Short-term ship roll motion prediction using the encoder–decoder Bi-LSTM with teacher forcing

Shiyang Li [*], Tongtong Wang, Guoyuan Li, Robert Skulstad, Houxiang Zhang

Department of Ocean Operations and Civil Engineering, Norwegian University of Science and Technology, 6009 Ålesund, Norway

ARTICLE INFO

ABSTRACT

The safety of maritime operations has become a paramount concern with the advancement of intelligent ships. Ship stability and safety are directly impacted by roll motion, making the prediction of short-term ship roll motion pivotal for assisting navigators in making timely adjustments and averting hazardous roll conditions. However, predicting ship roll motion poses challenges due to nonlinear dynamics. This study aims to predict short-term ship roll motion by leveraging the encoder–decoder structure of Bidirectional Long Short-Term Memory Networks (Bi-LSTM) with teacher forcing. The model is accomplished by employing an encoder–decoder structure to map input sequences to output sequences of varying lengths, and employing teacher forcing to enhance the network's ability to extract information. To refine and analyze the prediction model, aspects such as the quantity of training data to guarantee model generalization, establishing apposite length relationships between input and output sequences, and assessing the model performance in various sea states are investigated. Additionally, comparative experiments assessing roll motion prediction for intervals of 10s, 30 s, 60 s, and 120 s are conducted to substantiate the necessity and effectiveness of the proposed network. The dataset originates from a commercial professional simulator developed by the Norwegian company Offshore Simulator Center AS (OSC).

1. Introduction

The installation of subsea equipment, pipe replacement, seismic streamer deployment, and other offshore operations significantly contribute to the global energy and resource supply (Major et al., 2021). Achieving sustainable and responsible development of offshore resources necessitates the implementation of effective safety management (Skulstad et al., 2021; Han et al., 2023). Monitoring the rolling state of ships is one critical aspect as excessive rolling motion would lead to severe consequences, including human injuries and damages to ships. Ship operators typically rely on onboard motion sensors to monitor the vessel responses. However, relying solely on the available information does not always ensure a safe operational window, particularly considering potential fluctuations in environmental conditions. Therefore, accurate short-term predictions of ship roll motion during operations are crucial for enhancing offshore safety. Such predictions offer valuable insights, aiding in the analysis of ship roll motion, and thereby facilitating the formulation of precautionary plans to ensure the safety of offshore operations.

Ship roll motion prediction methods are commonly categorized into two types: physics-based and data-driven approaches. Physics-based models rely on mathematical equations and fundamental principles of physics to describe interactions between ships and external forces for precise predictions (Kanazawa et al., 2023). However, constructing a reliable model is challenging due to the determination of numerous coefficients, including damping coefficients, restoring coefficients, and environmental disturbances. To obtain and analyze the roll damping coefficients, computational fluid dynamics (CFD) (Liu et al., 2021), orthogonal design and variance analysis (Gu et al., 2015), finite element method (Chen et al., 2022), and system identification (Sun et al., 2021) have been investigated.

Hou and Zou (2015) proposed a novel system identification method based on support vector regression for identifying the parameters of the nonlinear equations governing the roll motion equations of a floating production storage and offloading vessel in regular waves. Yu et al. (2019) employed a five degrees-of-freedom (DOF) nonlinear time domain model based on potential flow to predict the roll angle of a KCS container ship quantitatively. Kianejad et al. (2019) proposed a numerical simulation method based on a harmonic excitation roll motion technique to determine the roll-added mass moment of inertia using CFD simulations. Rodríguez et al. (2020) proposed a hybrid approach to estimate roll damping coefficients in waves combining

experimental results from model tests in waves with numerical simulations. Zhang et al. (2022) established a high-precision prediction method by CFD of roll damping for the trimaran vessel and verified the roll decay motion of multi-forward speeds and multi-degrees of freedom by experiment. Nevertheless, mathematical models exhibit limited generalizability, particularly in the presence of significant environmental disturbances attributed to the extensive inclusion of nonlinear dynamics in ship roll motion (Bu et al., 2019). The ship roll mathematical model is only suitable for a particular ship and cannot be adapted to other vessels. Additionally, the six DOFs are interconnected, and the impact of environmental forces on every DOF varies, making it challenging to establish a relationship (Lyu et al., 2022). Identifying roll damping typically involves computationally intensive tests, such as the calm water and seakeeping tests. However, data-driven models excel in capturing intricate nonlinear relationships and patterns directly from historical data.

Data-driven models based on neural networks have gained prominence in ship motion prediction, due to their ability to automatically discern the underlying patterns and relationships from extensive data without the need for prior knowledge and comprehensive mathematical analysis (Li et al., 2016b). Numerous researchers have investigated ship roll motion prediction based on data-driven methods, aiming to enhance safety and stabilize the vessel, such as Elman network (Li et al., 2016a), Radial Basis Function network (Bahmyari et al., 2017; Yin et al., 2018), and Recurrent Neural Network (RNN) (Su et al., 2020; Xu et al., 2021). RNNs are characterized by feedback connections between nodes, enabling the output of each node to depend not only on the current inputs but also on the past inputs (Fang et al., 2021), a design particularly well-suited for processing time series data. However, RNNs often encounter issues of vanishing or diverging gradients during backpropagation. To mitigate these issues, LSTM networks incorporate recurrent connections into memory blocks, housing memory cells capable of retaining the network's temporal states. Additionally, LSTM introduces gate structures to regulate gradient flow (Alzubaidi et al., 2021). The bidirectional long short-term memory recurrent neural network, shortened as Bi-LSTM, expanding upon the LSTM architecture, employs two RNNs functioning in opposite directions, enabling simultaneous processing of input sequences in both forward and backward directions. Wang et al. (2021a) proposed single-input single-output and multiple-input single-output ship roll prediction methods based on Bi-LSTM and studied the influence of input variables on the ship roll prediction model.

Some researchers combine the advantages of various methods to establish a hybrid model to predict ship roll motion. Wang et al. (2021b) proposed a ship roll angle prediction method based on Bi-LSTM and temporal pattern attention mechanism combined deep learning model, aiming at the problem of low accuracy of ship roll angle prediction by traditional prediction algorithms and single neural network. Wei et al. (2021) proposed a new hybrid multi-step forecasting model, including adaptive empirical wavelet transform, multi-step forecasting under the multi-input multi-output strategy of the Bi-LSTM model, and hybrid particle swarm optimization to predict ship roll motion. Wei et al. (2022) proposed an ensemble multi-step forecasting model for ship roll motion under different environmental conditions, including adaptive secondary decomposition, deep belief network under MIMO strategy, multi-objective optimization, and adaptive error correction. Zhang et al. (2023) proposed a hybrid neural network that combines LSTM and convolutional neural network in parallel to extract the nonlinear dynamic characteristics and the hydrodynamic memory information through the advantage of Convolutional Neural Network (CNN) and LSTM, respectively to predict multi-step ship roll motion in high sea states. Despite recent advancements, most researchers have only managed to achieve one-step or multi-step roll motion prediction, which is not practical in real-world maritime applications.

Current research on ship roll motion prediction primarily focuses on multi-step forecasts, commonly restricted to a maximum prediction horizon of 10 s. Nevertheless, this constraint makes the current approaches impractical for real-world maritime applications. In this study, the ship undergoes dynamic position control, facilitating various operations, making it practically applicable. The succinct nature of current predictions provides limited information, posing a potential risk to the safety of operations extending over several minutes to hours. Consequently, this paper aims to bridge this gap by investigating extended prediction horizons, providing a more thorough understanding crucial for ensuring the safety and effectiveness of maritime operations in dynamic conditions. So in this paper, we propose a data-driven sequence-to-sequence model using the encoder–decoder architecture with Bi-LSTM and employ a technique known as teacher forcing to predict short-term ship roll motion. LSTM and Bi-LSTM models typically require input and output sequences of equal lengths, making them unsuitable for sequences of varying lengths. The encoder–decoder architecture overcomes this limitation by utilizing two RNNs to process sequences with diverse lengths. Moreover, teacher forcing is employed to enhance prediction models. This technique involves providing the decoder with its past output, educating it on the forecasting imprecision, and guiding the necessary adjustments for a quicker and more effective model convergence, especially in cases of limited training data. The data used in this study were gathered from a commercial professional simulator developed by the Norwegian company Offshore Simulator Center AS (OSC, 2023). A series of experiments are conducted to optimize and analyze the model. Firstly, the inclusion of a sufficient number of cases in the training data is imperative to enhance the model's generalization capabilities, enabling it to effectively discern the underlying patterns and relationships between input and output sequences. The quantity of training data is evaluated and verified using the validation loss. Second, the periodicity of ship roll motion affects the relationship between the lengths of the input and output sequences in prediction tasks. Therefore, appropriate lengths for the input and output sequences are identified to improve the performance of the ship prediction model. Additionally, model performance is assessed across different sea states to analyze the model's generalization capabilities. The contributions of this paper are as follows:

- A sequence-to-sequence method based on the encoder–decoder Bi-LSTM with teacher forcing is proposed to predict short-term ship roll motion.
- The impact of the training data quantity, input and output sequence length, and model performance in various sea states are investigated.
- Comparative experiments are conducted to verify the necessity of our network.

The paper is organized as follows. Section 2 elaborates on the methodology, mainly the experimental process, including the Bi-LSTM model, encoder–decoder architecture, and teacher forcing. Section 3 presents the experiment results encompassing model optimization, analysis, and the comparative experiment. Then, Section 4 discusses the limitations of the proposed method. Finally, Section 5 offers a summary of the paper.

2. Methodology

This paper investigates short-term ship roll motion prediction using a Bi-LSTM encoder–decoder architecture with the integration of teacher forcing. The choice of a data-driven model stems from its aptness in handling nonlinear dynamics and intricate coupling effects, effectively capturing complex relationships and behaviors that challenge traditional conventional methods. The schematic overview of the paper is depicted in Fig. 1. Firstly, the generated data undergoes data processing, encompassing time domain sampling, input feature selection, and data normalization. Subsequently, the processed data serves as the input for our proposed network. The Bi-LSTM is a widely acknowledged and effective approach for time series prediction, particularly proficient in managing long-term dependencies within sequential
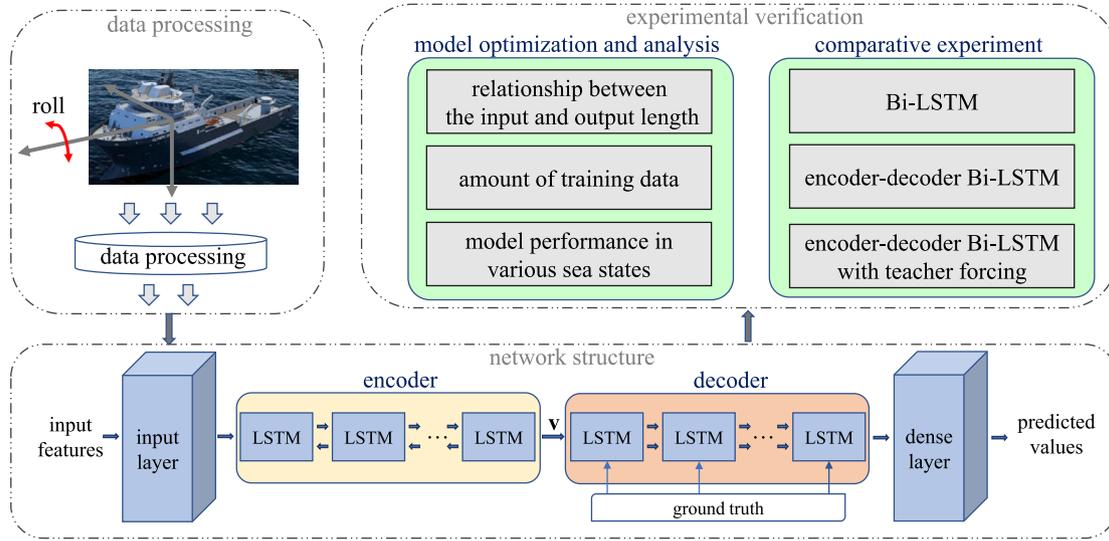
Fig. 1. The schematic of the overall paper.

data. Short-term prediction necessitates a higher emphasis on feature extraction compared to multi-step prediction scenarios. Incorporating an encoder–decoder structure not only enhances the network's feature extraction capability but also accommodates variations in input and output lengths, proving advantageous in addressing the uncertainty of the impact period. Convergence challenges faced by the network, owing to the system's nonlinear and complex dynamics, lead to the introduction of teacher forcing to facilitate network convergence during the training process. Subsequently, verification experiments are conducted, which include model optimization and comparative analyses. The investigation explores the relationship between the input and output length, the volume of train data, and model performance across various sea states, along with a comparison of the proposed network with other networks.

2.1. Bidirectional long short-term memory model

LSTM, a specialized form of RNN, is designed to capture temporal relationships, handle long-term dependencies, and efficiently preserve information in time-series data. The development of LSTM was initially proposed by Hochreiter and Schmidhuber (1997). An integral part of LSTM is its a RNN module with an integrated memory cell, fortifying the network's capability to determine the information to retain or discard while processing input sequences. Within LSTM, three gates are pivotal: the input gate, forget gate, and output gate. These gates selectively retain or discard sequence information and are fundamental to prolonging information storage within the network. The forget gate f_t facilitates the discarding of irrelevant or outdated information by calculating the element-wise multiplication of the current input x_t and the previous hidden state h_{t-1} via the *sigmoid* activation function σ, as shown in Eq. (1). Conversely, the input gate i_t decides information to reserve within the memory cell. It combines the current input x_t with the previous hidden state h_{t-1} and employs the *sigmoid* activation function σ, as shown in Eq. (2), (3). A candidate cell state $\tilde{C}_t$ is calculated by the *tanh* function and used as the proposed updating content. The updated cell state C_t is calculated by the *tanh* activation function, transforming the cell state from the previous moment C_{t-1} combined with the output $\tilde{C}_t$, as shown in Eq. (4). Finally, the output gate o_t determines information from the current memory cell state C_t that should be conveyed to the subsequent time step. It combines the current input x_t and the previous hidden state h_{t-1} while employing

the *sigmoid* activation function σ to generate the current hidden state h_t and the output o_t, as shown in Eq. (5), (6).

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

$$i_t = \sigma \left(W_i \cdot [h_{t-1}, x_t] + b_i \right) \tag{2}$$

$$\tilde{C}_t = \tanh \left(W_C \cdot [h_{t-1}, x_t] + b_C \right) \tag{3}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{4}$$

$$o_t = \sigma \left(W_o [h_{t-1}, x_t] + b_o \right) \tag{5}$$

$$h_t = o_t * \tanh \left(C_t \right) \tag{6}$$

where W_f, W_i, W_c, W_o are the weights of the forget gate, input gate, state update unit, and output gate, respectively. b_f, b_i, b_c, b_o are the biases of the forget gate, input gate, state update unit, and output gate, respectively.

Bi-LSTM represents an advancement over traditional LSTMs by processing input data in both forward and backward directions sequentially. The bidirectional approach enables the network to capture context from both past and future time steps simultaneously, resulting in better performance. Analyzing earlier and later time steps enables the network to identify dependencies more accurately and comprehend the context comprehensively. This attribute proves especially advantageous in tasks heavily reliant on information from both past and future time steps. The core equation of Bi-LSTM can be summarized to Eq. (7) at kth time step.

$$h_k = f(x_k, h_{k-1}) \tag{7}$$

where f is a non-linear activation function. By iterating these equations over time steps, the Bi-LSTM captures the temporal dynamics in the data and retains important information in the cell state while selectively outputting relevant information in the hidden state. This chain-like structure enables the Bi-LSTM to naturally capture the temporal behavior of sequences, making it suitable for tasks where temporal dependencies are important.

2.2. Encoder–decoder architecture

To enhance the understanding of the temporal context within the data and enable it to handle sequences with varying input and output lengths, an encoder–decoder architecture is incorporated into Bi-LSTM. The encoder employs the Bi-LSTM, and the decoder employs the LSTM.

The encoder–decoder architecture is a type of neural network model commonly employed for sequence-to-sequence mapping and comprises two main components: an encoder and a decoder (Cho et al., 2014; Sutskever et al., 2014). The encoder–decoder model has the ability to manage sequence-to-sequence mapping, whereby it receives a sequence of input data and produces a corresponding sequence of output data. By leveraging contextual information more effectively, it is anticipated that the model will make more accurate predictions, especially in scenarios that involve long-term dependencies or complex patterns. Additionally, the model becomes more adaptable to diverse time series data because it can handle input and output lengths that vary.

The encoder takes an input sequence $\mathbf{x} = (x_1, x_2, \ldots, x_t)$, where t represents the length of the input sequence. It processes the input sequence step by step and updates the hidden state $\mathbf{h} = (h_1, h_2, \ldots, h_t)$ at every time step using Bi-LSTM to capture the temporal information of the input sequence, as shown in Eq. (8).

$$h_k = Bi - LSTM_{enc}(x_k, h_{k-1}) \tag{8}$$

In the end, the encoder compresses the input sequence into a single vector, and the ultimate hidden state of the encoder called context vector $\mathbf{v}$ represents a compressed summary of the entire input sequence (Ghimire et al., 2022). The decoder includes autoregressive connections from the previous time step's output to the next time step's cell input and utilizes the LSTM layer to process input at each time step (Hewamalage et al., 2021). The decoder is trained to generate the predicted output sequence $\hat{\mathbf{y}} = (\hat{y}_{t+1}, \hat{y}_{t+2}, \ldots, \hat{y}_{t+p})$, where p represents the length of the output sequence, given the hidden state $h(t)$. Both $\hat{y}_t$ and $\hat{h}_t$ are also conditioned on $\hat{y}_{t-1}$ and the summary $\mathbf{v}$ of the input sequence. Hence, the hidden state of the decoder at time k is computed by Eq. (9).

$$h_k = LSTM_{dec}(h_{k-1}, \hat{y}_{k-1}, \mathbf{v}) \tag{9}$$

The output at each time step serves as the input for the succeeding time step, enabling the decoder to consecutively generate the output sequence. The decoder produces the output sequence continuously until it hits a designated end-of-sequence token or a predetermined length.

The encoder–decoder architecture is trained in a joint manner to maximize the conditional log-likelihood, as shown in Eq. (10).

$$\max_{\theta} \frac{1}{N} \sum_{n=1}^{N} log(p_{\theta}(\mathbf{y}_n | \mathbf{x}_n)) \tag{10}$$

where θ is the set of the model parameters and each $(\mathbf{x}_n, \mathbf{y}_n)$ is an (input sequence, output sequence) pair from the training set. Once the encoder–decoder model is trained, the model can generate a target sequence given an input sequence.

2.3. Teacher forcing

The ship roll motion prediction involves a regression problem, where the output corresponds to a continuous value. Teacher forcing is well-suited to regression problems due to its efficacy in enabling the model to effectively discern underlying data patterns (Kucherenko et al., 2020). In the training process, teacher forcing implies feeding the ground truth sequence from the previous time step y_{k-1} as input to the decoder model instead of using the predicted output $\hat{y}_{k-1}$. This approach ensures that the model learns from the actual output, leading to improved results over the long term. Consequently, this alteration results in a modification of Eq. (9) to Eq. (11), signifying the integration of the teacher forcing technique within the decoder equation.

$$h_k = LSTM_{dec}(h_{k-1}, y_{k-1}, \mathbf{v}) \tag{11}$$

During the inference process, ground truth data becomes unavailable when decoding unfamiliar input sequences, necessitating a modified approach. Initially, the encoder processes the input sequence to generate the context vector $\mathbf{v}$. To initialize the target sequence, the decoder

Table 1
Test ship specifications.

Specifications	Value
length between perpendiculars	82.7 m
breadth	23.058 m
draught	7.5 m
mass	1.0179×10^7 kg

then utilizes a target sequence of size one, which corresponds to the first ground truth value. To predict subsequent elements in the output sequence, the decoder relies on the previous state vectors and a one-step predicted sequence. The decoder selects and appends to the target sequence the prediction with the highest probability, determined using the *argmax* function. This process iterates, where the previously predicted value becomes input for predicting the succeeding value. Repetitively applying this iterative procedure concludes when the entire output sequence is generated. By leveraging the encoded information and its own predictions, the decoder effectively deciphers unknown input sequences.

3. Experimental results

3.1. Experiment setting

The experiment data come from a commercial professional simulation platform developed by the Norwegian company Offshore Simulator Center AS. The platform features a simulated environment in which users may manipulate the wind, waves, and ocean currents to simulate real-life conditions and offers a library of virtual vessels to choose from. Table 1 shows the ship specifications we use. A 3 DOF dynamic position controller is utilized for station-keeping. Each DOF is controlled through a single proportional–integral–derivative controller. The output of the motion controller is then connected to a basic generalized inverse-control allocator, which distributes the generalized force vector into individual commands for each thruster.

This paper investigates a stable environment characterized by constant environmental forces, including wind direction, wind velocity, wave direction, and wave height. This choice is based on the assumption that the natural environment does not typically experience sudden or frequent changes. In each scenario, the environmental forces can be characterized as random constants, while the wind and wave directions remain the same. Specifically, the wind direction ranges from 0 to 360 *degrees*, the wind velocity ranges from 0 to 13.5 m/s, the wave direction spans from 0 to 360 *degrees*, and the wave height ranges from 0 to 4 m, as shown in Fig. 2. The time interval is downsampled to 1 s. The data is transformed into a three-dimensional array, with a shape represented by (sample number, time step, feature number). The model predicts the upcoming roll angle based on historical data information. The input is the historical data with a specified input length, and the output is the subsequent roll angle with a specified output length. The network employs twelve features to extract relevant information, including roll angle, roll velocity, pitch angle, pitch velocity, yaw angle, yaw velocity, surge velocity, sway velocity, heave velocity, north position, east position, and down position. To prevent biases and enhance the performance of certain algorithms, the features are scaled and standardized before input into the network.

3.2. Model optimization and analysis

Developing a data-driven model with high accuracy typically requires meticulous adjustment of various parameters and settings. The LSTM kernel size hyperparameter is optimized using Optuna (Akiba et al., 2019). The hyperparameter candidates for the neural units in
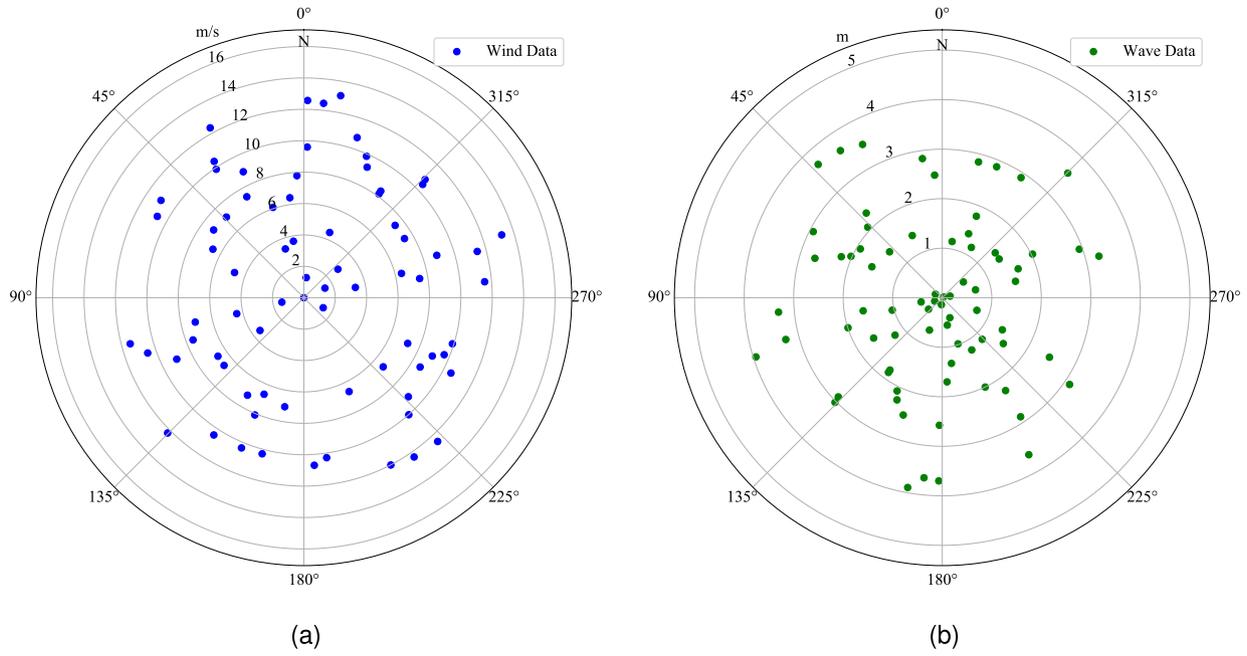
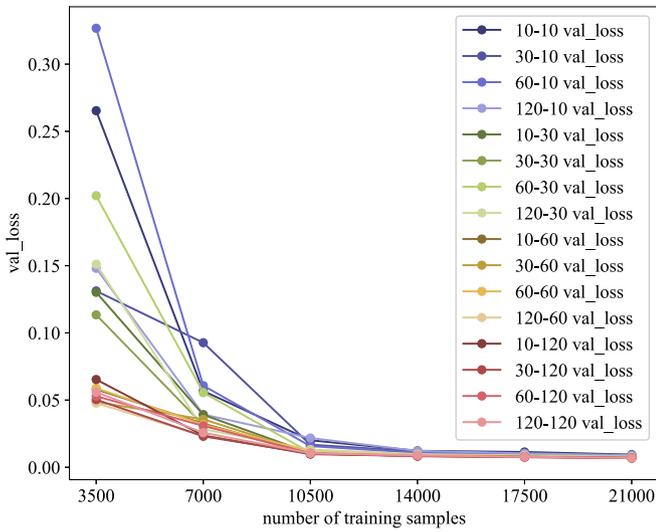**Fig. 2.** (a) Wind data in polar coordinates. (b) Wave data in polar coordinates.

Fig. 3. The validation loss influenced by the number of training samples for different input and output lengths.

the hidden layer of the LSTM range from one to 32 times 16, with the optimal parameter fixed at 256. The performance and generalization ability of the model can be significantly affected by the relationship between the input and output length, as well as the quantity of available training data. To evaluate the performance of the proposed model, the results are analyzed and evaluated by the index of the mean absolute error (MAE), as shown in Eq. (12) and the visualization of the predicted performance.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \qquad (12)$$

where y denotes the benchmark obtained by the original mathematical model, $\hat{y}$ denotes the estimated value of the parameters, and n denotes the length of time.

- **The amount of the training data**
  The performance of the model is greatly affected by the quantity of available training data. A larger dataset usually provides a more diverse range of examples, allowing the model to learn enhanced representations and patterns. By increasing the amount of data, the risk of overfitting is reduced and the model becomes better at generalizing to novel and unknown examples. However, training machine learning models entails longer and more compute-intensive training times, as well as increased resource consumption. The validation loss aids in assessing the extent to which the model generalizes. That is because the validation loss is calculated using a distinct dataset for validation, which the model has not encountered during the training process.

  In this experiment, the training and validation datasets consist of 30 cases each, with each case containing 700 samples. For each iteration, an additional 3500 samples, equivalent to the inclusion of 5 cases, are added. Six experiments were conducted, including training samples of 3500, 7000, 10,500, 14,000, 17,500, and 21,000. Fig. 3 illustrates the relationship between the validation loss and the increase in the number of training data cases for varying input and output lengths. It is evident that as the number of training data increases, the validation loss decreases. This indicates that the model's performance improves with a larger amount of training data. However, it should be noted that after exceeding 14,000 training samples, additional increments in the volume of training data do not yield a substantial decrease in the validation loss. So 17,500 training samples are selected for the experiment.

- **The relationship between the input and output length**
  Investigating the input and output length aims to investigate how prediction accuracy varies with different prediction horizons, thus aiding in understanding the ability of the model to capture long-term dependencies and make accurate long-term predictions. Determining the suitable relationship between input and output lengths can enhance the ability of the model to capture and utilize different levels of temporal information, providing insights into how the model understands past observations, and how performance varies with different amounts of historical information.

Table 2
Classification of wave level.

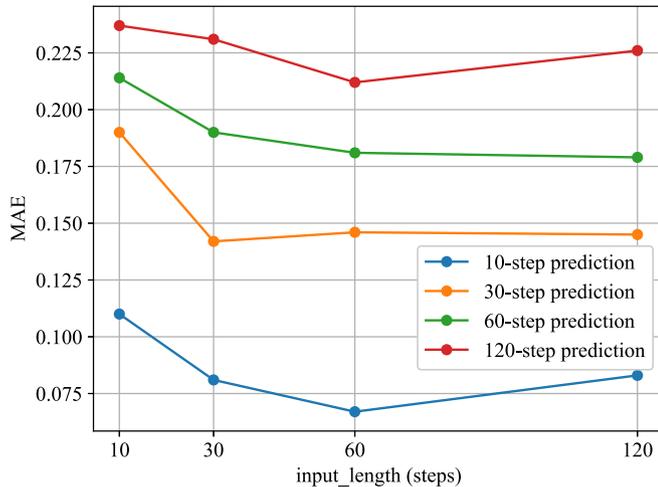| WMO sea state code	Wave height (m)	Characteristics	Character of the sea swell
1	0–0.1	calm (rippled)	low
2	0.1–0.5	smooth (wavelets)	
3	0.5–1.25	slight	
4	1.25–2.5	moderate	moderate
5	2.5–4	rough	

Fig. 4. The effect of the MAE influenced by variations in input and output lengths.

Additionally, it can reveal the trade-off between prediction accuracy and computational efficiency, assisting in finding an optimal balance that achieves accurate predictions while considering computational constraints. A total of 75 cases were generated from the simulator. Of these, 30 cases are allocated for training, 30 cases for validation, and 15 cases for testing. To mitigate overfitting during training caused by the high similarity of data within each case, diverse cases are selected as the validation set which enhances the model's capacity for generalization and prevents overfitting. The experiment employed a training dataset size of 17,500 samples for the training dataset, which is the same as the validation dataset. The test set consists of 750 samples taken from 25 test cases with a spacing of 10 steps between each sample.

The effect of the MAE influenced by variations in input and output lengths is shown in Fig. 4. Each line represents MAE for a specific output length, distinguished by different colors. It can be seen that the length of the input sequence has a direct impact on the accuracy of the prediction model. The period of ship roll motion is around 10 s, but it is not very regular due to the influence of environmental forces and the DP controller. Predicting the future roll motion for a specific time period based on just one cycle is inadequate, resulting in low prediction accuracy. This is likely due to the low similarity between successive cycles of the rolling motion despite the periodic changes. A shorter input length may not provide adequate information for the model, whereas longer input lengths have the potential to introduce irrelevant data, ultimately compromising the accuracy of the predictions. Based on the achieved highest performance, the optimal input length for the subsequent prediction models is determined to be 60 steps. While predicting for a duration of 30 s, the performance is optimal when employing a 30-step input; however, the disparity in MAE is negligible. The optimal input length for succeeding prediction models is determined to be 60 steps, attaining the highest performance.

The results demonstrate the MAE of the model concerning varied input and output time lengths, concurrently illustrating a notable

surge in error proportional to the augmentation in the number of output steps. As the prediction horizon extends, the model encounters greater complexity in capturing intricate data patterns. This complexity stems from the necessity to forecast more distant future states, entailing a higher level of uncertainty, variability, and interdependencies within the dataset. Consequently, the model's accuracy diminishes over extended forecasting periods due to the augmented difficulty in extrapolating and precisely predicting future data points. Predicting 10 steps yields more accurate results on the MAE metric compared to predicting 120 steps.

- **The generalization ability under different sea conditions**
  To study the generalization ability of proposed models, the test dataset is categorized by sea state level, according to the World Meteorological Organization (WMO) sea state code , as shown in Table 2. The wave height in our data ranges from 1 m to 4 m. The DP operation is deemed unsafe in high-sea states, making it infeasible to carry out. There are 3 cases in code 2, 3 cases in code 3, 7 cases in code 4, and 2 cases in code 5 for the sea state code. The results are depicted in a box plot, as presented in Fig. 5. Output lengths of 10, 30, 60, and 120 are displayed for different sea states, with input lengths derived from preceding results. The box plot illustrates the interquartile range (IQR) by means of a box, spanning from the first quartile (Q1) to the third quartile (Q3) of the dataset. Inside the box, a horizontal line represents the median (Q2). Extending from the box, the whiskers encompass the minimum and maximum values within a specific range, commonly 1.5 times the IQR. While disregarding outliers, the whiskers provide an overview of the data's range, portrayed as dotted lines. Any data points that exceed the whiskers are individually plotted as small circles, indicating potential extreme values or anomalies within the dataset.

  The models perform better in lower sea states, as smaller environmental forces have less impact on roll motions. As the sea conditions increase, the dispersion of MAE increases, indicating that the models become less stable, causing their predictions to be less accurate and more varied under high sea conditions. Conversely, an increase in the number of prediction steps leads to a proportional increase in the minimum value of MAE. This might indicate that predicting further into the future is more challenging for the model, resulting in higher prediction errors. The 120-step prediction model has fewer outliers compared to other models, indicating greater stability in its predictions. However, the MAE for this model is higher than that of other prediction models, suggesting that while it may be stable, it is less accurate in its predictions.

3.3. The experiment results

To assess the effectiveness of the proposed model, the proposed model undergoes comparisons with various models for short-term roll motion predictions across different sea states. The comparative models include the CNN, the CNN–Bi-LSTM, the Bi-LSTM encoder–decoder without teacher forcing, and the Bi-LSTM model, as depicted in Fig. 6, Fig. 7, and Table 3. It can be seen that our proposed model, the encoder–decoder Bi-LSTM with teacher forcing model, demonstrates
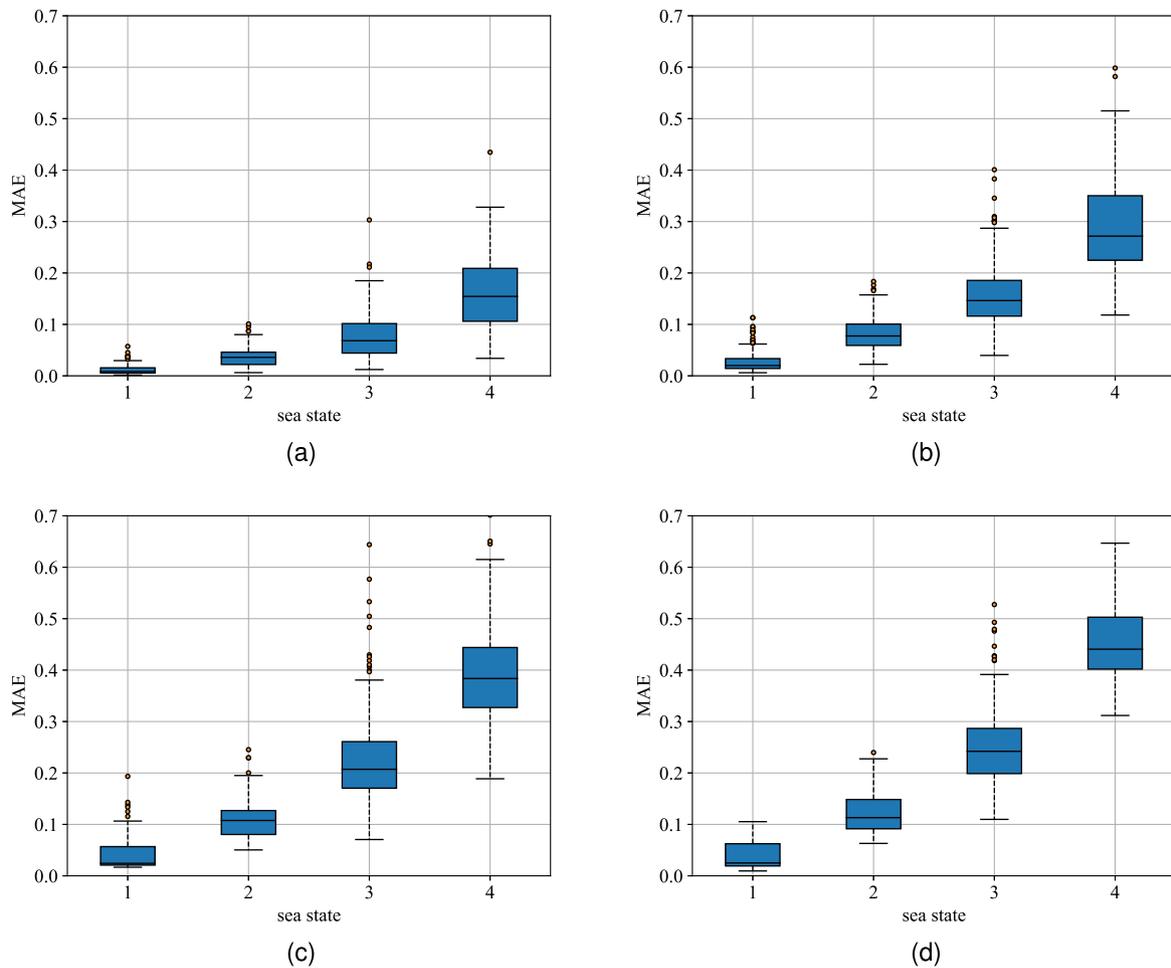
Fig. 5. The box plot of the MAE for multi-step prediction models. (a) 10-step prediction model. (b) 30-step prediction model. (c) 60-step prediction model. (d) 120-step prediction model.

Table 3
The MAE of the prediction results.

Time steps	CNN	CNN–Bi-LSTM	Encoder–decoder Bi-LSTM	Bi-LSTM	Our proposed model
10-step	0.125	0.130	0.074	0.142	0.066
30-step	0.215	0.166	0.183	0.202	0.146
60-step	0.219	0.197	0.207	0.198	0.181
120-step	0.234	0.227	0.253	0.269	0.212

optimal performance, particularly in the initial stage, yet is susceptible to cumulative error in later stages. Although there is a slight deviation in the predicted values, the overall trend remains accurate. The predicted values are impacted by cumulative errors, particularly concerning the amplitude. The comparative models are trained using the identical dataset but undergo individual parameter optimizations. Without the encoder–decoder structure, it is essential for the length of the input sequence to be equivalent to the length of the output sequence. In the case of an encoder–decoder structure, the input length is set to 60 steps for the 10-step, 30-step, 60-step, and 120-step roll predictions, which matches the configuration of our proposed model.

In low sea conditions, ships exhibit a relatively restricted range of roll. Therefore, a noticeable discrepancy arises between the predicted values and the ground truth data. The performance of the Bi-LSTM model is significantly compromised in these conditions, nearly rendering accurate predictions infeasible. Although CNN can predict 10 steps, it encounters challenges in short-term prediction. The encoder–decoder Bi-LSTM prediction results align with the trend of the real values, yet there is a considerable numerical deviation. In contrast, the CNN–Bi-LSTM model exhibits a relatively enhanced predictive capability when

compared to other models, but its accuracy falls short of our proposed model. In moderate sea conditions, the variation range of roll is more extensive than in low sea conditions, and the disparity between the model prediction results is not as conspicuous. However, it can be seen that our model performs best, followed by the encoder–decoder model and the CNN–Bi-LSTM model, with the Bi-LSTM and CNN models exhibiting the least favorable performance. This observation underscores that the encoder–decoder structure contributes to the enhanced accuracy of models, and the implementation of teacher forcing further improves their short-term prediction ability. The MAE serves as a statistical measure across different sea conditions, as shown in Table 3. Calculated by averaging the absolute differences between predicted and actual values, the MAE remains small when there is minimal fluctuation in its value. While the numerical distinction between the encoder–decoder Bi-LSTM model and our proposed model seems slight, the encoder–decoder Bi-LSTM model demonstrates the difference in comparison to our proposed model. Additionally, even a marginal improvement in such engineering applications can be immensely valuable for operators, aiding in the mitigation of potential risks and hazards. These nuanced variations can signify critical shifts in system behavior,
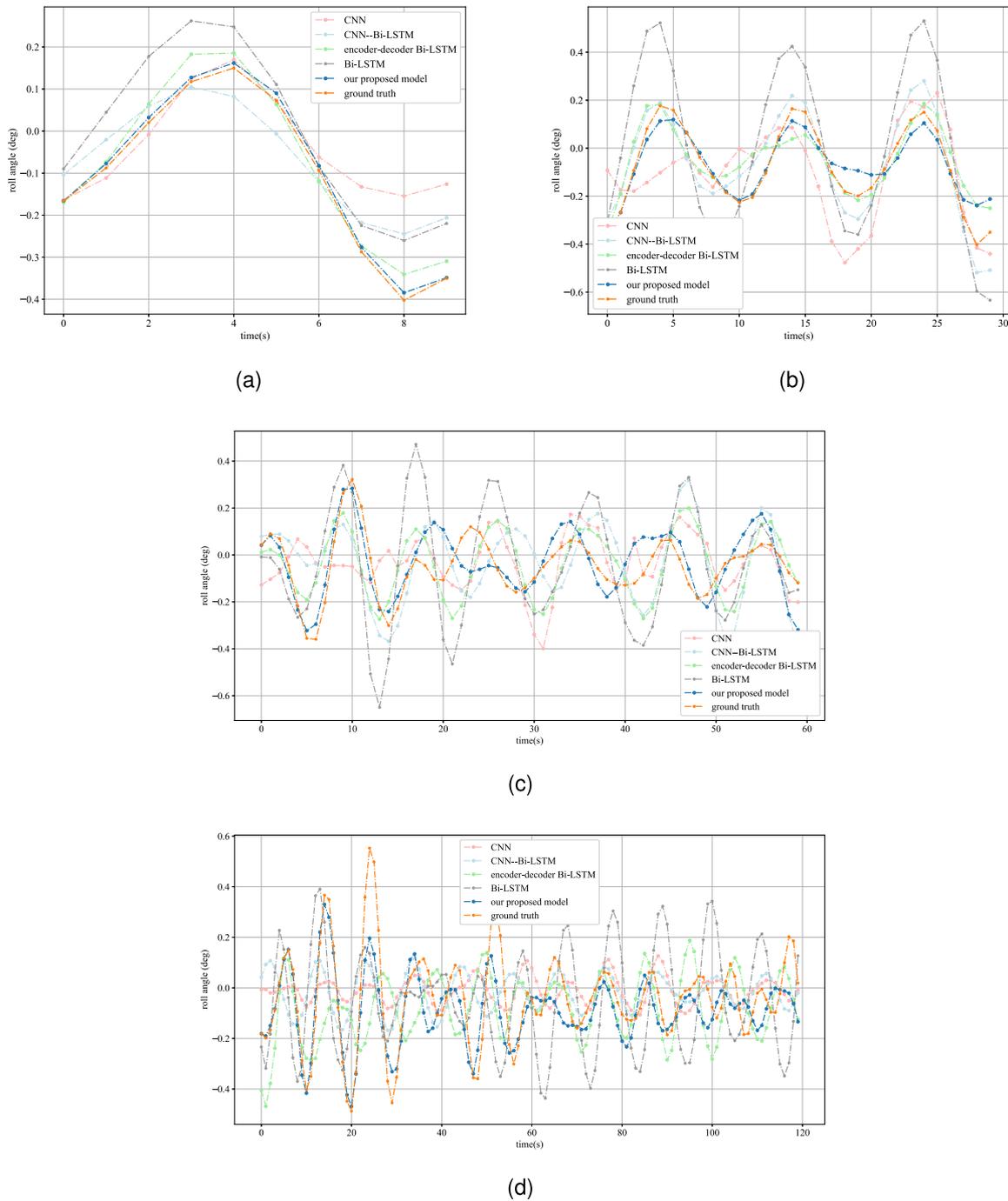
(a)

(b)

(c)

(d)
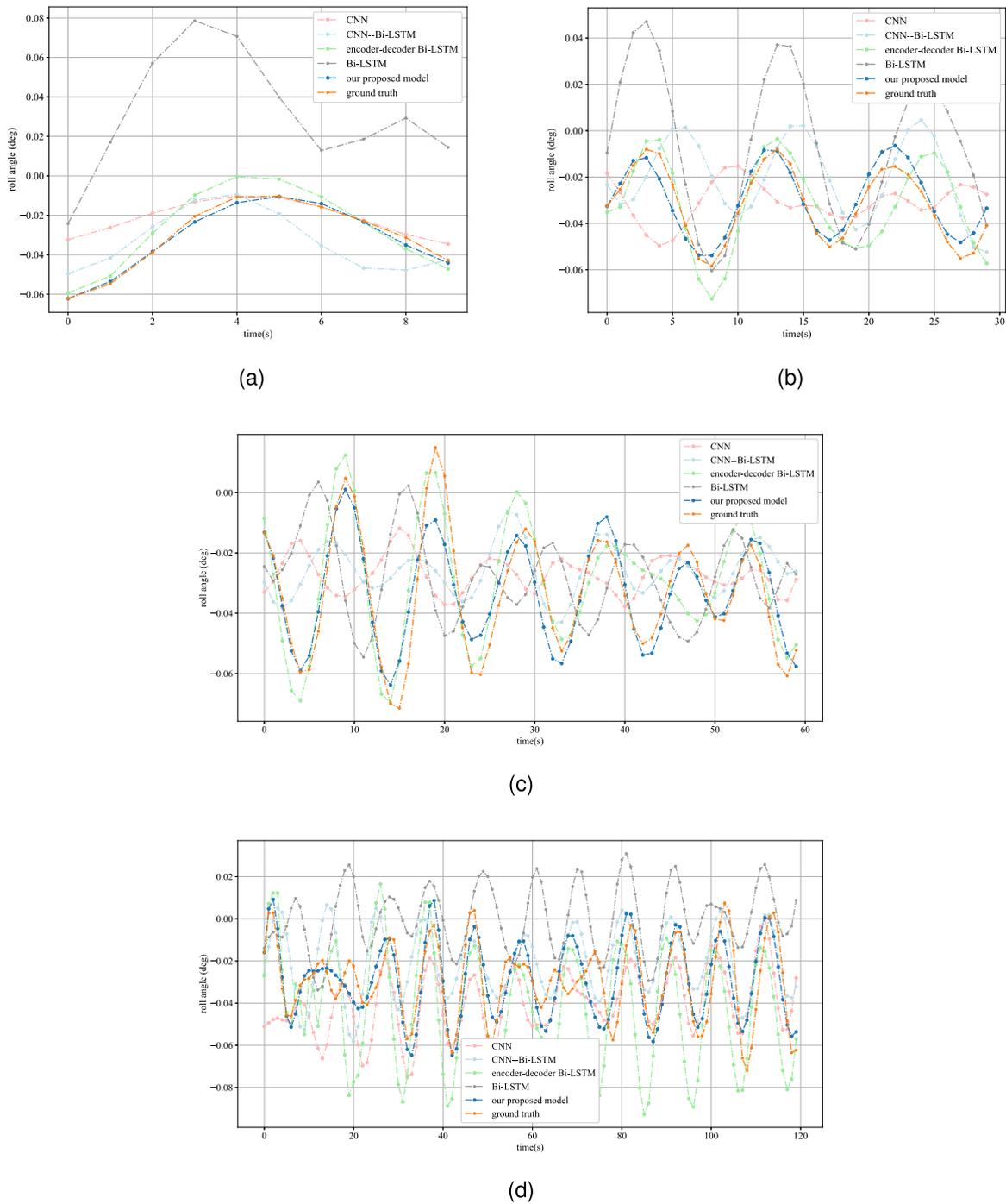
Fig. 6. The outcomes of roll prediction and comparative experiments conducted under low sea conditions. (a) The result of the 10-step ship roll prediction. (b) The result of the 30-step ship roll prediction. (c) The result of the 60-step ship roll prediction. (d) The result of 120-step ship roll prediction. The light orange line represents the results of the CNN, the light blue represents the results of the CNN–Bi-LSTM, the green line represents the results of the encoder–decoder Bi-LSTM, the gray line represents the results of the Bi-LSTM, the dark blue line represents the results of the proposed model, and the dark orange line represents the actual roll values.

particularly in situations where precise adjustments or limited oscillations are essential, indicating potential advancements or optimizations in ship stability or control under specific environmental conditions. This improvement, no matter how small, contributes to ensuring safer and more reliable ship operations in challenging maritime conditions. The persistence of this issue is evident regardless of the sea state—whether it is categorized as low or moderate. Notably, the predictions exhibit higher accuracy in the initial stages in comparison to later stages. This disparity in accuracy might potentially be linked to the

accumulation of errors over time, exerting an impact on the model's predictive performance.

4. Discussions

In this paper, the specific characteristics of the network input are impacted by the experimental setup. In this experiment, the ship operates under DP control, employing a control system or simulation to offset the forces caused by environmental factors. In other words, the

(a)

(b)

(c)

(d)

Fig. 7. The outcomes of roll prediction and comparative experiments conducted under moderate sea conditions. (a) The result of the 10-step ship roll prediction. (b) The result of the 30-step ship roll prediction. (c) The result of the 60-step ship roll prediction. (d) The result of 120-step ship roll prediction.

ship's movement adheres to predetermined rules, thereby not employing the command as input for the network. The data-driven model endeavors to discern patterns and relationships within the ship's motion data without relying on the ship's command. Throughout the entire process, the simulator maintains constant environmental forces. Nonetheless, employing these constant environmental forces as network inputs detrimentally impacts the network's performance. This issue arises due to the challenge posed when a feature term remains constant throughout the time series. The model confronts difficulty in capturing the dynamic relationship between this unvarying feature and other variables. As this feature remains constant across time steps, the model may perceive no correlation between it and the output, leading to potential inaccuracies in predicting the target. Therefore, the

input needs to be modified when applying the ship roll motion model to different scenarios. Additionally, future research should address feature selection and weighting. It is crucial to evaluate the significance of input data and whether implementing weighting mechanisms is advantageous.

While our method exhibits promise in the specific context of ship roll prediction, it is essential to recognize its broad applicability in time series prediction. Its applications can be extended beyond the specific domain of ship roll prediction encompassing areas such as financial data analysis and weather forecasting. In essence, the approach's flexibility stems from the nature of the neural network itself, which prioritizes analyzing data characteristics rather than being explicitly tailored to specific data types. While a theoretical foundation

supports the method's potential applicability in various fields, empirical validation is imperative to substantiate its efficacy in specific contexts.

5. Conclusions

The paper introduces a data-driven ship roll motion prediction model based on the Bi-LSTM encoder–decoder architecture with teacher forcing, which facilitates precise short-term ship roll motion prediction. To enhance the efficiency and accuracy of the model, the appropriate amount of training data is allocated, and the input and output lengths are adjusted through the encoder–decoder structure. Shorter input lengths lack sufficient features for accurate prediction, while longer lengths include irrelevant information that impedes prediction accuracy. The study assesses the model's performance across various sea states and conducts comparative experiments to identify the efficiency of the proposed model. Future research endeavors should concentrate on enhancing the prediction accuracy of the model for longer time periods, particularly in the high-sea states.

CRediT authorship contribution statement

Shiyang Li: Writing – original draft, Methodology, Investigation, Conceptualization. **Tongtong Wang:** Writing – review & editing, Supervision, Conceptualization. **Guoyuan Li:** Writing – review & editing, Supervision. **Robert Skulstad:** Writing – review & editing, Data curation. **Houxiang Zhang:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

References

Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M., 2019. Optuna: A next-generation hyperparameter optimization framework. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. pp. 2623–2631.

Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M., Farhan, L., 2021. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. J. Big Data 8, 1–74.

Bahmyari, E., Khedmati, M.R., Soares, C.G., 2017. Stochastic analysis of coupled heave-roll ship motion using the domain decomposition chaotic radial basis function. Ocean Eng. 140, 322–333. http://dx.doi.org/10.1016/j.oceaneng.2017.05.033, URL: https://www.sciencedirect.com/science/article/pii/S0029801817302925.

Bu, S.-X., Gu, M., Lu, J., Abdel-Maksoud, M., 2019. Effects of radiation and diffraction forces on the prediction of parametric roll. Ocean Eng. 175, 262–272. http://dx.doi.org/10.1016/j.oceaneng.2019.02.006, URL: https://www.sciencedirect.com/science/article/pii/S0029801818313568.

Chen, J., Yang, J., Shen, K., Zheng, Z., Chang, Z., 2022. Stochastic dynamic analysis of rolling ship in random wave condition by using finite element method. Ocean Eng. 250, 110973.

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y., 2014. Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078.

Fang, W., Chen, Y., Xue, Q., 2021. Survey on research of RNN-based spatio-temporal sequence prediction algorithms. J. Big Data 3 (3), 97–110. http://dx.doi.org/10.32604/jbd.2021.016993.

Ghimire, S., Deo, R.C., Wang, H., Al-Musaylh, M.S., Casillas-Pérez, D., Salcedo-Sanz, S., 2022. Stacked LSTM sequence-to-sequence autoencoder with feature selection for daily solar radiation prediction: A review and new modeling results. Energies 15 (3), 1061.

Gu, M., Lu, J., Bu, S., Wu, C., Qiu, G., 2015. Numerical simulation of the ship roll damping. In: Proceedings of STAB. pp. 341–348.

Han, P., Zhu, M., Zhang, H., 2023. Interaction-aware short-term marine vessel trajectory prediction with deep generative models. IEEE Trans. Ind. Inform. 1–9. http://dx.doi.org/10.1109/TII.2023.3302304.

Hewamalage, H., Bergmeir, C., Bandara, K., 2021. Recurrent neural networks for time series forecasting: Current status and future directions. Int. J. Forecast. 37 (1), 388–427. http://dx.doi.org/10.1016/j.ijforecast.2020.06.008, URL: https://www.sciencedirect.com/science/article/pii/S0169207020300996.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780.

Hou, X.-R., Zou, Z.-J., 2015. SVR-based identification of nonlinear roll motion equation for FPSOs in regular waves. Ocean Eng. 109, 531–538. http://dx.doi.org/10.1016/j.oceaneng.2015.08.068, URL: https://www.sciencedirect.com/science/article/pii/S0029801815004758.

Kanazawa, M., Wang, T., Skulstad, R., Li, G., Zhang, H., 2023. Physics-data cooperative ship motion prediction with onboard wave radar for safe operations. In: 2023 IEEE 32nd International Symposium on Industrial Electronics. ISIE, pp. 1–8. http://dx.doi.org/10.1109/ISIE51358.2023.10228113.

Kianejad, S.S., Enshaei, H., Duffy, J., Ansarifard, N., 2019. Prediction of a ship roll added mass moment of inertia using numerical simulation. Ocean Eng. 173, 77–89. http://dx.doi.org/10.1016/j.oceaneng.2018.12.049, URL: https://www.sciencedirect.com/science/article/pii/S002980181831429X.

Kucherenko, T., Jonell, P., Van Waveren, S., Henter, G.E., Alexandersson, S., Leite, I., Kjellström, H., 2020. Gesticulator: A framework for semantically-aware speech-driven gesture generation. In: Proceedings of the 2020 International Conference on Multimodal Interaction. pp. 242–250.

Li, Z., Dong, J., Song, Z., Zhang, H., Guo, Y., 2016a. Prediction of ship roll motion based on combination of phase space reconstruction theory and Elman network. In: 2016 IEEE International Conference on Information and Automation. ICIA, IEEE, pp. 686–689.

Li, G., Zhang, H., Kawan, B., Wang, H., Osen, O.L., Styve, A., 2016b. Analysis and modeling of sensor data for ship motion prediction. In: OCEANS 2016-Shanghai. IEEE, pp. 1–7.

Liu, L., Chen, M., Wang, X., Zhang, Z., Yu, J., Feng, D., 2021. CFD prediction of full-scale ship parametric roll in head wave. Ocean Eng. 233, 109180. http://dx.doi.org/10.1016/j.oceaneng.2021.109180, URL: https://www.sciencedirect.com/science/article/pii/S0029801821006132.

Lyu, W., el Moctar, O., Schellin, T.E., 2022. Ship motion-sloshing interaction with forward speed in oblique waves. Ocean Eng. 250, 110999. http://dx.doi.org/10.1016/j.oceaneng.2022.110999, URL: https://www.sciencedirect.com/science/article/pii/S0029801822004231.

Major, P., Zhang, H., Petter Hildre, H., Edet, M., 2021. Virtual prototyping of offshore operations: A review. Ship Technol. Res. 68 (2), 84–101. http://dx.doi.org/10.1080/09377255.2020.1831840.

OSC, 2023. Offshore Simulation Center, URL: https://osc.no/.

Rodríguez, C.A., Ramos, I.S., Esperança, P.T.T., Oliveira, M.C., 2020. Realistic estimation of roll damping coefficients in waves based on model tests and numerical simulations. Ocean Eng. 213, 107664.

Skulstad, R., Li, G., Fossen, T.I., Vik, B., Zhang, H., 2021. A hybrid approach to motion prediction for ship docking—Integration of a neural network model into the ship dynamic model. IEEE Trans. Instrum. Meas. 70, 1–11. http://dx.doi.org/10.1109/TIM.2020.3018568.

Su, Y., Lin, J., Zhao, D., Guo, C., Wang, C., Guo, H., 2020. Real-time prediction of large-scale ship model vertical acceleration based on recurrent neural network. J. Mar. Sci. Eng. 8 (10), 777.

Sun, J., Hu, S.-L.J., Li, H., 2021. Nonlinear roll damping parameter identification using free-decay data. Ocean Eng. 219, 108425.

Sutskever, I., Vinyals, O., Le, Q.V., 2014. Sequence to sequence learning with neural networks. In: Advances in Neural Information Processing Systems, vol. 27.

Wang, Y., Wang, H., Zhou, B., Fu, H., 2021a. Multi-dimensional prediction method based on Bi-LSTMC for ship roll. Ocean Eng. 242, 110106.

Wang, Y., Wang, H., Zou, D., Fu, H., 2021b. Ship roll prediction algorithm based on Bi-LSTM-TPA combined model. J. Mar. Sci. Eng. 9 (4), http://dx.doi.org/10.3390/jmse9040387.

Wei, Y., Chen, Z., Zhao, C., Tu, Y., Chen, X., Yang, R., 2021. A BiLSTM hybrid model for ship roll multi-step forecasting based on decomposition and hyper-parameter optimization. Ocean Eng. 242, 110138. http://dx.doi.org/10.1016/j.oceaneng.2021.110138, URL: https://www.sciencedirect.com/science/article/pii/S0029801821014591.

Wei, Y., Chen, Z., Zhao, C., Tu, Y., Chen, X., Yang, R., 2022. An ensemble multi-step forecasting model for ship roll motion under different external conditions: A case study on the South China Sea. Measurement 201, 111679. http://dx.doi.org/10.1016/j.measurement.2022.111679, URL: https://www.sciencedirect.com/science/article/pii/S0263224122008879.

Xu, W., Maki, K.J., Silva, K.M., 2021. A data-driven model for nonlinear marine dynamics. Ocean Eng. 236, 109469. http://dx.doi.org/10.1016/j.oceaneng.2021.109469, URL: https://www.sciencedirect.com/science/article/pii/S0029801821008726.

Yin, J.-C., Perakis, A.N., Wang, N., 2018. A real-time ship roll motion prediction using wavelet transform and variable RBF network. Ocean Eng. 160, 10–19. http://dx.doi.org/10.1016/j.oceaneng.2018.04.058, URL: https://www.sciencedirect.com/science/article/pii/S0029801818305523.

Yu, L., Ma, N., Wang, S., 2019. Parametric roll prediction of the KCS containership in head waves with emphasis on the roll damping and nonlinear restoring moment. Ocean Eng. 188, 106298. http://dx.doi.org/10.1016/j.oceaneng.2019.106298, URL: https://www.sciencedirect.com/science/article/pii/S002980181930469X.

Zhang, Y., Wang, P., Yu, P., Hu, J., 2022. Numerical and experimental study on nonlinear roll damping characteristics of trimaran vessel. Ocean Eng. 247, 110778.

Zhang, D., Zhou, X., Wang, Z.-H., Peng, Y., Xie, S.-R., 2023. A data driven method for multi-step prediction of ship roll motion in high sea states. Ocean Eng. 276, 114230. http://dx.doi.org/10.1016/j.oceaneng.2023.114230, URL: https://www.sciencedirect.com/science/article/pii/S0029801823006145.