Improved PER-DDPG based nonparametric modeling of ship dynamics with uncertainty

Zhu, Man; Tian, Kang; Wen, Yuan Qiao; Cao, Ji Ning; Huang, Liang

DOI
10.1016/j.oceaneng.2023.115513

Publication date
2023

Document Version
Final published version

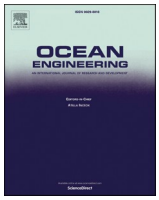
Published in
Ocean Engineering

Ocean Engineering 286 (2023) 115513

Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng



Improved PER-DDPG based nonparametric modeling of ship dynamics with uncertainty

Man Zhu [a,b], Kang Tian [a,b,*], Yuan-Qiao Wen [a,b], Ji-Ning Cao [a,b], Liang Huang [a,b,c]

[a] State Key Laboratory of Maritime Technology and Safety, Wuhan University of Technology, Wuhan, 430063, China
[b] National Engineering Research Center for Water Transport Safety, Wuhan University of Technology, Wuhan, 430063, China
[c] Faculty of Technology, Policy and Management, Delft University of Technology, Delft, 2628 BX, Netherlands

ARTICLE INFO

Handling Editor: Prof. A.I. Incecik

Keywords:
Ship dynamic model
Nonparametric modeling
Deep deterministic policy gradient (DDPG)
Prioritized experience replay (PER)
Uncertainty

ABSTRACT

This study contributes to addressing the challenge of quickly obtaining an effective and accurate nonparametric model for describing ship maneuvering motion in three degrees of freedom (3-DOF). To achieve this, an intelligent ship dynamics nonparametric modeling method named improved PER-DDPG is proposed. This method leverages the deep deterministic policy gradient algorithm (DDPG) and prioritized experience replay mechanism (PER) and analyzes the characteristics between the goal of deep reinforcement learning (DRL) and the modeling process of the nonparametric model. The PER mechanism is utilized to enhance the agent's understanding of the overall mechanism of ship motion by improving the utilization of samples. The meaning of target value is redefined due to transforming DRL aiming at maximizing cumulative rewards into maximizing the set of immediate rewards at each time step. To validate the performance of the proposed modeling method, we conduct simulation studies using a benchmark ship model i.e., a Mariner cargo ship dynamic model, and experimental studies using a real unmanned surface vehicle (USV). In the simulation test, we demonstrate the effectiveness and generalization of the proposed method through zigzag and turning circle tests. Furthermore, we verify the robustness and applicability of the proposed method by using datasets with uncertain environmental disturbances and datasets with different sampling frequencies. Additionally, the experimental tests conducted on the USV indicate the consistency of the proposed approach.

1. Introduction

Building a high-precision ship dynamic model is essential for analyzing ship maneuvering motion characteristics, constructing a motion control system, and realizing intelligent navigation. However, obtaining such a model is a highly complex process due to the characteristics of ship motion systems, which exhibit high coupling, strong nonlinearity, and multi-uncertainties. Therefore, it is critical to develop a reliable and easy-to-implement ship dynamics modeling method. Currently, ship modeling methods are generally classified into two types including mechanism-based modeling and data-driven modeling. Mechanism-based modeling involves establishing a mathematical model that describes ship motion based on traditional kinematics and dynamics theories. It considers hydrodynamic forces, control forces, and environmental disturbance forces acting on the ship (He et al., 2022). The unknown parameters of this model are usually determined through captive model test, Computational Fluid Dynamics (CFD), and empirical formulae or database method. However, these methods have drawbacks such as the requirement to identify numerous model parameters and the model accuracy being dependent on the experience and prior knowledge. To reduce the complexity of the modeling process and mitigate the reliance on prior knowledge, data-driven modeling provides an applicable solution for constructing ship dynamics in recent years.

Data-driven modeling is described as a system identification process involving the acquisition of input-output data, selection of a model class, estimation of model parameters, and validation of the identified model (Habib et al., 2021). This method offers advantages of low cost and high efficiency for modeling ship dynamics, particularly in the presence of time-varying and uncertain environmental disturbances. System identification (SI) presented many successes, and data-driven modeling became an enabling factor in modern design methods (Schoukens and Ljung, 2019). In ship dynamic model identification, the SI method relies on input-output data obtained from free-running model tests or full-scale trials. It can be categorized into parametric modeling and

* Corresponding author. Intelligent Transportation Systems Research Center, Wuhan University of Technology, Wuhan, 430063, China.
E-mail address: tiankang@whut.edu.cn (K. Tian).

https://doi.org/10.1016/j.oceaneng.2023.115513
Received 23 May 2023; Received in revised form 16 July 2023; Accepted 29 July 2023
Available online 3 August 2023

nonparametric modeling. Parametric modeling aims to identify the unknown parameters of ship dynamic models based on the input-output data (Zhu et al., 2018, 2019; Ouyang et al., 2022). This transforms the parameter identification problem into a multiple regression problem, which can be solved using various methods such as support vector machine (SVM), ridge regression (RR), Gaussian process regression (GPR), and their improved versions. The basic idea of these identification methods utilizes kernel functions to map the training data to high-dimensional feature space, allowing for the construction of ship dynamic models through implicit learning. For example, Zhu et al. (2017) simplified the 6 DOF model into the speed model and steering model to capture ship dynamics in six degrees of freedom (DOF) model by reasonable assumptions, and developed support vector machines (SVM) method optimized by the artificial bee colony algorithm (ABC) to identify the parameters of these two models. Wang et al. (2020) proposed a 'nu'-support vector regression method (v-SVR) with Gaussian kernel to establish a robust ship motion model and designed a parameter tuning scheme which combined hold-out validation and dynamic process simulation to prevent overfitting. Chen et al. (2023) proposed the LS-SVM with radial basis function and an error-based online modeling algorithm to identify an offline black-box model for maritime autonomous surface ships (MASS) motion under the effect of ocean waves. Meng et al. (2023) applied the grey wolf optimized support vector regression (GWO-SVR) and nonlinear innovation processed by hyperbolic tangent function to improve the accuracy of parameter identification of ship nonlinear motion model. It is important to note that these methods mentioned above require assuming a reasonable model in advance. However, the selected model structure may no longer be applicable due to changes in ship operating conditions (Wang et al., 2020). Therefore, these methods have significant limitations in practical application.

Nonparametric modeling aims to obtain the optimal mapping between input data and output data without requiring any prior knowledge of ship dynamics (Zhang et al., 2022a,b; Ouyang et al., 2023a). The accuracy of this method primarily relies on the input-output data collected through ship on-board sensors, making it suitable for modeling ship dynamics and predicting ship motion. With achievements in classical identification algorithms and neural network technology, numerous remarkable academic achievements in nonparametric modeling have emerged. For instance, recognizing the significance of kernel functions in the modeling of ship dynamics (Chen et al., 2022, 2023). Ouyang et al. (2023b) introduced an adaptive hybrid-kernel function with GPR to enhance its generalization and prediction accuracy. In addition, Ouyang et al. (2023a) proposed a local Gaussian process regression (LGPR) method to model the dynamics in 3-DOF for a KVLCC2 tanker ship and a real USV. The method demonstrated desirable performance in computational cost through two case studies. The aforementioned research works demonstrate notable capabilities in terms of model accuracy, generalization ability, and computational efficiency. However, their effectiveness relies heavily on the selection of a suitable kernel function, which remains a challenging task that lacks clear criteria (Francis and Raimond, 2021). In parallel, the utilization of neural network technology has garnered significant attention in ship dynamics modeling due to its ease of implementation and flexibility. Woo et al. (2018) applied the Long Short-Term Memory (LSTM) approach to the nonlinear component of the Nomoto model, albeit without considering environmental disturbances. Hao et al. (2022) proposed a Recurrent Neural Network (RNN) model for ship maneuvering motion prediction, highlighting the impact of historical motion states by incorporating past velocity values in the inputs. Nonetheless, their model solely addressed the linearization of sway and yaw motion, simplifying the influence of surge motion on transversal motion. Jiang et al. (2022) explored an LSTM-based modeling approach for constructing a dynamic model of the KVLCC2 ship. They verified the model's robustness against noise by utilizing a training dataset contaminated with varying levels of Gaussian white noise. While LSTM mitigated the gradient explosion problem of traditional RNNs at the algorithmic level, the model's actual performance still requires validation using experimental data from real ships. In a similar vein, He et al. (2022) employed a self-designed fully connected neural network with Bayesian optimization for nonparametric modeling of ship motion. They incorporated Gaussian processes and lower confidence bounds to estimate unobserved validation errors and strike a balance between exploration and exploitation. The modeling method was evaluated using both simulation and experimental data. However, to enhance the model's persuasiveness, further verification experiments under environmental disturbances such as wind, waves, and currents are necessary. Consequently, there is a need for further research to develop a nonparametric modeling approach that is well-suited for real environments with inherent uncertainty while enabling the rapid construction of ship dynamics models.

Currently, Deep Reinforcement Learning (DRL) has become one of the forefront areas of research in Artificial Intelligence (AI). It is an intelligent decision-making theory that emulates human thinking processes. The core idea of DRL is to obtain an optimal strategy through continuous interaction with environment, using trial-and-error and feedback learning. Due to its powerful feature learning and decision-making abilities, DRL has been extensively studied in domains such as autonomous driving, robot control, and intelligent ship control, resulting in fruitful research achievements (Zhao et al., 2021; Kiran et al., 2022; Sivaraj and Rajendran, 2022; Sun et al., 2023). Classical DRL algorithms are Deep Q Network (DQN) (Mnih et al., 2013) and DDPG (Lillicrap et al., 2015). Since the prediction issue in some domains can be transformed into a sequential decision problem, researchers have been exploring the use of DRL to address these prediction tasks. For example, Tang et al. (2021) achieved accurate forecasting of social-economic trends and assessment of current decisions by integrating DDPG with LSTM and dynamically adjusting the prediction parameters. Chen and Liu (2021) developed a dynamic ensemble model based on DQN that considered the time-varying characteristics of wind speed series to predict wind speed. The proposed dynamic ensemble model outperformed classic intelligent prediction models and six ensemble methods when evaluated on actual datasets. In Liang et al. (2023), a dynamic noisy proximal policy (DNPP) algorithm, by applying the dynamic noise exploration and dynamic approximation optimization into DDPG, is developed using the actor-critic (AC) frame. The prediction results indicate that DNPP has advantages in action exploration and network updating compared with other methods. Huang et al. (2023) proposed an accurate prediction of required virtual resources (APRR) approach based on DQN algorithm through analyzing the real service traffic, where the DQN-based matrix factorization designed was used to infer the missing elements for minimizing the prediction errors. The prediction performance of this approach well demonstrated through testing over real-world datasets.

Inspired by these studies especially the obvious advantages of DRL in modeling, DRL-based method is a good alternative to model ship dynamics with uncertainty induced by environmental disturbances or measurement noises without requirement of prior knowledge of ship motions. This paper is devoted to developing a novel nonparametric modeling approach by introducing DDPG and PER mechanism (Schaul et al., 2016) to describe ship dynamics with uncertainty mainly induced by disturbances in 3 degrees of freedom. The target value of DDPG is modified by analyzing the modeling process of the nonparametric model and aligning it with the ultimate goal of DRL. The PER mechanism is employed to improve the learning efficiency of the agent. Both the simulation tests on Mariner ship and experimental tests on a real USV are carried out to comprehensively evaluate the performance of the proposed modeling method. From the knowledge of authors, it is the first time using PER-DDPG method to investigate nonparametric modeling of ship dynamics with uncertainty within the ship dynamic modeling research field.

The structure of this paper is organized as follows. A 3-DOF ship

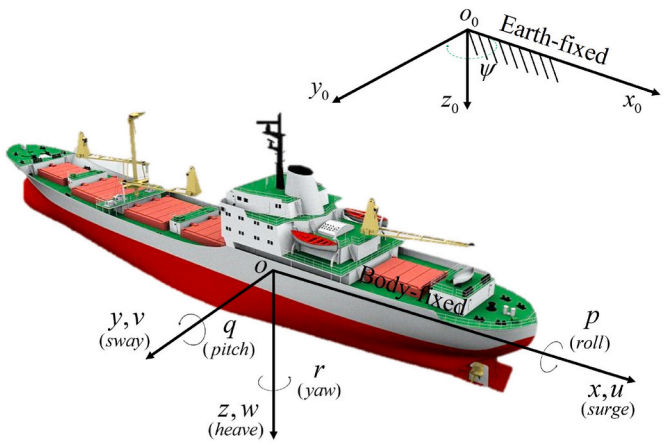


Fig. 1. Reference frame.

dynamic model is described in Section 2. The improved PER-DDPG identification method involving improved DDPG, improved PER-DDPG, and evaluation indexes is presented in Section 3. In Section 4, two case studies including the simulated Mariner maneuvers and a real USV experiment are conducted for comprehensively evaluating the performance of the proposed nonparametric modeling method. Finally, remarkable conclusions are presented in Section 5.

2. Problem formulation

2.1. Ship dynamic model

Generally speaking, a ship is typically considered as a rigid body with 6 degrees of freedom (DOF) in motion. However, when it comes to controller design for ships in the horizontal plane, the focus is often on the three degrees of freedom (3-DOF) ship dynamic model. This selective approach is primarily due to the limited benefits of using the full model for the development and testing of ship controllers. Notably, controlling the heading angle and horizontal translation speed has proven to be relatively effective, as highlighted in the works of Zhu et al. (2020) and Chen et al. (2021). Therefore, this paper specifically focuses on the motions of the 3 DOF (surge, sway, and yaw) in ships.

Two coordinate systems are utilized to describe the motion state of the ship. The first is the earth-fixed coordinate system, which remains fixed to the surface of the earth. The second is the body-fixed coordinate system, with its origin located at the mid-ship section, as depicted in Fig. 1. The ship's kinematic model is described in accordance with Fossen (2011).

$$\frac{dx_0}{dt} = u\cos\psi - v\sin\psi \tag{1-a}$$

$$\frac{dy_0}{dt} = u\sin\psi + v\cos\psi \tag{1-b}$$

$$\frac{d\psi}{dt} = r \tag{1-c}$$

where (x_0, y_0) is the position coordinates of the origin in the earth-fixed coordinate system, ψ is ship heading angle, u is ship surge speed, v is ship sway speed, and r is ship yaw rate.

The motion trajectory of the ship can be predicted through direct integration operation on Eqs. (1-a) and (1-b). However, if the predicted trajectory is too long, cumulative errors can arise and significantly impact the accuracy of the trajectory calculation. In this study, the simulated data encompasses a trajectory length of over 1000m as shown in Fig. 7. The ratio of the trajectory length to the hull length is approximately 22:1, indicating a significant span covered. The calculation was performed using a time step of 0.1 s. To address this issue, we calculate the measured position coordinates and heading angle with the predicted ship speed at the current moment. The calculation can be expressed using the following formula

$$x(t+1) = x(t) + [u_t\cos(\psi_t) - v_t\sin(\psi_t)]\cdot\Delta t \tag{2-a}$$

$$y(t+1) = y(t) + [u_t\sin(\psi_t) + v_t\cos(\psi_t)]\cdot\Delta t \tag{2-b}$$

where $(x(t+1), y(t+1))$ is the position coordinates predicted at the next moment; $(x(t), y(t))$ and ψ_t are the position coordinates and heading angle measured at the current moment; u_t and v_t are the predicted value of the model at the same time; Δt is the interval time of sampling.

According to Newton's second law, the ship kinetic model can be written as

$$\begin{aligned} &m(\dot{u} - vr - x_G r^2) = X \\ &m(\dot{v} + ur + x_G\dot{r}) = Y \\ &I_z\dot{r} + mx_G(\dot{v} + ur) = N \end{aligned} \tag{3}$$

where m is the mass of the ship, I_z is the moment of inertia about the z-axis, x_G is the longitudinal coordinate of the ship's gravity center in the body-fixed coordinate system, X, Y, N are the components of the hydrodynamic forces and moments acting on the ship.

Referring to the Abkowitz model proposed by Abkowitz (1964), the hydrodynamic forces and moments acting on the ship can be regarded as functions concerning the ship speed and rudder angle, i.e., it is expanded by third-order Taylor series about the state of straight-ahead motion with constant speed. Then Eq. (3) becomes

$$\begin{aligned} &(m - X_{\dot{u}})\dot{u} = f_1(u, v, r, \delta) \\ &(m - Y_{\dot{v}})\dot{v} + (mx_G - Y_{\dot{r}})\dot{v} = f_2(u, v, r, \delta) \\ &(mx_G - N_{\dot{v}})\dot{v} + (I_z - N_{\dot{r}})\dot{v} = f_3(u, v, r, \delta) \end{aligned} \tag{4}$$

where $X_{\dot{u}}, Y_{\dot{v}}, Y_{\dot{r}}, N_{\dot{v}}, N_{\dot{r}}$ are the hydrodynamic derivatives, f_1, f_2, f_3 are the nonlinear functions.

2.2. Problem statement

According to the study in (Wang et al., 2015), Eq. (4) can be rewritten as a black-box model where $u(t+1), v(t+1), r(t+1)$ can be described as functions of $u(t), v(t), r(t), \delta(t)$ in the following form

$$\begin{aligned} u(t+1) &= h_1[u(t), v(t), r(t), \delta(t)] \\ v(t+1) &= h_2[u(t), v(t), r(t), \delta(t)] \\ r(t+1) &= h_3[u(t), v(t), r(t), \delta(t)] \end{aligned} \tag{5}$$

where $h_i (i = 1, 2, 3)$ are the unknown nonlinear functions. Therefore, the black-box model also known as a nonparametric model is derived and needs to be identified using the PER-DDPG method to describe ship dynamics in the horizontal frame.

3. Improved PER-DDPG based modeling method

3.1. DDPG algorithm

The modeling problem of ship dynamics begins by formulating it as a Markov decision process (MDP), followed by the definition and artificial design of the state space, action space, and reward function.

According to Eq. (5), the state space (S) is defined as

$$S = \{u, v, r, \delta\} \tag{6}$$

The action space (A) is defined as

$$\mathrm{A} = \{u', v', r'\} \tag{7}$$

The purpose of the reward function (R) is to guide the agent in learning the optimal policy which involves minimizing the error between the true speed components and the predicted speed components. The specific function is as follows:

$$R = -|u-u^{'}| - |v-v^{'}| - |r-r^{'}| \tag{8}$$

where u, v, r represent the true surge speed, sway speed and yaw rate, respectively; $u^{'}, v^{'}, r^{'}$ represent the speed predicted; R represents the feedback from the environment on the predicted value of the model output.

DDPG mainly includes three parts: actor-critic structure, target network mechanism, and experience replay mechanism. The actor-critic structure consists of an execution part and an evaluation part, and each part has a neural network. The target network mechanism adds a neural network with the same structure in both the execution part and the evaluation part to enhance the stability of the algorithm, two neural networks in each part are called online network and target network respectively. Experience replay mechanism breaks the correlation between sequential data by randomly extracting training samples from the replay memory buffer, and thereby improving the convergence of the algorithm. The goal is to learn an optimal policy that maximizes the cumulative rewards from the start distribution.

In the DDPG algorithm, the input and the output of online network μ of execution part are the current state s_t and action $\mu(s_t|w_\mu)$. The input and the output of the target network $\mu^{'}$ are the next state s_{t+1} and action a_{t+1}. Meanwhile, the exploration ability of the agent is enhanced by adding Ornstein-Uhlenbeck (OU) noise (Uhlenbeck and Ornstein, 1930) to the actor policy

$$a_t = \mu(s_t|w_\mu) + \mathcal{N} \tag{9}$$

where $\mathcal{N}$ represents the OU function.

The inputs of online network η of evaluation part are the state s_t and actions a_t, and the output is the estimated value Q. The inputs of target network $\eta^{'}$ are the state s_{t+1} and actions a_{t+1}, and its output $Q^{'}$ is used to calculate the target value y_t by Eq. (10-a). The loss function $L(w)$ of online network is as follows

$$y_t = r_t + \gamma \cdot Q^{'}(s_{t+1}, a_{t+1}; w_{\eta^{'}}) \tag{10-a}$$

$$L(w) = \frac{1}{n}\sum_{i=1}^{n}(y_i - Q(s_i, a_i; w_\eta))^2 \tag{10-b}$$

where r_t is the immediate reward, γ is the discount factor, n is the number of samples extracted, y_i is the target value of the i-th sample.

The ultimate goal of this paper is to build a nonparametric model similar to an end-to-end model. In this model, the ship speed and rudder angle at the current moment serve as the input, and the output is the ship speed at the next moment. As seen, this process involves only one-step interaction, and the optimal strategy can then be characterized as a set of optimal solutions for each one-step interaction aligning with the goal of DRL, which aims to maximizing the corresponding immediate reward for each state-action pair. Therefore, the target value is redefined in a way that retains only the first term on the right side of Eq. (10-a).

$$y_t = r_t \tag{11}$$

The online network parameters of evaluation part are updated by loss function $L(w)$ using the gradient descent method. The online network parameters of execution part are updated by the gradient derived from the deterministic policy gradient principle with the following equation

$$\nabla_{w_\mu} J(\mu) \approx \frac{1}{n}\sum_{i=1}^{n} \nabla_{w_\mu}\mu(s|w_\mu)\big|_{s=s_t} \cdot \nabla_a Q(s, a|w_\eta)\big|_{s=s_t, a=\mu(s_t|w_\mu)} \tag{12}$$

The target network parameters for execution part and evaluation part are updated using the soft update method as follows

$$\begin{cases} w_\mu \leftarrow \tau \cdot w_\mu + (1-\tau)\cdot w_{\mu^{'}} \\ w_\eta \leftarrow \tau \cdot w_\eta + (1-\tau)\cdot w_{\eta^{'}} \end{cases} \tag{13}$$

where τ is the soft update factor.

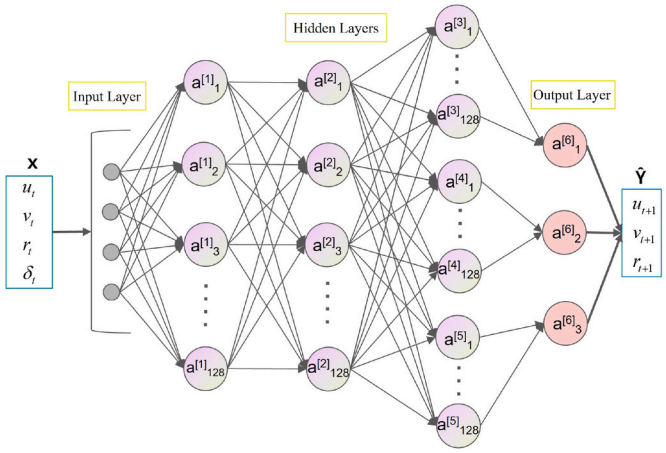


Fig. 2. Structure of the nonparametric model.

3.2. PER mechanism

The core idea of PER is to improve the learning efficiency of the agent by assigning higher priority to important samples, thereby increasing their likelihood of being selected during the training phase. The importance of the sample is determined by its TD-error, where a larger TD-error indicates a greater significance of the sample for the learning process of the agent. However, samples with larger TD-error tend to occur less frequently during training. By increasing the probability of selecting these high TD-error samples, the overall utilization of the samples is improved, leading to a higher quality of the learned strategy.

The TD-error (T_e) of the sample (e) is defined as

$$T_e \triangleq r_t + \gamma \cdot Q^{'}(s_{t+1}, a_{t+1}; w_{\eta^{'}}) - Q(s_t, a_t; w_\eta) \tag{14}$$

To address the limitations associated with using only TD-error such as the lack of diversity and bias issue, two methods are introduced, including the stochastic sampling method and the importance sampling method. The stochastic sampling method ensures that all samples in the replay memory buffer have a non-zero probability of being sampled while guaranteeing the diversity of training data. Concretely, the probability $p(e)$ of the sample (e) is defined as

$$p(e) = \frac{p_e^\alpha}{\sum_k p_k^\alpha} \tag{15-a}$$

$$p_e = |T_e| + \varepsilon \tag{15-b}$$

where the exponent α determines the degree of prioritization used, p_e is the priority of the sample (e) and ε is the small positive constant.

The importance sampling method is used to correct this bias through adding importance-sampling (IS) weights

$$w_{IS\_e} = \left(\frac{1}{N}\cdot\frac{1}{p(e)}\right)^\beta \tag{16}$$

where N is the total number of samples and the exponent β is annealed from its initial value to 1 in practice.

3.3. Improved PER-DDPG identification method

Due to the varying importance of samples during agent training, uniform sampling from the replay memory buffer in the original DDPG algorithm often leads to a low utilization rate. To address this issue, the combination of the DDPG algorithm and the Prioritized Experience Replay (PER) mechanism has been proven effective in enhancing the retrieval of valuable samples and delivering superior performance in related problem domains (Wu et al., 2018; Mo et al., 2019; Wei et al.,

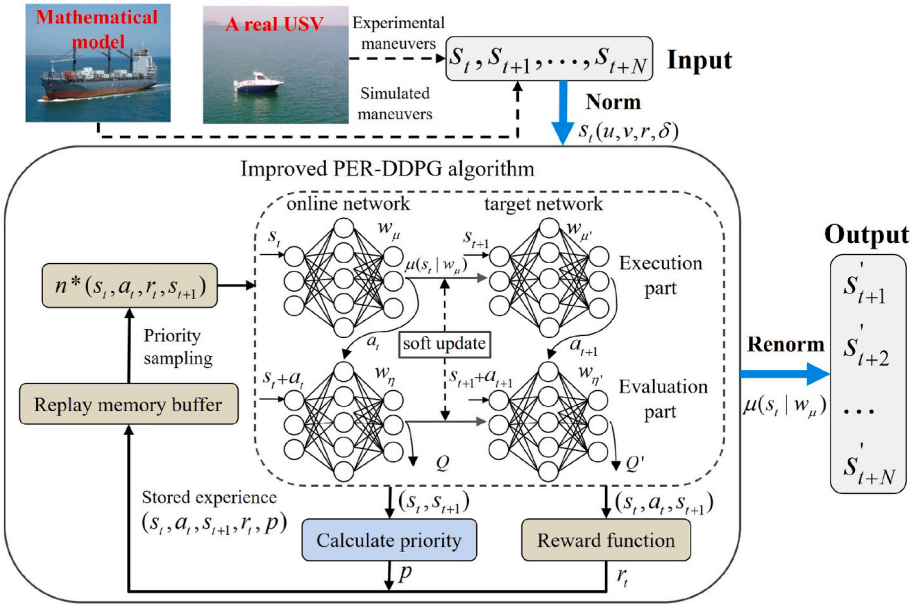


Fig. 3. Procedure of nonparametric modeling of ship dynamics applying the improved PER-DDPG method.

2021; Tang et al., 2022; Zhang et al., 2023). Therefore, this paper developed an improved PER-DDPG algorithm by taking advantages of the aforementioned improved DDPG algorithm and PER mechanism. The loss function Eq. (10-b) is also modified as follows

$$J(w)=\frac{1}{n}\sum_{i=1}^{n} w_{IS\_i}(y_i - Q(s,a;w_\eta))^2 \tag{17}$$

Fig. 2 illustrates the detailed model structure derived from the online network of the execution part in the improved PER-DDPG. The model comprises an input layer, a hidden layer, and an output layer. The hidden layer consists of multiple fully connected (FC) layers with 128 neurons, while the input and output layers contain four neurons and three neurons, respectively. The number of neurons in the input and output layers corresponds to the size of the state space and action space, respectively. The output layer employs the sigmoid and tanh functions as activation functions. Concerning the online network structure of the evaluation part, there are seven neurons in the input layer and one neuron in the output layer. The hidden layers consist of five fully connected (FC) layers, each with the same number of neurons as before. Specifically, the state and action information are separately processed by two FC layers and one FC layer to extract their features. The features from both parts are then concatenated to form a complete feature, which is subsequently passed through two FC layers to produce the final Q value. In terms of hyperparameters, the replay memory buffer size, batch update size, and soft update factor are set to 1000, 128, and 0.005, respectively. For the remaining hyperparameters, please refer to the works of Lillicrap et al. (2015) and Schaul et al. (2016). In addition, it is worth noting that the obtained nonparametric model is a multiple-input multiple-output (MIMO) model. This MIMO structure allows for the simultaneous prediction of motions across all degrees of freedom (DOFs), thus avoiding the accumulation of errors that can occur in single-output models incapable of predicting all DOFs simultaneously (Zhang et al., 2022a,b). Furthermore, Eq. (5) can be expressed in the following form

$$\left.\begin{aligned} u(t+1) &= h_1[u(t),v(t),r(t),\delta(t)] \\ v(t+1) &= h_2[u(t),v(t),r(t),\delta(t)] \\ r(t+1) &= h_3[u(t),v(t),r(t),\delta(t)] \end{aligned}\right\} \Rightarrow \widehat{\mathrm{Y}} = \begin{bmatrix} u(t+1) \\ v(t+1) \\ r(t+1) \end{bmatrix} = g(\mathrm{X}) \tag{18}$$

where $\widehat{\mathrm{Y}}$ is the output vector, X is the input vector, and $g(\cdot)$ is the nonlinear function.

Fig. 3 illustrates the step-by-step procedure of applying the improved PER-DDPG method to construct the nonparametric model of ship dynamics. The detailed process is as follows: Firstly, the agent begins feature learning once it receives the normalized state. Next, the output corresponding to the 3-DOF ship motion is obtained by applying the actor policy and undergoing renormalization. Simultaneously, an immediate reward is assigned, and a new state is generated for the subsequent one-step interaction. Finally, the output is evaluated by the evaluation part, which plays a crucial role in the algorithm's convergence. Additionally, the priority of the previous state is calculated. Concurrently, the output, previous state, priority, immediate reward, and new state are packaged together as a sample and stored in the replay memory buffer. Once the number of samples reaches a specified threshold, the neural networks are iteratively trained through batch updating, employing priority sampling. This iterative training process aims to obtain the optimal strategy.

3.4. Evaluation indexes

To measure the performance of the proposed method, the cumulative rewards, the training time (T-time), the time consumed for prediction (P-time), root mean square error (RMSE), and determination coefficient (R^2) are selected as evaluation indexes. Whether the algorithm converges is judged visually by the cumulative rewards curve, the prediction accuracy of models is reflected by RMSE and R^2 and relative formulas as follows

$$RMSE(y,\hat{y}) = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(y_i - \hat{y}_i)^2} \tag{19-a}$$

$$R^2(y,\hat{y}) = 1 - \frac{\sum_{i=1}^{k}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{k}(y_i - \bar{y})^2} \tag{19-b}$$

where y_i is the true value of the i-th data point, $\hat{y}_i$ is the corresponding predicted value and $\bar{y}$ is the average value of all data.

4. Case study

To demonstrate the effectiveness of the nonparametric model identified using the proposed improved PER-DDPG identification method, two case studies including a numerical simulation study on a Mariner cargo ship (Fossen, 2011) and an experimental study on a real USV are

Table 1
Main parameters of the Mariner cargo ship.

Parameters	Value	Unit
Length	171.80	m
Length between perpendiculars	160.93	m
Breadth	23.17	m
Design draft	8.23	m
Design displacement	18,541	m^3
Design speed	15	knots
Maximum rudder angle	±40	deg
Rudder rate	5	deg/s

carried out. The Mariner case is to verify the performance of the proposed method in several aspects, in which a Mariner cargo ship with known parameter values is used to simulate a series of zigzag maneuvers and turning circle maneuvers for which measure noises and different sampling frequencies are considered. The USV case is for more in-depth research on the proposed method by using experimental data from real sea conditions.

4.1. Simulation study

In this subsection, the feasibility of the proposed improved PER-DDPG method in identifying the model is proved by employing the model of the Mariner cargo ship, among which three groups of simulation experiments are conducted. The particulars are set as follows: maneuvering indices at a forward speed of 7.7175 m/s, v = 0 m/s, r = 0 deg/s, rudder deflection limitation $\delta \in [-20^\circ, +20^\circ]$. The main parameters of the model are listed in Table 1. All simulation studies are carried out in a uniform environment running PyCharm 2022 version with an Intel(R) Core(TM) i7-12700F, 2.10 GHz CPU.

The first group of experiments is to verify and validate the effectiveness and generalization of the proposed improved PER-DDPG method, from which conducting a 20°/20° zigzag test and a 20° turning circle test, and comparisons are made with the original simulated maneuvers, the improved DDPG and LSTM (Jiang et al., 2022). Each test obtained 7000 data points with a time interval of 0.1s, the first 71.4% of zigzag test data is classified as the training dataset and the remaining data is the validation dataset with referencing to (Wang et al., 2020; Ouyang and Zou, 2021; Ouyang et al. 2023b). The data of turning circle test is used as testing dataset to further validate the generalization ability of these identified models.

The second group is designed to test the robustness of the identified model against environmental disturbances and measurement noises. To simulate these uncertainties, Gaussian white noise with varying levels is added to the zigzag test data. This approach is inspired by the method used in Wang et al. (2019).

$$\zeta_i = \zeta_{oi} + \zeta^{\max} k_0 k_\zeta \xi_i \tag{20}$$

where ζ_i is the noise data, ζ_{oi} is the zigzag test data. $\zeta^{\max} k_0 k_\zeta \xi_i$ is the disturbance term, where $\zeta^{\max}$ is the maximum absolute value of the zigzag test data, k_ζ is the reduction factor of different responses, which is set to 0.03 for u, 0.01 for δ, and 0.25 for other responses. ξ_i is a randomly selected value from Gaussian white noise with the variance of 1 and mean of 0. k_0 is the reduction factor representing different noise levels, which is chosen as 1%, 5%, and 10%, and three cases are referred as Noise Level 1 (Nlv1), Noise Level 2 (Nlv2) and Noise Level 3 (Nlv3), respectively. The corresponding simulated data and noise data are shown in Fig. 4.

The third group is to prove the applicability of the proposed identification method on the premise of considering the different sampling frequencies problem induced by different levels of ships. Besides the sampling frequency of 10 Hz, this work plans to additionally set three different sampling frequencies in the above zigzag test, i.e. 1 Hz, 5 Hz, and 20 Hz. Three sets of data are stemmed from several time-series zigzag maneuvers conducted, the number of data points obtained are 700, 3500, and 14,000 respectively, and we can find that the differences of data in quantity and quality are obvious. The division of training dataset and validation dataset is set as same as the ones applied in the first group.

In order to solve the negative effects of different dimension caused problems such as algorithm instability or model accuracy degradation, the "MaxAbs" scaler (He et al., 2022) is used to normalize the dataset from the above tests, from which the key point is that this method has the advantage of not changing the data distribution. It is defined as

$$X_{\text{norm}} = \frac{X}{|X_{\max}|} \tag{21}$$

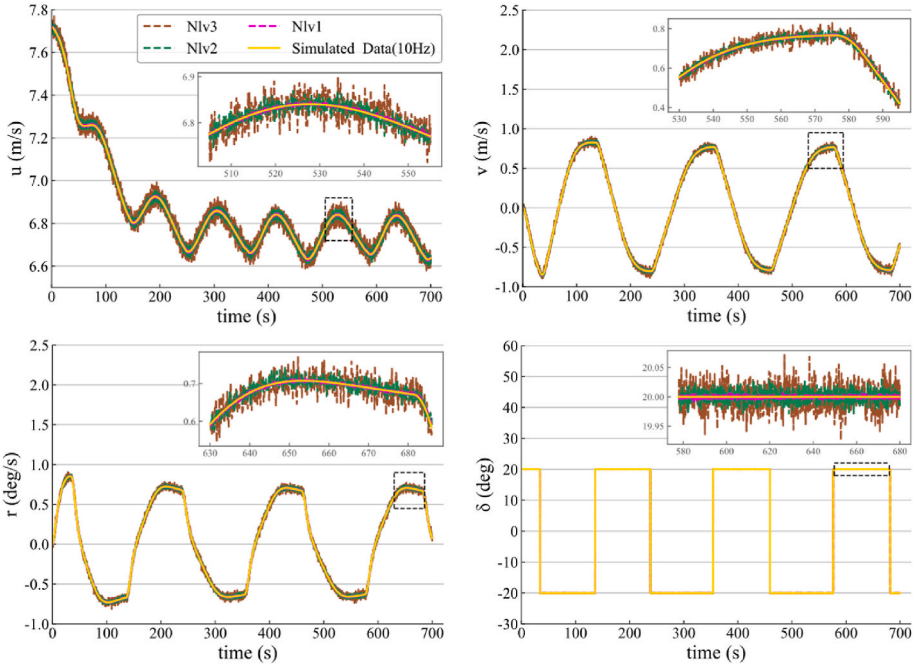


Fig. 4. Testing dataset with Gaussian white noise of different levels.

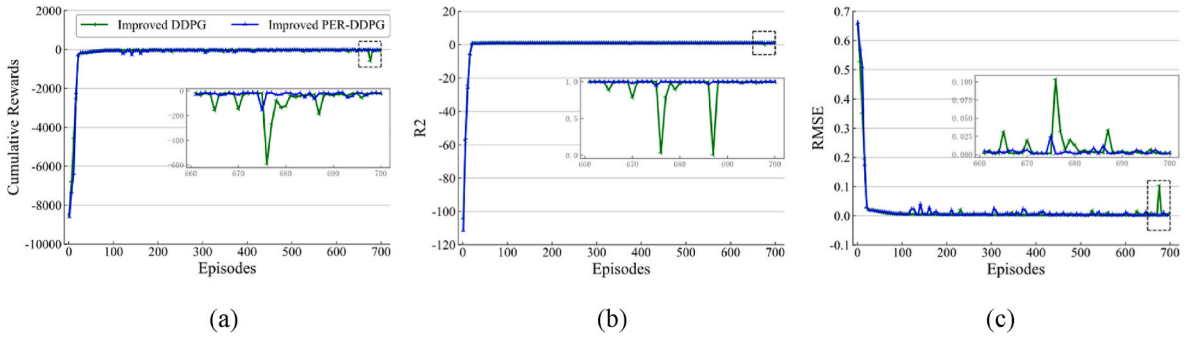


Fig. 5. Fluctuation of evaluation indexes of ship dynamic models identified by improved DDPG method and PER-DDPG method respectively.

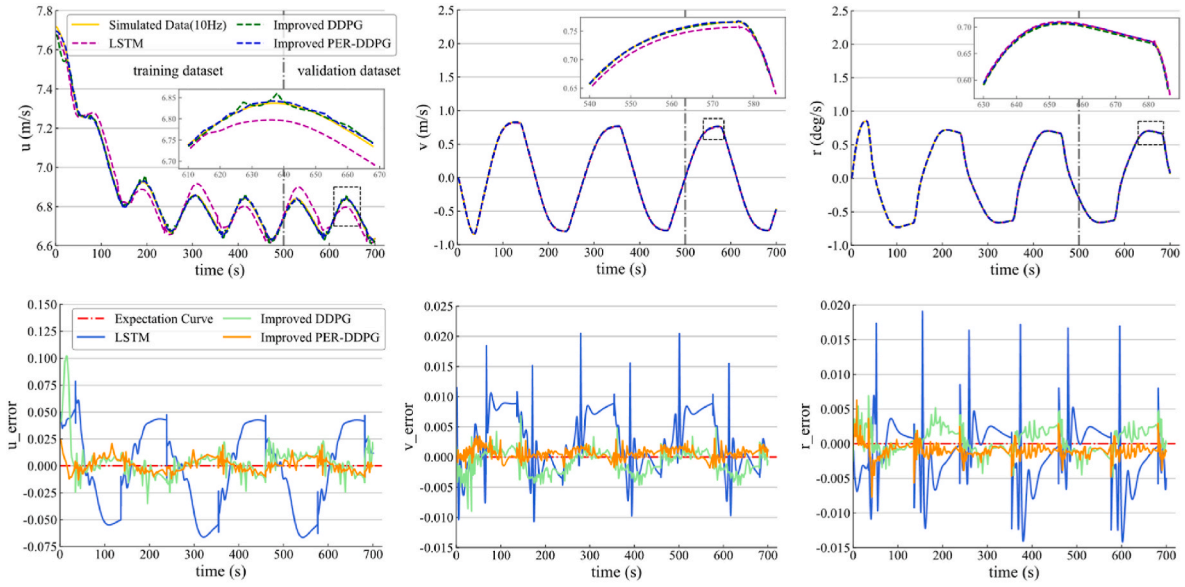


Fig. 6. Prediction results of 20°/20° zigzag maneuver.

Table 2
Accuracy analysis of simulation prediction with zigzag maneuver.

Methods	Cumulative rewards	Training dataset		Validation dataset		T-time (h)	P-time (s)
		RMSE	R^2	RMSE	R^2		
LSTM	–	0.016	0.993	0.017	0.874	**0.02↓**	3.00×10^{-3}
Improved DDPG	−30.523	0.008	0.998	0.005	0.992	5.76	9.24×10^{-4}
Improved PER-DDPG	**−18.938↑**	**0.003↓**	**1.000↑**	**0.003↓**	**0.997↑**	7.48	$\mathbf{9.04 \times 10^{-4}}$**↓**

where X_{norm} is the normalized data, X is the input data, $X_{\max}$ is the maximum absolute value of input data.

In the first group of experiments, after the processing of iterative training (700 episodes), the dynamic curves of cumulative rewards, R^2, and RMSE indexes are obtained with the use of two different identification methods, as shown in Fig. 5. We can find from Fig. 5a that both algorithms reach the convergence at the late stage of training but improved PER-DDPG algorithm performs better in terms of the stability and the accuracy, and the number of episodes needed is respectively about 400 (3.3 h) and 290 (3.1 h), it reveals that the training cost of these two different identification methods is considerably small. Fig. 5b and c illustrate the predictive effect of models is constantly improving because their RMSE value and R^2 value are gradually approaching 0 and 1.

The prediction results and relevant errors of the LSTM identified model, improved DDPG identified model and improved PER-DDPG identified model on the training dataset and validation dataset are visualized in Fig. 6, from which it can be found that the nonparametric model identified by the proposed improved PER-DDPG method approximates well the validation dataset with high accuracy. Besides, these three identification methods identify the ship dynamic model well with training dataset but the improved PER-DDPG method outperforms the other methods since the evaluation indexes presented in Table 2 are best except for the T-time, and the increase of its T-time is acceptable. It is deserved to be highlighted that the model identified by the LSTM and improved DDPG method has slightly larger errors in the surge motion, the two reasons for this phenomenon lie in that first the surge motion is more sensitive to the direction change of rudder angle, the other is that the agent has not learned full ship motion mechanism because the training dataset lacks sufficient data containing the surge speed at the peak or trough, and this also explains why we introduced PER mechanism. Therefore, the higher predictive performance of the improved PER-DDPG model is due to capturing more the characteristics of ship

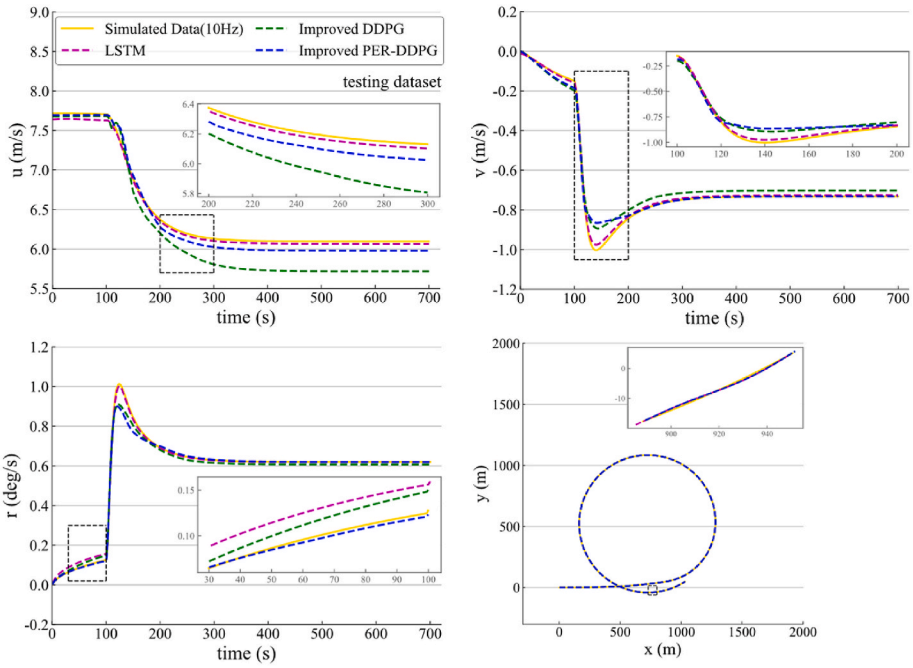


Fig. 7. Prediction results of 30° turning circle maneuver.

Table 3
Results analysis of simulation prediction of turning circle maneuver.

Methods	Cumulative rewards	Testing dataset		P-time (s)
		RMSE	R^2	
LSTM	–	**0.013**↓	**0.998**↑	3.00×10^{-3}
Improved DDPG	−607.398	0.121	0.907	9.36×10^{-4}
Improved PER-DDPG	**−230.730**↑	0.052	0.982	**9.21** × **10^{-4}**↓

motion by learning enough times of important samples.

Fig. 7 depicts the predictions of ship speed components and trajectory using a testing dataset extracted from 30° turning circle maneuvers for the three identified models. Table 3 records the P-time, RMSE, and R^2 indexes of these models. The LSTM model exhibits outstanding performance in terms of ship speed component and trajectory prediction, demonstrating low RMSE values of approximately 0.013 and high R^2 values around 0.998. The superior performance of the LSTM model can be attributed to its network structure, which effectively captures dependencies between adjacent states. Furthermore, the discrepancy between the LSTM model and the improved PER-DDPG model in terms of RMSE and R^2 indexes is minor. However, the improved PER-DDPG model offers higher computational efficiency. In contrast, the improved DDPG model's prediction accuracy is noticeably reduced in surge motion. This can be attributed to the training dataset being derived solely from single standard zigzag maneuvers, rather than a combination of multiple standard maneuvers that cover the entire input space of the steering angle. The zigzag tests with varying rudder angles encompass diverse ship motion characteristics, as discussed in Wang and Zou (2018) and Wang et al. (2019). Nevertheless, the discrepancies in

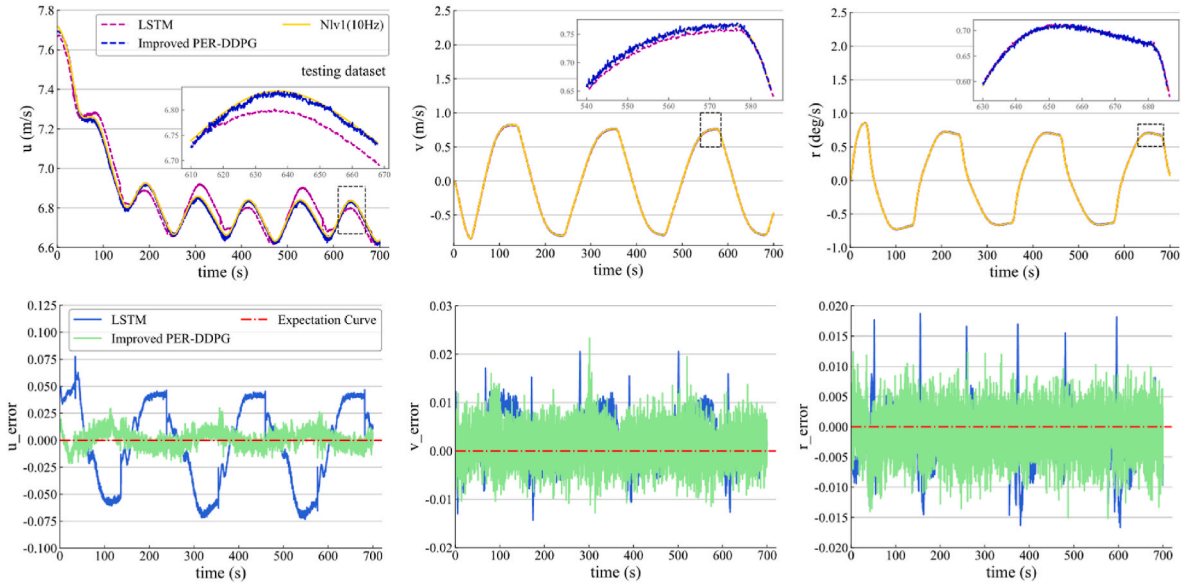


Fig. 8. Prediction results of the identified model on Nvl1.

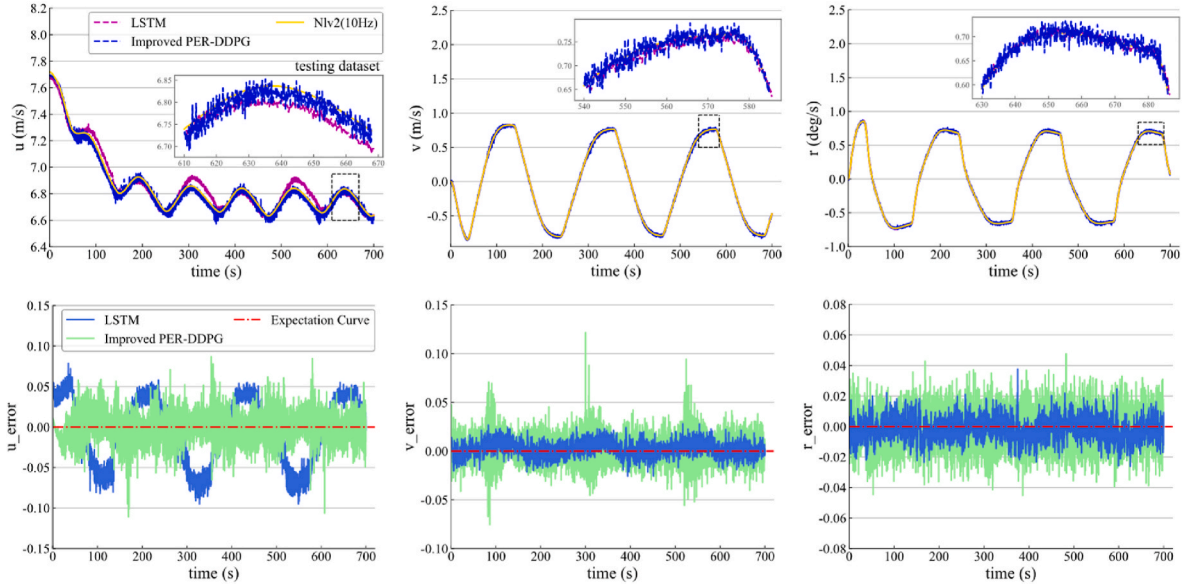


Fig. 9. Prediction results of the identified model on Nvl2.

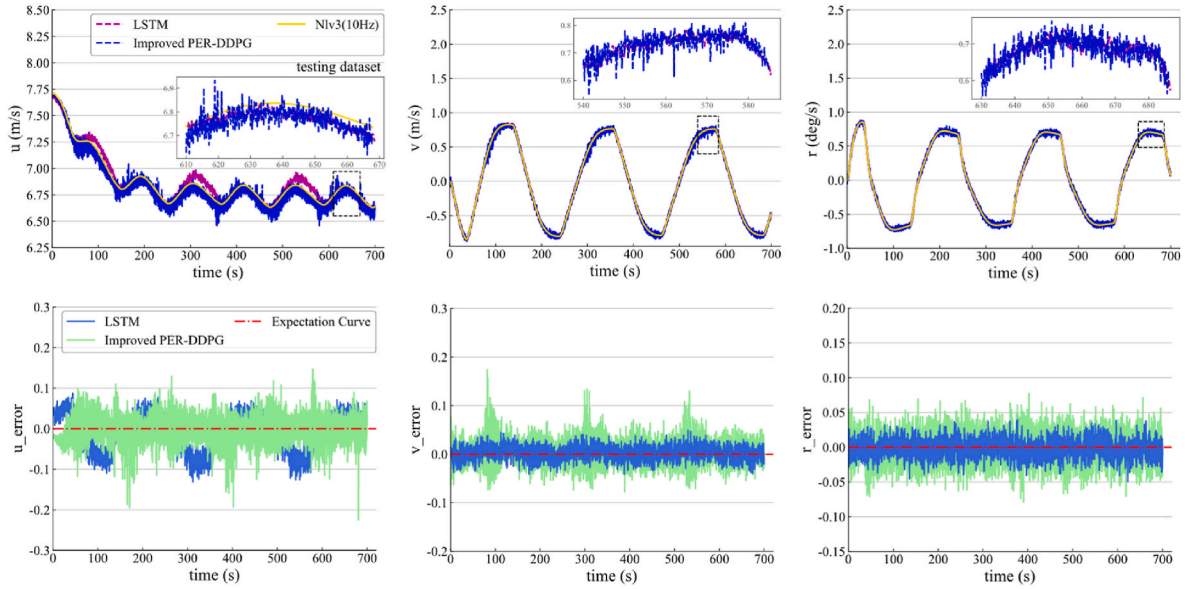


Fig. 10. Prediction results of the identified model on Nvl3.

Table 4
Results analysis of model identified under the condition of noise with different levels.

Methods	Cumulative rewards	Testing dataset		P-time (s)
		RMSE	R^2	
LSTM (Nlv1)	–	0.017	0.991	4.00×10^{-3}
Improved PER-DDPG (Nlv1)	−47.672	**0.004↓**	**1.000↑**	**8.83×10^{-4}↓**
LSTM (Nlv2)	–	0.019	0.990	2.00×10^{-3}
Improved PER-DDPG (Nlv2)	−213.284	**0.015↓**	**0.998↑**	**8.89×10^{-4}↓**
LSTM (Nlv3)	–	0.027	0.987	3.00×10^{-3}
Improved PER-DDPG (Nlv3)	−425.248	**0.025↓**	**0.991↑**	**8.92×10^{-4}↓**

the training dataset indicate that the improved PER-DDPG model possesses satisfactory reasoning and generalization abilities.

Therefore, the above results demonstrate the feasibility, effectiveness, and advantages of the improved PER-DDPG method in accurately predicting zigzag maneuvers and turning circle maneuvers. This paper will further focus on verifying and comparing the method in subsequent work.

In the second group of simulation tests, the ship dynamic model identified by the LSTM model and improved PER-DDPG are evaluated on the testing dataset with different noise levels. The prediction results, along with the corresponding errors, are presented in Figs. 8–10, and the evaluation indexes are summarized in Table 4. It can be observed that these two identified models accurately predict the ship speeds under Noise Level 1 and Noise Level 2 but the improved PER-DDPG model outperforms over the other one model since the evaluation indexes namely RMSE, R^2 and P-time are smallest. For the Noise Level 3 case where the noise level is increased, some deviation occurs, but the

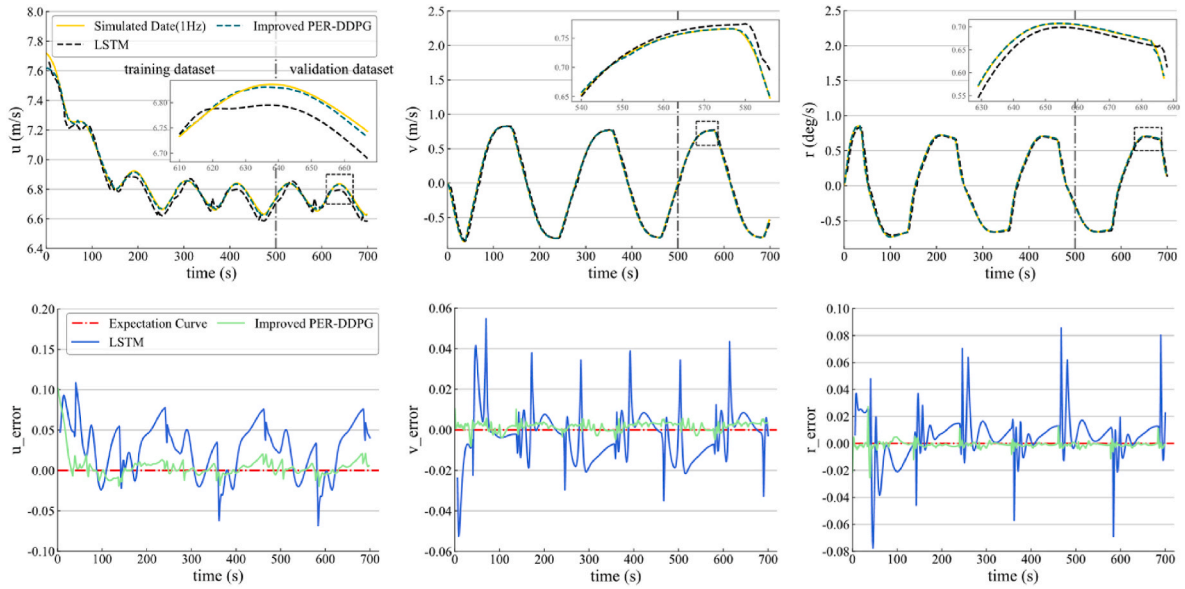


Fig. 11. Prediction results of ship dynamic model identified using dataset with 1Hz.

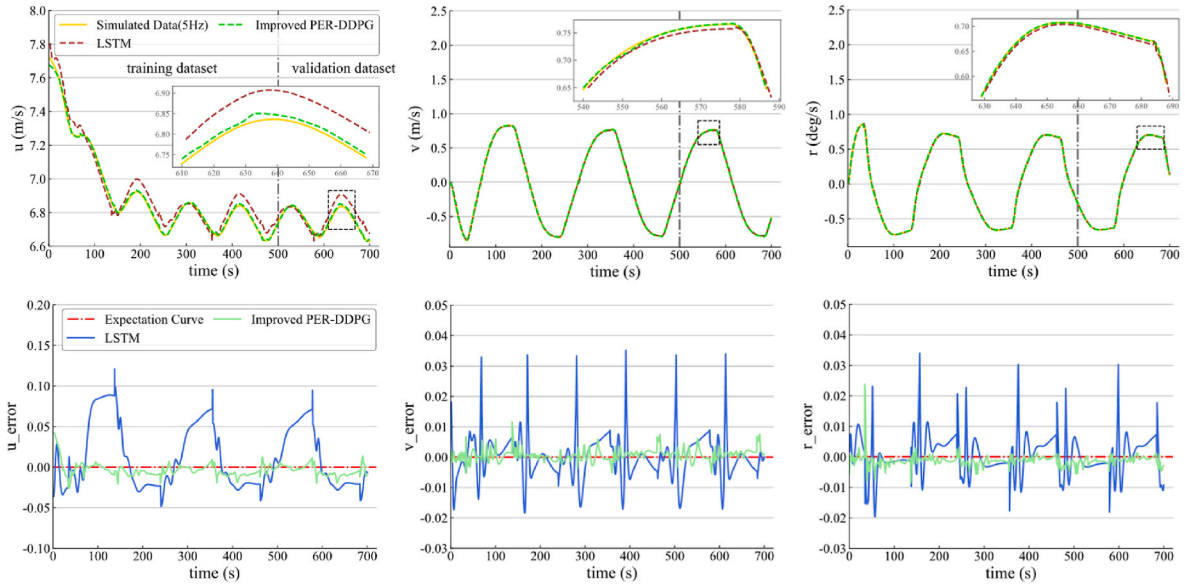


Fig. 12. Prediction results of ship dynamic model identified using dataset with 5Hz.

identified models still exhibit acceptable consistent trends of ship speed prediction. Comprehensively speaking, in the results, the largest RMSE of the improved PER-DDPG model is 0.025 which is small enough and demonstrates that the identified dynamic model of the ship is accurate to describe ship dynamics under the disturbance condition and the improved PER-DDPG based identification method is robust.

In the third group of simulation tests, the predictions of identified ship dynamic model and relevant errors of 20°/20° zigzag maneuvers with different sampling frequencies are shown in Fig. 6, Figs. 11–13 and their corresponding indexes are listed in Table 5. It can be seen that the improved PER-DDPG method enables the construction of high-precision ship dynamic models using ship maneuver data with different levels of sensors. This illustrates that the method is applicable to both a small number of samples (about 500 data points) and a large number of samples (about 10,000 data points) and shows applicability. Furthermore, as the training dataset becomes more diverse, the T-time cost of the model increases significantly since the off-line training mode needs to traverse the entire dataset in each training episode but the performance of the identified models is further enhanced. Comparatively, the larger prediction error on two different datasets with low sampling frequencies such as 5 Hz, and 1 Hz show that the identified models obtained by LSTM method rely on large amounts of data.

4.2. Experimental study

Experimental studies are also conducted to further demonstrate the effectiveness and robustness of the proposed improved PER-DDPG method in identifying the ship dynamics model for a real USV. Fig. 14 depicted the USV whose main particulars are shown in Table 6. The USV is equipped with various sensors such as GPS, IMU, compass, and rudder indicator to measure the surge speed, sway speed, yaw rate, and rudder angle, from which these data are extracted for the identification of the nonparametric model for the USV. The corresponding free-running model tests are carried out in the Weihai of China under less than

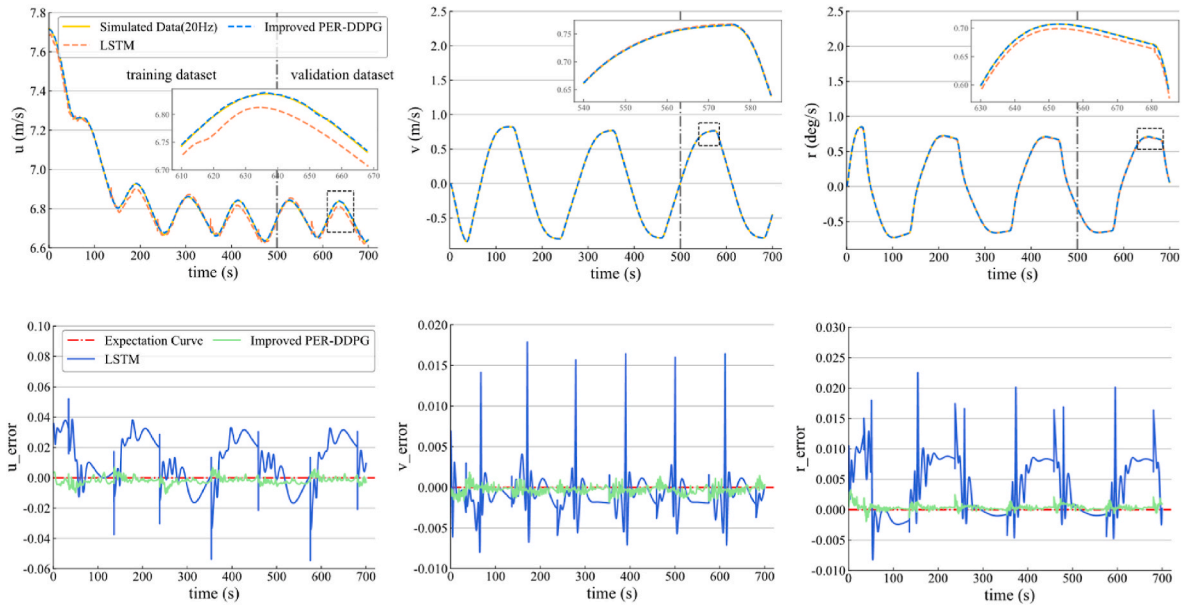


Fig. 13. Prediction results of ship dynamic model identified using dataset with 20 Hz.

Table 5
Evaluation indexes of ship dynamic models identified by the LSTM and improved PER-DDPG using datasets with different sampling frequencies.

Methods	Cumulative rewards	Training dataset		Validation dataset		T-time (h)	P-time (s)
		RMSE	R^2	RMSE	R^2		
LSTM (1Hz)	–	0.025	0.973	0.038	0.556	**0.005↓**	3.00×10^{-3}
Improved PER-DDPG (1Hz)	−3.483	**0.008↓**	**0.999↑**	**0.004↓**	**0.995↑**	0.52	$\mathbf{9.25 \times 10^{-4}}$**↓**
LSTM (5Hz)	–	0.019	0.992	0.017	0.892	**0.01↓**	4.00×10^{-3}
Improved PER-DDPG (5Hz)	−13.424	**0.005↓**	**1.000↑**	**0.004↓**	**0.996↑**	3.27	$\mathbf{9.28 \times 10^{-4}}$**↓**
LSTM (10Hz)	–	0.016	0.993	0.017	0.874	**0.02↓**	3.00×10^{-3}
Improved PER-DDPG (10Hz)	−18.938	**0.003↓**	**1.000↑**	**0.003↓**	**0.997↑**	7.48	$\mathbf{9.30 \times 10^{-4}}$**↓**
LSTM (20Hz)	–	0.009	0.998	0.009	0.975	**0.04↓**	3.00×10^{-3}
Improved PER-DDPG (20Hz)	−16.223	**0.001↓**	**1.000↑**	**0.001↓**	**0.999↑**	38.79	$\mathbf{10.4 \times 10^{-4}}$**↓**

Fig. 14. The Unmanned Surface Vessel used in the tests.

Table 6
Main particulars of the USV.

Elements	Values	Unit
Length	7.6	M
Width	2.58	M
Draught	0.37	M
Revolution	5300–6300	Rpm
Max power	183.9	kW
Max speed	39	Knot

level 3 sea conditions. Two sets of experimental data are stemmed from a time-series maneuver conducted at a steady forward speed varying around 2.5 m/s, and the sampling time is 0.1 s. Samples (1240 data points) obtained from 15°/15° zigzag maneuver are classified as training dataset and validation dataset which are presented in Fig. 15, and the division between the two datasets is set as same as before. Samples (400 data points) obtained from 10°/10° zigzag maneuver are used as testing dataset.

After training the proposed method and LSTM method to acquire two nonparametric models of the USV, the experimental results including the evaluation indexes, predictions, and relevant errors of 15°/15° zigzag maneuver are obtained, which are respectively presented in Table 7 and Fig. 16. From the validation results, it can be observed that these two identified models achieve a high level of predictive accuracy and have the advantages of low T-time and P-time cost but the improved PER-DDPG model is best compared to the LSTM model in terms of core evaluation indexes such as RMSE, R^2, and P-time, despite the slightly larger RMSE value compared to the simulation study with the same sampling frequency. This can be attributed to environmental disturbances caused by factors such as wind, currents, and waves, as well as the smaller number of samples in the experimental study. Furthermore, slight discrepancies between the prediction results and the validation dataset can be observed. However, these discrepancies remain within a tolerable range, which further indicates the effectiveness of the improved PER-DDPG method in identifying the ship dynamics model for the USV while considering uncertain environmental disturbances.

Fig. 17 illustrates the prediction results and associated errors

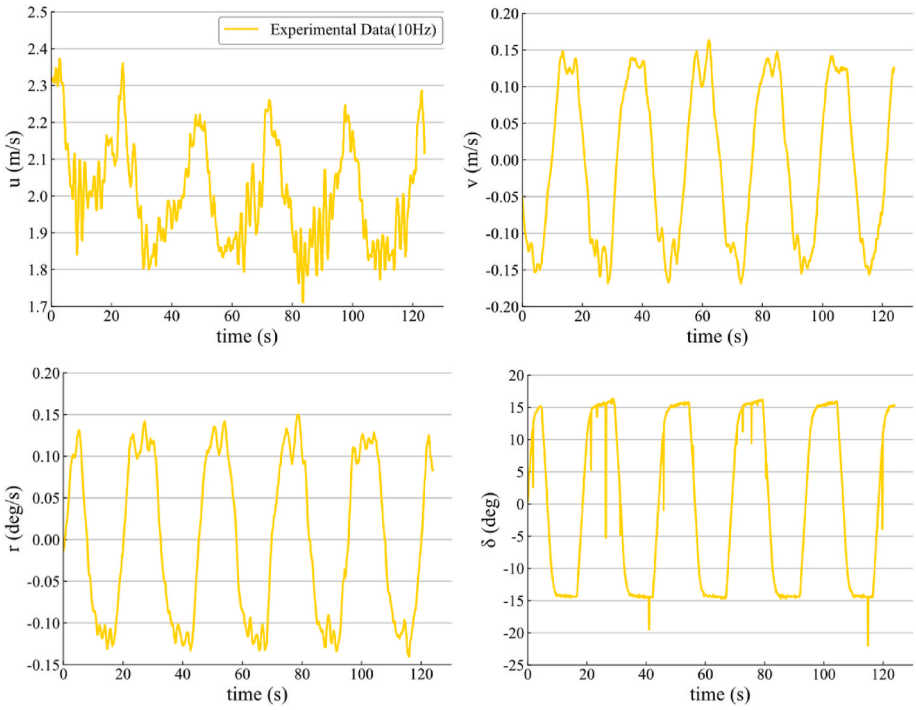


Fig. 15. Experimental data of 15°/15° zigzag maneuver.

Table 7
Evaluation indexes of the identified USV dynamic model in the 15°/15° zigzag test.

Methods	Cumulative rewards	Training dataset		Validation dataset		T-time (h)	P-time (s)
		RMSE	R^2	RMSE	R^2		
LSTM	–	0.022	0.962	0.026	0.942	**0.008↓**	4×10^{-3}
Improved PER-DDPG	−16.602	**0.010↓**	**0.992↑**	**0.012↓**	**0.984↑**	2.29	$\mathbf{9.02 \times 10^{-4}}$**↓**

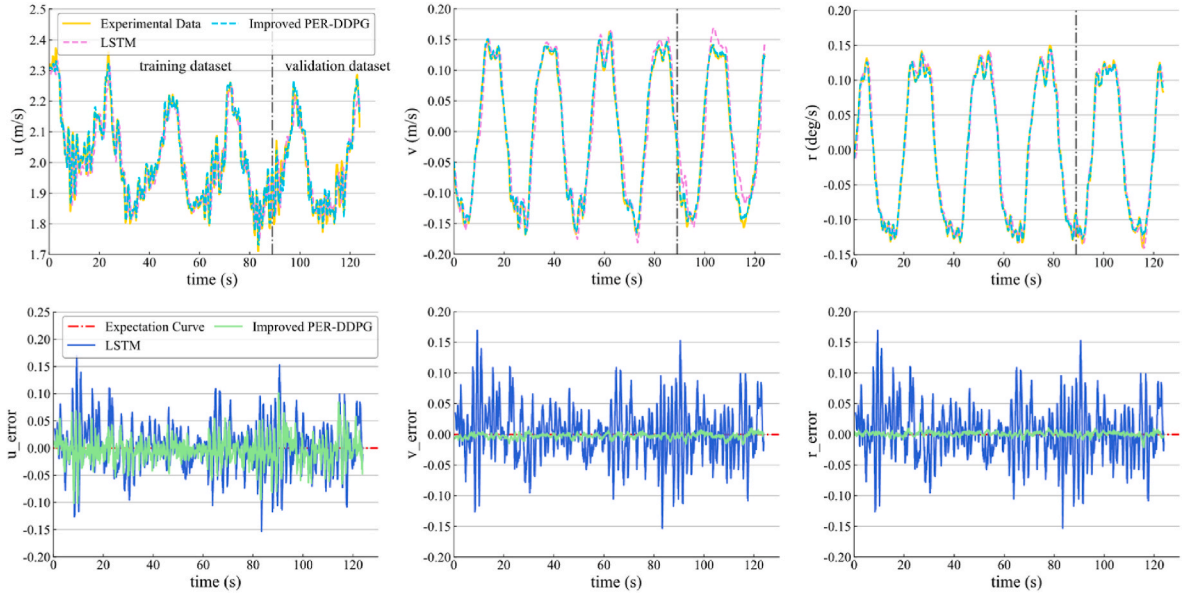


Fig. 16. Prediction results of 15°/15° zigzag maneuver using identified USV dynamic model.

obtained from two identified USV dynamic models, using experimental data from a 10°/10° zigzag maneuver as the testing dataset. The corresponding performance metrics are recorded in Table 8. It is evident from the results that the improved PER-DDPG identified model, trained with the 15°/15° zigzag maneuver, is capable of accurately and comprehensively predicting the 10°/10° zigzag maneuver. This finding confirms the excellent generalization ability of the improved PER-DDPG method, particularly when dealing with experimental data that

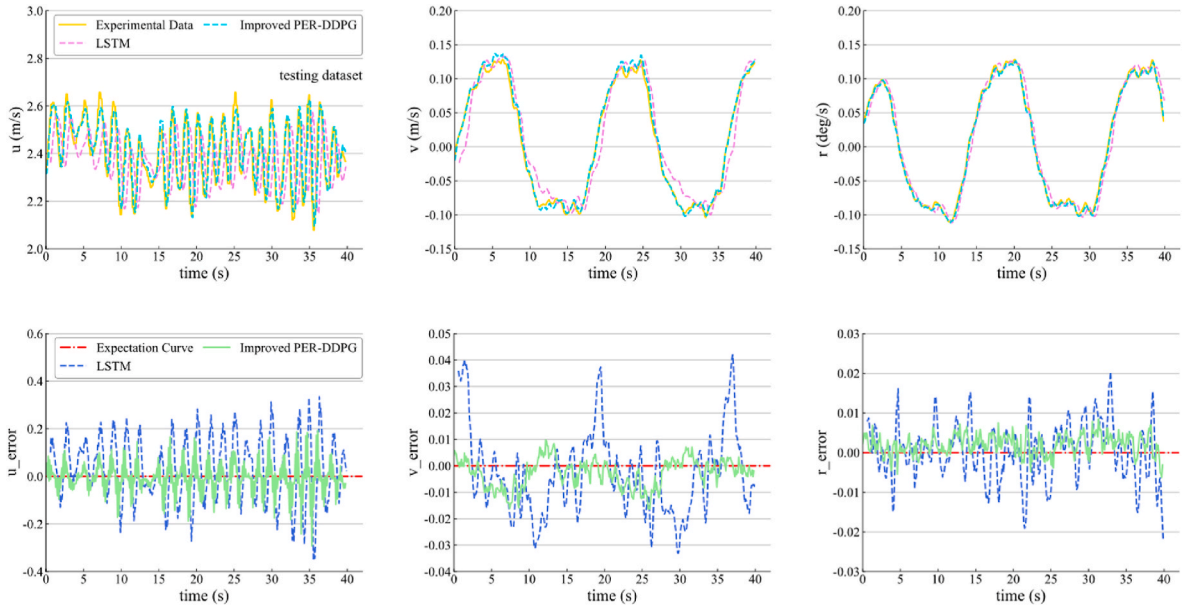


Fig. 17. Prediction results of 10°/10° zigzag maneuver using identified USV dynamic model.

Table 8
Evaluation indexes of the identified USV dynamic model in the 10°/10° zigzag test.

Methods	Cumulative rewards	Testing dataset		P-time (s)
		RMSE	R^2	
LSTM	–	0.053	0.652	3×10^{-3}
Improved PER-DDPG	−9.533	**0.025↓**	**0.919↑**	**9.24×10^{-4}↓**

incorporates uncertain environmental disturbances. However, the LSTM model exhibits lagged prediction results for each ship motion, failing to meet the requirements of real-world applications.

5. Conclusions

In this paper, a method of nonparametric modeling based on PER mechanism and DDPG algorithm for the ship dynamic model is proposed, which is used to capture and predict ship dynamics only using maneuver data. Firstly, a ship dynamic black-box model with maneuvering variables and ship responses is built, which is further modified into a form for identification. Then through the training of the improved PER-DDPG method with the use of pure data and noise data, the nonparametric dynamic models of ships including Mariner ship and a real USV are identified, and the model predictions are satisfactory. On the whole, the results of ship dynamic prediction are relatively accurate. It can be considered that the ship nonparametric dynamic model established by the proposed identification method has high robustness and can provide the corresponding degree of dynamic prediction for ships with different levels of disturbance.

Compared with the parametric modeling, this nonparametric modeling based on the improved PER-DDPG method is convenient and robust to capture and describe dynamics of ships with uncertainty induced by environmental disturbance or measurement noises. However, it also has certain limitations. The model can only be used for one-step deterministic prediction with the speed components and rudder angle as input. So it is worthy to test multi-step prediction. Additionally, the uncertainty caused by environmental disturbance or measurement noises is emulated in this study, but in fact, the variation of ship navigation conditions such as loading conditions or navigation speed leads to the change of ship dynamic model parameters. This point is deserved to be investigated to test the performance of the improved PER-DDPG based identification method in the upcoming research work.

CRediT authorship contribution statement

Man Zhu: Conceptualization, Data curation, Investigation, Writing – original draft. **Kang Tian:** Conceptualization, Methodology, Software, Writing – original draft. **Yuan-Qiao Wen:** Supervision, Writing – review & editing. **Ji-Ning Cao:** Writing – review & editing. **Liang Huang:** Writing – review & editing.

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.

Acknowledgements

This work is supported by the National Key Research and Development Program of China through Grant No. 2021YFC3101800, the National Natural Science Foundation of China (Grant No. 52001237, and No. U2141234), and Sanya Science and Education Innovation Park of Wuhan University of Technology (Grant No. 2021KF0030).

References

Abkowitz, M.A., 1964. Lectures on Ship Hydrodynamics - Steering and Manoeuvrability. Report No. Hy-5, Hydro-Og Aerodynamisk Laboratorium. Lyngby, Denmark.
Chen, C., Liu, H., 2021. Dynamic ensemble wind speed prediction model based on hybrid deep reinforcement learning. Adv. Eng. Inf. 48, 101290.
Chen, C.Y., Delefortrie, G., Lataire, E., 2021. Effects of water depth and speed on ship motion control from medium deep to very shallow water. Ocean Eng. 231, 109102.
Chen, L.J., Yang, P.Y., Li, S.W., Tian, Y.F., Liu, G.Q., Hao, G.Z., 2022. Grey-box identification modeling of ship maneuvering motion based on LS-SVM. Ocean Eng. 266, 112957.
Chen, L.J., Yang, P.Y., Li, S.G., Liu, K.Z., Wang, K., Zhou, X.W., 2023. Online modeling and prediction of maritime autonomous surface ship maneuvering motion under ocean waves. Ocean Eng. 276, 114183.

Fossen, T.I., 2011. Handbook of Marine Craft Hydrodynamics and Motion Control. John Wiley & Sons, Ltd.
Francis, D.P., Raimond, K., 2021. Major advancements in kernel function approximation. Artif. Intell. Rev. 54 (2), 843–876.
Habib, M.K., Ayankoso, S.A., Nagata, F., 2021. Data-driven modeling: concept, techniques, challenges and a case study. In: Proceedings of 2021 IEEE International Conference on Mechatronics and Automation. Takamatsu, Japan.
Hao, L.Z., Han, Y., Shi, C., Pan, Z.Y., 2022. Recurrent neural networks for nonparametric modeling of ship maneuvering motion. Int. J. Nav. Archit. Ocean Eng. 14, 100436.
He, H.W., Wang, Z.H., Zou, Z.J., Liu, Y., 2022. Nonparametric modeling of ship maneuvering motion based on self-designed fully connected neural network. Ocean Eng. 251, 111113.
Huang, H.J., Li, Z.X., Tian, J.L., Min, G.Y., Miao, W., Wu, D.O., 2023. Accurate prediction of required virtual resources via deep reinforcement learning. IEEE/ACM Trans. Netw. 31 (2), 920–933.
Jiang, Y., Hou, X.R., Wang, X.G., Wang, Z.H., Yang, Z.L., Zou, Z.J., 2022. Identification modeling and prediction of ship maneuvering motion based on LSTM deep neural network. J. Mar. Sci. Technol. 27 (1), 125–137.
Kiran, B.R., Sobh, I., Talpaert, V., Mannion, P., Sallab, A.A.A., Yogamani, S., Pérez, P., 2022. Deep reinforcement learning for autonomous driving: a survey. IEEE Trans. Intell. Transport. Syst. 23 (6), 4909–4926.
Liang, X.D., Luo, P., Li, X.Y., Wang, X., Shu, L.L., 2023. Crude oil price prediction using deep reinforcement learning. Resour. Pol. 81, 103363.
Lillicrap, T.P., Hunt, J.J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., Wierstra, D., 2015. Continuous Control with Deep Reinforcement Learning. https://doi.org/10.48550/arXiv.1509.02971.
Meng, Y., Zhang, X.K., Zhang, X.F., 2023. Identification modeling of ship nonlinear motion based on nonlinear innovation. Ocean Eng. 268, 113471.
Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., Riedmiller, M., 2013. Playing Atari with Deep Reinforcement Learning. https://doi.org/10.48550/arXiv.1312.5602.
Mo, S.J., Pei, X.F., Chen, Z.F., 2019. Decision-making for oncoming traffic overtaking scenario using double DQN. In: Proceedings of 3rd Conference on Vehicle Control and Intelligence. Hefei, China.
Ouyang, Z.L., Zou, Z.J., 2021. Nonparametric modeling of ship maneuvering motion based on Gaussian process regression optimized by genetic algorithm. Ocean Eng. 238, 109699.
Ouyang, Z.L., Liu, S.Y., Zou, Z.J., 2022. Nonparametric modeling of ship maneuvering motion in waves based on Gaussian process regression. Ocean Eng. 264, 112100.
Ouyang, Z.L., Chen, G., Zou, Z.J., 2023a. Identification modeling of ship maneuvering motion based on local Gaussian process regression. Ocean Eng. 267, 113251.
Ouyang, Z.L., Zou, Z.J., Zou, L., 2023b. Adaptive hybrid-kernel function based Gaussian process regression for nonparametric modeling of ship maneuvering motion. Ocean Eng. 268, 113373.
Schaul, T., Quan, J., Antonoglou, I., Silver, D., 2016. Prioritized experience replay. In: Proceedings of 4th International Conference on Learning Representations. Puerto Rico, USA.
Schoukens, J., Ljung, L., 2019. Nonlinear system identification: a user-oriented road map. IEEE Control Syst. Mag. 39 (6), 28–99.
Sivaraj, S., Rajendran, S., 2022. Heading control of a ship based on deep reinforcement learning (RL). In: Proceedings of OCEANS 2022-Chennai. Chennai, India.
Sun, Y.H., Yan, C., Xiang, X.J., Zhou, H., Tang, D.Q., Zhu, Y., 2023. Towards end-to-end formation control for robotic fish via deep reinforcement learning with non-expert imitation. Ocean Eng. 271, 113811.
Tang, X., Li, Z.X., Hu, X., Xu, Z.F., Peng, L.X., 2021. Self-correcting error-based prediction model for the COVID-19 pandemic and analysis of economic impacts. Sustain. Cities Soc. 74, 103219.
Tang, X.L., Zhou, H.T., Wang, F., Wang, W.D., Lin, X.K., 2022. Longevity-conscious energy management strategy of fuel cell hybrid electric Vehicle Based on deep reinforcement learning. Energy 238, 121593.
Uhlenbeck, G.E., Ornstein, L.S., 1930. On the theory of the brownian motion. Phys. Rev. 36 (5), 823–841.
Wang, X.G., Zou, Z.J., Yu, L., Cai, W., 2015. System identification modeling of ship manoeuvring motion in 4 degrees of freedom based on support vector machines. China Ocean Eng. 29 (4), 519–534.
Wang, Z.H., Xu, H.T., Xia, L., Zou, Z.J., Guedes Soares, C., 2020. Kernel-based support vector regression for nonparametric modeling of ship maneuvering motion. Ocean Eng. 216, 107994.
Wang, Z.H., Zou, Z.J., 2018. Quantifying multicollinearity in ship manoeuvring modelling by Variance Inflation Factor. In: ASME 2018 37th International Conference on Ocean, Offshore and Arctic Engineering. Madrid, Spain.
Wang, Z.H., Zou, Z.J., Guedes Soares, C., 2019. Identification of ship manoeuvring motion based on nu-support vector machine. Ocean Eng. 183, 270–281.
Wei, Z.B., Ruan, H.K., He, H.W., 2021. Battery thermal-conscious energy management for hybrid electric bus based on fully-continuous control with deep reinforcement learning. In: Proceedings of 2021 IEEE Transportation Electrification Conference & Expo. Chicago, IL, USA.
Woo, J., Park, J., Yu, C., Kim, N., 2018. Dynamic model identification of unmanned surface vehicles using deep learning network. Appl. Ocean Res. 78, 123–133.
Wu, X.G., Liu, S.W., Zhang, T.C., Yang, L., Li, Y.H., Wang, T.J., 2018. Motion control for biped robot via DDPG-based deep reinforcement learning. In: Proceedings of 1st WRC Symposium on Advanced Robotics and Automation. Beijing, China.
Zhang, Y.Y., Wang, Z.H., Zou, Z.J., 2022a. Black-box modeling of ship maneuvering motion based on multi-output nu-support vector regression with random excitation signal. Ocean Eng. 257, 111279.
Zhang, Z., Ren, J.S., Bai, W.W., 2022b. MIMO non-parametric modeling of ship maneuvering motion for marine simulator using adaptive moment estimation locally weighted learning. Ocean Eng. 261, 112103.
Zhang, W.X., Zhao, T., Zhao, Z.K., Ma, D., Liu, F.R., 2023. Performance analysis of deep reinforcement learning-based intelligent cooperative jamming method confronting multi-functional networked radar. Signal Process. 207, 108965.
Zhao, Y.J., Ma, Y., Hu, S.L., 2021. USV formation and path-following control via deep reinforcement learning with random braking. IEEE Transact. Neural Networks Learn. Syst. 32 (12), 5468–5478.
Zhu, M., Hahn, A., Wen, Y.Q., Bolles, A., 2017. Identification-based simplified model of large container ships using support vector machines and artificial bee colony algorithm. Appl. Ocean Res. 68, 249–261.
Zhu, M., Hahn, A., Wen, Y.Q., 2018. Identification-based controller design using cloud model for course-keeping of ships in waves. Eng. Appl. Artif. Intell. 75, 22–35.
Zhu, M., Hahn, A., Wen, Y.Q., Sun, W.Q., 2019. Optimized support vector regression algorithm-based modeling of ship dynamics. Appl. Ocean Res. 90, 101842.
Zhu, M., Sun, W.Q., Hahn, A., Wen, Y.Q., Xiao, C.S., Tao, W., 2020. Adaptive modeling of maritime autonomous surface ships with uncertainty using a weighted LS-SVR robust to outliers. Ocean Eng. 200, 107053.