanding and addressing these effects is critical for improving the robustness and reliability of ship motion forecasting systems in real-world scenarios.

Although previous studies have successfully applied deep learning models such as LSTM and GRU for ship motion prediction, most of them assumed complete time-series data without considering sensor failures or data gaps. However, few studies have systematically evaluated how different imputation strategies influence model performance. Therefore, this work focuses on examining the impact of missing data on LSTM-based prediction and assessing the effectiveness of common imputation methods under various missing rates. Combined with evidence from other fields—such as medicine, ecology, traffic forecasting, and financial markets, this indicates that the impact of missing values extends across application domains—yet similar review-level analyses remain limited in ship-motion prediction, where sensor dropout is common.

Accordingly, this study is devoted to a comprehensive investigation of the influence of feature-level missing data on the predictive accuracy of LSTM-based ship motion models. By designing controlled experiments with continuous missing rates and patterns under typical operating conditions, this paper evaluated the different effects of missing specific motion-related parameters on model performance. The model is then trained with motion data under complete database conditions, and feature importance analysis is performed to identify the most influential variables, whose missingness leads to a significant drop in prediction accuracy. In addition, this study conducts a comparative evaluation of several interpolation strategies, ranging from classical statistical techniques to machine learning algorithms and deep generative models. The goal is to determine the most effective way to mitigate the adverse effects of data incompleteness,

thereby improving the robustness and reliability of LSTM-based prediction systems in practical maritime applications. The workflow of this article is shown in Figure 1.

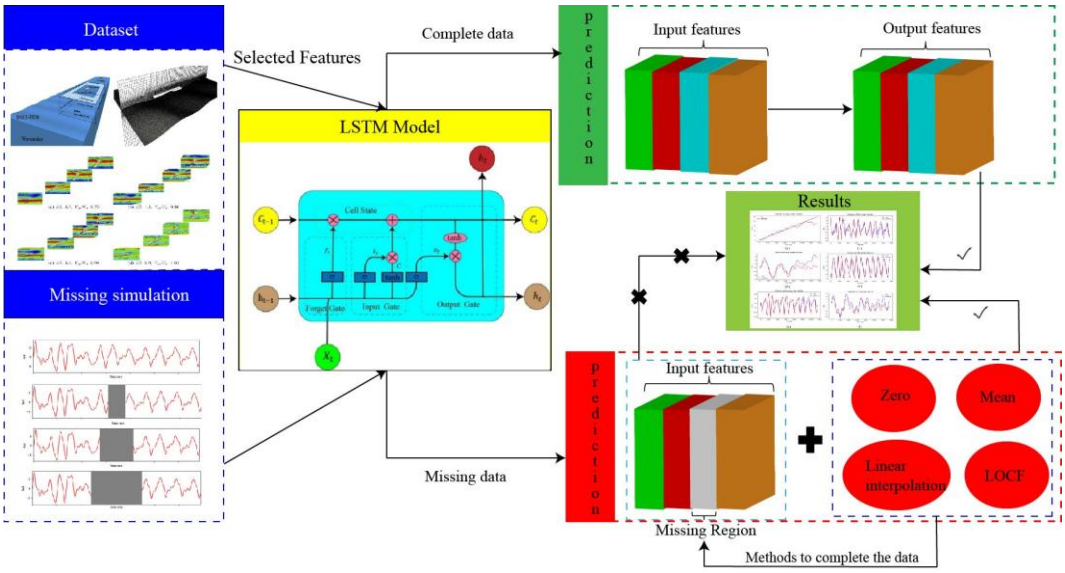


Fig. 1 Framework of LSTM-based ship motion prediction under normal and missing data conditions

The rest of this paper is organized as follows. Section 2 introduces the theoretical background of the LSTM model and describes the approach used to simulate missing data, including the zero- and mean-filling strategies. Section 3 presents the ship-motion dataset generated by Open field operation and manipulation (OpenFOAM) simulations and explains the relevant parameters under different operating conditions. Section 4 reports the prediction results obtained by the LSTM model, comparing the performance between complete and missing data cases, followed by a discussion of the effects of different filling methods. Finally, Section 5 concludes the paper and summarizes the main findings and implications of this study.

2. Methodology

This section mainly introduces the overall methodology, including the origin and basic principles of the LSTM-based prediction framework, the data imputation methods for handling missing values, and the evaluation metrics used to assess prediction performance such as mean absolute error (*MAE*), mean squared error (*MSE*), root mean squared error (*RMSE*), and coefficient of determination (R^2).

2.1 LSTM neural network

RNN is a type of artificial neural network for handling sequential data like time series and text. In 1991, Hochreiter discovered the long-term dependencies problem of recurrent neural networks, which means that when learning long sequences, recurrent neural networks will experience gradient vanishing and gradient explosion and cannot grasp nonlinear relationships over long time spans. To solve the long-term dependency problem, RNN improvements continue to emerge, the most important of which is LSTM networks. LSTM is a specially designed RNN for modeling long-term dependencies in time series data and has set accuracy records in multiple application areas [27]. Due to its unique design structure, LSTM is suitable for processing and predicting important events in time series with very long intervals and delays.

LSTM introduces a memory cell and a gated mechanism to regulate the information flow. The core of LSTM lies in three gates: the forget gate, the input gate, and the output gate, each of which learns to preserve or discard information based on the context of the current input and the past hidden state. The forget gate filters redundant past signals, the input gate updates memory selectively, and the output gate determines the relevant part to be propagated. This architecture enables LSTM to preserve long-term dependencies without suffering from gradient vanishing. The structure of LSTM networks is shown in Figure 2.

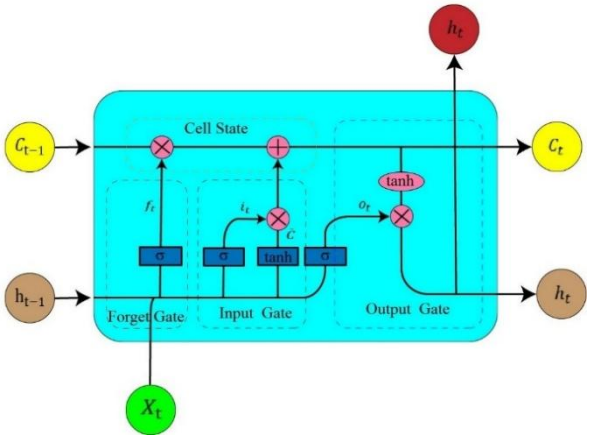


Fig. 2 Structure of a standard LSTM cell

Specifically, at each time step t, the LSTM performs the following operations:

$$\begin{aligned} f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ \tilde{c}_t &= \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \\ c_t &= f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \\ o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ h_t &= o_t \odot \tanh(C_t) \end{aligned} \tag{1}$$

where i_t is the input gate, f_t is the forget gate, o_t is the output gate, $\tilde{c}_t$ is candidate hidden state, $\tilde{c}_t$ is the cell state of LSTM, x_t is the input vector at time t, h_{t-1} is the hidden state from the previous step, σ denotes the sigmoid activation, and $\odot$ represents the element-wise multiplication. Through these computations, LSTM learns to selectively "forget" outdated information, "update" the memory cell with current relevant inputs, and "output" refined temporal features for downstream prediction tasks.

LSTM networks rely on the continuous flow of sequential information to maintain accurate memory states. However, in real-world scenarios, missing values often occur due to sensor failures or transmission delays. Che et al. [28] demonstrated that recurrent models like LSTM suffer performance degradation when missing values are present, especially when the missingness is not random.

2.2 Baseline models and prediction framework

To evaluate the robustness of ship-motion prediction under missing data conditions, several representative models were adopted for comparison, including both linear and nonlinear approaches. These baseline methods were selected to represent different levels of model complexity and physical interpretability.

Linear regression model: linear regression assumes a linear relationship between the input features and the target variable, expressed as:

$$y = wX + b \tag{2}$$

where w and b denote the regression coefficients and intercept. As a simple yet effective statistical model, it provides a fundamental benchmark for assessing whether the target variable exhibits linear dependence on past observations. However, due to its linear nature, the model cannot capture the nonlinear and coupled relationships often observed in ship motion dynamics.

Persistence model: The persistence model assumes that the future value of a variable is equal to its most recent observation. This approach requires no training and reflects a short-term temporal continuity assumption, which can yield acceptable performance in slowly varying systems. Nonetheless, for oscillatory ship motion signals influenced by external disturbances, persistence cannot account for phase shifts or nonlinear trends, making it a limited but interpretable baseline.

Dynamic mode decomposition: DMD is a data-driven technique originating from fluid dynamics and system identification. It decomposes the evolution of high-dimensional systems into a set of spatiotemporal modes governed by a linear operator A, satisfying $x_{t+1} = Ax_t$. In this study, the Hankel-DMD variant was employed to identify dominant temporal modes from multivariate ship-motion data. DMD provides a transparent linear-dynamics predictor that captures dominant oscillatory behavior.

2.3 Evaluation Metrics

To quantitatively evaluate the performance of the prediction and imputation models, four regression metrics are used for evaluation, including *MAE*, *MSE*, *RMSE*, and R^2. These metrics capture different aspects of prediction quality. The *MAE* is:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\left(y_i - \hat{y}_i\right)\right| \tag{3}$$

where, y_i represent the true value, $\hat{y}_i$ denote the predicted value, and n be the number of samples. In statistics, *MAE* is a measure of the error between pairs of observations expressing the same phenomenon. MAE measures the average magnitude of the absolute errors between predictions and ground truth values. It provides a direct interpretation of the average prediction error in the same unit as the target variable and is robust to outliers compared to MSE.

The *MSE* is:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{4}$$

where the *MSE* measures the average of the squared differences between predicted and true values. By squaring the errors, *MSE* penalizes larger errors more heavily, making it sensitive to outliers. It is widely used for model training due to its smooth differentiable nature. *MSE* is sensitive to outliers (because when the difference between outliers and normal values is large, the error will be greater than 1 and squaring the value will further increase the value), but they can reflect the distribution of prediction errors.

The *RMSE* is:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{5}$$

where the *RMSE* is the square root of *MSE* and shares its properties while keeping the error in the same unit as the target. *RMSE* is more intuitive than *MSE* and is often used in performance reporting, especially when large errors need to be emphasized.

The R^2 is:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{6}$$

where $\bar{y}$ is the mean value of the observed data. The R^2 score indicates the proportion of variance in the target variable explained by the model. An R^2 value of 1 means perfect prediction, while a value of 0 indicates that the model performs no better than simply predicting the mean of the target values.

In general, these four evaluation indicators provide a quantitative assessment of the prediction ability. According to their different characteristics and advantages, they capture the prediction error scale of the model and the overall model expression ability and can comprehensively evaluate the model prediction ability. In

this experiment, all four indicators were recorded to evaluate the accuracy and stability of the model under different working conditions and missing data backgrounds.

2.4 Missing data filling

In comparative studies of missing value estimation, Tuikkala et al. [29] employed two simple reference approaches: one where all missing entries were substituted with zeros (zero imputation), and another where they were replaced by the means of the available values within the same row (row average). In this study, three typical static imputation methods were adopted to handle artificially constructed missing data: zero-filling and mean-filling, which has been shown in Figure 3 and LOCF. The first one is the traditional zero-filling method, which refers to the method of filling the missing data of a static feature with the zero. This method is simple to operate and has low computational cost and is often used as a benchmark strategy for comparative analysis. However, in the ship motion characteristics, the value of zero often has no actual physical meaning and may deviate greatly from the actual state, resulting in a tendency for the model to learn errors and produce obvious errors. The mean-filling is to use the historical average value of the feature in the test data before it is missing to replace all the values of the missing data. This method can maintain the statistical characteristics of the data to a certain extent and reduce the interference of mutations on the model. However, this method is also a static completion method that cannot reflect the law of time series changes and has the risk of information loss within the missing segment. The LOCF method is a commonly used imputation technique in time-series data analysis, particularly for handling short gaps or sporadic missing values. The main idea of LOCF is to replace each missing value with the most recent observed (non-missing) value. This approach assumes that the system state remains approximately constant between two successive observations, which can be reasonable for variables with relatively slow temporal variation.

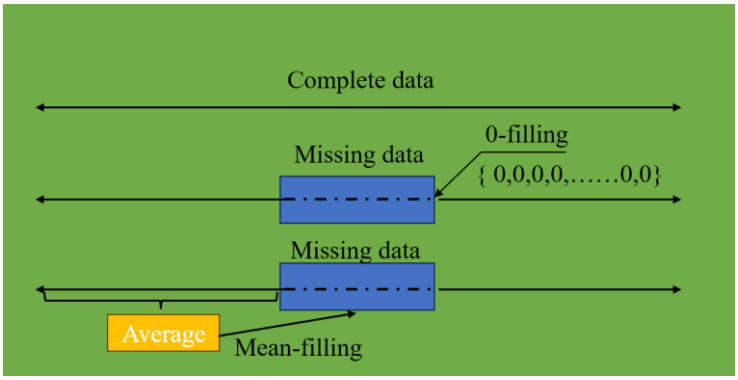


Fig. 3 Zero-filling and mean-filling method

All methods are applied in the test under the premise of keeping the model superstructure unchanged, aiming to study the impact of missing data location and completion method on prediction performance, and provide a benchmark and comparison for subsequent dynamic completion methods.

3. Dataset and experimental setup

This section presents a comprehensive description of the dataset source and its reliability, followed by an overview of the data conditions, feature composition, and patterns of missing data. Subsequently, the types of data missingness are categorized, and the model parameter configuration used in this study is detailed.

3.1 Dataset source

The dataset used in this study is generated from a series of high-fidelity numerical simulations based on a trimaran model. The simulations are conducted using OpenFOAM and follow the Reynolds averaged Navier–Stokes (RANS) equations with appropriate turbulence modeling [30, 31]. This computational setup has been previously validated for trimaran motions in waves and maneuvering scenarios [32, 33].

The simulation considers both viscous effects and nonlinear wave-body interactions and provides realistic 6DOF motion responses of the trimaran under various combinations of wave steepness and Froude numbers.

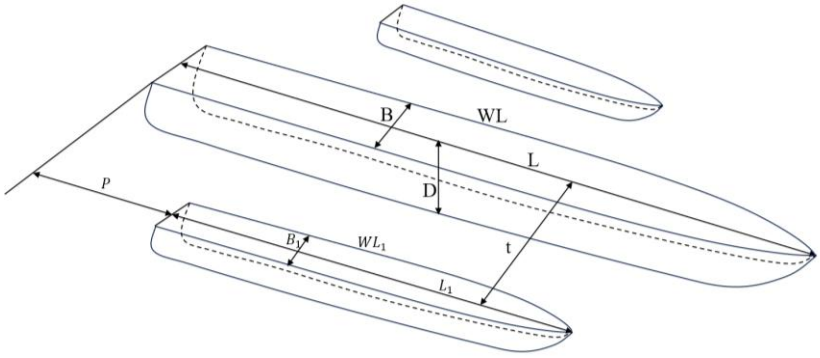


Fig. 4 Schematic diagram of trimaran hull geometry and characteristic dimensions

The model geometry and dimensions used for simulation are illustrated in Figure 4. The resulting time-series motion data under oblique stern waves are used to construct the training and evaluation dataset for this study.

3.2 Working conditions and data feature composition

To ensure a comprehensive evaluation of the model, the working conditions are selected to make the ship move in stern quartering seas with a speed close to the wave celerity, which is relatively dangerous for the ship safety and makes the ship vulnerable to severe motion [34]. Under these scenarios, accurate motion prediction becomes particularly critical for operational safety. The full configuration of working conditions used in this study is shown in Table 1. In the table, V_s stands for ship speed, V_w is the wave propagation speed, and F_r refers to the Froude number, defined by $F_r = V_s/\sqrt{gL}$, where g is gravitational acceleration and L is the waterline length of ship. It reflects the ratio between inertial and gravitational forces. β is the initial wave heading, λ is the wavelength, ak is a dimensionless wave steepness parameter, where a is wave amplitude and $k = 2\pi / \lambda$ is wave number. It reflects the nonlinearity of wave shape.

In this study, the selected datasets are chosen to sufficiently cover a wide range of realistic operating conditions where ship motion instability is likely to occur, such as surf-riding and broaching scenarios. Cases 1-5 vary the forward velocity of the ship under fixed wave conditions, while Cases a-e fix the ship speed and heading but vary wave characteristics. This combination ensures that both speed-sensitive and wave-sensitive dynamic responses are adequately captured. The dataset design is not aimed at analyzing the influence of specific parameters, but rather to ensure robustness and generalization of the model under diverse yet representative sea conditions.

Table 1 The data cases selected in this paper

Case	F_r	V_s/V_w	λ/H	$\beta(°)$
1	0.35	0.84	3.27	30
2	0.38	0.91		
3	0.40	0.96		
4	0.45	1.08		
5	0.50	1.20		
6	0.35	0.84	2.34	15
7	0.38	0.91		
8	0.40	0.96		
9	0.45	1.08		
10	0.50	1.20		
11	0.45	1.19	2.70	30
12	0.45	1.08	3.27	
13	0.45	0.99	3.90	
14	0.45	0.92	4.50	
15	0.45	0.87	5.10	

3.3 Data features and missing data type

For each sea condition, the collected data contains a full set of motion-related time series features, capturing both rigid-body dynamics and hydrodynamic loads. The 6DOF motion features of ship can be systematically categorized based on translational and rotational components. Specifically, translational motion along the three coordinate axes is grouped together, while rotational motion is further divided into three separate groups, each corresponding to rotation around one of the principal axes (roll, pitch, and yaw). This classification yields a structured set of motion feature combinations, as summarized in the accompanying table. The 6DOF and motion coordinate system of the ship are shown in Figure 5.

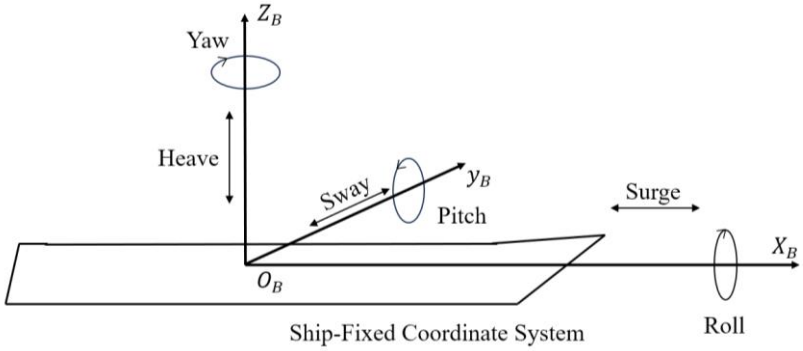


Fig. 5 6DOF of ship motion

Totally 18 features of each case are used for training and prediction in this paper, including heave, sway, surge, roll, pitch, yaw, and the corresponding velocity, angular velocity, force, and moment. Among these, the roll feature was deliberately selected to be missing in the experimental setup. Roll motion is a critical feature reflecting the ship's lateral stability, and its accurate prediction is essential for ensuring navigational safety. In this paper, the proportion of missing data is expressed using r_{miss}, which represents the fraction of missing data points relative to the total number for the roll feature. This indicator allows a direct comparison of missing data levels across different experimental configurations.

In this study, 200 consecutive time steps of selected feature data are employed as input to forecast the same feature data at the next, 201st, time step. The entire ship motion sequence lasts for 55 seconds with a sampling interval of 0.004 s, resulting in a total of 13,750 time steps. Among them, the first 80 % (11,000 time steps) are used for model training, while the remaining 20 % serve as the validation dataset. For consistency, all figures are presented in time steps, which are linearly related to physical time through this conversion. In Table 2, 18 physical features are selected as input variables to comprehensively characterize the trimaran's motion and hydrodynamic responses. These features include linear and angular displacements, linear and angular velocities (U_x, U_y, U_z, Ω_x, Ω_y, Ω_z), and the corresponding hydrodynamic forces and moments (F_x, F_y, F_z, M_x, M_y, M_z). Together, these features provide a complete 6DOF representation of the vessel's dynamic behavior, serving as comprehensive input for data-driven modeling and prediction.

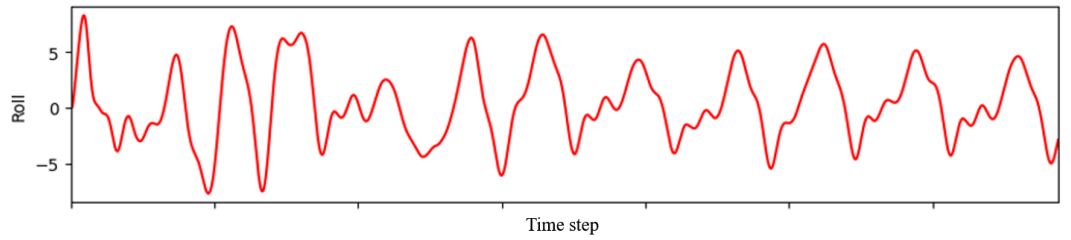
Table 2 Definition of input features for 6DOF motion

Displacement	Displacement	Linear V	Angular V	Force	Moment
Surge	Roll	U_x	Ω_x	F_x	M_x
Sway	Pitch	U_y	Ω_y	F_y	M_y
Heave	Yaw	U_z	Ω_z	F_z	M_z

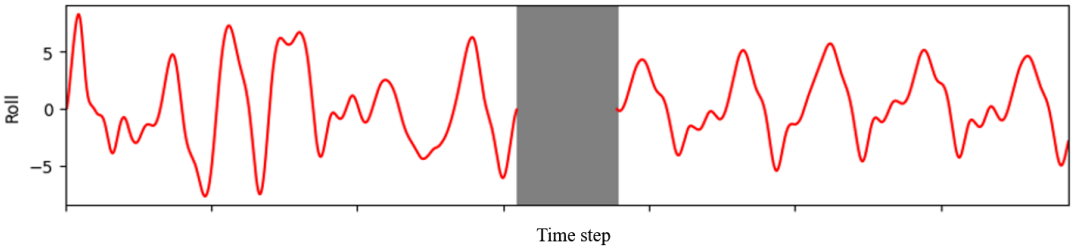
To simulate the data missing caused by sensor failure or other factors in the real world, this paper selects a data missing mode for subsequent analysis. The missing type is continuous missing, imitating large segments of uninterrupted missing. The missing rate uses three ratios: 10 %, 20 %, and 30 %. The continuous missing involves removing entire continuous segments of the time series, representing more severe sensor failures or data outages. In this study, missing blocks were inserted at random start points, with the length of each block adjusted to match the desired missing rate. An example of the missing pattern applied to the feature sequence is shown in Figure 6. As illustrated, continuous missing creates large contiguous voids, potentially disrupting temporal dependencies in the data and posing great challenges for imputation.

To illustrate the design of the missing data scenarios in this study, the roll feature from Case 1 is selected as a representative example for visualization, which is significant for ship sailing in quartering seas and relatively dangerous for safety. The same missing data strategies are consistently applied to all other motion features and are therefore not individually presented here. From top to bottom are the complete data and the data with the missing rate gradually increasing by 10 %. The gray shaded areas represent the missing data segments. As the missing rate increases, the proportion of the gray shaded area in the graph also becomes larger, resulting in a greater gap in data continuity.

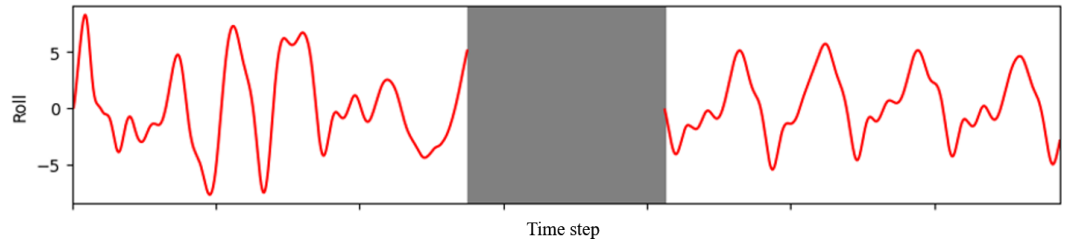
By evaluating model performance under this missing type and varying levels, this study investigates not only the sensitivity of different feature groups to missing data, but also the robustness of imputation methods under realistic maritime data conditions. To clearly differ the prediction under different scenarios, this paper uses the following symbols to represent: (True Value: ground truth values without missing data. Pc: model prediction using the complete input sequence. Pz: prediction using the input sequence with missing blocks filled by zeros. Pm: prediction using the input sequence with missing blocks filled by mean values. Pl: prediction using the input sequence with missing blocks filled by LOCF.)



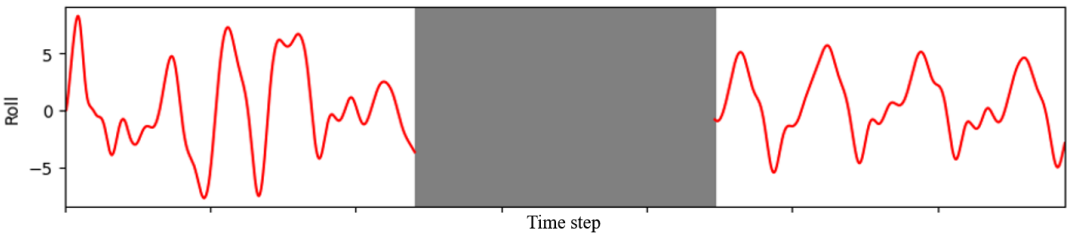
(a) r_{miss}=0%



(b) r_{miss}=10%



(c) r_{miss}=20%



(d) r_{miss}=30%

Fig. 6 The sketch of data missing used in this paper

4. Result and discussion

In this section, a typical working condition is used to test the performance of the LSTM model used in this paper, and the effect of the missing data on the prediction accuracy is tested. Then, the effect of hyperparameters and data of various working conditions is discussed.

4.1 The effect of missing data on the prediction

The data of Case 1 is used as a typical working condition to test the performance of the LSTM model first. The time history of the ship motion and force data is divided into training and test sets, with 80 % used for training and 20 % for testing, respectively. The parameter settings of LSTM model are shown in Table 3, which were determined based on our previous work.

Table 3 Hyperparameter configuration of the LSTM-based prediction model

Parameter	Value	Description
Hidden layer size	64, 64, 32	Number of hidden units per LSTM layer
Input sequence length	200	Number of past time steps as input
Number of LSTM layers	3	Depth of the LSTM network
Batch size	32	Mini-batch size during training
Learning rate	0.001	Step size for optimizer
Epochs	30	Training iterations
Activation function	ReLU	Introduces non-linearity to improve performance

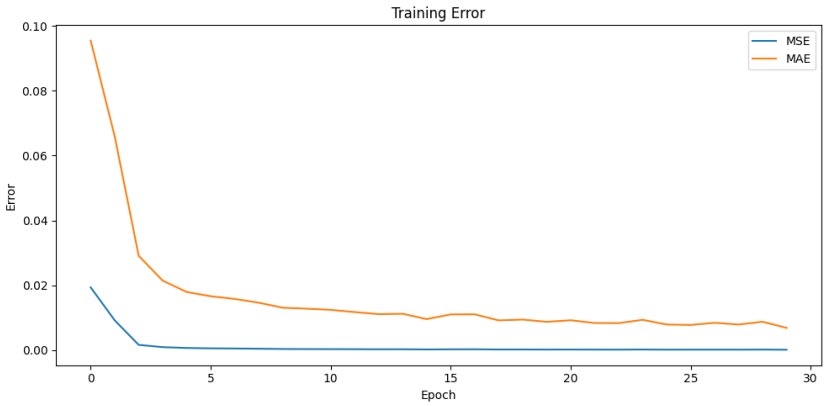
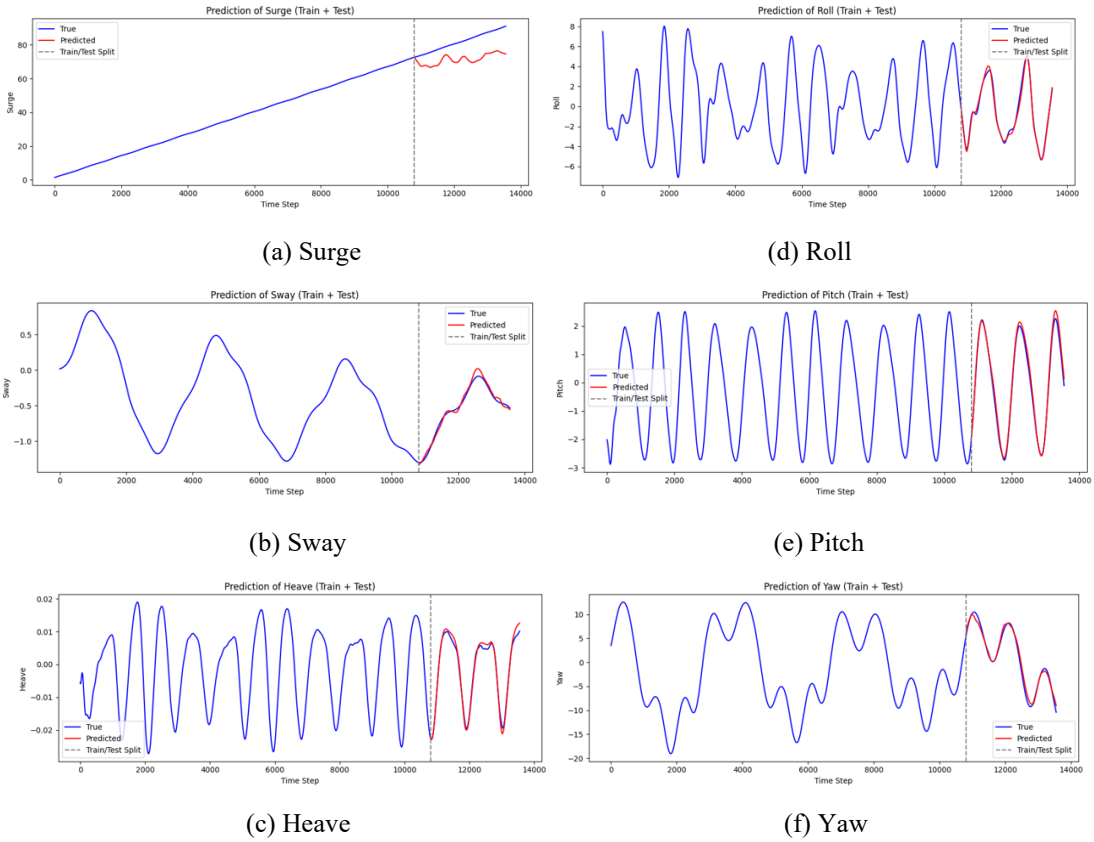


Fig. 7 Training error curves of the LSTM model in terms of *MAE* and *MSE*



(a) Surge (d) Roll

(b) Sway (e) Pitch

(c) Heave (f) Yaw

Fig. 8 Prediction results by LSTM with the data of Case 1 used for training and prediction

Figure 7 shows the *MAE* and *MSE* training error curves over 30 epochs. Both metrics drop sharply within the first 5 epochs, indicating quick learning of the temporal patterns of input data. Afterward, errors gradually stabilized, with minimal fluctuations beyond epoch 10. The *MAE* decreases smoothly, while the *MSE* drops more rapidly and approaches near-zero, reflecting the sensitivity of squared errors to large deviations. The convergence of both metrics suggests the LSTM model effectively captures the input-output mapping without overfitting, demonstrating efficient learning and stable error reduction in modeling ship motion data.

To demonstrate the performance of the LSTM model under self-supervised conditions, the data is split with 80 % used for training and 20 % for testing. Figure 8 shows the prediction results of the 6DOF motions with Figure 8 (a)-(f) representing surge, sway, heave, pitch, and yaw, respectively. In each plot, the blue line represents the true value, the red lines indicate the predicted results, and the vertical dashed line marks the train-test split. It shows that the predicted values closely follow the true motion trajectories in the test segment, indicating that the model has successfully captured the temporal dependencies of ship motion features. The smooth and continuous alignment between true and predicted curves suggests that the LSTM model is well-fitted and generalizes well within the same distribution. These examples serve as representative cases.

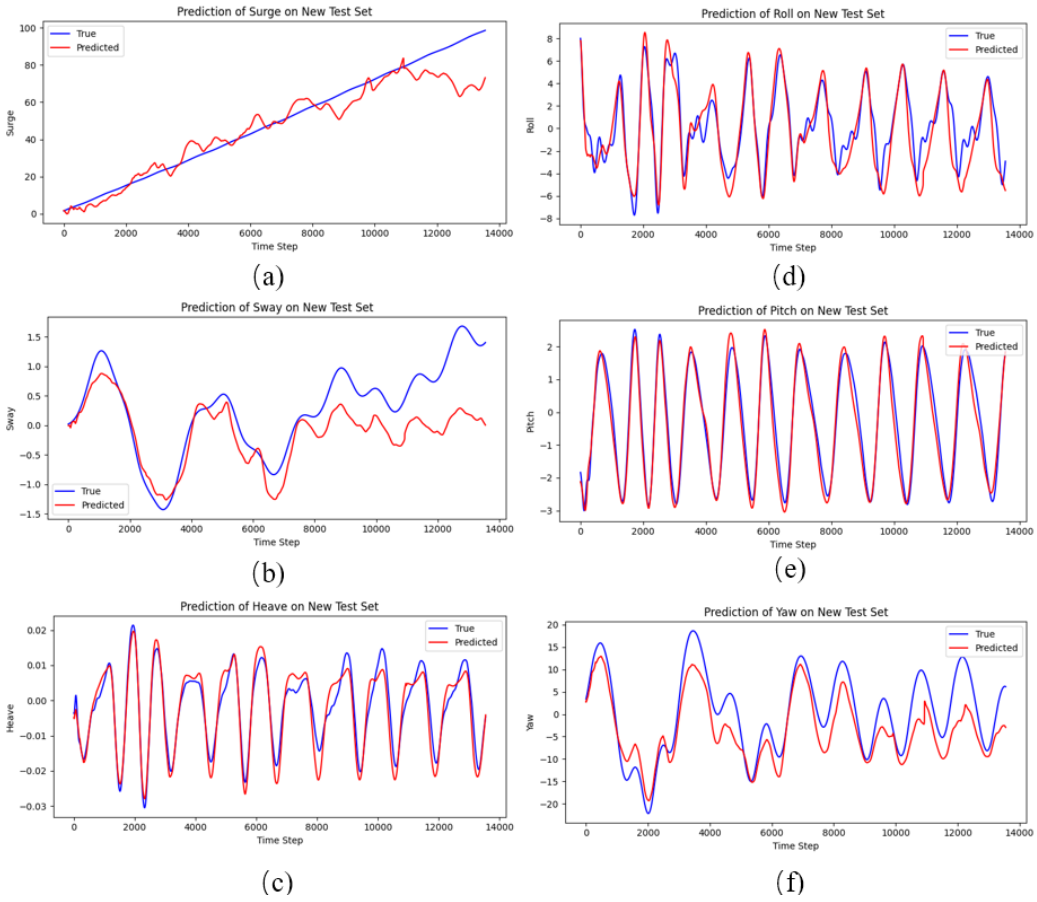
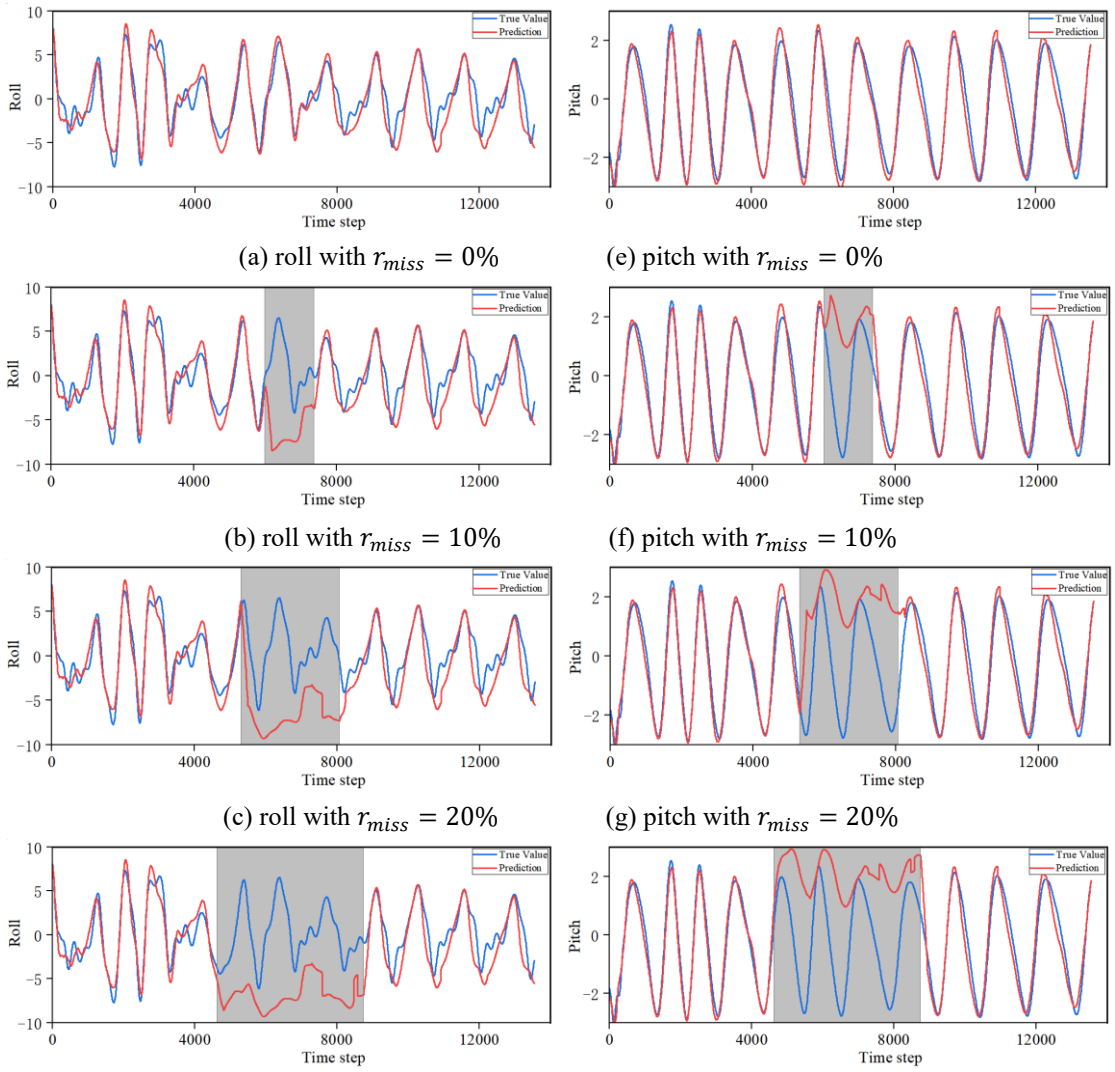


Fig. 9 Prediction results of Case 2 at r_{miss}=0 with the model trained by Case 1. (a) surge, (b) sway, (c) heave, (d) roll, (e) pitch, (f) yaw

To further study the LSTM model and the effect of missing data, two cases are used for tests, where Case 1 are used as training set and Case 2 as prediction set. The motion of ship in Case 2 is to some extent different from that in Case 1 as the ship speed increases. The predicted results with and without data being missing are compared with each other. Figure 9 shows that prediction results with r_{miss}=0. It is found that, when the data is intact, the model successfully captures the general trends and temporal patterns of most features, particularly for periodic angular motions, such as roll and pitch, where the predicted trajectories closely align with the true signals. In contrast, certain translational components, such as surge and sway, exhibit larger deviations, indicating a more complex or less transferable dynamic behavior across cases.



(a) roll with $r_{miss} = 0\%$ (e) pitch with $r_{miss} = 0\%$

(b) roll with $r_{miss} = 10\%$ (f) pitch with $r_{miss} = 10\%$

(c) roll with $r_{miss} = 20\%$ (g) pitch with $r_{miss} = 20\%$

(c) roll with $r_{miss} = 30\%$ (g) pitch with $r_{miss} = 30\%$

Fig. 10 Prediction results of Case 2 at various r_{miss} with the model trained by Case 1 by zero-filling method

Based on the result in Figure 9 the roll and pitch could be accurately predicted by current data set and model, hence, in Figure 10, the roll and pitch are used to preliminarily study the effect of missing data on the prediction, and the zero-filling method is applied, which means the missing data a filled with zero. Figure 10 shows that, when data is missing, there is a particularly large deviation between the predicted value and the true value, indicating that the missing data significantly impacts prediction accuracy, especially for the roll motion. This suggests that the traditional zero-filling method may not be suitable for predicting nonlinear ship motion. While this method allows the prediction model to continue functioning when data is missing, the lack of mapping relationship between motions leads to noticeable prediction errors. As the missing rate increases, the impact on prediction becomes more pronounced.

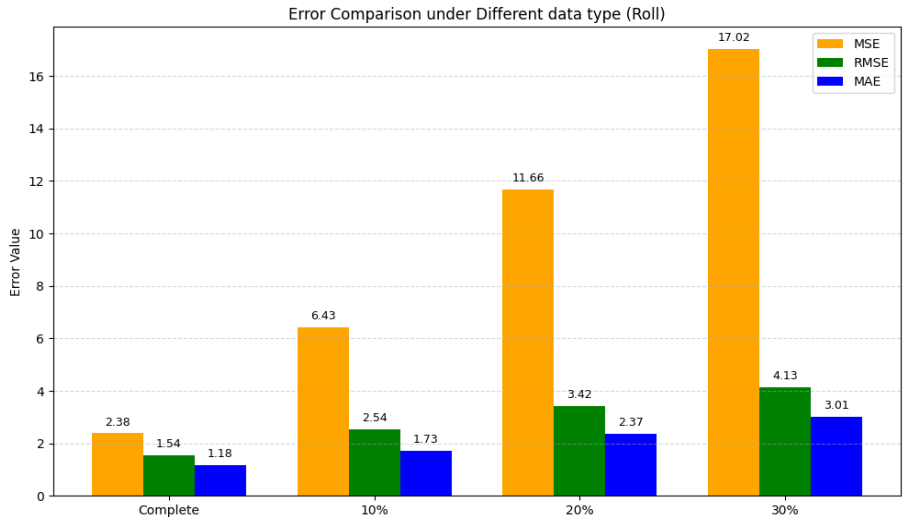


Fig. 11 Roll prediction error of Case 2 at various r_{miss} with the model trained by Case 1 by zero-filling method

Figure 11 compares the results of complete data and missing data; it can be clearly known that with the increase of missing rate the error increased. That is to say, the lack of data has a significant impact on the accuracy of the model. The reason for this phenomenon is that the model cannot capture continuous data for updates and thus cannot obtain information to complete the prediction.

In this study, the LSTM model performs one-step-ahead independent prediction, where an input window of 200 consecutive samples is used to predict the next step. During testing, each prediction uses the true observed sequence to update the input window, rather than feeding the model's previous outputs. Therefore, the prediction is non-recursive, and error accumulation does not occur. This design allows it to evaluate the model's intrinsic short-horizon predictive capability under missing data conditions.

Although the model was trained for one-step-ahead prediction, recursive rollout allows evaluating its performance over a prediction horizon, which corresponds to forecasting the ship's motion over a future time span. This horizon-based assessment is crucial for real-time control and decision-making, as it reflects how prediction errors accumulate when the model operates autonomously without future ground truth. Therefore, horizon-based evaluation (H = 1, 10, 30, 60, 100 time steps, physical time is 0.004 s, 0.04 s, 0.12 s, 0.24 s, 0.4 s) is included to quantify the model's robustness in practical, real-time ship motion monitoring scenarios.

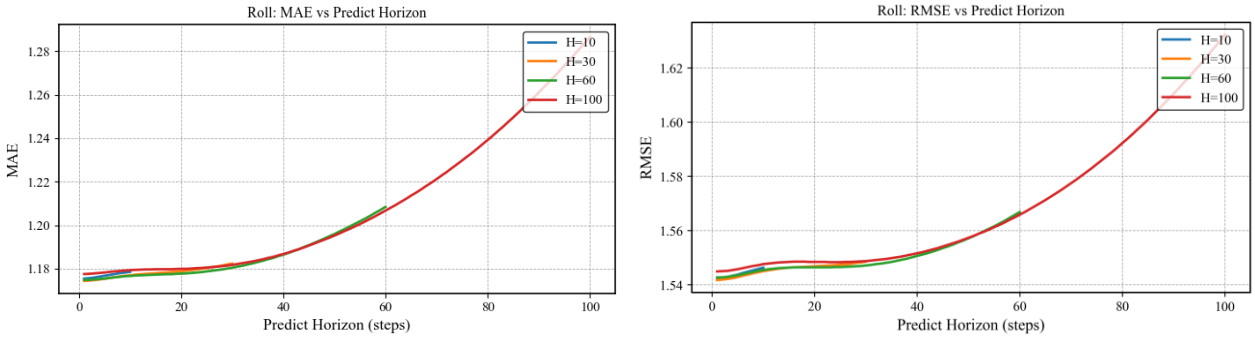


Fig. 12 The error accumulation over longer horizons

In Figure 12, the *MAE* and *RMSE* become greater with the increase of the predicted horizon. The results demonstrate that the one-step LSTM model provides accurate short-horizon predictions, while recursive rollout leads to gradual error accumulation over longer horizons. The limitation is inherent to one-step forecasting and can potentially be mitigated by adopting sequence-to-sequence or direct multi-step architectures, which are left for future exploration.

To further study the negative impact of the missing data, various hyperparameters of the model is studied under the assumption of r_{miss}=30 %, and the zero-filling method is still applied. Because the input sequence

length is crucial for the training stability and prediction performance, cases of three step size are used, as shown in Table 4. Besides, the effect of the key hyperparameters of LSTM model on robustness is studied, including input sequence length, hidden size and number of layers. The input sequence length refers to the number of consecutive time steps input into the LSTM model at one time, and the feature step size determines how much past information the LSTM can remember at one time.

Table 4 The hyperparameters used for test

Test	Input sequence length	hidden units number	number of LSTM layers
Test 1	50	64, 64, 32	3
Test 2	100		
Test 3	200		
Test 4	200	32, 32, 16	3
Test 5		64, 64, 32	
Test 6		128, 128, 64	
Test 7	200	64, 64, 32	1
Test 8			2
Test 9			3

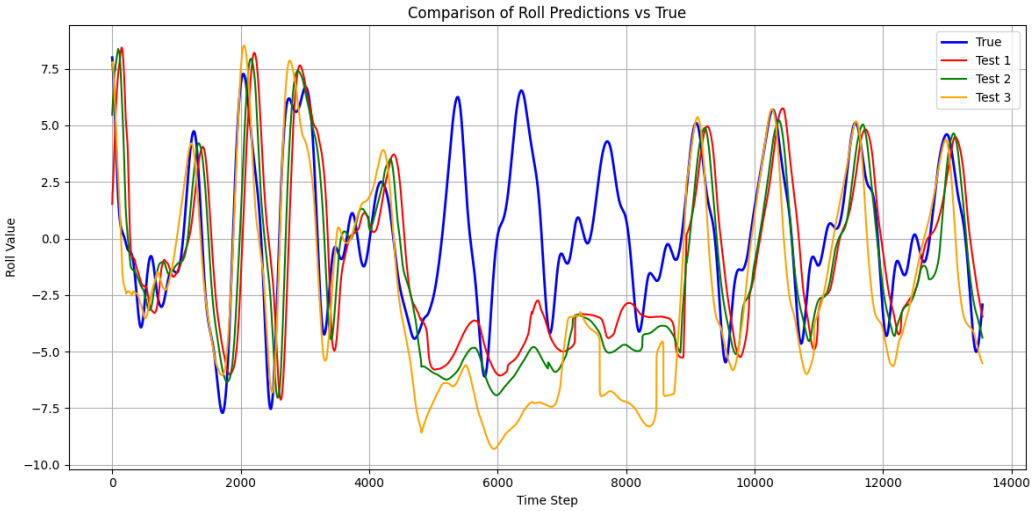


Fig. 13 Predicted roll for different input sequence length by zero-filling method (r_{miss}=30 %, Cases 1-2)

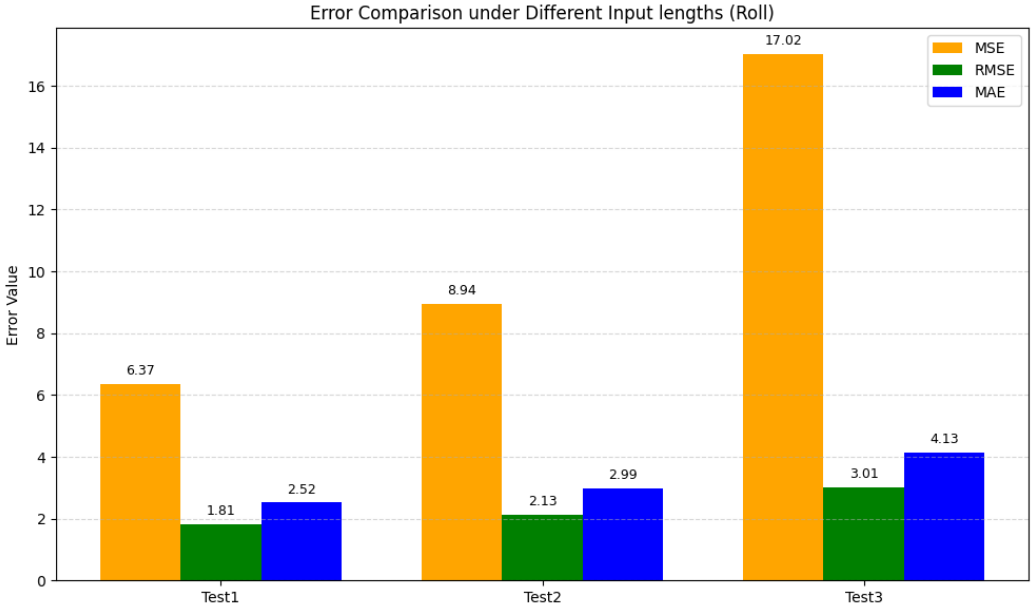


Fig. 14 Prediction error of roll for different input sequence length by zero-filling method (r_{miss}=30 %, Cases 1-2)

Figure 13 shows the predicted results of roll motion with different input sequence length under 30 % data missing. It shows that the prediction errors with three input sequence length are all obvious when the data missing begins, however, to decrease the input sequence length is valid to make the prediction more accurate. It means that adjusting the input sequence length can effectively decrease the negative impact of data missing on the memory effects of neural network and the disruption of mapping relationship. Figure 14 shows the comparison of *MSE*, *MAE*, and *RMSE* under different input sequence length. On one hand the prediction error gets relatively smaller with the decrease of input sequence length, on the other hand it indicates that simply adjusting the input sequence length has a very limited impact on the prediction performance for cases with missing data filled with zero.

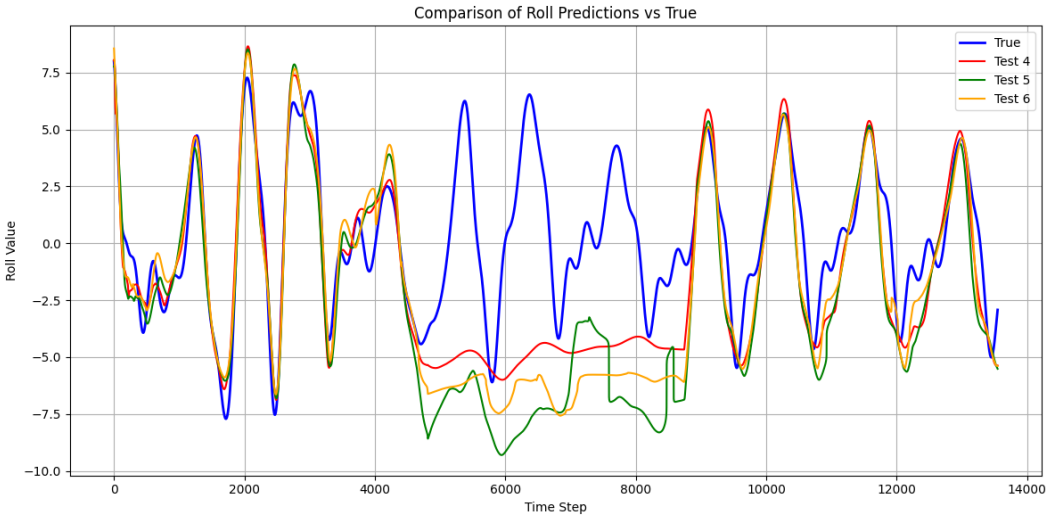


Fig. 15 Predicted roll for different numbers of hidden units by zero-filling method (r_{miss}=30 %, Cases 1-2)

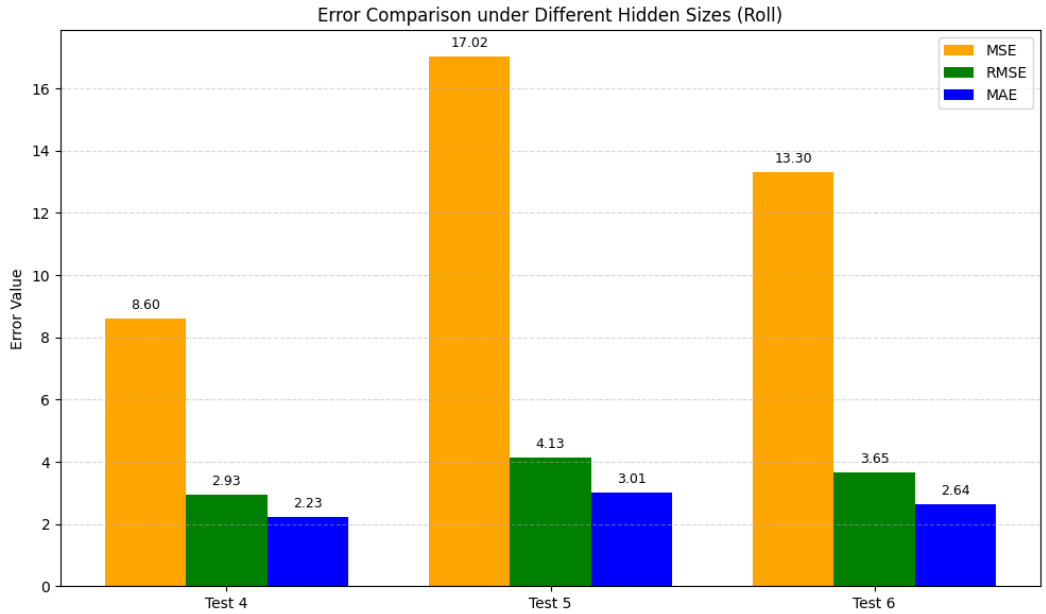


Fig. 16 Prediction error of roll for different hidden unit sizes by zero-filling method (r_{miss}=30 %, Cases 1-2)

Figures 15 and 16 shows the prediction results with 30 % data missing for different the number of hidden units, as shown as Tests 4-6 in Table 4. The number of hidden units is the dimension of the hidden state generated at each time step, which reflects the capacity of the LSTM to learn more complex temporal relationships. Figure 15 shows that the number of hidden units has little effect on improving the prediction accuracy, but more hidden units could make the variation trend of predicted values get closer to variation trend of true values. The comparison of the corresponding errors is shown in Figure 16, it can be found that, when the data missing begins, all the cases show a phenomenon of overall sinking, and the effect on the prediction accuracy is limited.

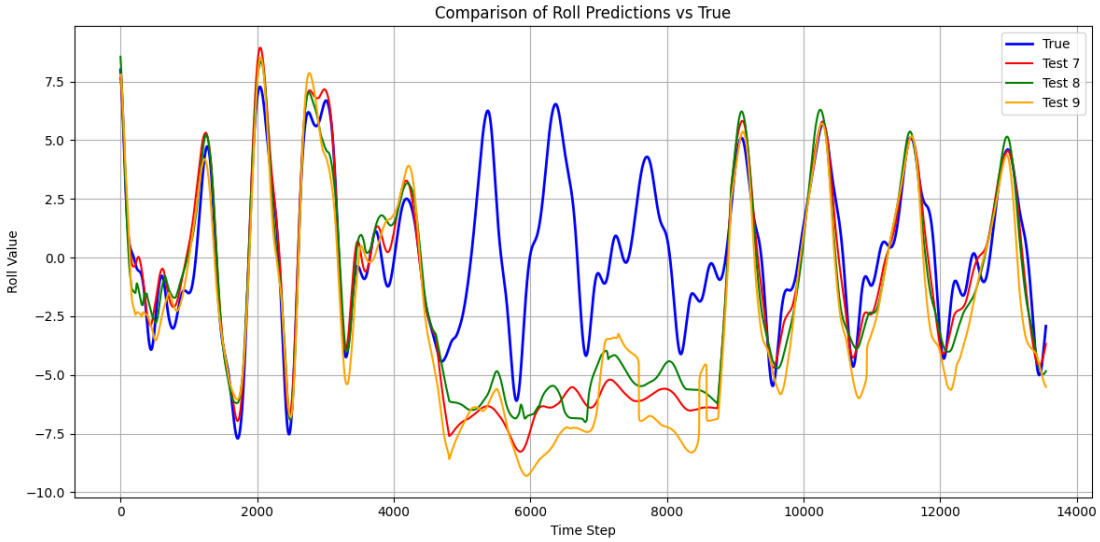


Fig. 17 Predicted roll for different LSTM layer numbers by zero-filling method (r_{miss}=30 %, Cases 1-2)

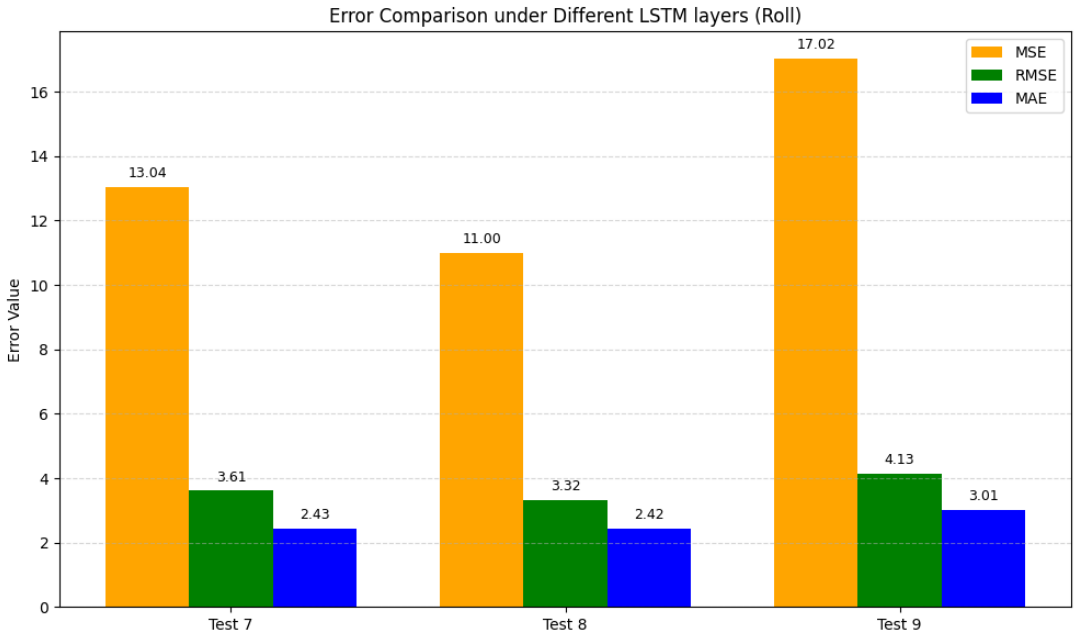


Fig. 18 Prediction error of roll for different LSTM layers by zero-fill method. (r_{miss}=30 %, Cases 1-2)

In Figures 17 and 18, The effect of LSTM layers number is tested by Tests 7-9 in Table 4. Theoretically, more layers can improve the temporal expression ability and abstract learning level of LSTM model, but it also introduces the risk of training instability and overfitting. As shown in Figure 17. The effect of the LSTM layers is to some extent similar to that of hidden unit number. It mainly impacts the ability of the LSTM model to capture the mapping relationship between sequential data. Figure 18 compares the prediction errors of roll for different LSTM layers. Test 8 of two LSTM layers achieved the lowest error and performed better than the other two cases in terms of stability.

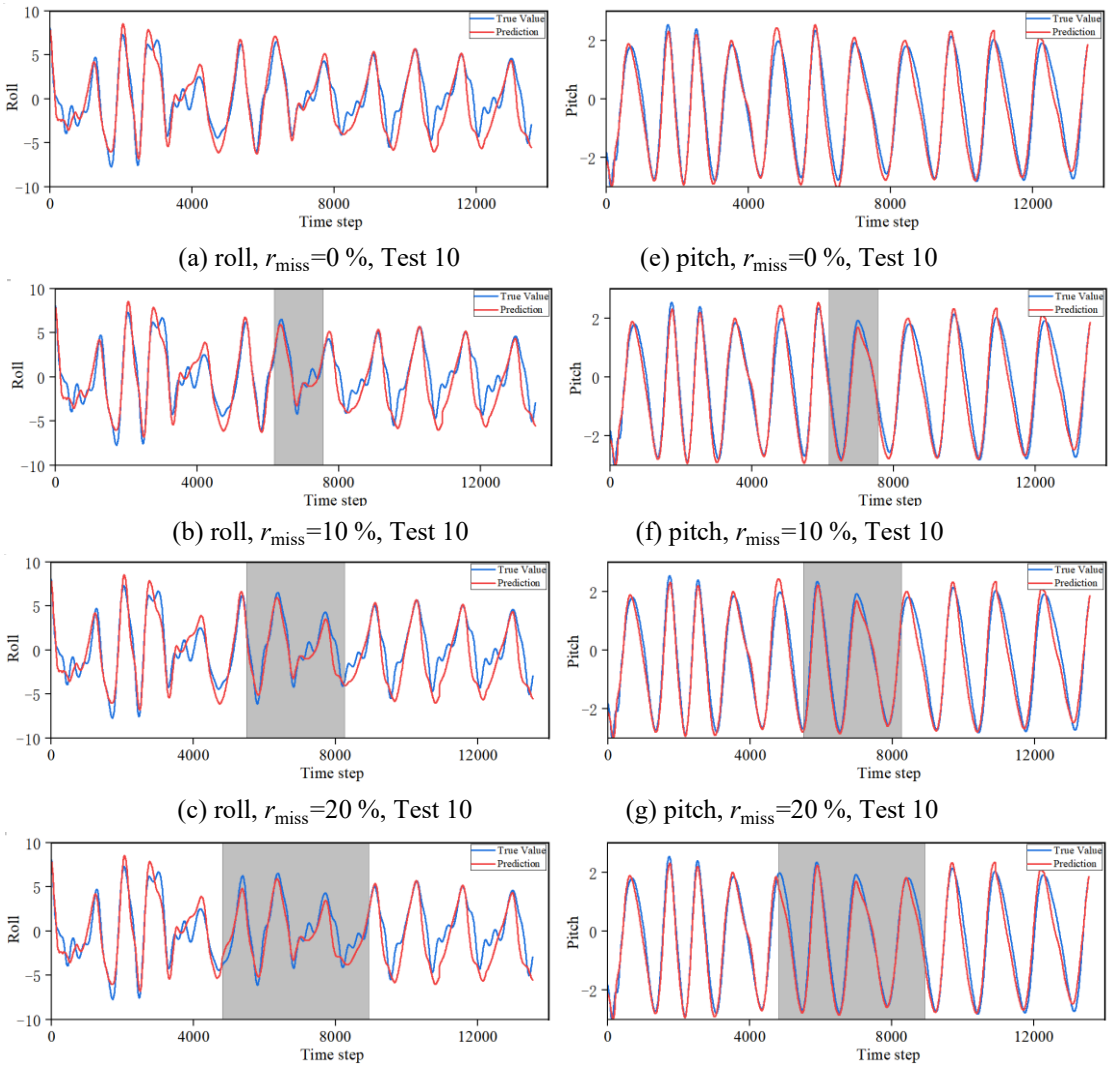
To sum up, the effects of input sequence length, hidden unit number, and LSTM layer on the prediction are obviously different under missing data condition. The smaller input sequence length, larger hidden unit number, and medium LSTM layer performs relatively better in prediction accuracy and capturing the mapping relationship, which will be used here after. It also shows that, when missing data occurs, it is not suitable to apply the traditional zero-filling method to improve the prediction accuracy.

4.2 The effect of mean filling method on the prediction

To solve the problem of the inaccurate prediction and the disruption in capturing the mapping relationship lead by the intermittent data missing, the mean-filling method is presented in this paper, where the average of existing data and predicted results is used as input in each time step when data is missing. Cases 1-2 and Cases 6-7 of Table 5 are used for train and prediction, which are data set of ship sailing in different wave heading angles.

Table 5 Training and prediction cases for the test of mean filling method

Test	Training case	Prediction case	r_{miss}
Test 10	Case 1	Case 2	30 %
Test 11	Case 1	Case 3	30 %
Test 12	Case 6	Case 7	30 %



(a) roll, r_{miss}=0 %, Test 10 (e) pitch, r_{miss}=0 %, Test 10

(b) roll, r_{miss}=10 %, Test 10 (f) pitch, r_{miss}=10 %, Test 10

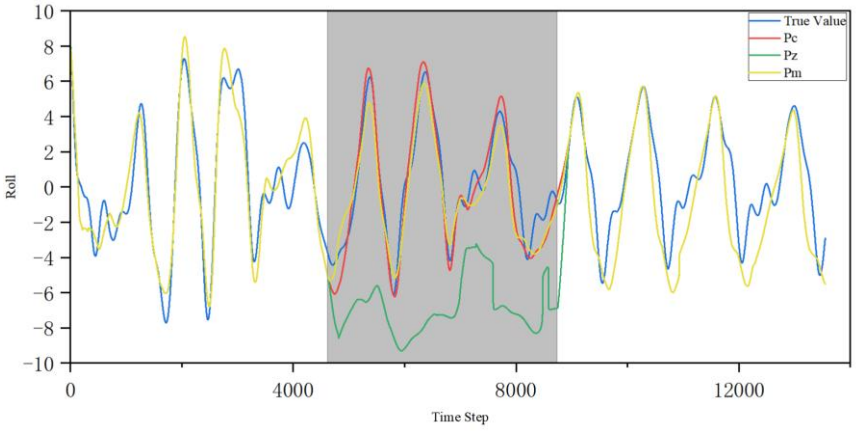
(c) roll, r_{miss}=20 %, Test 10 (g) pitch, r_{miss}=20 %, Test 10

(d) roll, r_{miss}=30 %, Test 10 (h) pitch, r_{miss}=30 %, Test 10

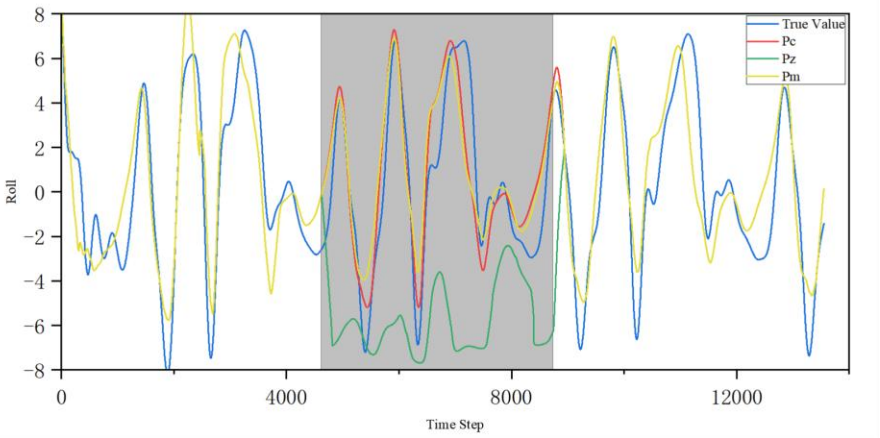
Fig. 19 Prediction results of roll motion for different r_{miss} by mean-filling method

Figure 19 illustrates the prediction results for the roll and pitch features after applying mean-value imputation to the missing regions of the input sequence. It could be seen that, compared with the zero-filling method, the application of mean-filling noticeably improves the prediction accuracy when data missing occurs, especially for the pitch motion. The predicted value remains closely aligned with the ground truth; even missing rat is as large as 30 %. For roll motion, although some phase shift and minor amplitude attenuation can be observed in cases with severe missingness, such as Figure 19 (d), the variation trend of the data is well preserved. These results suggest that, despite its simplicity, mean-fill method may be sufficiently robust for the prediction of ship motion with missing data.

Figure 20 shows the comparison of the predicted roll by zero-filling and mean-filling method. The prediction accuracy by the mean-filling method is in good agreement with the true value. Compared with the prediction results without missing data, the mean-filling method can significantly decrease the negative effect by the missing data. Figure 21 shows the error by different methods. The prediction error without missing data is taken as the baseline for comparison. It shows that the *MSE* of mean-filling method decreases significantly than the zero-filling method by about 83 %, while the *MAE* and *RMSE* decrease by about 35 % and 40 %, respectively.

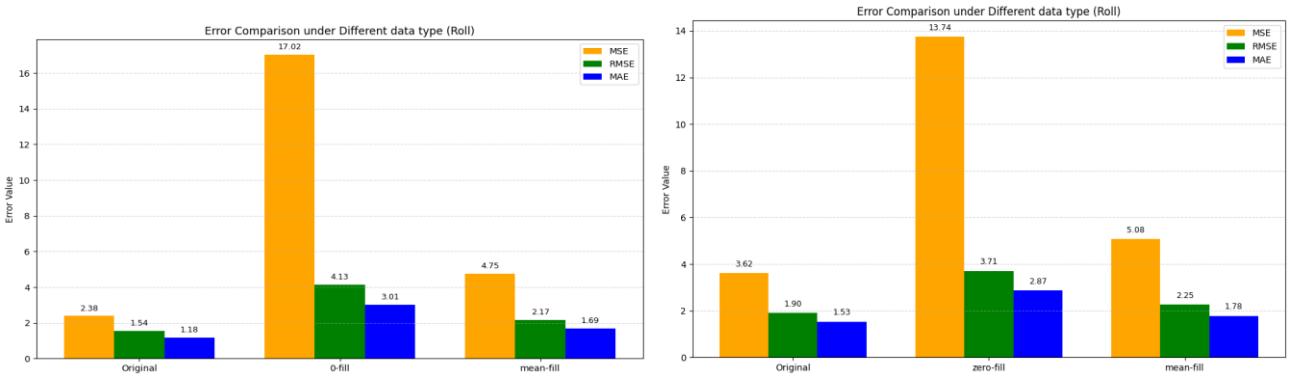


(a) r_{miss}=30 %, Test 10



(b) r_{miss}=30 %, Test 11

Fig. 20 Prediction of roll motion by zero-filling method and mean-filling method



(a) r_{miss}=30 %, Test 10 (b) r_{miss}=30 %, Test 11

Fig. 21 Error of Roll prediction by zero-filling method and mean-filling method

To further study the performance of mean-filling method in different working conditions, Cases 6-7 with different wave heading angles is used for prediction. The results are shown in Figure 22. It could be found that the predicted values by mean-filling method is still close to the predicted value without missing data, indicating that the mean-filling method has certain stability and availability in different scenarios.

It can be observed that the mean-filling method consistently achieves lower *MAE* and *RMSE* values than zero-filling across different missing data rates. This is because the mean-filling approach provides smoother and statistically representative replacements for missing samples, which helps maintain the local continuity of the time series and prevents abrupt changes in the input. In contrast, zero-filling introduces sharp discontinuities and non-physical spikes at the boundaries of missing segments. Such artificial distortions may

mislead the temporal feature extraction of the LSTM, resulting in degraded predictive accuracy. Therefore, the mean-filling method proves to be more robust and physically reasonable for ship-motion time-series data, where continuity and smooth variation are critical for capturing dynamic behavior.

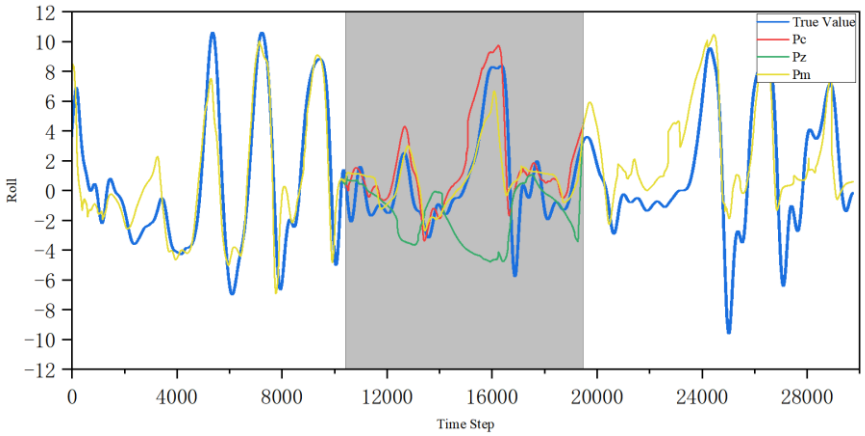


Fig. 22 Predicted roll by zero-filling and mean-filling methods (r_{miss}=30 %, Test 12)

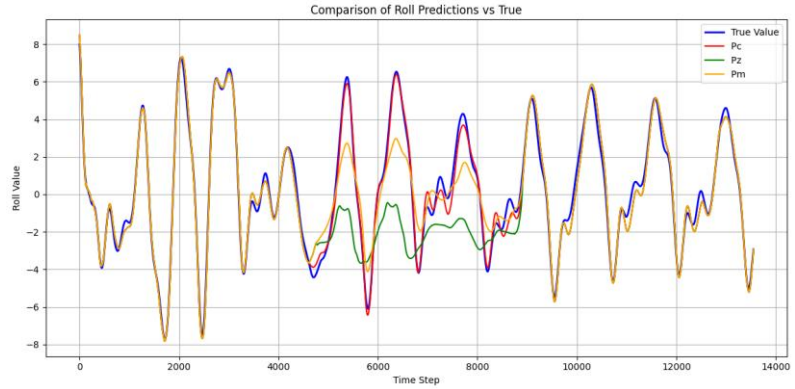
4.3 The performance of mean imputation at various working conditions

To further study the effect of data sets on prediction performance of the mean-filling method under missing data conditions, more cases are used for training and prediction to test its validity in complex working conditions. The combinations of training and prediction cases are shown in Table 6. Test B is used first, and the results are shown in Figure 23, where Case 1 and Cases 3-5 are used for training, and Case 2 for prediction. For data missing condition, the missing rate r_{miss}=30 % are used for both zero-filling and mean-filling methods.

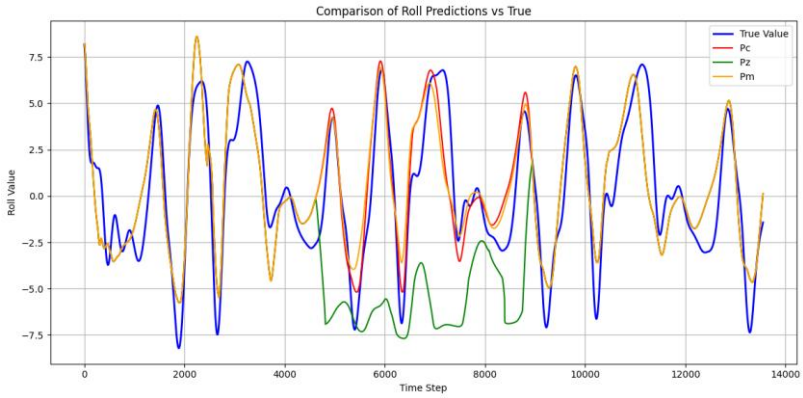
Table 6 Different combinations of training and prediction cases

Test	Training case	Test case	r_{miss}
Test 13	Case 1, Cases 3-5	Case 2	30 %
Test 14	Case 1-2, Cases 4-5	Case 3	30 %
Test 15	Case 11, Cases 13-15	Case 12	30 %
Test 16	Cases 11-12, Cases 14-15	Case 13	30 %
Test 17	Case 1, Cases 3-5, Cases 10-15	Case 2	30 %

Comparing Figures 20 and 23, it is found that, during time steps without missing data, the prediction accuracy of model trained by cases of various working conditions is obviously improved. When data missing occurs, the prediction by mean-filling method is still obviously more accurate than zero-filling method, but the prediction accuracy decreases with the working conditions used for training getting complex. It probably means that increasing the amount and diversity of training data can improve prediction accuracy under no missing data conditions, but it may also amplify the negative impact of mapping confusion when data is missing. To further study such phenomena, the combinations of training and prediction cases Test 15-16 with different wave conditions in Table 6 are used.



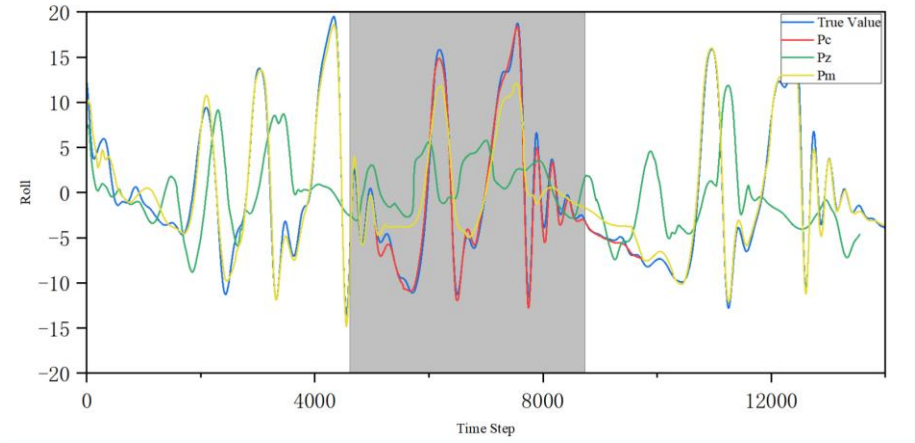
(a) r_{miss}=30 %, Test 13



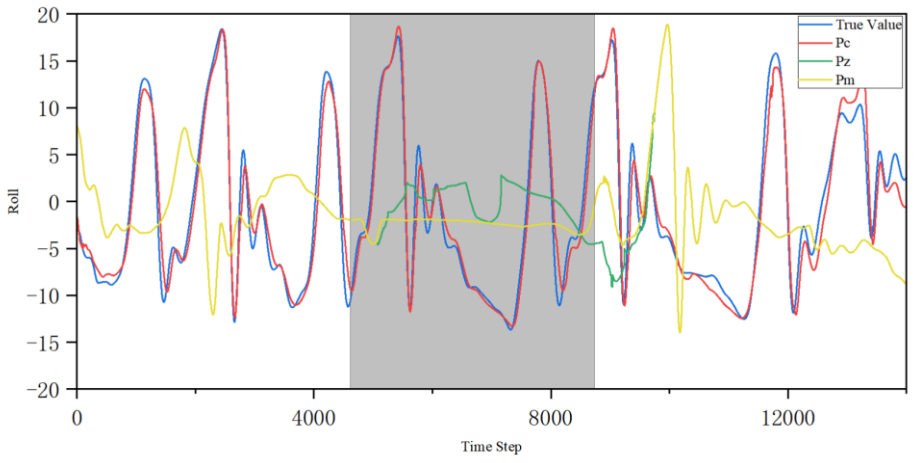
(b) r_{miss}=30 %, Test 14

Fig. 23 Predicted roll by zero-filling and mean filling methods

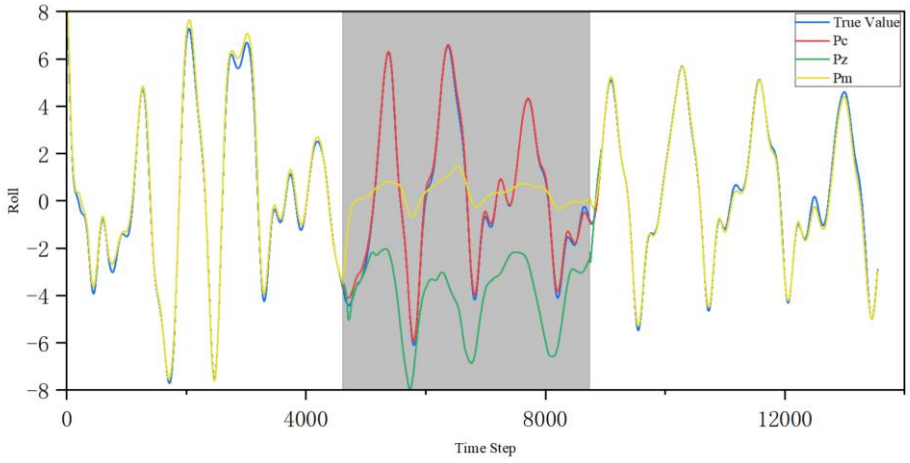
To sum up, the phenomenon of Figures 23 and 24 shows that increasing the training data set can improve the prediction performance of neural network for ship motion in the non-missing, whereas it makes the mean filling method unstable in the missing area. The main reason is that, when the training set becomes complicated, the missing local data can significantly affect the internal logical relationships of nonlinear data, then the negative impact of the data missing on the capture of the mapping relationship and logical consistency of the data is significantly amplified. Also, the increase of the data amount and variability amplifies the divergence between historical average imputation by mean filling method and true values, overwhelming gating mechanisms of LSTM model in capturing long-term dynamic dependencies [35]. Simultaneously, heterogeneous data disrupts the hidden-state mapping relationship, where input noise sensitivity inherently degrades feature representation coherence, triggering nonlinear error propagation in temporal predictions. Therefore, it is crucial to reduce the variability of heterogeneous data by grouping temporally coherent motion patterns while missing data occurs, such as cluster-based imputation. Thereby aligning LSTM input distributions with localized dynamics to suppress error cascades. This stratification preserves feature-target mapping consistency, enabling adaptive learning of intra-cluster dependencies while mitigating noise amplification across divergent ship behaviors.



(a) r_{miss}=30 %, Test 15



(b) r_{miss}=30 %, Test 16



(c) r_{miss}=30 %, Test 17

Fig. 24 Predicted roll by zero-filling and mean filling methods

4.4 The LOCF and linear methods comparison

Based on the comparison shown in Figure 10, the LOCF method was further investigated as an additional imputation strategy to assess its influence on prediction accuracy. In this experiment, missing data rates of 10 %, 20 %, and 30 % were introduced, and the LOCF approach was applied to fill the missing segments by carrying forward the last observed value.

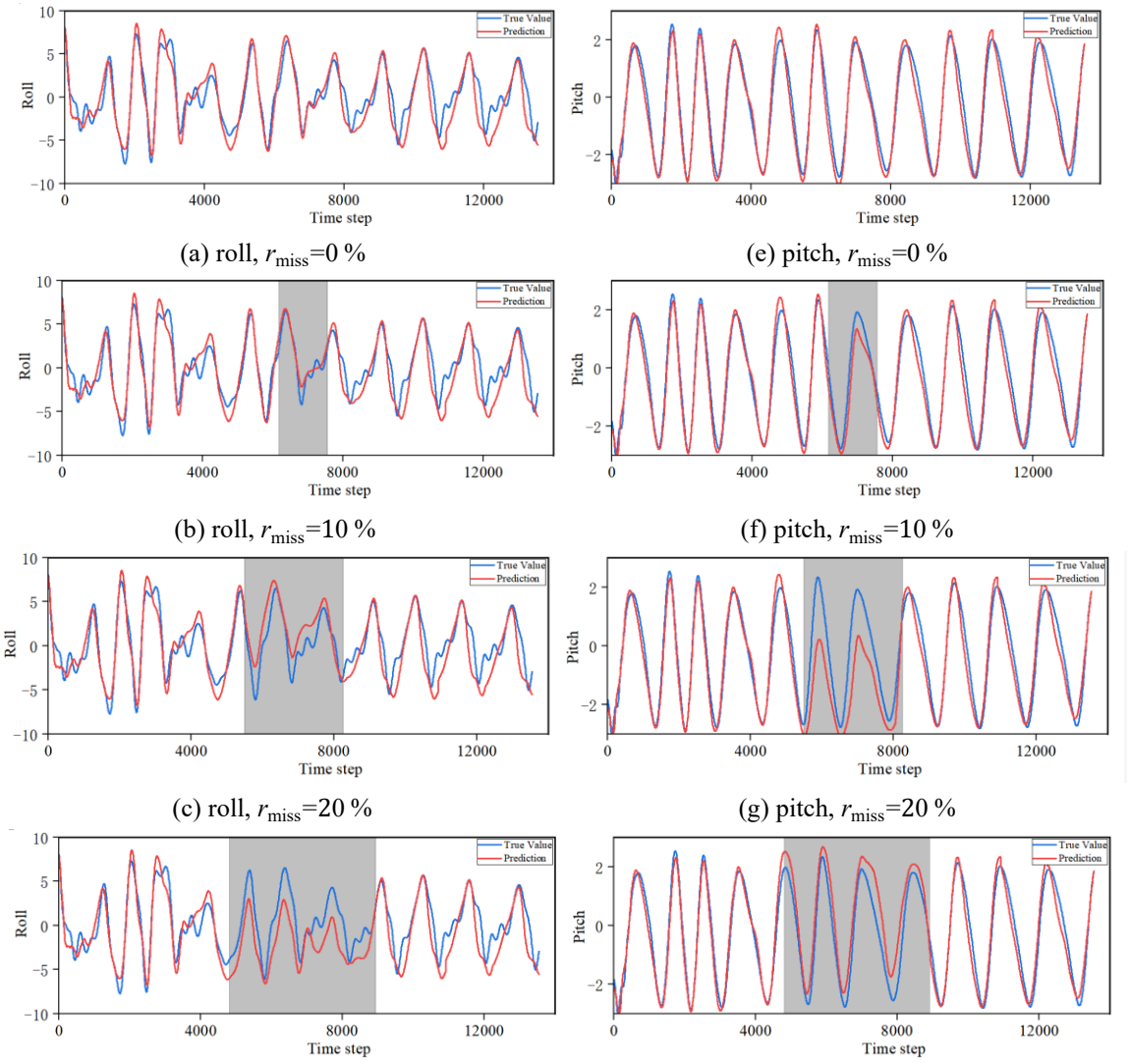


Fig. 25 Prediction results of Case 2 at various r_{miss} with the model trained by Case 1 by LOCF method

Figure 25 illustrates the prediction results obtained using the LOCF method under different missing data rates (10 %, 20 %, and 30 %). The imputed segments differ among missing rates because the last available values before each missing block vary with the specific missing data pattern. As a result, the reconstructed sequences show distinct local behaviors at different missing ratios. Overall, the LOCF method produces relatively poor prediction performance, as reflected by the noticeable deviation between the predicted and true values. It is mainly because LOCF maintains constant values within missing intervals, thereby breaking the temporal continuity and failing to capture dynamic variations in ship motion. Consequently, the LSTM model trained on continuous sequences struggles to adapt to such step-like imputations, leading to larger accumulated errors compared with smoother imputation strategies such as mean filling.

In addition, the LOCF strategy was incorporated into the Test 13 setup, which involved multi-condition training using ship-motion data under different speeds and wave parameters. As shown in Figure 26, the prediction results still exhibit large deviations between the true and predicted values. The performance degradation indicates that the limitation of LOCF is structural rather than data-dependent—its constant-value interpolation fails to represent dynamic variations across conditions, resulting in poor generalization even in a richer training environment.

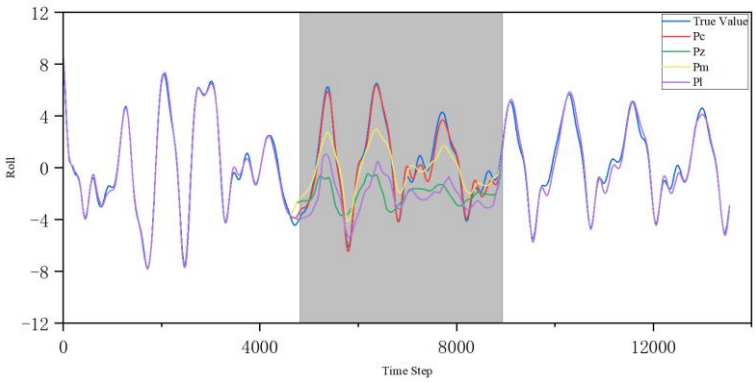
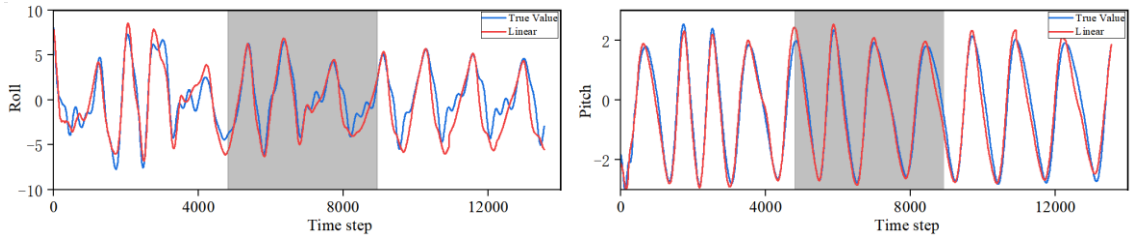


Fig. 26 Predicted roll by LOCF r_{miss}=30 %, Test 13



(a) roll, r_{miss}=30 % (b) pitch, r_{miss}=30 %

Fig. 27 Prediction results of Case 2 at r_{miss} 30 % with the model trained by Case 1 by linear interpolation method

As shown in Figure 27, the overall prediction trend and waveform shape remain highly consistent between the two cases, with only slight amplitude and phase differences in the interpolated region. The interpolation method effectively reconstructs missing segments while maintaining the temporal continuity of the input signal, allowing the model to achieve stable and accurate predictions. Although minor discrepancies can be observed at the edges of the interpolated region—mainly due to the smoothing nature of linear interpolation—the overall prediction accuracy remains close to that of the complete-data case. This indicates that the linear interpolation approach provides a simple yet sufficiently accurate solution for handling short-term data gaps in practical ship motion forecasting applications.

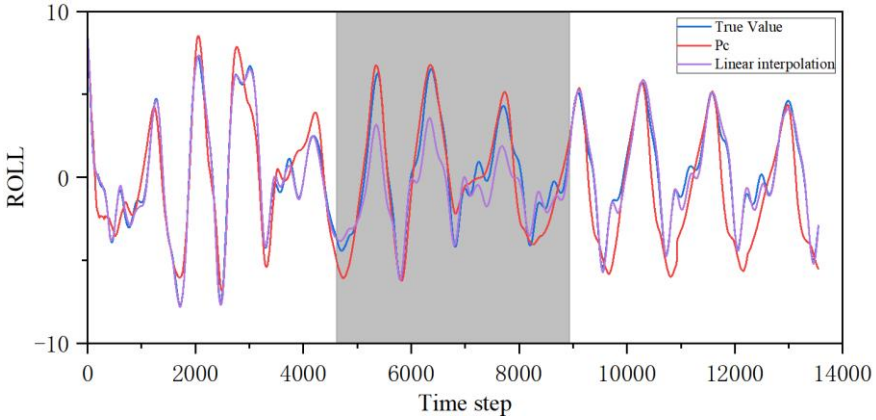


Fig. 28 Predicted roll by Linear interpolation r_{miss}=30 %, Test 13

Similarly, this paper further evaluates the prediction performance of the linear interpolation method under multiple working conditions in Figure 28. It is observed that the prediction accuracy of the linear interpolation method decreases compared with that under a single working condition, and the deviation between the predicted and true values becomes more significant. This degradation can be attributed to the increased variability of the input features caused by changes in ship speed, wave parameters, and flow conditions. Since linear interpolation assumes a locally linear temporal evolution, it cannot accurately reconstruct nonlinear transitions or dynamic coupling effects among different working conditions. As a result,

the interpolated data introduces additional bias into the model input, leading to larger prediction errors and reduced robustness.

4.5 Comparison with baseline models

To further evaluate the predictive performance of the proposed LSTM network, several baseline models were implemented, including a simple linear regression model, a persistence model, and a DMD model. These models were compared with the LSTM under both complete and missing data conditions in Subsection 4.1 test which uses Case 1 to predict Case 2.

Table 7 summarizes the prediction accuracy of the four models under complete data conditions. As can be seen, all models achieve very high accuracy, with *MAE* and *RMSE* values remaining within a small range and $R^2 > 0.99$. This indicates that, when the input sequence is continuous and complete, the underlying ship-motion dynamics can be reasonably captured even by simple linear models. In other words, the dataset exhibits partial linearity in its short-term behavior, allowing both regression-based and data-driven approaches to perform comparably.

Table 7 The evaluation error of different models under complete data

Model	*MAE*	*MSE*	*RMSE*	R^2
Linear regression	0.000005	0.001823	0.002260	0.9990
Persistence	0.000626	0.019778	0.02520	0.9991
DMD	0.0925	0.2042	0.3041	0.9912
LSTM	0.0110	0.0002	0.0148	0.9990

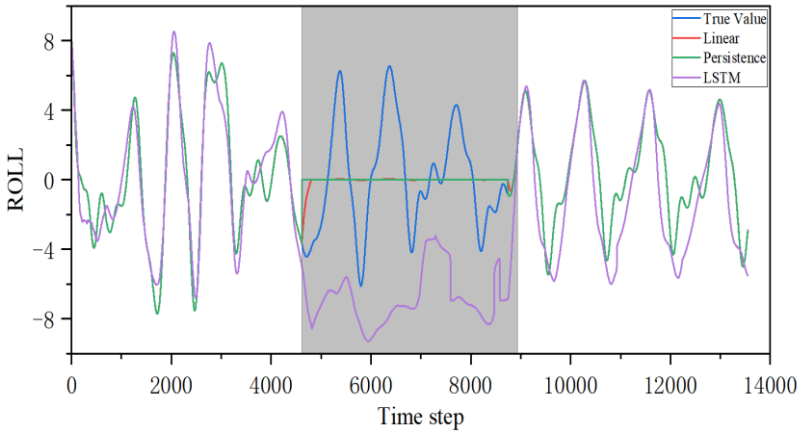


Fig. 29 Comparison of true and predicted roll motion using linear regression, persistence, and LSTM under missing data condition by zero-filling

However, once missing segments were introduced, the performance divergence became pronounced. In Figure 29, the linear regression and persistence models show a clear tendency to degenerate into static or linearly biased predictions once data gaps appear. Specifically, the persistence model maintains a constant value throughout the missing interval, while the linear regression output follows a nearly straight line determined by the imputed mean or zero values. Such behavior indicates that linear methods fail to reconstruct the dynamic variations of ship motion when the temporal continuity of the input sequence is disrupted. Although DMD performs well on complete and noise-free time series, its predictive capability becomes sensitive to data discontinuities. Under zero-filling or linear interpolation, the reconstructed trajectory deviates from the smooth dynamical structure assumed by DMD, leading to reduced forecasting accuracy. This is not a limitation of DMD itself but rather reflects a mismatch between simple imputation strategies and the intrinsic temporal coherence required by DMD-based models. The LSTM model demonstrates a markedly better capability to preserve the physical pattern of the motion signal. It successfully tracks the amplitude and phase

evolution even within and after the missing regions, reflecting its inherent ability to learn nonlinear temporal dependencies and internal memory representation.

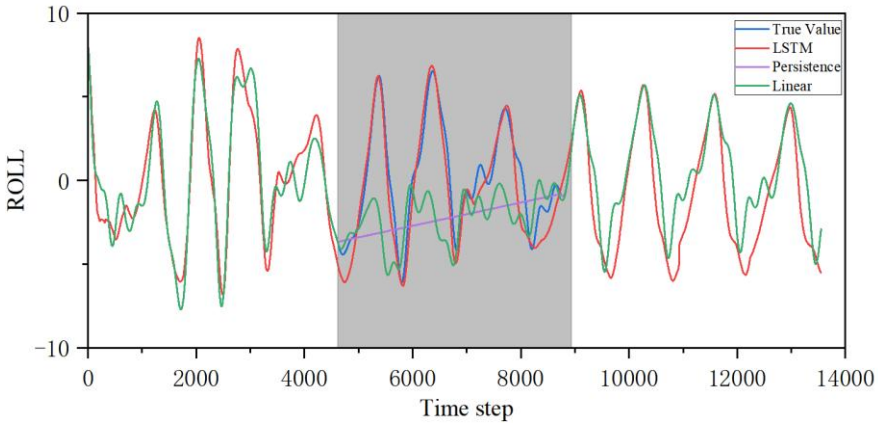


Fig. 30 Comparison of true and predicted roll motion using linear regression, persistence, DMD, and LSTM under missing data condition by linear interpolation

Figure 30 illustrates the comparison between the predicted and true roll motion under the linear interpolation imputation method using different predictive models, including LSTM, persistence, and linear regression. The shaded region indicates the time interval containing missing values in the original dataset. It can be observed that under the missing data condition, the LSTM model exhibits the best consistency with the true values, maintaining the oscillation phase and amplitude relatively well. The DMD model, while effective at extracting coherent temporal structures under fully observed conditions, becomes sensitive when data continuity is disrupted by interpolation.

Overall, the results demonstrate that while interpolation provides a continuous input sequence, the predictive performance still depends strongly on the model's ability to represent nonlinear and time-dependent dynamics. The LSTM model remains the most robust among all tested baselines under linear interpolation filling. While DMD and linear models provide interpretable and computationally efficient baselines, their effectiveness rapidly deteriorates under missing data scenarios. The LSTM network remains the most reliable and robust predictor, capable of maintaining temporal consistency and suppressing error amplification caused by imputation distortions.

5. Conclusion

This paper studies the performance of LSTM model in the prediction of ship motion in waves with missing data. The dataset of ship motion in one working condition of ship speed and wave is used to verify the feasibility of the model, and the effect of hyperparameters on the prediction accuracy is tested. Then, the effect of traditional zero-filling and mean filling method is used for the prediction under data missing condition. Finally, ship motion data under different workings conditions were used for training and prediction to analyze the impact of data missing and the effectiveness of the mean filling method. Based on these results, the following conclusions can be drawn:

(1) Data missing will cause the model to lose data information, the model prediction accuracy is distorted, and the prediction task cannot be completed.

(3) The traditional zero-filling method is not suitable to predict the ship motion under data missing condition. The mean-filling method has certain stability and availability for ship motion prediction in different scenarios.

(3) Increasing the diversity of training data set has negative impact on the prediction stability for the mean filling method under data missing condition. Grouping coherent motion patterns can reduce the error propagation and preserve model robustness.

(4) The mean filling method for data missing requires high classification and recognition accuracy of training data. Exploring the use of mapping relationships between different ship motion data to fill missing

data may offer greater practical value. Future work will focus on investigating this approach to improve the robustness and applicability of data imputation methods in ship motion prediction.

FUNDING

This project was supported by the National Natural Science Foundation of China (Grant number 52401385), Open Research Project of State Key Laboratory of Maritime Technology and Safety (Grant number W25CG000073), and MTIC-JUST Joint Innovation Center Development Fund.

APPENDIX

A minimal pseudo-code and a code snippet.

Inputs: Multivariate time series $X \in \mathbb{R}^{T\times 18}$ with ordered channels $\mathcal{C} = \{c_1, \dots, c_{18}\}$; sampling step $\Delta t = 0.004$s; window length L(200); random seed.

Outputs: Per-channel MAE/MSE/RMSE/R^2 on the test set; matrices $\hat{Y}_{\text{test}}$and Y_{test}(prediction vs. ground truth) for reproducibility.

Set seeds 42 for Python/NumPy/PyTorch; enable deterministic flags.

Chronological split: first 80% samples →train; last 20% →test

Normalization (train-only stats): compute per-channel scaler on train (Min-Max scaler); apply to train and test.

Windowing (seq-to-one, multivariate): for each $t = L, \dots, T_{\text{train}} - 1$, form $x_t = X[t - L{:}\, t, :] \in \mathbb{R}^{L\times 18}$, label $y_t = X[t, :] \in \mathbb{R}^{18}$.

Model: LSTM mapping $\mathbb{R}^{L\times 18} \to \mathbb{R}^{18}$; loss = MSE & MAE; optimizer = Adam (lr 0.001).

Train: iterate epochs; minimize MSE over training windows (teacher forcing on by construction of seq-to-one).

(Optional for missing data study): create test-only block-missing mask, impute masked test inputs by a chosen method (zero/mean/LOCF).

One-step test evaluation:

Feed forward through the test window x_t to obtain $\hat{y}_t \in \mathbb{R}^{18}$

Calculate MAE, MSE, RMSE, R^2 for each channel, and save the predicted/real CSV for reproduction.

Recursive multi-step rollout

Inputs: normalized test sequence $X_{\text{test}} \in \mathbb{R}^{T_{\text{test}}\times 18}$ (intact and imputed version); window L; prediction horizons $H = \{10,30,60,100\}$.

Outputs: Per-channel MAE, RMSE, and R^2as functions of H; error-growth curves and exported prediction tables.

REFERENCES

[1] Nguyen, A., Yosinski, J., Clune, J., 2015. Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 7-12 June, Boston, Massachusetts, USA, 427-436. https://doi.org/10.1109/CVPR.2015.7298640

[2] Markey, M. K., Tourassi, G. D., Margolis, M., DeLong, D. M., 2006. Impact of missing data in evaluating artificial neural networks trained on complete data. *Computers in Biology and Medicine*, 36(5), 516-525. https://doi.org/10.1016/j.compbiomed.2005.02.001

[3] Hua, V., Nguyen, T., Dao, M. S., Nguyen, H. D., Nguyen, B. T., 2024. The impact of data imputation on air quality prediction problem. *PLOS ONE*, 19(9), e0306303. https://doi.org/10.1371/journal.pone.0306303

[4] Junninen, H., Niska, H., Tuppurainen, K., Ruuskanen, J., Kolehmainen, M., 2004. Methods for imputation of missing values in air quality data sets. *Atmospheric Environment*, 38(18), 2895-2907. https://doi.org/10.1016/j.atmosenv.2004.02.026

[5] Jiang, M., Liu, Z., Xu, Y., 2025. A traffic prediction method for missing data scenarios: Graph convolutional recurrent ordinary differential equation network. *Complex & Intelligent Systems*, 11(2), 1-24. https://doi.org/10.1007/s40747-024-01768-7

[6] Takami, T., Nielsen, U. D., Jensen, J. J., 2021. Real-time deterministic prediction of wave-induced ship responses based on short-time measurements. *Ocean Engineering*, 221, 108503. https://doi.org/10.1016/j.oceaneng.2020.108503

[7] Cosgun, T., Esenkalan, M., Kinaci, O. K., 2024. Four-quadrant propeller hydrodynamic performance mapping for improving ship motion predictions. *Brodogradnja*, 75(3), 75306. https://doi.org/10.21278/brod75306

[8] Cheng, X., Li, G., Skulstad, R., Major, P., Chen, S., Hildre, H. P., Zhang, H., 2019. Data-driven uncertainty and sensitivity analysis for ship motion modeling in offshore operations. *Ocean Engineering*, 179, 261-272. https://doi.org/10.1016/j.oceaneng.2019.03.014

[9] Halder, P., Liu, S., 2025. Numerical investigation of added resistance of a container ship in short regular waves using unsteady RANS simulations. *Brodogradnja*, 76(2), 76204. https://doi.org/10.21278/brod76204

[10] Xu, J., Gong, J., Li, Y., Fu, Z., Wang, L., 2024. Surf-riding and broaching prediction of ship sailing in regular waves by LSTM based on the data of ship motion and encounter wave. *Ocean Engineering*, 297, 117010. https://doi.org/10.1016/j.oceaneng.2024.117010

[11] Zhang, T., Zheng, X. Q., Liu, M. X., 2021. Multiscale attention-based LSTM for ship motion prediction. *Ocean Engineering*, 230, 109066. https://doi.org/10.1016/j.oceaneng.2021.109066

[12] D'Agostino, D., Serani, A., Stern, F., Diez, M., 2022. Time-series forecasting for ships maneuvering in waves via recurrent-type neural networks. *Journal of Ocean Engineering and Marine Energy*, 8(4), 479-487. https://doi.org/10.1007/s40722-022-00255-w

[13] Diez, M., Gaggero, M., Serani, A., 2024. Data-driven forecasting of ship motions in waves using machine learning and dynamic mode decomposition. *International Journal of Adaptive Control and Signal Processing* 39(10), 2119-2142. https://doi.org/10.1002/acs.3835

[14] Hoang, A. T., Bui, T. A. E., Nguyen, X. P., Bui, V. H., Nguyen, Q. C., Truong, T. H., Chung, N., 2025. Explainable machine learning-based prediction of fuel consumption in ship main engines using operational data. *Brodogradnja*, 76(4), 76405. https://doi.org/10.21278/brod76405

[15] Niako, N., Melgarejo, J. D., Maestre, G. E., Vatcheva, K. P., 2024. Effects of missing data imputation methods on univariate blood pressure time series data analysis and forecasting with ARIMA and LSTM. *BMC Medical Research Methodology*, 24(1), 320. https://doi.org/10.1186/s12874-024-02448-3

[16] Zhang, C., Ding, W., Zhang, L., 2024. Impacts of missing buoy data on LSTM-based coastal chlorophyll-a forecasting. *Water*, 16(21), 3046. https://doi.org/10.3390/w16213046

[17] Gao, J., Cai, Z., Sun, W., Jiao, Y., 2023. A novel method for imputing missing values in ship static data based on generative adversarial networks. *Journal of Marine Science and Engineering*, 11(4), 806. https://doi.org/10.3390/jmse11040806

[18] Graham, J. W., 2009. Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, 60(1), 549-576. https://doi.org/10.1146/annurev.psych.58.110405.085530

[19] Loh, W. Y., Zhang, Q., Zhang, W., Zhou, P., 2020. Missing data, imputation and regression trees. *Statistica Sinica*, 30(4), 1697-1722. https://doi.org/10.5705/ss.202019.0122

[20] Zainuddin, A., Hairuddin, M. A., Yassin, A. I. M., Abd Latiff, Z. I., Azhar, A., 2022. Time series data and recent imputation techniques for missing data: A review. *Proceedings of the International Conference on Green Energy, Computing and Sustainable Technology (GECOST)*, Virtual event, 346-350. https://doi.org/10.1109/GECOST55694.2022.10010499

[21] Festag, S., Spreckelsen, C., 2023. Medical multivariate time series imputation and forecasting based on a recurrent conditional Wasserstein GAN and attention. *Journal of Biomedical Informatics*, 139, 104320. https://doi.org/10.1016/j.jbi.2023.104320

[22] Yoon, J., Jordon, J., Schaar, M., 2018. GAIN: Missing data imputation using generative adversarial nets. *Proceedings of the Thirty-Fifth International Conference on Machine Learning (ICML)*, 10-15 July, Stockholmsmässan, Stockholm, Sweden, 5689-5698. https://doi.org/10.48550/arXiv.1806.02920

[23] Gao, H., Shen, W., Qiu, X., Xu, R., Yang, B., Hu, J., 2025. SSD-TS: Exploring the potential of linear state space models for diffusion models in time series imputation. *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 25-29 August, Toronto, Canada, 2, 649-660. https://doi.org/10.1145/3711896.3737135

[24] Cao, K., Zhang, T., Huang, J., 2024. Advanced hybrid LSTM-transformer architecture for real-time multi-task prediction in engineering systems. *Scientific Reports*, 14(1), 4890. https://doi.org/10.1038/s41598-024-55483-x

[25] Ribeiro, S. M., de Castro, C. L., 2022. Missing data in time series: A review of imputation methods and case study. *Learning and Nonlinear Models*, 20(1), 31-46. https://doi.org/10.21528/lnlm-vol20-no1-art3

[26] Zhang, M., Taimuri, G., Zhang, J., Hirdaris, S., 2023. A deep learning method for the prediction of 6-DoF ship motions in real conditions. *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment*, 237(4), 887-905. https://doi.org/10.1177/14750902231157852

[27] Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. *Neural Computation*, 9(8), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

[28] Che, Z., Purushotham, S., Cho, K., Sontag, D., Liu, Y., 2018. Recurrent neural networks for multivariate time series with missing values. *Scientific Reports*, 8(1), 6085. https://doi.org/10.1038/s41598-018-24271-9

[29] Tuikkala, J., Elo, L. L., Nevalainen, O. S., Aittokallio, T., 2008. Missing value imputation improves clustering and interpretation of gene expression microarray data. *BMC Bioinformatics*, 9(1), 202. https://doi.org/10.1186/1471-2105-9-202

[30] Dai, K., Li, Y., Gong, J., Fu, Z., Li, A., Zhang, D., 2022. Numerical study on propulsion factors in regular head and oblique waves. *Brodogradnja*, 73(1), 37-56. https://doi.org/10.21278/brod73103

[31] Liang, L., Baoji, Z., Hao, Z., Jiaye, G., Zheng, T., Shuhui, G., Yuanbiao, B., Xu, Z., 2024. Study on numerical simulation and mitigation of parametric rolling in a container ship under head waves. *Brodogradnja*, 75(3), 75305. https://doi.org/10.21278/brod75305

[32] Gong, J., Li, Y., Yan, S., Ma, Q., 2022. Numerical simulation of turn and zigzag maneuvers of trimaran in calm water and waves by a hybrid method. *Ocean Engineering*, 253, 111239. https://doi.org/10.1016/j.oceaneng.2022.111239

[33] Li, Y. B., Tang, Z. Y., Gong, J. Y., 2023. The effect of PID control scheme on the course-keeping of ship in oblique stern waves. *Brodogradnja*, 74(4), 155-178. https://doi.org/10.21278/brod74408

[34] Jong, P. D., Renilson, M. R., Walree, F. V., 2013. The broaching of a fast rescue craft in following seas. *Proceedings of the 12th International Conference on Fast Sea Transportation*, December 2-5, Amsterdam, Netherlands.

[35] Kazijevs, M., Samad, M. D., 2023. Deep imputation of missing values in time series health data: A review with benchmarking. *Journal of Biomedical Informatics*, 144, 104440. https://doi.org/10.1016/j.jbi.2023.104440