

# Probabilistic Interval Prediction of Ship Roll Motion Using Multi-Resolution Decomposition and Non-Parametric Kernel Density Estimation

Dongxing Xu<sup>1,2,3</sup> and Jianchuan Yin<sup>1,2,3</sup>

Received: 07 March 2025 / Accepted: 17 April 2025

© Harbin Engineering University and Springer-Verlag GmbH Germany, part of Springer Nature 2026

## Abstract

The prediction of ship roll motion is important for the safety of marine vessels and the efficiency of offshore operations. This study aims to address the issues of low prediction accuracy and incapability to illustrate the uncertainty of ship roll motion in traditional deterministic prediction models. Thus, a probabilistic interval prediction scheme for ship roll motion is established utilizing the novel data preprocessing strategy and non-parametric kernel density estimation (KDE). The novel data preprocessing method based on multi-resolution decomposition and reconstruction strategy is used to reduce the non-stationarity of the time series and number of mode components. This method is also utilized to extract trends from the mode components for preserving more effective information. Accordingly, the dynamic recurrent neural network is adopted to enable the identification and deterministic prediction of reconstruction data for ship roll motion. The deterministic prediction values of ship motion status can be obtained by reconstructing the predicted outcomes of the mode components. Non-parametric KDE is applied to predict the fluctuation range of ship roll status by combining deterministic prediction results. The feasibility and effectiveness of the probability prediction scheme are exhibited through the measured data of M.V. *YuKun*. Results show that the deterministic prediction accuracy of the proposed scheme is superior to those of traditional deterministic prediction models, with maximum improvements of 22.7% and 27.5% for single-and multi-step prediction accuracies, respectively. The minimum prediction interval coverage probability at the 85% confidence level is 98%. The reliable fluctuation range of ship roll can be realized through the non-parametric KDE, which can offer more effective reference information for ship operators. The proposed model not only satisfies the requirements of deterministic prediction accuracy but also produces reliable uncertainty estimates of ship roll motion status.

**Keywords** Ship roll motion probabilistic prediction; Data preprocessing strategy; Neural network; Non-parametric kernel density estimation

## 1 Introduction

Ships at sea will undergo six degrees of freedom (6-DOF)

### Article Highlights

- A novel data preprocessing scheme is proposed to minimize nonstationary features of the ship status data by integrating the ensemble empirical mode decomposition, energy entropy reconstruction strategy, and singular spectrum analysis.
- The dynamic recurrent neural network is utilized to reveal the characteristics of ship roll status with nonlinearity, uncertainty, and time-varying dynamic.
- The non-parametric kernel density estimation based on optimal bandwidth is used to efficiently express the fluctuation range of ship roll angle.

✉ Jianchuan Yin  
yinjianchuan@gdou.edu.cn

<sup>1</sup> Naval Architecture and Shipping College, Guangdong Ocean University, Zhanjiang 524005, China

<sup>2</sup> Guangdong Provincial Engineering Research Center for Ship Intelligence and Safety, Zhanjiang 524005, China

<sup>3</sup> Guangdong Provincial Key Laboratory of Intelligent Equipment for South China Sea Marine Ranching, Zhanjiang 524088, China

in swaying motion as a result of wind and waves (Fossen, 2011). When confronting severe weather, ships will encounter violent roll, pitch, and heave movements, which can easily cause the ship to capsize and lead to significant economic losses (Wang et al., 2021). Highly accurate prediction of ship motion status can help the ship control platform efficiently enhance ship stability through compensation and correction mechanisms (Zhou et al., 2021). It can further guarantee the safety of the crew members and increase the efficiency of offshore operations (Zhang et al., 2021). Therefore, the prediction of ship motion status has always been a major topic in the field of ship motion control.

The model for ship status prediction based on analytical expressions is difficult to solve due to the influence of random waves (Chung et al., 1990). Therefore, scholars have converted the problem of ship status prediction into a time series of ship motion. Autoregressive recursive (AR) model and Kalman filtering theory have been applied for time series prediction of ship motion status (Huang et al., 2014; Jiang et al., 2020; Lin et al., 2011; Peng et al., 2001; Reis et al., 2023; Song et al., 2022). Given the nonlinear and nonstationary characteristics of ship motion status, scholars have shifted their focus from recursive models to

the field of machine learning. Data-driven modeling schemes based on the measured data of systems can address the incapability to establish efficient mathematical models in practical engineering, which cannot be resolved using traditional methods. Artificial neural networks have nonlinear mapping abilities and have been extensively utilized in addressing problems in ship motion status prediction; examples include the radial basis function (RBF) neural network (Yin et al., 2013) and the nonlinear autoregressive exogenous neural network model (Li et al., 2017). Deep learning models can be integrated with the attention mechanism to enhance the prediction accuracy of the model (Fu et al., 2022; Mao et al., 2025; Wang et al., 2024), which can help further investigate the spatial and temporal characteristics of ship status. Improper setting of the structure and parameters of artificial neural networks may result in local optima when dealing with complex problems. The generalization capability of the network model will also be weakened. The development of artificial intelligence has facilitated the use of swarm intelligence optimization algorithms in optimizing the structure and parameters of neural networks, which can boost the performance of neural network models (Luo et al., 2020; Tarkhaneh and Shen, 2019). Therefore, swarm intelligence optimization algorithms can also be adopted in the field of ship motion status prediction. The neural network model trained by intelligent optimization algorithms can considerably enhance the accuracy of ship motion status prediction (Ding et al., 2023; Geng et al., 2023; Peng et al., 2019; Wei et al., 2021; Yuan et al., 2025; Zhang et al., 2023; Zhang et al., 2020a; Zhang et al., 2020b). In addition, the time series of ship motion status exhibits chaotic and grey characteristics. Grey system and chaos theories can be used to construct prediction models for ship motion status, and satisfactory prediction outcomes have been obtained (Li et al., 2016; Yin et al., 2014).

According to forecasting theory, hybrid prediction can mine the effective information of multiple single-prediction models to find effective combination modes and completely leverage the advantages of single models for high-precision prediction. Therefore, combination models have been proposed to boost the accuracy of ship motion status prediction; examples include the convolutional neural network and bidirectional long short-term memory model (Wei et al., 2023a), the deep learning and autoregressive integrated moving average (ARIMA) model (Suhermi et al., 2018), the extreme learning machine and grey model (Yin et al., 2014), and the Gaussian process regression and long short-term memory model (Sun et al., 2022). The results showed that the prediction accuracy of these models has been improved.

The development of data mining technology has enabled fusion data decomposition methods to effectively solve major issues in different fields, such as signal denoising (Zhang et al., 2018), wind power prediction (An et al., 2012; Mi

et al., 2019; Yu et al., 2017), and financial market analysis (Li et al., 2020). The preprocessing of ship motion status data to reduce the nonstationary characteristic is specifically crucial for advancing the accuracy of model prediction. Noise interference typically occurs in the time series of ship motion status due to the influence of the external environment and ship data acquisition devices in the field of ship motion control. Therefore, data preprocessing techniques such as wavelet decomposition and empirical mode decomposition (EMD) have also been widely applied to models for ship motion status prediction to produce satisfactory results (Duan et al., 2015; Gao et al., 2023; Li et al., 2022b; Xia et al., 2022; Yin et al., 2018; Zhang et al., 2020a). However, the traditional deterministic prediction models for ship status have their applicable conditions and drawbacks. Linear recursive models and Kalman filtering algorithms are appropriate for stationary random processes, but they have difficulty in accurately describing the nonstationary characteristics of ship motion status. Moreover, improper parameter settings of intelligent optimization algorithms can directly influence the reliability of prediction results for neural network models. Deep learning models need large ship motion datasets for learning, which can consume more computational resources (Wei et al., 2023b). Furthermore, the grey prediction model cannot easily reflect the patterns of ship sway. Wavelet decomposition and EMD techniques (Huang et al., 1998) have problems such as weak adaptation and mode mixing, respectively. In addition, the current deterministic prediction model based on decomposition technology not only needs to predict numerous mode components but also ignores the dominant role of noise in each mode component and the effective information in the residual component.

Traditional deterministic prediction models can offer valuable reference information for ship operators. However, the incapacity to determine the error information contained in the prediction results introduces certain risks for ship operators when making decisions, which limits the practical application of ship motion prediction models. The probability prediction model not only considers the errors caused by the time-varying characteristics of ship roll motion but also delivers clearer and more reliable reference information for ship operators by providing fluctuation intervals at different confidence levels. Non-parametric kernel density estimation (KDE) is a non-parametric testing method in probability theory; it estimates unknown density functions, which can offer reliable estimation intervals in practical engineering applications (Gramacki, 2018; Zhou et al., 2019; Li et al., 2022a). The non-parametric KDE method does not require the observation of the prior distribution of ship roll data in estimating the actual distribution of ship motion status. Therefore, a novel probability prediction hybrid model (HPM) is developed for ship motion status prediction to improve prediction accuracy and properly describe the

uncertainty of ship roll status with time-varying dynamics. This model integrates the novel data preprocessing technology, recurrent neural network, and non-parametric KDE. The time series of ship motion status can be decomposed and reconstructed based on novel data preprocessing technology. The reconstructed mode components are predicted individually using a dynamic recurrent neural network. Then, the deterministic predicted outcomes of each mode component are reconstructed to gain the expected values of ship motion status. Non-parametric KDE is used to estimate the distribution range of prediction errors in ship roll angle, which can generate a reliable fluctuation range of ship roll angle. To validate the feasibility and effectiveness of the proposed probability prediction model, ship roll prediction experiments have been conducted using the measured data of sea trials of M.V. *YuKun*.

The remainder of the paper is organized as follows. The structure of the probability prediction model is illustrated in Section 2. The key technologies of probability prediction scheme for ship roll motion are presented in Section 3. Section 4 presents the simulation experiments and discussions for ship roll motion prediction. Conclusions are derived in Section 5.

## 2 Structure of the probability prediction model

The time series of ship motion status exhibits strong nonlinearity, randomness, and non-stationarity under complex external environments. A novel probability hybrid prediction model is proposed in this study to comprehensively analyze the characteristics of ship motion status. Three key methodologies, namely, the novel data preprocessing scheme, the dynamic recurrent neural network, and non-parametric KDE, are applied in the proposed model. A novel data preprocessing scheme is proposed to reduce the nonstationary features of ship status data by integrating the ensemble EMD (EEMD), energy entropy reconstruction strategy, and singular spectrum analysis (SSA). EEMD is utilized to reduce the nonstationary features of time series data and facilitate identification and prediction using the dynamic recurrent neural network. Moreover, the energy entropy reconstruction strategy is applied to adaptively find the reconstruction boundaries of mode components based on the characteristics of energy distribution and decrease the number of mode components and model complexity for ship status prediction. Furthermore, SSA is adopted to extract trends from high-frequency reconstructed mode components for retaining more effective information regarding mode components. The non-parametric KDE based on optimal bandwidth is also employed to address the incapacity of traditional deterministic prediction methods to accurately express the fluctuation range of roll angle, which

can provide more reliable decision-making information for ship operators. The probabilistic interval prediction model for ship motion status is presented in Figure 1.

Ship roll motion status can be acquired by shipborne data acquisition equipment. First, EEMD is utilized to stabilize the time series of ship motion status and decompose it into a series of intrinsic mode functions (IMFs) to reduce the complexity of ship motion status. Then, the energy entropy method is employed to adaptively choose and reconstruct high-frequency mode components for easing the complexity of the prediction model. The trend of the reconstructed high-frequency mode components is extracted using the SSA algorithm. The Elman network prediction model is developed for the reconstructed mode components. The deterministic prediction values of ship roll motion status can be obtained by reconstructing the forecast results of each mode component. Finally, the error distribution range of the test set is estimated using a non-parametric KDE model with the optimal bandwidth value through the error distribution of the training set data. The deterministic prediction results and error estimation are fused to generate the prediction range of the ship roll angle.

## 3 Key methodology in the proposed model

### 3.1 Mode component reconstruction-based energy entropy

The value of energy entropy indicates the uncertainty and complexity of signals and can characterize the distribution of signal energy; this method has been extensively utilized in fault diagnosis and signal denoising (Yang et al., 2006; Zhang et al., 2018). The energy entropy method is adopted to reconstruct the high-frequency mode components decomposed by EEMD (Wu and Huang, 2009), which helps minimize the complexity of the model prediction component and enhance the prediction accuracy. The EEMD method can overcome the problems of manual selection of appropriate wavelet basis functions and decomposition orders in the wavelet transform, along with mode mixing and endpoint effects existing in EMD. Therefore, the sub-series of ship motion status decomposed by EEMD has stronger regularity than the original time series with nonlinear, randomness, and time-varying dynamics. The EEMD method can decrease the non-stationarity of ship motion time series and increase prediction accuracy. In the  $K$  mode components from high frequencies to low frequencies obtained by EEMD, the dominant effect of noise on each component gradually decreases. Therefore, the boundary point for the reconstruction of mode components is determined as follows:

- 1) The energy entropy of IMFs is calculated.
- 2) The energy entropy of IMFs is traversed. When the local maximum value is taken, the corresponding IMF <sub>$i$</sub>  is

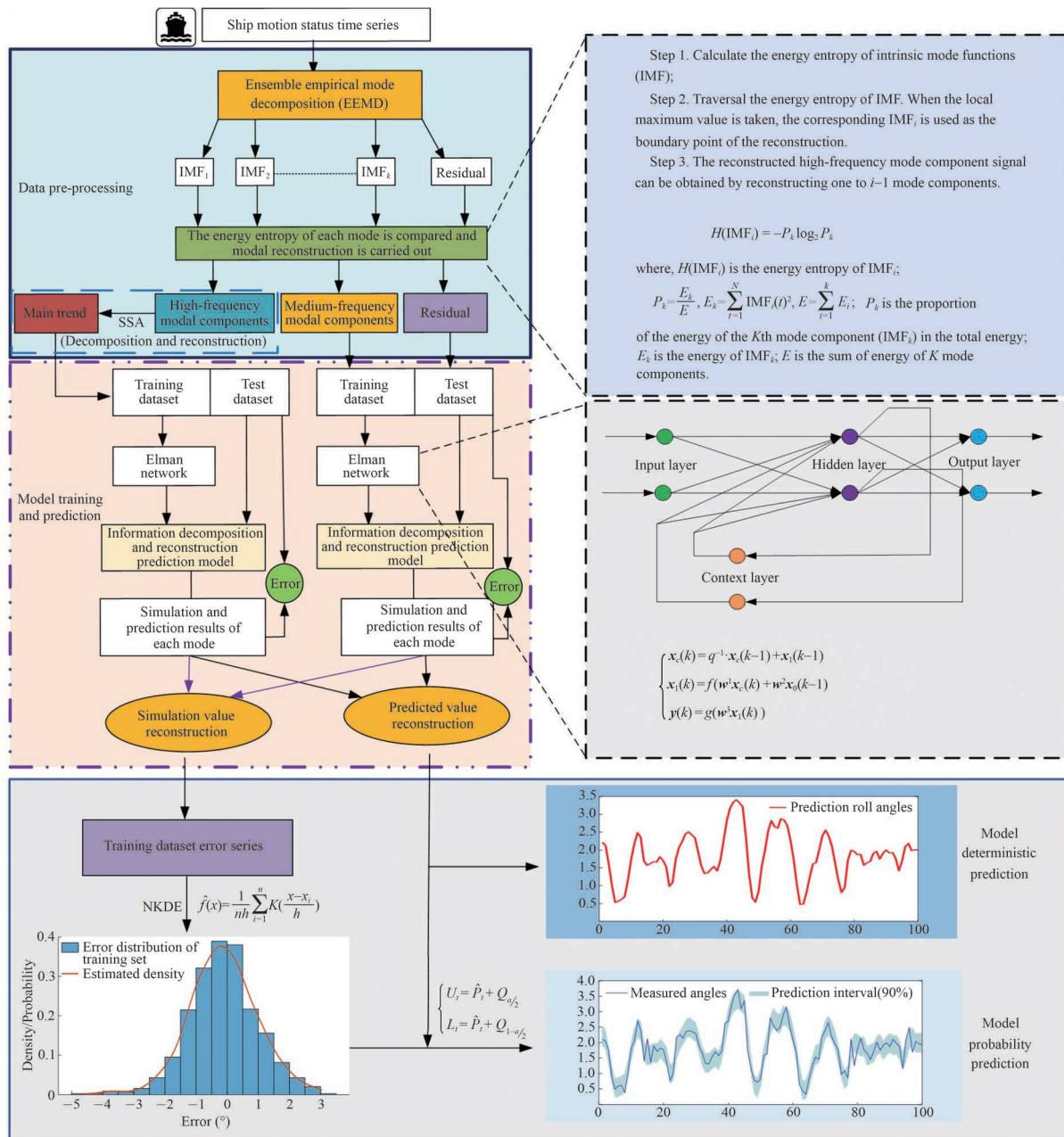


**Figure 1** Structure of the probability prediction model for ship motion status

employed as the boundary point of the reconstruction.

3) The reconstructed high-frequency mode component signal can be acquired by reconstructing one to  $i-1$  mode components.

$$H(\text{IMF}_i) = -P_k \log_2 P_k \quad (1)$$

where  $H(\text{IMF}_i)$  is the energy entropy of IMF<sub>i</sub>;  $P_k = \frac{E_k}{E}$ ,  $E_k = \sum_{t=1}^N \text{IMF}_i(t)^2$ ,  $E = \sum_{i=1}^k E_i$ ;  $P_k$  is the proportion of the energy of the  $K$ th mode component (IMF<sub>k</sub>) in the total energy;  $E_k$  is the energy of IMF<sub>k</sub>;  $E$  is the sum of energy of  $K$  mode components.

The measured ship motion status data are commonly influenced by noise from recording systems and external environments when the ship sails at sea. After the ship motion status data are decomposed by EEMD, the mode components contain not only the effective signal details but also the noise components. In addition, the high-frequency mode components of ship motion status are usually covered by noise. Therefore, detailed information needs to be further extracted and retained to achieve the ideal decomposition effect. SSA is an efficient approach for time series analysis and prediction; this technique is primarily used for trend extraction, detection, and denoising of time series with potential structure (Moreno and dos Santos Coelho,

2018; Afshar and Bigdeli, 2011).

### 3.2 Recurrent neural network

Artificial neural networks exhibit nonlinear mapping capability and do not require precise mathematical models. They can solve certain problems that are difficult to handle for traditional methods. According to the flow of information during the operation of neural networks, they can be classified into two basic types: feedforward and feedback. Feedforward neural networks can achieve complex nonlinear mapping functions by using hidden layers and nonlinear transfer functions. However, the output of a feedforward neural network is ascertained only by the current input and weight matrix and is independent of the previous output state of the network (Han and Shi, 2023). The input of the feedback neural network includes feedback with delayed input or output data. Thus, it is considered a feedback dynamics system that directly reflects the dynamic process characteristics of the system. The feedback neural network can be applied to reflect the time-varying dynamic characteristic of the ship roll motion more efficiently. The Elman neural network is a typical dynamic recurrent neural network. Therefore, the Elman network can be used to indicate the characteristics of ship roll status with nonlinearity, uncertainty, and time-varying dynamics, as illustrated in Figure 2. The network structure comprises four parts: input layer, hidden layer, output layer, and context layer. The context layer realizes the feedback connection between layers or within layers, which can be considered a one-step delay operator; this way enables the network to store internal states for realizing dynamic feature mapping, as well as boosts the ability of the network to process dynamic information (Elman, 1990; 1993; Li et al., 2021).

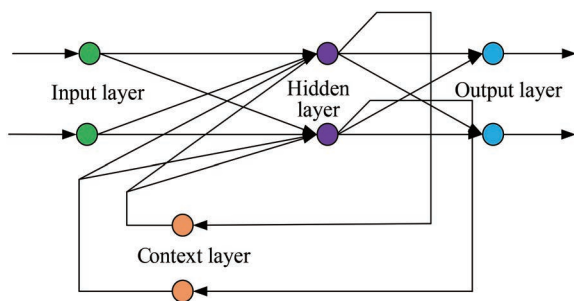


**Figure 2** Structure of the Elman neural network

The Elman neural network mathematical model can be described as follows:

$$\begin{cases} \mathbf{x}_c(k) = \mathbf{q}^{-1} \cdot \mathbf{x}_c(k-1) + \mathbf{x}_1(k-1) \\ \mathbf{x}_1(k) = f(\mathbf{w}^1 \mathbf{x}_c(k) + \mathbf{w}^2 \mathbf{x}_0(k-1)) \\ \mathbf{y}(k) = g(\mathbf{w}^3 \mathbf{x}_1(k)) \end{cases} \quad (2)$$

where  $\mathbf{q}^{-1}$  is the delay operator;  $\mathbf{w}^1$ ,  $\mathbf{w}^2$  and  $\mathbf{w}^3$  are the weight matrices between the context layer-hidden layer, input layer-hidden layer, and hidden layer-output layer, respectively.  $f(\cdot)$  and  $g(\cdot)$  are activation functions of hidden layer neurons and output layer neurons, respectively.  $\mathbf{x}_0$ ,  $\mathbf{x}_1$ ,  $\mathbf{x}_c$  and  $\mathbf{y}$  are the input layer input vectors, the hidden layer output vectors, the context layer output vectors, and the output node vectors, respectively.

### 3.3 Non-parametric KDE

Probability density estimation involves estimating the probability density function of a random variable based on a set of observation samples. It mainly includes two approaches: parametric and non-parametric methods. The parameter estimation requires the assumption that the data distribution obeys some parametric models. In practice, certain problems exist, such as model selection, unobservable variables, and dimensionality catastrophe. Ship roll motion status exhibits randomness and uncertainty. Therefore, obtaining the prior distribution of ship roll data samples is difficult. However, the non-parametric density estimation does not require the assumption that the data obey certain distribution; it approximates the probability density function of the data by separating the sample space into regions and approximating the probability of each region (Qiu, 2020; Harvey and Oryshchenko, 2012). KDE is a non-parametric test method that can directly estimate the probability distribution information of the deterministic prediction error of ship roll angle and properly reflect the real distribution of prediction errors. It can acquire a reliable range of error variations in ship roll angle, which provides more comprehensive decision-making support for ship operators.

The basic principle of non-parametric KDE is as follows: first, a sample set of data  $\mathbf{x} = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n\}$  is assumed. Then, the probability density function of the non-parametric KDE of the sample set  $\mathbf{x}$  can be expressed as

$$\hat{f}(\mathbf{x}) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right) \quad (3)$$

where  $n$  is the number of data samples,  $h$  denotes the bandwidth,  $K(\cdot)$  represents the kernel function, and  $\mathbf{x}_i$  signifies the  $i$ th sample value.

Notably, the choice of kernel function negligibly affects the fitting error. Therefore, the smooth Gaussian function is selected as the probability density estimation model for ship roll motion prediction errors in this study.

$$K\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right) = \frac{1}{\sqrt{2\pi} h} \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_i\|^2}{2h^2}\right) \quad (4)$$

In non-parametric KDE estimation, the choice of parameters for bandwidth directly influences the estimation accuracy of the model (Zambom and Dias, 2013). If the band-

width value is too large, then the probability density function will be smoother, which will mask the structural characteristics of the data sample and result in significant fitting errors. If the bandwidth value is small, then the probability density function has greater volatility, which can cause overfitting and cannot reflect the actual characteristics of the ship roll. To better estimate the non-parametric distribution of ship roll angle errors, the optimal bandwidth is selected using the following equation (Gramacki, 2018):

$$h = 1.06 \min(\sigma, Q/1.34) n^{(-1/5)} \quad (5)$$

where  $Q$  is the difference between the quantiles of 75% and 25% in the sample set, and  $\sigma$  denotes the standard deviation of the sample set.

## 4 Simulation experiments and discussion

In this study, the software MATLAB2023b is used to perform the simulation experiments on the operating system Windows 11 with 12th Gen Intel (R) Core (TM) i9-12900H, 2.5 GHZ (CPU) and 32.0 GB (RAM).

### 4.1 Model evaluation metrics

The effectiveness and accuracy of the prediction model proposed in this study are validated using the measured data of sea trials of M. V. *YuKun* for prediction analysis. Root mean square error (RMSE) can determine the deviation between the predicted value and the measured value. Mean absolute error (MAE) can indicate the actual situation of predicted value error. Thus, RMSE and MAE are utilized to assess the prediction accuracy of the model, as presented in Eq. (6).

$$\begin{cases} \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \\ \text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \end{cases} \quad (6)$$

where  $y_i$  and  $\hat{y}_i$  are the measured value and predicted value of the  $i$ th sample, respectively;  $n$  means the number of samples.

The reliability evaluation is vital for probability prediction of ship roll motion using the non-parametric KDE, which can directly influence the decision making of ship operators. Therefore, two evaluation metrics are used to analyze the reliability of model probability prediction. These metrics are referred to as prediction interval coverage probability (PICP) and prediction interval average width (PINAW). PICP can denote the coverage of the predicted range on the measured roll angle of the ship. PINAW signifies the average width of the model's predicted interval. The smaller the PINAW value, the better the model's interval

prediction performance. The expression is given as follows:

$$\begin{cases} \text{PICP} = \frac{1}{n} \sum_{i=1}^n \zeta_i, \quad \zeta_i = \begin{cases} 1 & y_i \in [L_i, U_i] \\ 0 & y_i \notin [L_i, U_i] \end{cases} \\ \text{PINAW} = \frac{1}{nR} \sum_{i=1}^n (U_i - L_i) \end{cases} \quad (7)$$

where  $n$  is the number of prediction samples;  $y_i$  refers to the  $i$ th measured value;  $\zeta_i$  denotes whether  $y_i$  falls in the prediction interval;  $L_i$  and  $U_i$  signify the upper and lower bounds of the prediction interval, respectively;  $R$  means the range of measured values.

### 4.2 Ship roll prediction based on sea trials of M. V. *YuKun*

In this study, the measured roll data from M. V. *YuKun* when sailing in Bohai Bay are chosen to validate the reliability of the model. M. V. *YuKun* is a scientific and training ship from Dalian Maritime University, as displayed in Figure 3 (<https://image.baidu.com>). The principal parameters of M. V. *YuKun* are presented in Table 1 (Yin et al., 2013).



**Figure 3** Scientific and training ship M.V. *YuKun*

**Table 1** Principal dimensions of M.V. *YuKun*

| Parameters                                  | Values |
|---------------------------------------------|--------|
| $L_{oa}$ (Length overall, m)                | 116    |
| $L_{bp}$ (Length between perpendiculars, m) | 105    |
| $B$ (Breadth molded, m)                     | 18     |
| $D$ (Depth to main deck, m)                 | 8.35   |
| $T$ (Design draft, m)                       | 5.4    |
| GT (Gross tonnage, t)                       | 6106   |
| $V$ (Design speed, kn)                      | 16.9   |

The shipborne ADU2 device is employed to collect ship roll data. The data sample interval is set to 1 s. The ship roll data for the experiments are displayed in Figure 4. When the sampling data of ship roll status are divided for training neural networks, the autocorrelation of the series needs to be considered to determine the input and output orders of the model. At present, the partial autocorrelation

function (PACF) is frequently used to determine the autocorrelation characteristics of stationary series. When the Kwiatkowski–Hills–Schmidt–Shin method is adopted to test the stationarity of the ship roll status, the test result shows that ship roll status exhibits a nonstationary characteristic due to the presence of unit roots (Kwiatkowski et al., 1992). Therefore, PACF is employed to ascertain the autocorrelation of differential time series of ship roll data, as presented in Figure 5.

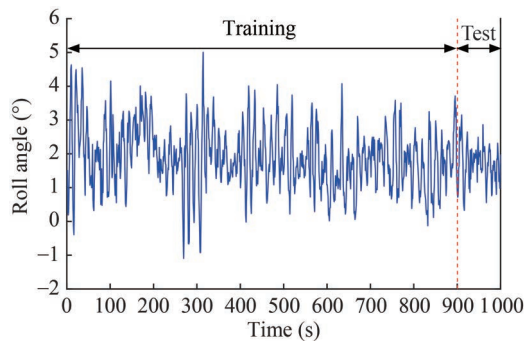


**Figure 4** Measured roll angles of the scientific and training ship M.V. YuKun

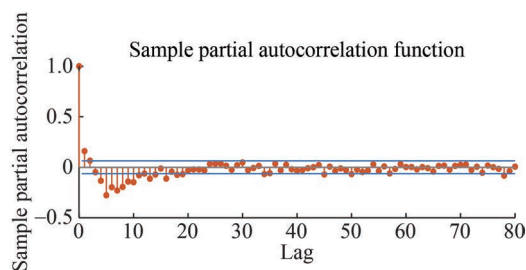


**Figure 5** PACF from differential time series of ship roll status

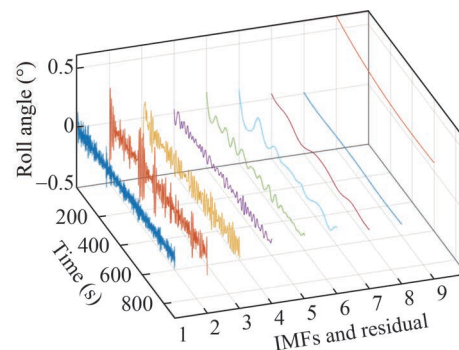
Figure 5 shows that the estimated values of PACF primarily fall within the 95% confidence interval after a lag order of 16. Considering the effects of ship maneuvering and the marine environment, a lag order of 16 is selected in this study. Therefore, when predicting future ship roll angles, the 16 previous roll angles are correlated with the current roll angle and contain more useful roll motion information. Therefore, the neural network model structure is set as 16 input nodes and 1 output node for the prediction models constructed for the ship roll status of the full-scale trial. The number of hidden nodes in the neural network model is set using Eq. (8). The training epochs and errors are chosen as 200 and 0.001, respectively. The number of roll data used in the experiment is 1 000, with the first 900 samples as the training dataset and the last 100 samples as the test dataset for prediction simulation.

$$H = 2 \times I + 1 \quad (8)$$

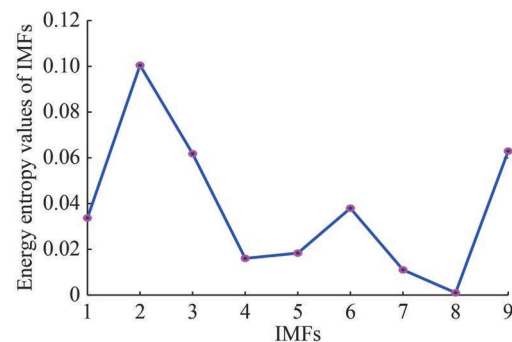
where  $I$  and  $H$  are the number of input layer nodes and

output nodes, respectively.

Accurately predicting ship roll status is difficult due to its nonstationary and time-varying nature. Thus, the EEMD method is selected to adaptively decompose the series to reduce the nonstationary characteristics and improve the model's prediction accuracy. The IMFs and residual series obtained from EEMD decomposition are presented in Figure 6(a). The roll status series is decomposed into nine IMFs and one residual component using the EEMD method. The frequencies of mode components are arranged from high frequencies to low frequencies. The high-frequency mode component can be considered the random interference of external factors on ship motion. The high-frequency mode component exhibits not only high fluctuation but also strong noise dominance, which greatly influences the prediction accuracy of the model. As the frequency and amplitude of the mode components gradually decrease, the IMFs tend to stabilize. The nonstationary of the IMF components decomposed by EEMD is decreased, and IMFs present a certain variation pattern. The energy entropy method is used to adaptively reconstruct the high-frequency components, which helps enhance prediction accuracy and minimize computational complexity. According to the principle of energy entropy reconstruction, the boundary of mode component reconstruction is IMF<sub>2</sub>, as displayed in Figure 6(b). Thus, IMF<sub>1</sub> can be a high-frequency mode component. IMF<sub>2</sub>–IMF<sub>8</sub> can be reconstructed as medium-frequency mode components, as illustrated in Figure 7.



(a) Multi-resolution decomposition using the EEMD method



(b) Energy entropy distribution of mode components

**Figure 6** Multi-resolution decomposition and energy entropy estimation in the modeling region

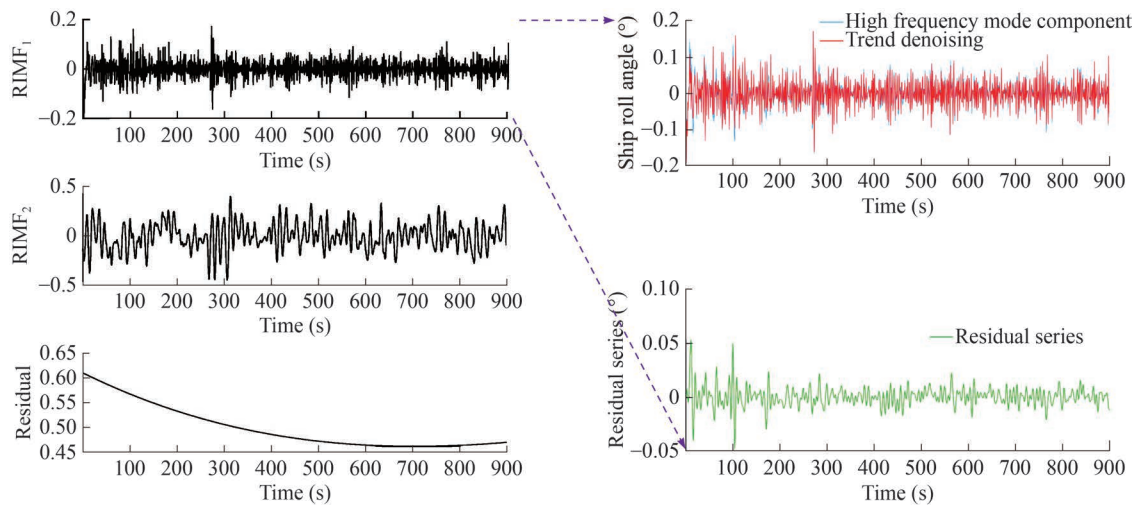


**Figure 7** Reconstruction and denoising of mode components in the modeling region

However, the high-frequency mode component is usually covered by noise, which is a key influencing factor of prediction accuracy. Therefore, the trend of the high-frequency component is further extracted using SSA to retain more effective information regarding the mode component. The contribution threshold (TR) of SSA is set to 95%, and the window length is  $L = 80$ . The effects of SSA for decomposition and reconstruction mode components are shown in Figure 7. The RBF, Elman, EMD–Elman, EEMD–Elman, and proposed prediction models are constructed for ship motion status prediction. A single hidden layer network structure is used for the Elman model, in which the activation function of the hidden layer is set as an S-function. The prediction results and accuracy of the models are presented in Figure 8 and Figure 9.

The prediction curve of HPM is displayed in Figure 8, and the prediction evaluation metrics are illustrated in Figure 9. As shown in Figure 8, the prediction curve for the proposed model presents the changing trend of the measured roll status of M.V. *YuKun*. The dynamic tracking ability of the proposed model for ship roll motion status is superior to those of RBF, Elman, EMD–Elman, and EEMD–Elman. Specifically, the proposed model can effectively indicate the dynamic characteristics of ship motion status. The MSE and RMSE of the proposed model are the smallest compared with those of the four other models (RBF, Elman, EMD–Elman, and EEMD–Elman), as displayed in Figure 9. The RMSE values of the proposed model in 1-step-ahead and 3-steps-ahead prediction are 0.316 59 and 0.364 15, respectively. The single- and multi-step prediction accuracies of the proposed model are as high as 22.7% and 27.5%, respectively, compared with those of other models. In addition, the prediction result of the Elman network is superior to that of RBF. The prediction accuracies of EEMD–Elman and EMD–Elman models are better than those of RBF and Elman models. The prediction accuracy of EEMD–Elman

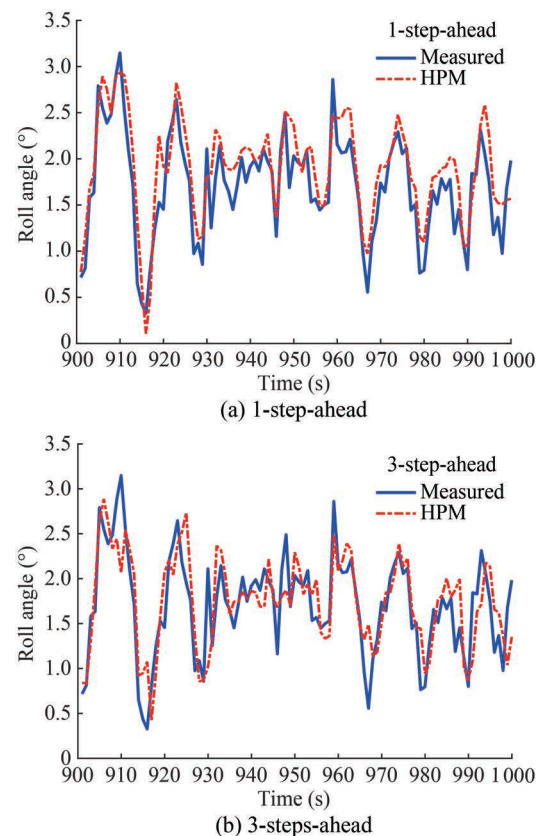


**Figure 8** Deterministic prediction results for roll motion status of M.V. *YuKun*

is also better than that of EMD–Elman. In summary, the abovementioned simulation results show that the prediction accuracy and generalization ability of the proposed model have been enhanced.

Non-parametric KDE is conducted on the ship roll status errors obtained from the deterministic prediction method to reflect the uncertainty of ship roll motion. In general, probability prediction models with different confidence levels can

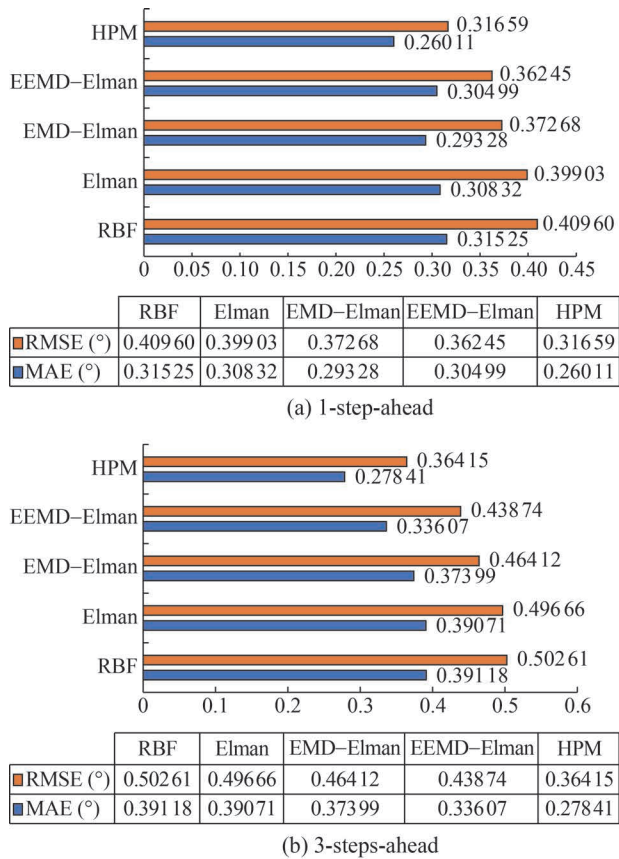


**Figure 9** Comparison of prediction accuracy for roll motion status of M.V. YuKun

provide prediction results at different intervals of reliability. A high confidence level corresponds to a larger prediction interval, as well as more uncertainty information regarding the ship's roll angle. The lower the confidence level, the smaller the prediction interval, and the smaller the uncer-

tainty information contained in the ship's roll angle. In this study, the 95%, 90%, and 85% confidence levels are chosen to derive the fluctuation range of ship roll angle prediction errors, which can better guarantee the reliability and accuracy of probability prediction for ship roll motion. In non-parametric KDE, the Gaussian kernel function is used to estimate the distribution of prediction errors in ship roll angle. The probability distribution function fitting curve of the errors of ship roll angle is shown in Figure 10. The figure shows that the probability density function can accurately estimate the actual distribution of prediction errors in ship roll angle.

In addition, the probability prediction interval of the ship roll angle is obtained by the superposition of error distribution estimation under different confidence levels and deterministic prediction results. Notably, a lower confidence level corresponds to a smaller prediction interval. The evaluation metric values of model prediction accuracy under different confidence levels are presented in Table 2.

Table 2 shows that, under the confidence levels of 95%, 90%, and 85%, the minimum PICP value for the predicted intervals of the model is 98% in multi-step prediction. The PINAW values obtained by the model for multi-step prediction are relatively small, with values of 1.2273, 0.9354, and 0.7499. Some measured values have exceeded the interval estimation boundary. As a result, the PICP value at the 85% confidence level drops to 98%. However, the prediction interval of the model contains most of the measured roll angle information of the ship. This result indicates that the prediction interval quality of the model is relatively high and reliable. Thus, the probability prediction method for ship roll motion based on non-parametric KDE and the novel data preprocessing strategy is feasible.

The simulation results reveal that the ship roll motion probability prediction model integrating novel data prepro-

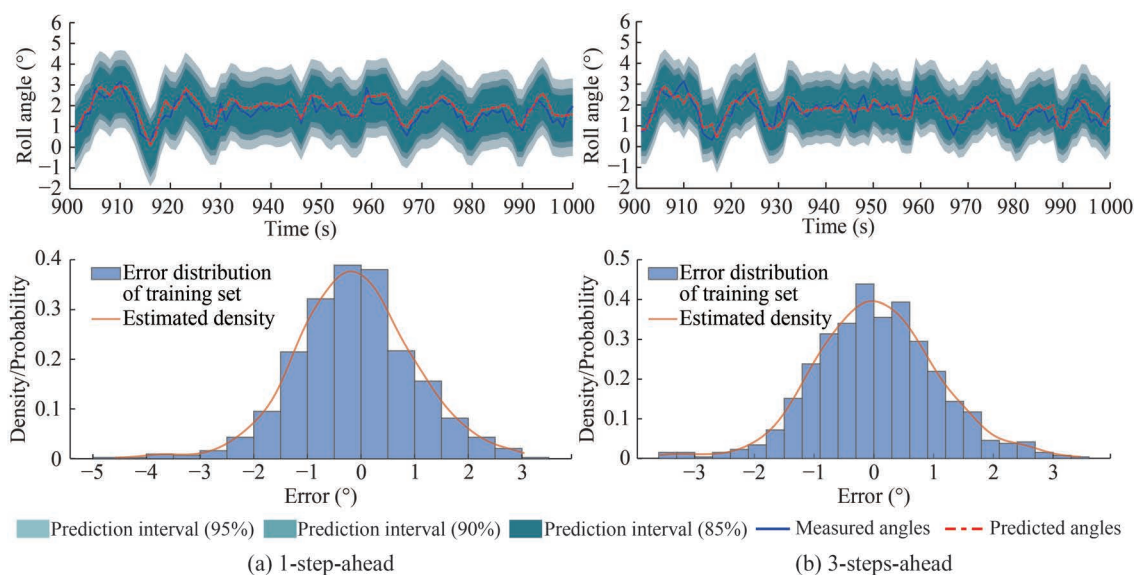


**Figure 10** Probability prediction results for ship roll status under different confidence levels

**Table 2** Model evaluation metrics under different confidence levels

| Prediction steps | PICP (%) |        |        | PINAW  |        |        |
|------------------|----------|--------|--------|--------|--------|--------|
|                  | 95% CL   | 90% CL | 85% CL | 95% CL | 90% CL | 85% CL |
| 1-step           | 100      | 100    | 100    | 1.3207 | 0.9998 | 0.7993 |
| 3-steps          | 100      | 100    | 98     | 1.2273 | 0.9354 | 0.7499 |

cessing technology, dynamic recursive neural network, and non-parametric KDE method can enhance the predictive accuracy and accurately describe the uncertainty of ship roll status with nonlinear time-varying dynamic characteristics. A major advantage of this study is the use of two prospective methods to explore the internal laws of ship roll motion and improve prediction accuracy. Figure 6 shows that EEMD can effectively decompose the roll status and overcome mode mixing and endpoint effects while preserving the time-varying dynamic characteristics of raw ship roll motion status. The proposed strategy for energy entropy reconstruction can adaptively determine the reconstruction boundary of mode components to decrease the number of mode components and the complexity of the prediction model. This method can also overcome the shortcomings of subjective selection for reconstructing mode components. As shown in Figure 8, Figure 9, and Table 1, the deterministic prediction results of ship roll motion using the proposed model are superior to those of the other prediction models (i.e., RBF, Elman, EMD–Elman, and EEMD–Elman). The Elman network is a locally recursive dynamic neural network with the capability to adapt to the time-varying characteristics of the system. It can more effectively reflect the dynamic characteristics of ship roll motion status than the RBF neural network. Thus, the dynamic recursive neural network is more suitable for defining the dynamic characteristics of ship roll motion than the feedforward neural network with respect to the prediction of ship roll motion status. The prediction accuracy of a neural network model based on the decomposition strategy is better than that of a single model. The fluctuation of ship roll motion status can be effectively decreased using the decomposition strategy. Moreover, the EEMD decomposition method can efficiently avoid mode aliasing and endpoint effects generated by the EMD method. The comparison of the proposed model with the four other models shows that the SSA algorithm can reduce the non-stationarity of the high-frequency mode component reconstructed by the energy entropy method while preserving more effective information. Figure 10 and Table 2 indicate that the fluctuation range of ship roll angle under different confidence levels can be determined using the estimated probability density function. The probability prediction interval of the model can precisely convey the fluctuation trend of ship roll angle, which can comprehensively reveal the uncertainty information of ship roll motion. The non-parametric KDE based on optimal bandwidth is an effective method to approximate the uncertainty of ship roll motion.

The proposed probability prediction scheme not only can provide a potential solution for the identification and prediction of complex dynamic systems but also can generate more reliable information to support decision making for ship operators.

## 5 Conclusions

A novel probabilistic interval hybrid prediction model is proposed to describe the uncertainty of ship roll motion. This model integrates the novel data preprocessing technology, the dynamic recursive neural network, and non-parametric KDE. The feasibility and efficiency of the proposed model are confirmed using full-scale test data. The main conclusions of this study are presented as follows. 1) Recursive neural networks can accurately identify the characteristics of ship roll motion. 2) The novel data preprocessing strategy integrating EEMD, the energy entropy reconstruction method, and SSA can enhance the deterministic prediction accuracy of ship roll motion status. 3) The non-parametric KDE based on optimal bandwidth can appropriately reflect the distribution of prediction errors in ship roll motion. The proposed probability prediction model can precisely capture the fluctuation range of ship roll motion. Thus, it can offer more reliable decision-making information for ship operators.

The prediction model for the ship motion quiescent period will be explored in our future research. This analysis is also important for enhancing the effectiveness of offshore operations and safeguarding operational security.

**Funding** Supported by the National Natural Science Foundation of China (under Grants of 52271361 and 52231014), the Natural Science Foundation of Guangdong Province of China (under Grant 2023A1515010684) and the Special Projects of Key Areas for Colleges and Universities of Guangdong Province (under Grant 2021ZDZX1008).

**Competing interest** The authors have no competing interests to declare that are relevant to the content of this article.

## References

- Afshar K, Bigdeli N (2011) Data analysis and short term load forecasting in Iran electricity market using singular spectral analysis (SSA). *Renew Energ* 36(5): 2620–2627. <https://doi.org/10.1016/j.energy.2011.02.003>
- An XL, Jiang DX, Zhao MH, Liu C (2012) Short-term prediction of wind power using EMD and chaotic theory. *Commun Nonlinear Sci Numer Simulat* 17(2): 1036–1042. <https://doi.org/10.1016/j.cnsns.2011.06.003>
- Chung JC, Bien Z, Kim YS (1990) A note on ship-motion prediction based on wave-excitation input estimation. *IEEE J Oceanic Eng* 15(3): 244–250. DOI: 10.1109/48.107153

- Ding SF, Ma Q, Zhou L, Han S, Dong WB (2023) Multipoint heave motion prediction method for ships based on the PSO-TGCN model. *China Ocean Eng* 37(6): 1022-1031. DOI: 10.1007/s13344-023-0085-4
- Duan WY, Huang LM, Han Y, Zhang YH, Huang S (2015) A hybrid AR-EMD-SVR model for the short-term prediction of nonlinear and non-stationary ship motion. *J Zhejiang Univ-Sc A (Appl Phys Eng)* 16(7): 562-576. DOI:10.1631/jzus.A1500040
- Elman JL (1990) Finding structure in time. *Cognitive Sci* 14(2): 179-211. [https://doi.org/10.1207/s15516709cog1402\\_1](https://doi.org/10.1207/s15516709cog1402_1)
- Elman JL (1993) Learning and development in neural networks: The importance of starting small. *Cognition* 48(1): 71-99. [https://doi.org/10.1016/0010-0277\(93\)90058-4](https://doi.org/10.1016/0010-0277(93)90058-4)
- Fossen TI (2011) *Handbook of marine craft hydrodynamics and motion control*. John Wiley & Sons Ltd, Chichester, West Sussex, United Kingdom
- Fu HX, Gu ZQ, Wang YC (2022) Ship pitch prediction based on Bi-ConvLSTM-CA model. *J Mar Sci Eng* 10(7): 840. <https://doi.org/10.3390/jmse10070840>
- Gao N, Hu AK, Hou LX, Chang X (2023) Real-time ship motion prediction based on adaptive wavelet transform and dynamic neural network. *Ocean Eng* 280: 114466. <https://doi.org/10.1016/j.oceaneng.2023.114466>
- Geng XY, Li YB, Sun Q (2023) A novel short-term ship motion prediction algorithm based on EMD and adaptive PSO-LSTM with the sliding window approach. *J Mar Sci Eng* 11: 466. <https://doi.org/10.3390/jmse11030466>
- Gramacki A (2018) *Nonparametric kernel density estimation and its computational aspects*. Springer International Publishing, Cham, Switzerland. <https://doi.org/10.1007/978-3-319-71688-6>
- Han LQ, Shi Y (2023) *Artificial neural network theory and application*. China Machine Press, Beijing, China
- Harvey A, Oryshchenko V (2012) Kernel density estimation for time series data. *Int J Forecasting* 28(1): 3-14. <https://doi.org/10.1016/j.ijforecast.2011.02.016>
- Huang LM, Duan WY, Han Y, Chen YS (2014) A review of short-term prediction techniques for ship motions in seaway. *J Ship Mech* 18(12): 1534-1542. DOI: 10.3969/j.issn.1007-7294.2014.12.013
- Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng QA, Yen NC, Tung CC, Liu HH (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc R Soc London Ser A Math Phys Eng Sci* 454(1971): 903-995. <https://doi.org/10.1098/rspa.1998.0193>
- Jiang H, Duan SL, Huang LM, Han Y, Yang H, Ma QW (2020) Scale effects in AR model real-time ship motion prediction. *Ocean Eng* 203: 107202. <https://doi.org/10.1016/j.oceaneng.2020.107202>
- Kwiatkowski D, Phillips PCB, Schmidt P, Shin Y (1992) Testing the null hypothesis of stationarity against the alternative of a unit root. *J Econometrics* 54(1): 159-178. [https://doi.org/10.1016/0304-4076\(92\)90104-Y](https://doi.org/10.1016/0304-4076(92)90104-Y)
- Li GY, Kawan B, Wang H, Zhang HX (2017) Neural-network-based modelling and analysis for time series prediction of ship motion. *Ship Technol Res* 64(1): 30-39. <http://dx.doi.org/10.1080/09377255.2017.1309786>
- Li L, Wang TT, Yin HR, Niu J, Liang YB, He JN, Miao SH (2022a) Point-interval joint prediction method of dissolved gas volume fraction in transformer oil based on GWO-LSTM and NKDE. *High Volt Appar* 58(11): 88-97. (in Chinese) DOI:10.13296/j.1001-1609.hva.2022.11.012
- Li MW, Geng J, Han DF, Zheng TJ (2016) Ship motion prediction using dynamic seasonal RvSVR with phase space reconstruction and the chaos adaptive efficient FOA. *Neurocomputing* 174: 661-680. <https://doi.org/10.1016/j.neucom.2015.09.089>
- Li MW, Xu DY, Geng J, Hong WC (2022b) A ship motion forecasting approach based on empirical mode decomposition method hybrid deep learning network and quantum butterfly optimization algorithm. *Nonlinear Dyn* 107: 2447-2467. <https://doi.org/10.1007/s11071-021-07139-y>
- Li SQ, Liu XM, Lin AJ (2020) Fractional frequency hybrid model based on EEMD for financial time series forecasting. *Commun Nonlinear Sci Numer Simulat* 89: 105281. <https://doi.org/10.1016/j.cnsns.2020.105281>
- Li XM, Han ZW, Zhao TY, Zhang JL, Xue D (2021) Modeling for indoor temperature prediction based on time-delay and Elman neural network in air conditioning system. *J Build Eng* 33: 101854. <https://doi.org/10.1016/j.jobe.2020.101854>
- Lin Z, Yang Q, Guo ZQ, Li J (2011) An improved Autoregressive method with Kalman Filtering Theory for vessel motion predication. *Int J Intell Eng Syst* 4(4): 11-18. DOI:10.22266/ijies2011.1231.02
- Luo QF, Li J, Zhou YQ, Liao L (2020) Using spotted hyena optimizer for training feedforward neural networks. *Cogn Syst Res* 65: 1-16. <https://doi.org/10.1016/j.cogsys.2020.09.001>
- Mao ZY, Zhao Y, Zhao JX, Shao LH, Zuo H (2025) A dual-channel ultra-short-term ship motion prediction model using temporal convolutional network, frequency-enhanced channel attention and informer. *Phy Fluids* 37(2): 022145. <https://doi.org/10.1063/5.0257465>
- Mi XW, Liu H, Li YF (2019) Wind speed prediction model using singular spectrum analysis, empirical mode decomposition and convolutional support vector machine. *Energ Convers Manage* 180: 196-205. <https://doi.org/10.1016/j.enconman.2018.11.006>
- Moreno SR, dos Santos Coelho L (2018) Wind speed forecasting approach based on singular spectrum analysis and adaptive neuro fuzzy inference system. *Renew Energ* 126: 736-754. <https://doi.org/10.1016/j.renene.2017.11.089>
- Peng XY, Zhang B, Zhou HG (2019) An improved particle swarm optimization algorithm applied to long short-term memory neural network for ship motion attitude prediction. *T I Meas Control* 41(15): 4462-4471. DOI: 10.1177/0142331219860731
- Peng XY, Zhao XR, Wei NX, Xie N (2001) AR algorithm for extremely short-term prediction of large ship's motion. *Ship Eng* (5): 5-7, 10. (in Chinese)
- Qiu XP (2020) *Neural networks and deep learning*. China Machine Press, Beijing, China
- Reis J, Batista P, Oliveira P, Silvestre C (2023) Discrete-time Kalman filter for heave motion estimation. *Ocean Eng* 277: 114240. <https://doi.org/10.1016/j.oceaneng.2023.114240>
- Song CY, Zhang XK, Zhang GQ (2022) Attitude prediction of ship coupled heave-pitch motions using nonlinear innovation via full-scale test data. *Ocean Eng* 264: 112524. <https://doi.org/10.1016/j.oceaneng.2022.112524>
- Sun Q, Tang Z, Gao JP, Zhang GC (2022) Short-term ship motion attitude prediction based on LSTM and GPR. *Appl Ocean Res* 118: 102927. <https://doi.org/10.1016/j.apor.2021.102927>
- Suhermi N, Suhartono, Prastyo DD, Ali B (2018) Roll motion prediction using a hybrid deep learning and ARIMA model. *Proced Comput Sci* 144: 251-258. <https://doi.org/10.1016/j.procs.2018.10.526>
- Tarkhaneh O, Shen HF (2019) Training of feedforward neural networks for data classification using hybrid particle swarm optimization, Mantegna Lévy flight and neighborhood search. *Heliyon* 5(4): e01275. <https://doi.org/10.1016/j.heliyon.2019.e01275>
- Wang YC, Dou YB, Yang ZQ, Fu HX (2024) Ship pitch prediction method based on LSTM and multi-head attention. *Ocean Eng*

- 309: 118236. <https://doi.org/10.1016/j.oceaneng.2024.118236>
- Wang YC, Wang H, Zou DX, Fu HX (2021) Ship roll prediction algorithm based on Bi-LSTM-TPA combined model. *J Mar Sci Eng* 9(4): 387. <https://doi.org/10.3390/jmse9040387>
- Wei YY, Chen ZZ, Zhao C, Chen X (2023a) Deterministic ship roll forecasting model based on multi-objective data fusion and multi-layer error correction. *Appl Soft Comput* 132: 109915. <https://doi.org/10.1016/j.asoc.2022.109915>
- Wei YY, Chen ZZ, Zhao C, Tu YH, Chen X, Yang R (2021) A BiLSTM hybrid model for ship roll multi-step forecasting based on decomposition and hyperparameter optimization. *Ocean Eng* 242: 110138. <https://doi.org/10.1016/j.oceaneng.2021.110138>
- Wei YY, Chen ZZ, Zhao C, Chen X, He JH, Zhang CY (2023b) A three-stage multi-objective heterogeneous integrated model with decomposition-reconstruction mechanism and adaptive segmentation error correction method for ship motion multi-step prediction. *Adv Eng Inform* 56: 101954. <https://doi.org/10.1016/j.aei.2023.101954>
- Wu ZH, Huang NE (2009) Ensemble empirical model decomposition: A noise assisted data analysis method. *Adv Adapt Data Anal* 1(1): 1-41. DOI: 10.1142/S1793536909000047
- Xia JD, Zheng WL, Wang ZH, Li J (2022) Short-term prediction of ship motion attitude based on EMD-LSTM. *Comput Digit Eng* 50 (7): 1434-1438. DOI: 10.3969/j.issn.1672-9722.2022.07.007
- Yang Y, Yu DJ, Cheng JS (2006) A roller bearing fault diagnosis method based on EMD energy entropy and ANN. *Journal of Sound and Vibration* 294(1-2): 269-277. <https://doi.org/10.1016/j.jsv.2005.11.002>
- Yin JC, Zou ZJ, Xu F (2013) On-line prediction of ship roll motion during maneuvering using sequential learning RBF neural networks. *Ocean Eng* 61: 139-147. <https://doi.org/10.1016/j.oceaneng.2013.01.005>
- Yin JC, Perakis NA, Wang N (2018) A real-time ship roll motion prediction using wavelet transform and variable RBF network. *Ocean Eng* 160: 10-19. <https://doi.org/10.1016/j.oceaneng.2018.04.058>
- Yin JC, Zou ZJ, Xu, F, Wang NN (2014) Online ship roll motion prediction based on grey sequential extreme learning machine. *Neurocomputing* 129: 168-174. <https://doi.org/10.1016/j.neucom.2013.09.043>
- Yu CJ, Li YL, Zhang MJ (2017) An improved wavelet transform using singular spectrum analysis for wind speed forecasting based on Elman Neural Network. *Energ Convers Manage* 148: 895-904. <https://doi.org/10.1016/j.enconman.2017.05.063>
- Yuan PY, Zhao WG, Zhao Y (2025) Intelligent prediction algorithm of ship roll and pitch motion based on SSA-optimized BiLSTM network. *Ocean Eng* 320: 120331. <https://doi.org/10.1016/j.oceaneng.2025.120331>
- Zambom AZ, Dias R (2013) A review of kernel density estimation with applications to econometrics. *Int Economet Rev* 5(1): 20-42. <https://doi.org/10.48550/arXiv.1212.2812>
- Zhang B, Wang S, Deng LW, Jia MQ, Xu JZ (2023) Ship motion attitude prediction model based on IWOA-TCN-Attention. *Ocean Eng* 272: 113911. <https://doi.org/10.1016/j.oceaneng.2023.113911>
- Zhang B, Peng XY, Gao J (2020a) Ship motion attitude prediction based on ELM-EMD-LSTM integrated model. *J Ship Mech* 24(11): 1413-1421. DOI: 10.3969/j.issn.1007-7294.2020.11.005
- Zhang GY, Tan F, Wu YX (2020b) Ship motion attitude prediction based on an adaptive dynamic particle swarm optimization algorithm and bidirectional LSTM Neural Network. *IEEE Access* 8: 90087-90098. DOI: 10.1109/ACCESS.2020.2993909
- Zhang T, Zheng XQ, Liu MX (2021) Multiscale attention-based LSTM for ship motion prediction. *Ocean Eng* 230(13): 109066. <https://doi.org/10.1016/j.oceaneng.2021.109066>
- Zhang XL, Lu XM, Jia RS, Kan ST (2018) Micro-seismic signal denoising method based on variational mode decomposition and energy entropy. *J China Coal Soc* 43(2): 356-363. (in Chinese) DOI: 10.13225/j.cnki.jccs.2017.4153
- Zhou BW, Ma XJ, Luo YH, Yang DS (2019) Wind power prediction based on LSTM networks and nonparametric kernel density estimation. *IEEE Access* 7: 165279-165292
- Zhou L, Duan YX, Ren ZR, An S (2021) Active heave compensation control system and motion prediction research. *J Huazhong Univ of Sci & Tech (Natural Science Edition)* 49(3): 98-104. (in Chinese) DOI: 10.13245/j.hust.210318