Journal of *Marine Science and Engineering*

MDPI

Article

Deep Reinforcement Learning-Based Energy Management Strategy for Green Ships Considering Photovoltaic Uncertainty

Yunxiang Zhao [1], Shuli Wen [2,*], Qiang Zhao [3,*], Bing Zhang [1] and Yuqing Huang [4]

[1] Ocean College, Jiangsu University of Science and Technology, Zhenjiang 212003, China; zhaoyunxiang@stu.just.edu.cn (Y.Z.); zhbing0114@163.com (B.Z.)
[2] Key Laboratory of Control of Power Transmission and Conversion, Ministry of Education, Shanghai Jiao Tong University, Shanghai 200240, China
[3] College of Automation, Jiangsu University of Science and Technology, Zhenjiang 212100, China
[4] Shanghai Marine Equipment Research Institute, Shanghai 200031, China; huangyuging2017@126.com
* Correspondence: wenshuli@sjtu.edu.cn (S.W.); arzq@just.edu.cn (Q.Z.)

Abstract: Owing to the global concern regarding fossil energy consumption and carbon emissions, the power supply for traditional diesel-driven ships is being replaced by low-carbon power sources, which include hydrogen energy generation and photovoltaic (PV) power generation. However, the uncertainty of shipboard PV power generation due to weather changes and ship motion variations has become an essential factor restricting the energy management of all-electric ships. In this paper, a deep reinforcement learning-based optimization algorithm is proposed for a green ship energy management system (EMS) coupled with hydrogen fuel cells (HFCs), lithium batteries, PV generation, an electric power propulsion system, and service loads. The focus of this study is reducing the total operation cost and improving energy efficiency by jointly optimizing power generation and voyage scheduling, considering shipboard PV uncertainty. To verify the effectiveness of the proposed method, real-world data for a hybrid hydrogen- and PV-driven ship are selected for conducting case studies under various sailing conditions. The numerical results demonstrate that, compared to those obtained with the Double DQN algorithm, the PPO algorithm, and the DDPG algorithm without considering the PV system, the proposed DDPG algorithm reduces the total economic cost by 1.36%, 0.96%, and 4.42%, while effectively allocating power between the hydrogen fuel cell and the lithium battery and considering the uncertainty of on-board PV generation. The proposed approach can reduce energy waste and enhance economic benefits, sustainability, and green energy utilization while satisfying the energy demand for all-electric ships.

Keywords: energy management; deep reinforcement learning; green ship; shipboard PV uncertainty

check for updates

Academic Editor: Kostas Belibassakis

Received: 14 February 2025
Revised: 9 March 2025
Accepted: 9 March 2025
Published: 14 March 2025

Citation: Zhao, Y.; Wen, S.; Zhao, Q.; Zhang, B.; Huang, Y. Deep Reinforcement Learning-Based Energy Management Strategy for Green Ships Considering Photovoltaic Uncertainty. *J. Mar. Sci. Eng.* **2025**, *13*, 565. https://doi.org/10.3390/jmse13030565

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The shipping industry accounts for 80–90% of global trade and is also responsible for high greenhouse gas emissions [1]. Given the increasing concern for environmental protection and energy efficiency worldwide, the shipping industry faces significant pressure to reduce pollutants and emissions, forcing the shipping industry to find new energy supply solutions [2]. Zero-carbon power sources, represented by HFCs and PV power generation, are gradually being integrated into electric ships. However, with the increasing penetration of new energy systems into ships, the uncertainty of ship-board PV systems poses a critical challenge in ship energy management, which can deteriorate generation scheduling

and system stability [3]. Accordingly, intelligent algorithms are necessary to develop an EMS in a shipboard microgrid to improve energy performance, ensuring maximum energy utilization and carbon emission reduction. Therefore, developing more efficient and intelligent EMSs, particularly for managing energy use on hybrid ships, will be crucial for the green transformation of the shipping industry.

Traditional diesel-driven ships face immense pressure to reduce emissions and improve energy efficiency, making green transformation imperative. Previous studies have investigated a variety of EMSs for hybrid shipboard power systems. Magkouris et al. [4] proposed a new boundary element method for analyzing the hydrodynamic behavior and resistance of twin-hull ships at low speeds, with a focus on integrating solar energy systems to enhance energy efficiency and reduce carbon emissions. Lu et al. [5] proposed a distributionally robust optimization model to efficiently schedule power generation and voyage planning for hybrid ships, aiming to minimize operational costs and reduce greenhouse gas emissions while addressing environmental uncertainties like waves, wind, and variable PV outputs. Balsamo et al. [6] proposed a semitransparent cadmium telluride PV window system for offshore passenger ships. The system increased the PV area through multidimensional PV windows to increase both the PV output and efficiency. Dawoud et al. [7] proposed an efficient hybrid renewable energy system model based on perovskite solar cells, with the goal of improving the energy efficiency of offshore oil ships and reducing greenhouse gas emissions. Wu et al. [8] evaluated the application of an onboard hybrid PV power system on dredging ships, utilizing both grid-connected and off-grid modes. The study revealed that the system reduced greenhouse gas emissions, energy consumption, and maintenance costs. Zhu et al. [9] proposed a probabilistic optimization-based design method for a wind-sail PV hybrid power system. The method was used to establish a joint distribution model for wind speed, wind direction, solar radiation, and environmental temperature to optimize greenhouse gas emissions and enhance the system lifecycle. Wen et al. [10] proposed a hybrid ensemble model for shipboard PV output prediction. The model combines various machine learning techniques with particle swarm optimization to predict solar power outputs for ship power systems. Gaber et al. [11] proposed an intelligent EMS for a renewable-energy-based hybrid ship power system. The system integrates PV and fuel cell technologies to optimize the power distribution, adjust power generation, improve operational efficiency, and reduce fuel consumption. Tang et al. [12] proposed an optimal power scheduling strategy for a hybrid energy system on ships that combines PV, batteries, diesel, and cold sources to maximize solar energy utilization and minimize electricity costs. Maaruf et al. [13] introduced a hybrid energy system that combines solar power, PEM fuel cells, and electrolyzers for all-electric ships. This system is designed to balance load and power generation, improve overall system efficiency, and achieve zero-emission propulsion. Hasan et al. [14] developed a design strategy for solar-powered electric ships to optimize the size of the PV system. The strategy enhances PV power and reduces the battery capacity. Abdelrahman et al. [15] introduced a rule-based approach for power and energy allocation in shipboard microgrids equipped with PVs and hybrid energy storage systems (ESSs). This approach considers the specific fuel consumption of diesel engine generators and the characteristics of the hybrid ESS to ensure efficient power distribution within the microgrid. Song et al. [16] proposed an EMS based on a hybrid penalty proximal policy optimization algorithm to address power generation scheduling and demand-side regulation issues for fully electric ships in uncertain marine environments. Yiğit et al. [17] proposed a new ship energy management algorithm that selects the optimal energy source for a ship on the basis of the economic and environmental standards at the associated port. Hein et al. [18] proposed a data-driven multi-objective range and energy management scheduling method to optimize energy scheduling for all-electric ships

equipped with ESSs and PV systems. Firdaus et al. [19] explored the utilization of wind and solar energy for thrust and speed control on catamarans to optimize an EMS. Fang et al. [20] proposed a data-driven PV generation uncertainty co-ordination management method on the basis of an extreme learning machine to optimize the operation of fully electric ships and reduce fuel costs. Igder et al. [21] proposed a system reliability-centered maintenance approach for all-electric ships that combines Markov processes and an enhanced JAYA algorithm to increase reliability. Wen et al. [22] proposed a data-driven optimization approach that combines deep learning prediction methods to jointly optimize the energy storage system size and voyage scheduling for all-electric ships, which minimized the total cost and greenhouse gas emissions. Dolatabadi et al. [23] explored the technical feasibility of developing hybrid propulsion systems with different sizes of bulk carriers that combine green hydrogen with wind and solar energy to achieve zero emissions. Iqbal et al. [24] improved the microgrid system of all-electric ships and reported that the combination of lithium battery systems with photovoltaic/wind energy/battery yielded the lowest net present value and electricity cost. The results showed that this combination offers superior performance and environmental benefits compared to other methods. Wen et al. [25] proposed a stochastic co-ordination framework for all-electric ships, combining the Taguchi orthogonal method with adaptive multi-objective particle swarm optimization to address uncertainties related to solar energy, waves, and ship motion. Lan et al. [26] proposed an interval forecasting framework that combines neural networks and kernel density estimation to address the challenges posed by multiple factors affecting onboard PV power generation. Wang et al. [27] focused on optimizing the operation and energy scheduling of hydrogen-driven ships. They proposed designs for two states, PV power surplus and constant power, to optimize the energy balance and improve the utilization of renewable energy. Hou et al. [28] proposed a real-time EMS based on a data-driven stochastic predictive control approach to address the uncertainty of solar energy and loads in ship microgrids. Although numerous studies have focused on optimization designs and EMSs for hybrid ship power systems, the real-time performance of EMSs in practical applications remains underexplored, and the uncertainty of on-board PV generation outputs has been ignored.

Since the EMS plays a significant role in a ship power system, with the goal of improving operational performance, a novel real-time onboard EMS framework that jointly optimizes the outputs of HFCs, lithium batteries, and PV generation while considering voyage scheduling is proposed to reduce costs and improve energy efficiency. The primary contributions of this study can be summarized as follows:

(1) To increase the range and energy efficiency of hybrid-powered ships, a deep reinforcement learning-based EMS approach that optimizes the output of HFCs and lithium batteries while considering PV uncertainty is proposed. The framework utilizes an intelligent adaptive method to effectively manage the complex energy dynamics of hybrid-powered ships under various operational conditions, thus improving the flexibility and adaptability of the system.

(2) The aim of this paper is to minimize operational costs while considering multiple constraints with respect to both generation and voyages, such as the power balance, hydrogen fuel cell output limits, state of charge (SOC) of shipboard lithium batteries, PV generation limits, ship speed, and route length. A DDPG algorithm is designed to address this joint optimization problem, specifically considering PV uncertainty and battery degradation. The proposed method significantly improves the operational performance and system reliability of ships.

(3) Unlike previous studies of diesel-powered ships, this paper explores the integration and optimization management of multi-energy systems, particularly considering PV uncertainty. The proposed EMS solution offers high robustness and adaptability, enabling

effective energy dispatching under dynamic voyage conditions and significantly improving the sustainability and energy efficiency of ship power systems.

The subsequent sections of this paper are organized as follows. Section 2 provides details of the hybrid ship powertrain model. In Section 3, the optimization problem is formulated. Section 4 introduces the shipboard EMS solution based on the DDPG algorithm. Section 5 presents the simulation experiments conducted with a hybrid ship, and the conclusions are given in Section 6.

2. Modeling of Hybrid Ship Power System

In contrast to a land-based power grid, the entirety of the hybrid ship can be construed as a mobile microgrid, which incorporates multiple energy sources, namely, HFCs, lithium batteries, and PV generation systems. In addition, the electrical system of the ship must be designed considering the energy requirements of the propulsion system and the service loads. In a ship, the propulsion system provides power for navigation, whereas service loads encompass auxiliary equipment and facilities, such as lighting, air conditioning, and galley equipment, which are essential for the stability and energy efficiency of the power system. The layout of the integrated power system of a hybrid ship is shown in Figure 1.

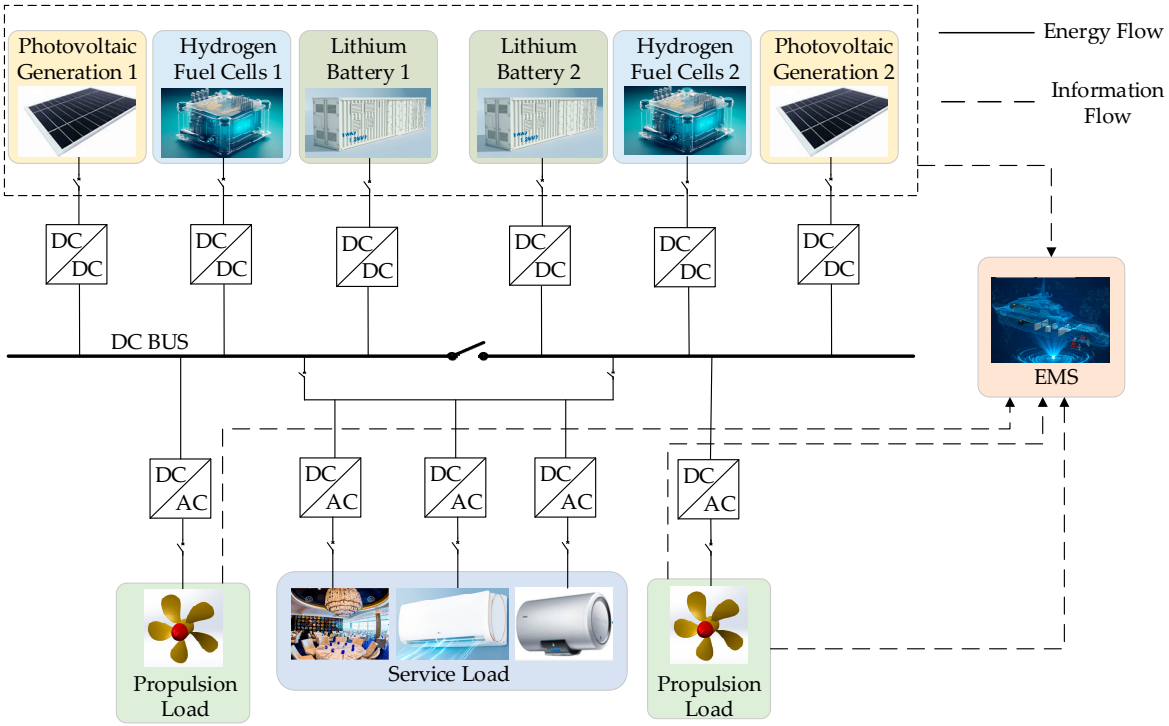


Figure 1. Topology diagram of a hybrid ship power system.

To achieve efficient energy management and system optimization, accurate mathematical models must be developed for each energy source and their interactions while considering the demands of the propulsion system and the service loads. These models are organically integrated into the ship power network to ensure the co-ordination and interaction among different energies, loads, and systems. Through appropriate modeling methods, energy utilization efficiency can be improved and system stability and reliability can be ensured under various operating conditions. Therefore, the goal of this research is to optimize energy management and jointly optimize the navigation and generation processes of a hybrid power system to enhance overall energy efficiency and system stability.

2.1. Photovoltaic Generation Model

The PV generation system in the electrical system of a hybrid ship differs from that of a land-based PV power generation system and is influenced primarily by numerous factors in the maritime environment. The motion of the ship causes the orientation and tilt angle of the solar panels to change constantly, thus affecting their power generation efficiency. In addition, factors such as the sea state, weather conditions, geographic location, and temperature also affect PV power generation. Therefore, to accurately estimate the power generation of a shipboard PV system, these factors must be considered. The output power of the shipboard PV is detailed in Equation (1). The generator efficiency of the shipboard PV is detailed in Equation (2). The global horizontal irradiance and the direct horizontal irradiance are detailed in Equation (3) and Equation (4), respectively [29].

$$P_{PV}^{t} = \eta_{PV} \cdot A_{PV} \cdot I_{Gh}^{t} \tag{1}$$

$$\eta_{PV} = \eta_{PV,ref} \cdot \eta_{MPPT}[1 - \delta(T_{PV}^{t} - T_{PV,ref})] \tag{2}$$

$$I_{Gh}^{t} = I_{Bh}^{t}[\cos(o) + \lambda \cos^2(\frac{\varphi}{2}) + \rho(\cos\omega + \lambda) + \sin^2(\frac{\varphi}{2})] \tag{3}$$

$$I_{Bh}^{t} = I_{Gh}^{t} - I_{Dh}^{t} \tag{4}$$

2.2. Hydrogen Fuel Cell Model

HFCs are electrochemical devices that convert hydrogen fuel and oxygen in the air into electrical energy. Compared with conventional internal combustion engines, HFCs offer higher energy conversion efficiency, lower operating noise, and longer service life, making them a key technology for achieving the goals of zero emissions and green shipping. The power output of HFCs is influenced by various factors, including temperature, the partial pressures of hydrogen and oxygen, and overall efficiency. These factors affect the electrochemical reaction rate, cell voltage, and energy conversion efficiency, leading to variations in performance and potential degradation over time. Additionally, losses caused by activation, ohmic, and concentration polarizations must be considered in power estimation. The actual output voltage, open-circuit voltage, activation loss voltage, ohmic loss voltage, and concentration loss voltage of the HFCs are represented by Equations (5) to (9), respectively. The hydrogen mass flow rate is given by Equation (10), while the total chemical power and actual output power of the HFCs are represented by Equations (11) and (12), respectively [27].

$$V_{cell}^{t} = V_{oc}^{t} - V_{act}^{t} - V_{ohm}^{t} - V_{conc}^{t} \tag{5}$$

$$V_{oc}^{t} = V_{ref} - \frac{R_{gas}T_{FC}^{t}}{2F}\ln\left(\frac{p_{H_2}p_{O_2}}{p_{H_2O}^2}\right) \tag{6}$$

$$V_{act}^{t} = \xi_1 + \xi_2 T_{fc}^{t} + \xi_3 T_{FC}^{t}\ln I_{FC}^{t} + \xi_4 T_{FC}^{t}\ln\frac{p_{O_2}}{5.08\times10^6\exp\left(-\frac{458}{T_{FC}^{t}}\right)} \tag{7}$$

$$V_{ohm}^{t} = I_{FC}^{t}R_{\text{int}} \tag{8}$$

$$V_{conc}^{t} = \alpha\exp(\beta I_{FC}^{t}) \tag{9}$$

$$\dot{m}_{H_2}^{t} = \frac{I_{FC}^{t}N_{cell}}{2F\eta_{FC,eff}} \tag{10}$$

$$P_{FC,H_2}^{t} = \dot{m}_{H_2}^{t}H_{HV,H_2} \tag{11}$$

$$P_{FC}^{t} = \mu_{FC}\frac{V_{cell}^{t}}{H_{HV,H_2}}P_{FC,H_2}^{t} \tag{12}$$

2.3. Shipboard Lithium Battery Model

The shipboard lithium battery is a critical component of hybrid ships, serving as an ESS with flexible energy management capabilities, such as energy storage and release. When the load demand increases or there is an insufficient energy supply, the lithium battery can quickly release stored energy to provide supplementary power for the operation of the ship, ensuring system stability and reliability. Conversely, when the load demand decreases or excess energy is available, the lithium battery efficiently absorbs and stores the surplus energy. The power of the ESS is influenced by the variations in the SOC, the capacity, and the charge/discharge efficiency. The SOC fluctuations impact the available discharge power and depth of discharge, thereby affecting the operational stability and lifespan of the system. The capacity determines the maximum energy storage capability, directly influencing the duration and magnitude of power supply. The variations in the SOC, capacity, and charge/discharge power are shown in Equations (11)–(15).

$$SOC^t = SOC^{t-1} + \frac{E^t_{ESS}}{E_{ESS}} \tag{13}$$

$$E^t_{ESS} = \begin{cases} E^{t-1}_{ESS} - P^t_{ESS}\Delta t, & t > 1 \\ E_{ESS}SOC^0 - P^t_{ESS}\Delta t, & t = 1 \end{cases} \tag{14}$$

$$P^t_{ESS} = \begin{cases} P^t_{ESS,theor}/\eta_{dch}, & P^t_{ESS} \geq 0 \\ P^t_{ESS,theor}\eta_{ch}, & P^t_{ESS} < 0 \end{cases} \tag{15}$$

2.4. Propulsion Load Model

Note that the ship propulsion load model is utilized to describe the energy demand of the propulsion system under different operating conditions, which has a direct effect on the energy management and navigation optimization of the ship. To evaluate the performance of hybrid ship navigation, the propulsion load is modeled via Equation (16).

$$P^t_{pro} = c_1\left(v^t\right)^{c_2} \tag{16}$$

2.5. Service Load

In addition to the propulsion load, the service load refers to the power consumption required to meet the basic living needs of the crew and passengers, including lighting, air conditioning, kitchen equipment, navigation, communication, and cargo refrigeration. It has dynamic characteristics that change with the ship operating status and environmental conditions. The propulsion load and service load together constitute the total load of the ship, as shown in Equation (17).

$$P^t = P^t_{pro} + P^t_{sl} \tag{17}$$

3. Optimal Energy Management for Hybrid Ship

3.1. Objective Function

The EMS for the hybrid ship aims to optimize the power outputs of HFCs and lithium batteries given the uncertainty of PV generation, thus ensuring efficient energy distribution to meet the dynamically changing load demands during navigation. This strategic coordination of energy and voyage scheduling prioritizes the utilization of available energy to ensure stable ship operation while minimizing overall operational costs. The total costs include the expenses of the HFCs, shipboard ESS, and PV system, as shown below.

$$\min C^t_{total} = C^t_{H_2} + C^t_{ESS} + C^t_{PV} \tag{18}$$

$$\begin{cases} C_{H_2}^t = \sum_{n}^{N_{FC}} \sum_{t}^{T} (C_{H_2,pr}^t + C_{H_2,maint}^t) \\ C_{H_2,pri}^t = \dot{m}_{H_2}^t M_{H_2} p_{H_2,pr} \Delta t \\ C_{H_2,maint}^t = \chi P_{FC}^t \Delta t \end{cases} \tag{19}$$

$$\begin{cases} C_{ESS}^t = \sum_{n}^{N_{ESS}} \sum_{t}^{T} C_{ESS,maint}^t \\ C_{ESS,maint}^t = \varsigma |P_{ESS}^t| \Delta t \end{cases} \tag{20}$$

$$\begin{cases} C_{PV}^t = \sum_{n}^{N_{PV}} \sum_{t}^{T} (C_{PV,cap}^t + C_{PV,maint}^t) \\ C_{PV,maint}^t = \phi P_{PV}^t \Delta t \end{cases} \tag{21}$$

3.2. Constraints

Optimal energy management for hybrid ships is related to power generation and navigation limits, such as PV generation limits, power limits for HFCs, SOC limits for onboard ESSs, and charging and discharging power limits. Maximum speed limits must also be considered to ensure the safe operation of the ship. These limits are critical to maintaining the energy balance of the ship and directly influence how the ship can meet safety and efficiency requirements while achieving optimal energy scheduling and stable system operation.

3.2.1. Generation Constraints

- PV generation system limits:

The output power of the PV generation system is limited by weather conditions and system performance and must not exceed the maximum generating capacity, which is represented by Equation (22) [30].

$$P_{PV}^{\min} \le P_{PV}^t \le P_{PV}^{\max} \tag{22}$$

- HFCs limits:

To ensure the stable operation of the shipboard HFCs, the power must be constrained within a safe range, as shown below [31]. The constraints on the total chemical power and actual power of the HFCs are represented by Equations (23) and (24).

$$P_{FC,H_2}^{\min} \le P_{FC,H_2}^t \le P_{FC,H_2}^{\max} \tag{23}$$

$$P_{FC}^{\min} \le P_{FC}^t \le P_{FC}^{\max} \tag{24}$$

- Shipboard ESS limits:

The stable operation of the ESS is crucial for the EMS of the ship. To ensure the safety, reliability, and efficient energy scheduling of the battery system, a series of constraints must be considered [32]. The constraints on the variations in the SOC and the initial SOC of the ESS are represented by Equations (25) and (26). The charge/discharge power constraints are represented by Equation (27).

$$SOC^{\min} \le SOC^t \le SOC^{\max} \tag{25}$$

$$SOC^0 \le SOC^{t_0} \le (1 + v^{soc}) SOC^0 \tag{26}$$

$$-P_{ESS,cha}^{\max} \le P_{ESS}^t \le P_{ESS,dis}^{\max} \tag{27}$$

3.2.2. Voyage Constraints

To ensure safe and timely operation, the speed of the ship must be constrained within a specified range to balance safety, efficiency, and punctuality, as represented by Equation (28) [29] because excessive speeds can lead to increased fuel consumption, structural stress, and environmental impact, while excessively low speeds may result in delayed arrivals and inefficient operation. The traveled distance of the ship, updated iteratively based on its velocity and time step, is represented by Equation (29). The allowable variations in the sailing distance due to uncertainties or operational deviations are represented by Equation (30). The constraints on the total sailing distance, ensuring that the ship adheres to acceptable deviation limits during its journey, are represented by Equation (31).

$$v^{\min} \leq v^t \leq v^{\max} \tag{28}$$

$$Dis^t = Dis^{t-1} + v^t \Delta t \tag{29}$$

$$(1 - \zeta_{\min}) Dis_n \leq Dis^t \leq (1 + \zeta_{\min}) Dis_n \tag{30}$$

$$Dis_{term} \leq Dis_n^T \leq (1 + \kappa_{term}) Dis_{term} \tag{31}$$

3.2.3. Power Balance Constraint

In a hybrid ship, ensuring the stable operation of the power system requires maintaining a continuous dynamic equilibrium between power generation and load consumption. Consequently, the combined outputs of the PV generation system, hydrogen fuel cells, and shipboard ESS must collectively satisfy the load demand to maintain system reliability and efficiency, as expressed in Equation (32). To further enhance system resilience against unexpected failures, power fluctuations, and the uncertainty of the PV generation, redundancy has been incorporated into the system by configuring additional capacity in the HFCs and the ESS. This allows the system to compensate for variations in power generation, sudden load changes, and fluctuations in solar power availability, ensuring stable operation under different conditions.

$$P_{FC}^t + P_{ESS}^t + P_{PV}^t = P_{pro}^t + P_{sl}^t \tag{32}$$

4. DDPG-Algorithm-Based EMS

4.1. DDPG Algorithms

Reinforcement learning [33] is a machine learning approach in which an agent optimizes its decision-making strategy through continuous interaction with the environment. Its foundation is based on the Markov decision process [34], where an agent observes the state of their environment at each time step and selects an action according to a predefined policy. The environment then updates its state on the basis of the agent's decision and provides corresponding feedback. The objective of the agent is to refine its policy to maximize long-term cumulative rewards, thereby achieving optimal decision making in various scenarios. This process follows the Markov approach, where future states depend only on the current state and the selected action, without influence from past history. The overall framework of reinforcement learning is illustrated in Figure 2.

As reinforcement learning applications have become increasingly complex, deep reinforcement learning (DRL) [35] has emerged as a powerful approach that integrates the feature extraction capabilities of deep neural networks with the decision-making optimization of reinforcement learning. This combination enables DRL methods to be effectively applied in high-dimensional state spaces and continuous action spaces. Within the DRL framework, deep neural networks are utilized to approximate policy and value functions, allowing agents to learn and execute efficient decisions in complex environments. More-

over, this approach enhances the ability to extract critical features from valuable sensory data, improving adaptability and generalization to diverse tasks. The overall process is illustrated in Figure 3.

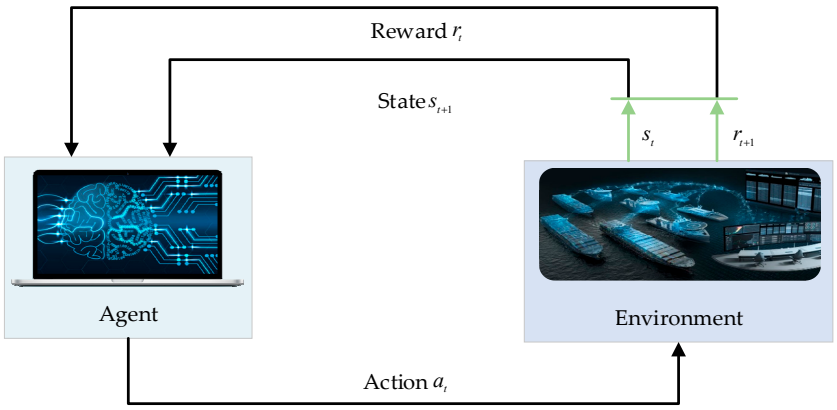


Figure 2. The architecture of reinforcement learning.

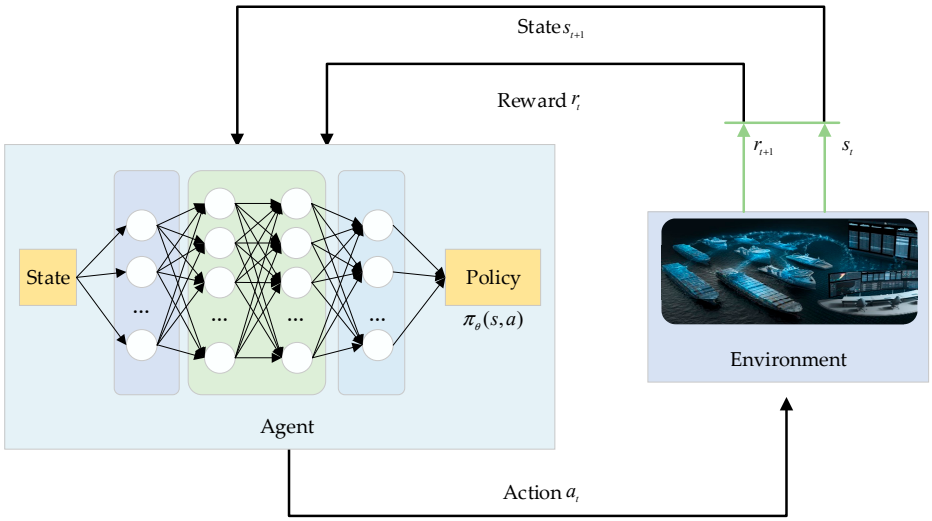


Figure 3. The DRL architecture.

The DDPG algorithm, a deep reinforcement learning approach, is tailor-made for making decisions in high-dimensional continuous action spaces. It makes use of the actor–critic architecture to fine-tune policies and assess values [36]. This algorithm integrates deterministic policy gradient methods with deep neural networks, enabling agents to efficiently make optimal decisions in complex environments. Compared to traditional reinforcement learning methods, DDPG exhibits superior sample efficiency by reusing past experiences through experience replay, significantly reducing the number of interactions required with the environment. Moreover, its off-policy nature allows for the use of a replay buffer, which enhances learning stability and prevents catastrophic forgetting. To increase training stability and ensure policy convergence, the DDPG algorithm employs four interdependent neural networks, namely, the actor network $\mu(s_t;\theta^\mu)$, which is responsible for generating actions; the critic network $Q(s_t,a_t;\theta^Q)$, which evaluates the value of these actions; and two target networks (the target actor network $\mu'(s_t;\theta^{\mu'})$ and the target critic network $Q'(s_t,a_t;\theta^{Q'})$), which serve as delayed versions of the main networks to stabilize learning. In the DDPG algorithm, the ReLu activation function is used in the actor network because it provides nonlinear mapping, avoids the vanishing gradient problem,

and improves computational efficiency. Additionally, the sparse activation property of ReLu helps reduce computational load and accelerates the training process.

During execution, the actor network $\mu(s_t;\theta^\mu)$ produces a deterministic action based on the current environment state, whereas the critic network $Q(s_t,a_t;\theta^Q)$ receives the action–state pair and computes the corresponding Q value to assess the long-term return of the decision. The target networks, which are delayed versions of the main networks, are introduced to improve learning stability. Their parameters are updated gradually through a soft update mechanism, allowing network weights to change smoothly and preventing abrupt variations that could lead to unstable training. Additionally, to further enhance training performance, the DDPG algorithm incorporates an experience replay mechanism, which stores past experiences and randomly samples them during updates. This approach helps reduce data correlation effects, ultimately improving policy optimization. The DDPG algorithm flowchart is shown in Figure 4.

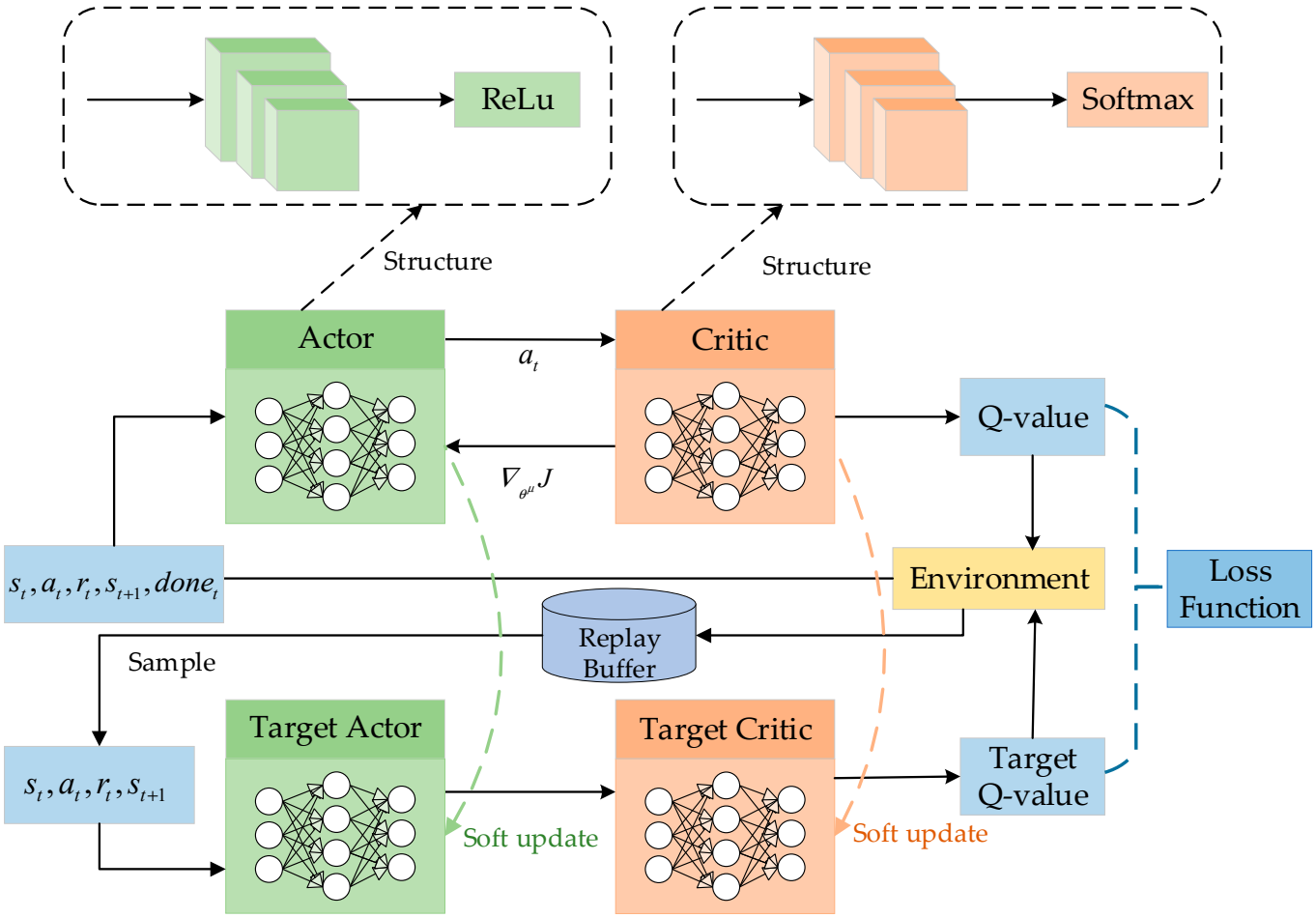


Figure 4. The flowchart of the DDPG algorithm.

As shown in Figure 4, the critic network $Q(s_t,a_t;\theta^Q)$ evaluates the action–value of the current policy. Its updates are driven by the temporal difference target, which minimizes the error between the current Q value and the target Q value. The target Q value is calculated by the target network, allowing the critic to assess the performance of the chosen actions and adjust by comparing predicted rewards with actual outcomes. This process helps the model refine its decision making over time, as shown in Equation (33).

$$y_t = r_t + \gamma Q'(s_{t+1}, \mu'(s_{t+1};\theta^{\mu'});\theta^{Q'}) \tag{33}$$

The loss function of the critic network $Q(s_t, a_t; \theta^Q)$ is the mean squared error, as shown in Equation (34).

$$L_Q = \frac{1}{N} \sum_{t=1}^{N} \left(y_t - Q(s_t, a_t; \theta^Q) \right)^2 \tag{34}$$

The actor network $\mu(s_t; \theta^\mu)$ updates the policy on the basis of the Q-value gradients provided by the critic network $Q(s_t, a_t; \theta^Q)$. The goal is to find a policy that generates actions in a given state that maximize the Q value, the gradient of which is approximated as shown in Equation (35).

$$\nabla_{\theta^\mu} J \approx \frac{1}{N} \sum_{t=1}^{N} \nabla_a Q(s_t, a; \theta^Q)|_{a=\mu(s_t;\theta^\mu)} \nabla_{\theta^\mu} \mu(s_t; \theta^\mu) \tag{35}$$

The parameters of the target network are updated in a smooth manner to reduce instability during training, as shown in Equation (36).

$$\theta' \leftarrow \tau\theta + (1 - \tau)\theta' \tag{36}$$

4.1.1. State Space Definition

The state space represents the key environmental variables observed by the agent at each time step, enabling informed decision making on the basis of the current system status. In a hybrid ship energy management system that uses the DDPG algorithm, the design of the state space must comprehensively capture the system's operational characteristics and dynamic variations. Specifically, it should include the power output of the OV generation system, the generation level of the HFCs, the charging/discharging status and stage of charge (SOC) of the onboard energy storage system (ESS), the ship's sailing speed, and the current load demand. This ensures that the decision-making process is effectively adapted to the system's real-time operating conditions. The state space S_t of the system is expressed in Equation (37).

$$S_t = \left\{ P_{PV}^t, P_{FC}^t, P_{ESS}^t, v^t, P_{pro}^t, P_{sl}^t, SOC^t, t \right\} \tag{37}$$

4.1.2. Action Space Definition

In the DDPG algorithm, the agent generates decisions through a continuous policy network on the basis of the current state space information. The DDPG algorithm can directly output continuous action values, making it highly suitable for addressing complex problems with continuous control requirements, such as hybrid ship energy management. In this study, the continuous hydrogen fuel cell load rate coefficient is defined as the decision variable. The action space A_t of the system is defined in Equation (38).

$$A_t = \left\{ ratio^t = P_{FC}^t / P_{FC}^{\max} \right\} \tag{38}$$

4.1.3. Reward Function Definition

In the DDPG-algorithm-based EMS, the reward function is utilized to guide the agent to make the optimal decision at each time step by evaluating its performance in learning optimization. For the EMS of hybrid ships, the reward function needs to balance multiple objectives, which include the total economic cost, the power fluctuation penalties, and the SOC safety constraint penalties for the ESS. The reward function of the system is shown below.

$$R_t = \tanh(R_{cost} + R_{bal} + R_{saf}) \tag{39}$$

$$R_{cost} = -(C_{H_2}^t + C_{ESS}^t + C_{PV}^t) \tag{40}$$

$$R_{bal} = -\vartheta_{bal} \cdot \max(0, \left| P_{sup}^{t} - P_{dem}^{t} \right|) \tag{41}$$

$$R_{saf} = \begin{cases} -4, SOC^{t} < 0 \ or \ SOC^{t} > 1 \\ \vartheta_{saf}, \ 0 \le SOC^{t} \le 1 \end{cases} \tag{42}$$

The EMS flowchart based on the DDPG algorithm is shown in Figure 5.

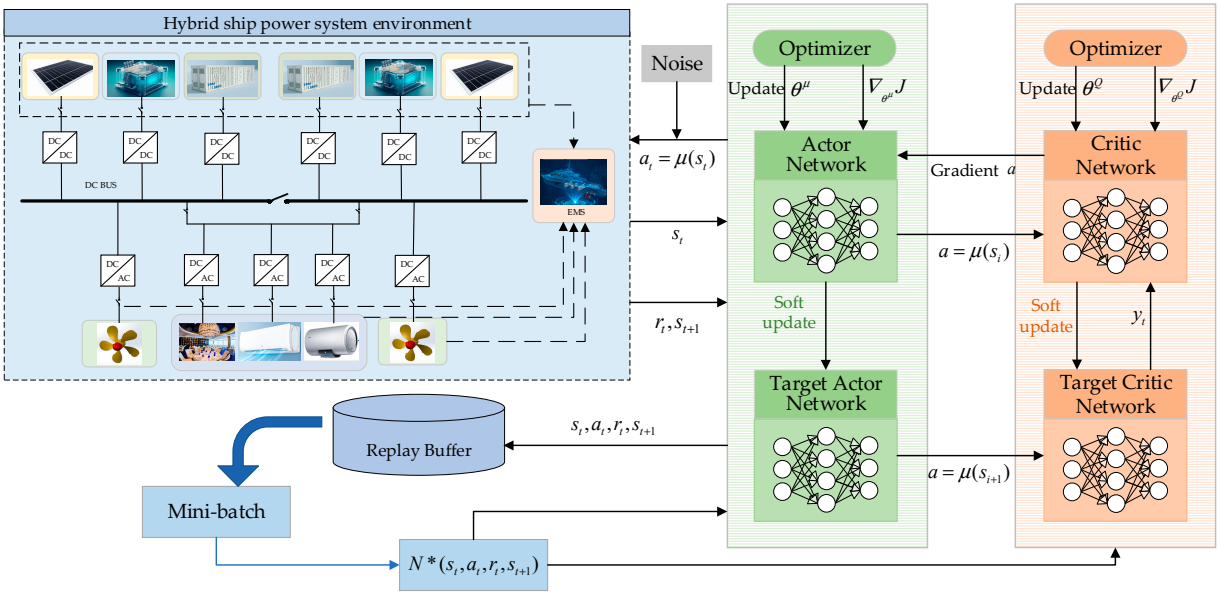


Figure 5. The EMS flowchart based on the DDPG algorithm.

5. Results and Discussion

5.1. Data Set

In this study, the hybrid ship is propelled by alternating-current electric power and is configured with a hybrid energy system consisting of two HFCs, two lithium batteries, and a PV generation system. The HFCs are utilized as the main sources of electricity, the lithium batteries are utilized to balance the load fluctuations, and the PV generation system provides additional clean energy for the ship. The detailed parameters are shown in Table 1. The global horizontal irradiance (GHI) and temperature of the PV system throughout the year are shown in Figures 6 and 7. The service load is shown in Figure 8.

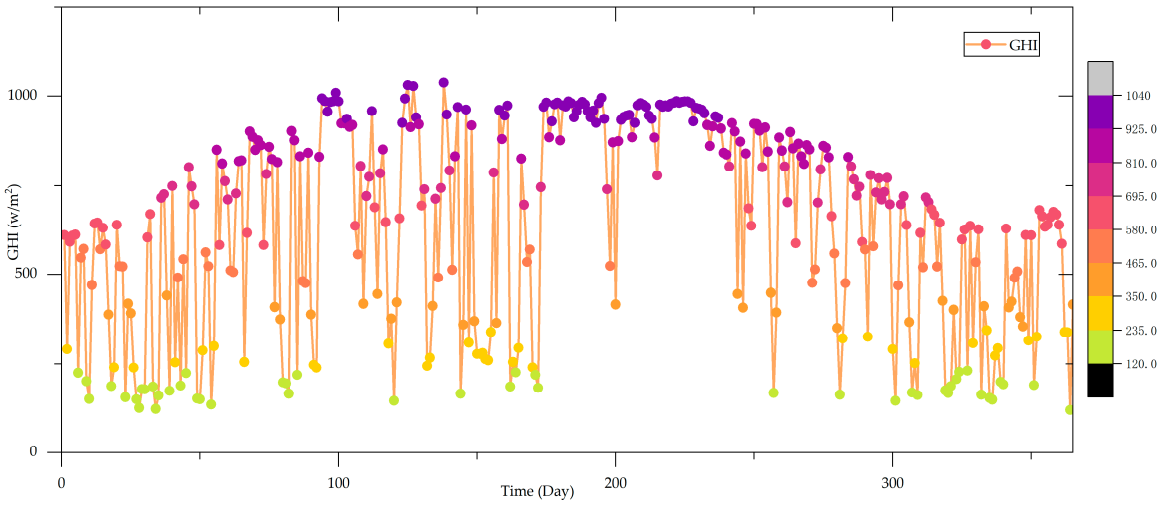


Figure 6. The GHI of the PV system.

Table 1. Parameters of the hybrid ship.

Component	Type	Numerical	Unit
PV [30]	η_{PV}	0.18	/
	A_{PV}	300	m^2
	η_{MPPT}	1	/
	$T_{PV,ref}$	25	°C
	δ	0.0048	/
	$C^t_{PV,cap}$	180	USD/m^2
HFCs [31]	$P^{\max}_{FC}$	500	kW
	$\eta_{FC,eff}$	0.5	/
	μ_{FC}	0.75	/
	$p_{H_2,pr}$	5	USD/kg
	H_{HV,H_2}	142	MJ/kg
ESS [32]	$P^{\max}_{ESS}$	200	kW
	η_{ch}/η_{dch}	0.99/0.99	/
	$SOC^{\min}/SOC^{\max}$	0.2/0.8	/
Ship speed [32]	$v^{\max}$	20	knot

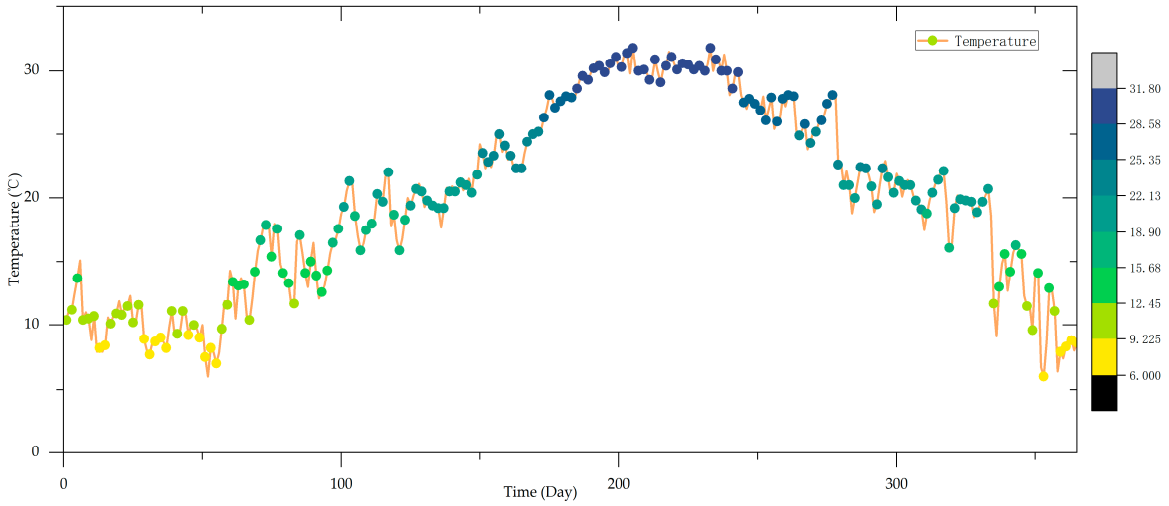


Figure 7. The temperature of the PV system.

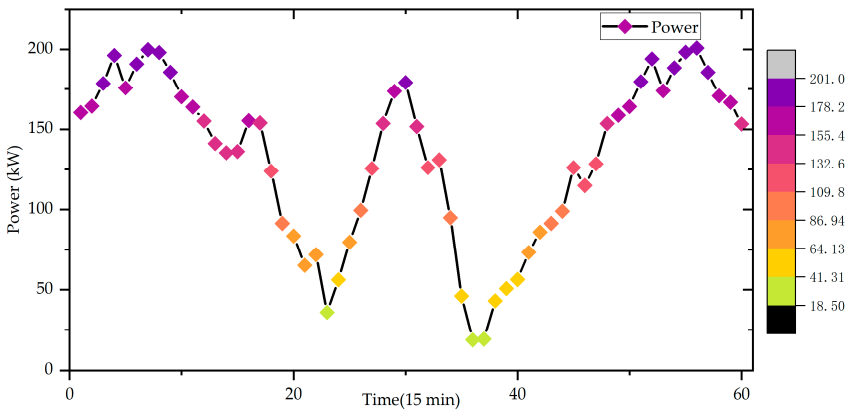


Figure 8. The service load.

As shown in Table 1, the dataset includes parameters such as the generator efficiency, effective area, efficiency of the maximum power point tracking equipment, reference temperature, temperature coefficient of efficiency, and installation cost of the shipboard PV

system for the PV system; the maximum actual power, efficiency, fuel utilization rate, unit price of hydrogen fuel, and the maximum heating value of hydrogen for the HFCs; the maximum charge/discharge power, charging/discharging efficiency, and minimum/maximum SOC of the ESS; and the ship speed parameter.

5.2. Case Study

To comprehensively validate the effectiveness of the proposed method, four representative cases under different scenarios are investigated.

Case 1: Double DQN-algorithm-based operation analysis of a ship power system considering HFCs, lithium batteries, and a PV system.

Case 2: PPO-algorithm-based operation analysis of a ship power system considering HFCs, lithium batteries, and a PV system.

Case 3: DDPG-algorithm-based operation analysis of a ship power system considering only HFCs and lithium batteries (excluding PV system).

Case 4: DDPG-algorithm-based operation analysis of a ship power system considering HFCs, lithium batteries, and a PV system.

The parameters for training the specific models for these cases are given in Table 2.

Table 2. Parameters for training the specific models for different cases.

Parameters	Type	Numerical
Actor learning rate	Actor_lr	0.0003
Critic learning rate	Critic_lr	0.003
Soft update parameter	τ	0.005
Discount factor	γ	0.98
Standard deviation for gaussian noise	σ	0.01
Replay buffer size	RB	10,000
Minimum number of training starts	MS	1000
Number of samples	NS	64
Dimensions of the hidden layer	Hidden_dim	64

The reward values for the four cases are shown in Figure 9.

Figure 9a–d show the cumulative reward curves for Case 1, Case 2, Case 3, and Case 4, respectively. In Figure 9a, the cumulative reward curve exhibits significant fluctuations during episodes 0–100, indicating that the Double DQN algorithm is learning new strategies to explore optimal action selection. The model converges after 400 episodes. Due to the adoption of the ϵ-greedy strategy in Double DQN, minor oscillations in the reward curve persist even after convergence. The model converges after 400 episodes, with minor oscillations in the reward curve persisting due to the adoption of the ϵ-greedy strategy in Double DQN. In Figure 9b, the cumulative reward curve exhibits significant fluctuations during 0–400 episodes, indicating that the policy is still exploring and some decisions are suboptimal. After 400 episodes, the curve stabilizes overall, suggesting that the policy has been optimized, though minor oscillations persist due to the PPO clipping mechanism and exploration strategy. In Figure 9c, Case 2 considers only the HFCs and the ESS while excluding the PV system, resulting in reduced model complexity compared to that in other cases. Consequently, the cumulative reward converges quickly, after approximately 350 episodes. Although the models in Case 1 and Case 2 achieve convergence, the DDPG algorithm continuously introduces Gaussian noise to maintain policy exploration, which results in slight fluctuations in the reward in the stable region. In Figure 9d, the cumulative reward remains low during the first 0–300 episodes, as the model relies on random exploration for action selection before the policy is effectively learned. Between episodes 300 and 400, the cumulative reward increases rapidly, suggesting that

the model starts to learn how to achieve higher rewards. After 400 epochs, the model converges, suggesting that the policy becomes stable.

Figures 10 and 11 present the energy scheduling results and SOC of the ESS variations for the four different cases, respectively.

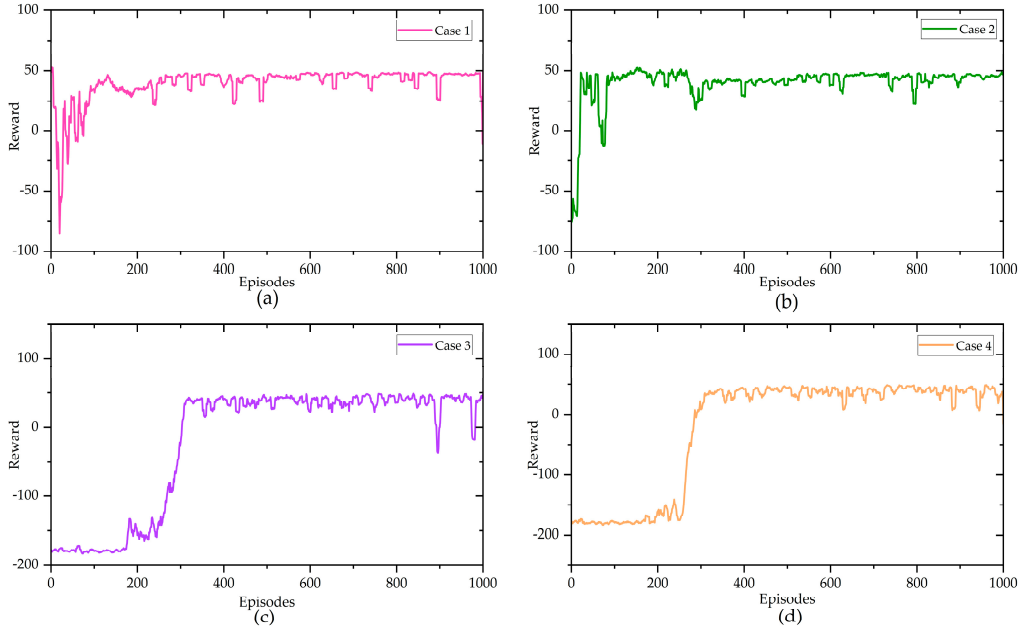


Figure 9. The cumulative reward curves in the four cases. (**a**) The cumulative reward curves in Case 1. (**b**) The cumulative reward curves in Case 2. (**c**) The cumulative reward curves in Case 3. (**d**) The cumulative reward curves in Case 4.



Figure 10. The energy scheduling results of the four cases. (**a**) The energy scheduling results of Case 1. (**b**) The energy scheduling results of Case 2. (**c**) The energy scheduling results of Case 3. (**d**) The energy scheduling results of Case 4.

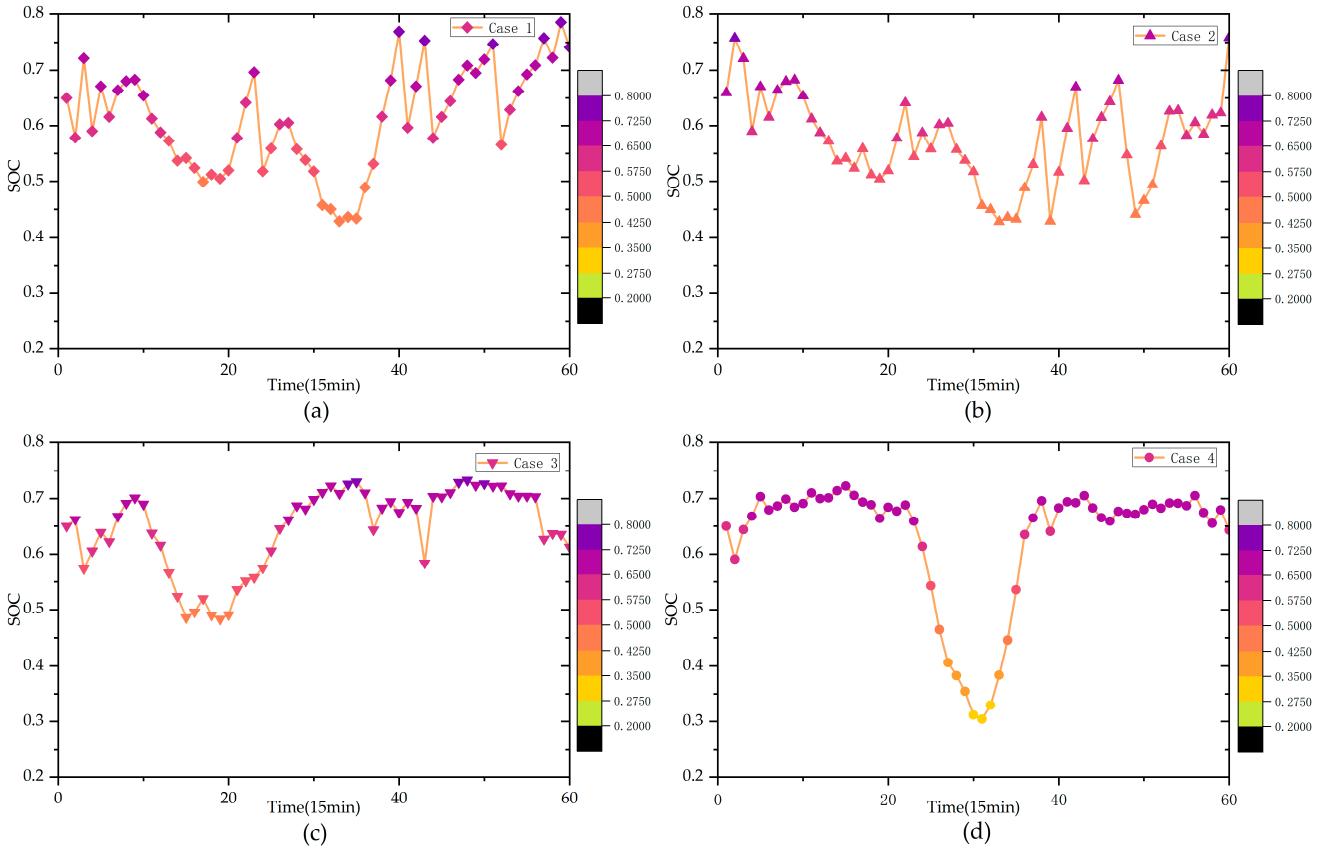


Figure 11. Variations in the SOC of the ESS in the four cases. (**a**) Variations in the SOC of the ESS in Case 1. (**b**) Variations in the SOC of the ESS in Case 2. (**c**) Variations in the SOC of the ESS in Case 1. (**d**) Variations in the SOC of the ESS in Case 4.

Figure 10a shows the energy scheduling results for Case 1, Figure 10b shows those for Case 2, Figure 10c shows those for Case 3, and Figure 10d shows those for Case 4. Figure 10 shows that HFCs dynamically adjust their output power in real time on the basis of the variations in the PV system output and load power. At the same time, the ESS charges and discharges according to the total ship load power demand, ensuring that the ship load requirements are met and achieving a dynamic balance between energy generation and consumption for the hybrid-power ship. In Figure 10a,b, it is evident that both the double DQN and PPO models fail to effectively implement peak shaving and valley filling with the ESS, resulting in more frequent high-power charging and discharging cycles than are ideal and significantly increasing the ESS maintenance costs. Figure 10c reveals that, in Case 3, in which the PV system is not considered, the ship power is provided solely by HFCs and the ESS, leading to a significant increase in hydrogen fuel costs and ESS maintenance costs. As shown in Figure 10d, to achieve optimal economic performance, hydrogen fuel costs must be fully considered, and the ESS should ideally charge at high power during low load periods and discharge at high power during peak load periods to perform peak shaving and valley filling. Additionally, the maintenance costs of HFCs, the PV system, and the ESS should be accounted for to optimize the overall economic efficiency and energy utilization of the system.

Figure 11a illustrates the SOC variation of the ESS during the energy scheduling period for Case 1, and Figure 11b–d show the results for Case 2, Case 3, and Case 4, respectively. In Figure 11a, the SOC of the ESS ranges from 0.42 to 0.78, which is due to the value obtained with the double DQN algorithm in Case 1. Due to the frequent activation of the ESS, the charging–discharging strategy appears unstable. Similarly, in Figure 10b, the

SOC of the ESS ranges from 0.43 to 0.76, and the excessive scheduling of the ESS leads to instability in the charging–discharging strategy. Figure 11b shows that the SOC of ESS fluctuates between 0.48 and 0.73, indicating that the scheduling of the available capacity of the ESS is relatively conservative and fails to fully leverage its scheduling potential. In Figure 11a, the SOC of the ESS fluctuates between 0.30 and 0.72, indicating that the agent, while ensuring that the ESS is fully utilized, performs effective scheduling. By reducing the scheduling amplitude and frequency, the scheduling function of the ESS is fully utilized, ensuring stable system operation within a reasonable range, thereby achieving optimal economic performance.

Table 3 presents the total economic costs in the four cases. In Case 1 and Case 2, the frequent activation of the ESS significantly increases its maintenance cost, leading to total economic costs of USD 95,186.37 and USD 94,266.18, respectively. In Case 3, due to the absence of a PV system, the ship is powered entirely by the HFCs and the ESS. As a result, the hydrogen fuel cost and HFC maintenance cost are relatively high, and the total economic cost reaches USD 98,232.04. In contrast, although the ESS maintenance cost in Case 4 is slightly higher than that in Case 3, the total economic cost is the lowest. Specifically, the total economic cost in Case 4 is reduced by 1.36%, 0.96%, and 4.42% compared with those in Cases 1, 2, and 3, respectively.

Table 3. Total economic costs in the four cases.

Component	Economic Costs (USD)	Case 1	Case 2	Case 3	Case 4
PV	Installation	2400	2400	0	2400
	Maintenance	121.64	121.64	0	121.64
HFCs	Hydrogen fuel	85,507.76	85,035.12	90,886.68	84,804.01
	Maintenance	5937.50	5905.22	6711.58	5889.17
ESS	Maintenance	1219.47	1339.95	633.78	676.93
Total economic costs		95,186.37	94,801.93	98,232.04	93,891.75

5.3. Sensitivity Analysis

To further investigate the impact of hyperparameters on the overall economic performance of the DDPG algorithm, a comparative analysis of different actor learning rates and critic learning rates combinations has been conducted. The tall economic cost results are shown in Figure 12.

As shown in Figure 12, as the critic learning rate increases from 0.001 to 0.003, the economic cost decreases from USD 98,015.36 to USD 93,891.75, indicating that a moderate increase in the critic learning rate improves the accuracy of Q-value estimation, thereby enhancing policy convergence. However, when the critic learning rate reaches 0.004, the economic cost rises to USD 94,256.15, which may be attributed to excessive Q-value updates, leading to policy instability or overfitting.

Similarly, for a fixed critic learning rate, the actor learning rate exhibits a non-monotonic trend, where an increase from 0.0001 to 0.0003 leads to a reduction in economic cost from USD 98,015.36 to USD 93,891.75, indicating improved decision-making efficiency. However, when the actor learning rate increases to 0.0004, the cost rises to USD 95,886.61, suggesting that excessively large updates may introduce instability in the exploration process.

Overall, the lowest economic cost of USD 93,891.75 is observed when the actor learning rate is set at 0.0003 and the critic learning rate at 0.003, beyond which further increases in the learning rate result in cost escalation. This phenomenon can be explained by the impact of learning rates on policy optimization: a lower critic learning rate around 0.001 leads to insufficient Q-value updates, resulting in higher economic costs, whereas a higher critic

learning rate around 0.004 may cause overfitting and Q-value fluctuations, increasing the cost. Similarly, a lower actor learning rate around 0.0001 slows down policy learning, leading to suboptimal decisions, whereas a higher actor learning rate around 0.0004 may cause policy oscillations or divergence, thereby increasing costs.

Therefore, careful selection of learning rates plays a critical role in ensuring effective policy optimization in the DDPG algorithm, where avoiding extreme values significantly enhances system performance and minimizes overall economic cost.

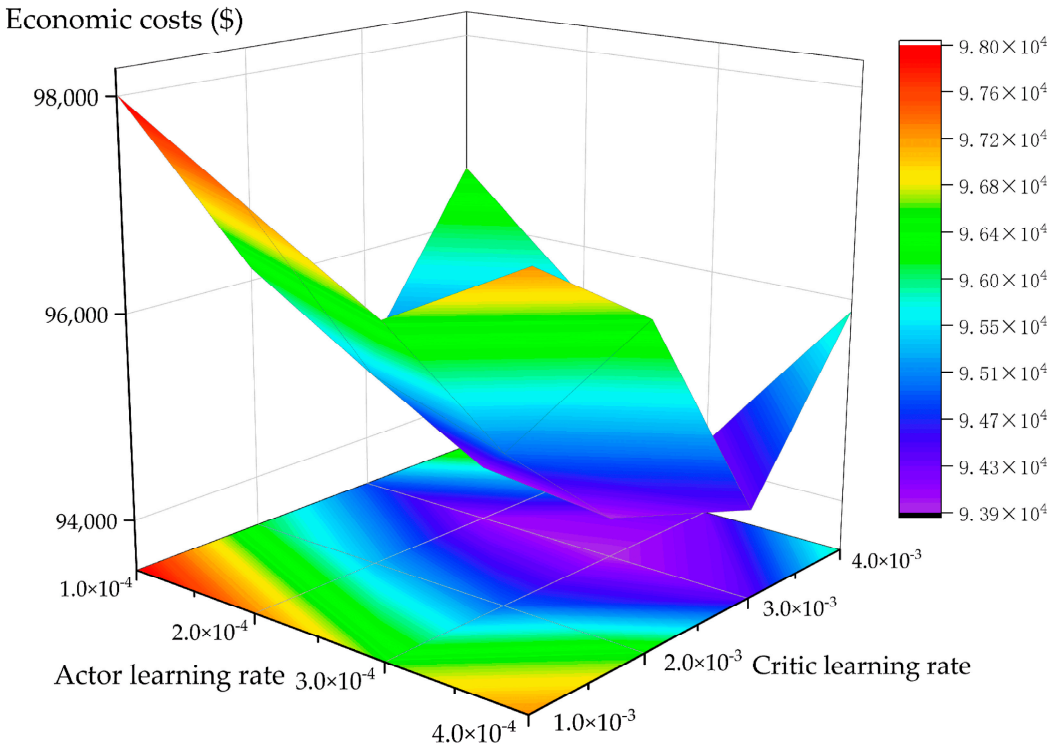


Figure 12. Economic costs at different learning rates.

6. Conclusions

With the rapid development of shipboard power systems toward greener and smarter designs, EMSs have become significant in both ship navigation and power generation optimization. In this paper, a deep reinforcement learning-based EMS framework is designed for a hybrid hydrogen- and PV-powered ship. The framework accounts for the uncertainty of PV generation and aims to jointly optimize the HFCs on a hybrid ship, the power output of the ESS, and the ship navigation speed. In addition, to evaluate the operational performance of the proposed framework, a mathematical model of the hybrid ship, consisting of the shipboard PV generation system, HFCs, shipboard lithium batteries, propulsion loads, and service loads, is established. Through experimental validation, the proposed DDPG algorithm is compared with other classic methods under various operating conditions. The simulation results show that the proposed DDPG algorithm is able to optimally co-ordinate energy scheduling and ship navigation, achieving a reduction in the total economic cost of 1.36%, 0.96%, and 4.42% compared to those obtained with the Double DQN algorithm, the PPO algorithm, and the DDPG algorithm without considering the PV system, respectively. These research results provide a valuable reference for the stable and economic operation of hybrid green energy ships, which will be helpful for the sustainable development of the shipping industry. In future work, the recovery of heat from the HFCs will be further studied to enhance the efficiency of the HFCs.

Author Contributions: Conceptualization, Y.Z.; methodology, Y.Z. and S.W.; formal analysis, S.W.; investigation, Y.H.; resources, Y.H.; data curation, S.W. and Q.Z.; writing—original draft preparation, Y.Z.; writing—review and editing, Y.Z. and S.W.; supervision, B.Z.; project administration, Y.H.; funding acquisition, S.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China under grant 52177101 and Postgraduate Research & Practice Innovation Program of Jiangsu Province, grant number SJCX24_2607. The authors would like to thank the Shanghai Marine Equipment Research Institute for providing valuable data for the research.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy and confidentiality concerns.

Conflicts of Interest: Author Yuqing Huang was employed by the company Shanghai Marine Equipment Research Institute. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Nomenclature

Acronyms:	
DDPG	Deep deterministic policy gradient
DRL	Deep reinforcement learning
DQN	Deep Q network
HFCs	Hydrogen fuel cells
GHI	Global horizontal irradiance
PV	Photovoltaic
EMS	Energy management system
SOC	State of charge
ESS	Energy storage system
Parameters:	
Δt	Time step
$\eta_{PV}/\eta_{PV,ref}$	Generator efficiency/reference generator efficiency of the shipboard PV system
η_{MPPT}	Efficiency of the maximum power point tracking equipment
A_{PV}	Effective area of the PV module
δ	Temperature coefficient of efficiency of the PV system
$T_{PV,ref}$	Reference temperature of the PV system
$\lambda/\rho/\omega$	Diffuse component constant/reflection index/zenith angle
o/φ	Angle between the panel and the sunlight/tilt angle relative to the horizontal plane
V_{ref}	Standard voltage of the HFCs
R_{gas}	Gas constant
F	Faraday constant
$p_{H_2}/p_{O_2}/p_{H_2O}$	Partial pressure of hydrogen/oxygen/water vapor
$\xi_1/\xi_2/\xi_3/\xi_4$	Empirical coefficients for ohmic loss
R_{int}	Output current and internal resistance of the HFCs
α/β	Empirical coefficients for concentration loss of the HFCs
N_{cell}	The number of individual cells in an HFC stack
$\eta_{FC,eff}$	Efficiency of the HFCs
H_{HV,H_2}	Maximum heating value of hydrogen
μ_{FC}	Fuel utilization rate
E_{ESS}	Maximum capacity of the ESS
η_{ch}/η_{dch}	Charging/discharging efficiency of the ESS
SOC^0	SOC initial value

SOC^{t_0}	SOC at time t_0
c_1/c_2	Propulsion coefficients
M_{H_2}	Molar mass
$p_{H_2,pr}$	Unit price of hydrogen fuel
$\chi/\varsigma/\phi$	Maintenance factors of the HFCs/ESS/PV system
$P_{PV}^{\min}/P_{PV}^{\max}$	Minimum/maximum output power of the PV system
$P_{FC,H_2}^{\min}/P_{FC,H_2}^{\max}$	Minimum/maximum total chemical power of the HFCs
$P_{FC}^{\min}/P_{FC}^{\max}$	Minimum/maximum actual power of the HFCs
$SOC^{\min}/SOC^{\max}$	Minimum/maximum SOC of the ESS
$P_{ESS,cha}^{\min}/P_{ESS,dis}^{\max}$	Minimum/maximum charge/discharge power of the ESS
v^{soc}	Maximum SOC deviation
$v^{\min}/v^{\max}$	Minimum/maximum speed of the ship
$\zeta_{\min}$	Minimum allowable cruising speed of the ship
Dis_n	Target voyage of the ship at time t
Dis_{term}	Terminal voyage of the ship
Dis^T	Total voyage of the ship
κ_{term}	Maximum allowable voyage variation of the ship
$R_{cost}/R_{bal}/R_{saf}$	Total economic cost reward/power fluctuation penalty/SOC safety constraint penalty
$\vartheta_{bal}/\vartheta_{saf}$	Penalty coefficients
γ	Discount factor
$\mu(s_t;\theta^{\mu})$	Actor network
$\mu'(s_t;\theta^{\mu'})$	Target actor network
$Q(s_t,a_t;\theta^{Q})$	Critic network
$Q'(s_t,a_t;\theta^{Q'})$	Target critic network
τ	Soft update parameter
$\theta^{Q}/\theta^{Q'}$	Parameters of the critic/target critic networks
$\theta^{\mu}/\theta^{\mu'}$	Parameters of the actor/target actor networks
θ/θ'	Parameters of the actor/target networks
N	Batch size
∇	Gradient
$\vartheta_{bal}/\vartheta_{saf}$	Penalty coefficient of the power balance/safety
Variables:	
P_{PV}^t	Output power of the shipboard PV system at time t
$I_{Gh}^t/I_{Bh}^t/I_{Dh}^t$	Global horizontal irradiance, direct horizontal irradiance and diffuse irradiance at time t
T_{PV}^t/T_{FC}^t	Temperature of the PV system/HFCs at time t
$V_{FC}^t/V_{oc}^t/V_{act}^t/V_{ohm}^t/V_{conc}^t$	Actual output voltage/open-circuit voltage/activation loss voltage/ohmic loss voltage/concentration loss voltage of the HFCs at time t
I_{FC}^t	Output current of the HFCs at time t
$\dot{m}_{H_2}^t$	Hydrogen mass flow rate at time t
$P_{FC,H_2}^t/P_{FC}^t$	Total chemical power/actual output power of the HFCs at time t
E_{ESS}^t	ESS capacity at time t
SOC^t	SOC at time t
$P_{ESS,theor}^t$	Theoretical charging/discharging power of the ESS at time t
P_{ESS}^t	Charging/discharging power of the ESS at time t
P_{pro}^t	Propulsion power demand at time t
v^t	Ship speed at time t
Dis^t	Voyage of the ship at time t
P_{sl}^t	Service load at time t
P^t	Total load of the ship at time t
C_{total}	Total cost
$C_{H_2}^t/C_{ESS}^t/C_{PV}^t$	Cost of the HFCs/ESS and/or the PV system at time t

$C^t_{H_2,pr}$	Cost of hydrogen fuel at time t
$C^t_{PV,cap}$	Installation cost of the PV system at time t
$C^t_{H_2,maint}/C^t_{ESS,maint}/C^t_{PV,maint}$	Maintenance cost of the HFCs/ESS/PV systems at time t
$ratio^t$	Scaling factor for the HFC load at time t
$C^t_{ESS,cap}/C^t_{PV,cap}$	Installation cost of the ESS/PV systems at time t
P^t_{sup}/P^t_{dem}	Supply/demand of total power at time t
y_t	Reward of the DDPG algorithm at time t
L_Q	Loss function of the critic network at time t
$\nabla_{\theta^\mu} J$	Gradient of the actor network at time t
L_Q	Loss function of the critic network at time t
S_t	State space of the system
A_t	Action space of the system
R_t	Reward function of the system
$R_{cost}/R_{bal}/R_{saf}$	Reward of the cost/power balance/safety
P^t_{sup}/P^t_{dem}	Power of the supply/demand at time t

References

1. Soni, G.; Neto, R.C.; Moreira, L. Hydrodynamic Simulation of Green Hydrogen Catamaran Operating in Lisbon, Portugal. *J. Mar. Sci. Eng.* **2023**, *11*, 2273. [CrossRef]
2. Inal, O.B.; Charpentier, J.-F.; Deniz, C. Hybrid Power and Propulsion Systems for Ships: Current Status and Future Challenges. *Renew. Sustain. Energy Rev.* **2022**, *156*, 111965. [CrossRef]
3. Díaz-Secades, L.A. Enhancement of Maritime Sector Decarbonization through the Integration of Fishing Vessels into IMO Energy Efficiency Measures. *J. Mar. Sci. Eng.* **2024**, *12*, 663. [CrossRef]
4. Magkouris, A.; Belibassakis, K. A Novel BEM for the Hydrodynamic Analysis of Twin-Hull Vessels with Application to Solar Ships. *J. Mar. Sci. Eng.* **2024**, *12*, 1776. [CrossRef]
5. Lu, F.; Tian, Y.; Liu, H.; Ling, C. Distributionally Robust Optimal Scheduling of Hybrid Ship Microgrids Considering Uncertain Wind and Wave Conditions. *J. Mar. Sci. Eng.* **2024**, *12*, 2087. [CrossRef]
6. Balsamo, F.; Capasso, C.; Lauria, D.; Veneri, O. Optimal Design and Energy Management of Hybrid Storage Systems for Marine Propulsion Applications. *Appl. Energy* **2020**, *278*, 115629. [CrossRef]
7. Dawoud, S.M.; Selim, F.; Lin, X.; Zaky, A.A. Techno-Economic and Sensitivity Investigation of a Novel Perovskite Solar Cells Based High Efficient Hybrid Electric Sources for Off-Shore Oil Ships. *IEEE Access* **2023**, *11*, 41635–41643. [CrossRef]
8. Wu, C.-H.; Wang, H.-C.; Chang, H.-Y. Greenhouse Gas Emissions Reduction and Energy Savings for a Dredger at Port Area Using Hybrid Photovoltaic Power System Onboard. *Energy Sustain. Dev.* **2024**, *78*, 101354. [CrossRef]
9. Zhu, J.; Chen, L. A Probabilistic Multi-Objective Design Method of Sail-Photovoltaic-Hybrid Power System for an Unmanned Ocean Surveillance Trimaran. *Appl. Energy* **2023**, *350*, 121604. [CrossRef]
10. A Hybrid Ensemble Model for Interval Prediction of Solar Power Output in Ship Onboard Power Systems. Available online: https://ieeexplore.ieee.org/abstract/document/8946708 (accessed on 25 December 2024).
11. Gaber, M.; Hamad, M.S.; El-banna, S.H.; El-Dabah, M. An intelligent energy management system for ship hybrid power system based on renewable energy resources. *J. Al-Azhar Univ. Eng. Sect.* **2021**, *16*, 712–723. [CrossRef]
12. Tang, R.; Li, X.; Lai, J. A Novel Optimal Energy-Management Strategy for a Maritime Hybrid Energy System Based on Large-Scale Global Optimization. *Appl. Energy* **2018**, *228*, 254–264. [CrossRef]
13. Maaruf, M.; Khalid, M. Hybrid Solar/PEM Fuel Cell/and Water Electrolyzer Energy System for All-Electric Ship. In Proceedings of the 2022 IEEE Kansas Power and Energy Conference (KPEC), Manhattan, KS, USA, 25–26 April 2022; pp. 1–5.
14. Study and Analysis of a Solar Electric Boat with Dynamic Design Strategy in Efficient Way by Tawheed Hasan, Shahrizan Jamaludin, WB Wan Nik, Mehedi Hasan Rajib: SSRN. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4446339 (accessed on 25 December 2024).
15. Abdelrahman, M.S.; Hussein, H.; Mohammed, O.A. Rule-Based Power and Energy Management System for Shipboard Microgrid with HESS to Mitigate Propulsion and Pulsed Load Fluctuations. In Proceedings of the 2023 IEEE Green Technologies Conference (GreenTech), Denver, CO, USA, 19–21 April 2023; pp. 224–228.
16. Song, T.; Fu, L.; Zhong, L.; Fan, Y.; Shang, Q. HP3O Algorithm-Based All Electric Ship Energy Management Strategy Integrating Demand-Side Adjustment. *Energy* **2024**, *295*, 130968. [CrossRef]
17. Yiğit, K.; Acarkan, B. A New Ship Energy Management Algorithm to the Smart Electricity Grid System. *Int. J. Energy Res.* **2018**, *42*, 2741–2756. [CrossRef]

18. Hein, K. Emission-Aware and Data-Driven Many-Objective Voyage and Energy Management Optimization of Solar-Integrated All-Electric Ship. *Electr. Power Syst. Res.* **2022**, *213*, 108718. [CrossRef]
19. Identifying Hybrid Renewable Energy Power Management and Speed Control on Catamaran Ship | IEEE Conference Publication | IEEE Xplore. Available online: https://ieeexplore.ieee.org/document/10335546 (accessed on 27 December 2024).
20. Fang, S.; Xu, Y.; Wen, S.; Zhao, T.; Wang, H.; Liu, L. Data-Driven Robust Coordination of Generation and Demand-Side in Photovoltaic Integrated All-Electric Ship Microgrids. *IEEE Trans. Power Syst.* **2020**, *35*, 1783–1795. [CrossRef]
21. Igder, M.A.; Rafiei, M.; Boudjadar, J.; Khooban, M.-H. Reliability and Safety Improvement of Emission-Free Ships: Systemic Reliability-Centered Maintenance. *IEEE Trans. Transp. Electrif.* **2021**, *7*, 256–266. [CrossRef]
22. Wen, S.; Zhao, T.; Tang, Y.; Xu, Y.; Zhu, M.; Huang, Y. A Joint Photovoltaic-Dependent Navigation Routing and Energy Storage System Sizing Scheme for More Efficient All-Electric Ships. *IEEE Trans. Transp. Electrif.* **2020**, *6*, 1279–1289. [CrossRef]
23. Dolatabadi, S.H.; Ölçer, A.I.; Vakili, S. The Application of Hybrid Energy System (Hydrogen Fuel Cell, Wind, and Solar) in Shipping. *Renew. Energy Focus* **2023**, *46*, 197–206. [CrossRef]
24. Comparative Study Based on Techno-Economics Analysis of Different Shipboard Microgrid Systems Comprising PV/Wind/Fuel Cell/Battery/Diesel Generator with Two Battery Technologies: A Step toward Green Maritime Transportation—ScienceDirect. Available online: https://www.sciencedirect.com/science/article/abs/pii/S0960148123015859 (accessed on 27 December 2024).
25. Probabilistic Coordination of Optimal Power Management and Voyage Scheduling for All-Electric Ships. Available online: https://ieeexplore.ieee.org/abstract/document/10246416 (accessed on 27 December 2024).
26. Lan, H.; Gao, J.; Hong, Y.-Y.; Yin, H. Interval Forecasting of Photovoltaic Power Generation on Green Ship under Multi-Factors Coupling. *Sustain. Energy Technol. Assess.* **2023**, *56*, 103088. [CrossRef]
27. Wang, Z.; Liao, P.; Liu, S.; Ji, Y.; Han, F. Scenario-Based Energy Management Optimization of Hydrogen-Electric-Thermal Systems in Sustainable Shipping. *Int. J. Hydrog. Energy* **2025**, *99*, 566–578. [CrossRef]
28. Hou, H.; Gan, M.; Wu, X.; Xie, K.; Fan, Z.; Xie, C.; Shi, Y.; Huang, L. Real-Time Energy Management of Low-Carbon Ship Microgrid Based on Data-Driven Stochastic Model Predictive Control. *CSEE J. Power Energy Syst.* **2023**, *9*, 1482–1492. [CrossRef]
29. Huang, Y.; Lan, H.; Hong, Y.-Y.; Wen, S.; Fang, S. Joint Voyage Scheduling and Economic Dispatch for All-Electric Ships with Virtual Energy Storage Systems. *Energy* **2020**, *190*, 116268. [CrossRef]
30. Lan, H.; Wen, S.; Hong, Y.-Y.; Yu, D.C.; Zhang, L. Optimal Sizing of Hybrid PV/Diesel/Battery in Ship Power System. *Appl. Energy* **2015**, *158*, 26–34. [CrossRef]
31. Liang, H.; Pirouzi, S. Energy Management System Based on Economic Flexi-Reliable Operation for the Smart Distribution Network Including Integrated Energy System of Hydrogen Storage and Renewable Sources. *Energy* **2024**, *293*, 130745. [CrossRef]
32. Ship Energy Scheduling with DQN-CE Algorithm Combining Bi-Directional LSTM and Attention Mechanism—ScienceDirect. Available online: https://www.sciencedirect.com/science/article/pii/S0306261923007420 (accessed on 28 February 2025).
33. Drungilas, D.; Kurmis, M.; Senulis, A.; Lukosius, Z.; Andziulis, A.; Januteniene, J.; Bogdevicius, M.; Jankunas, V.; Voznak, M. Deep Reinforcement Learning Based Optimization of Automated Guided Vehicle Time and Energy Consumption in a Container Terminal. *Alex. Eng. J.* **2023**, *67*, 397–407. [CrossRef]
34. Shang, C.; Fu, L.; Bao, X.; Xu, X.; Zhang, Y.; Xiao, H. Energy Optimal Dispatching of Ship's Integrated Power System Based on Deep Reinforcement Learning. *Electr. Power Syst. Res.* **2022**, *208*, 107885. [CrossRef]
35. Long, L.N.B.; You, S.-S.; Cuong, T.N.; Kim, H.-S. Optimizing Quay Crane Scheduling Using Deep Reinforcement Learning with Hybrid Metaheuristic Algorithm. *Eng. Appl. Artif. Intell.* **2025**, *143*, 110021. [CrossRef]
36. Li, W.; Cui, H.; Nemeth, T.; Jansen, J.; Ünlübayir, C.; Wei, Z.; Feng, X.; Han, X.; Ouyang, M.; Dai, H.; et al. Cloud-Based Health-Conscious Energy Management of Hybrid Battery Systems in Electric Vehicles with Deep Reinforcement Learning. *Appl. Energy* **2021**, *293*, 116977. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.