Journal of *Marine Science and Engineering*

MDPI

Article

A New Energy Management Strategy Supported by Reinforcement Learning: A Case Study of a Multi-Energy Cruise Ship

Xiaodong Guo [1,2,3], Daogui Tang [1,2,4,*], Yupeng Yuan [1,2,4,*], Chengqing Yuan [1,2,4], Boyang Shen [5] and Josep M. Guerrero [6,7,8]

[1] State Key Laboratory of Maritime Technology and Safety, Wuhan University of Technology, Wuhan 430062, China; xiaodongguo@whut.edu.cn (X.G.); ycq@whut.edu.cn (C.Y.)
[2] National Engineering Research Center for Water Transport Safety, Wuhan University of Technology, Wuhan 430062, China
[3] School of Naval Architecture, Ocean and Energy Power Engineering, Wuhan University of Technology, Wuhan 430062, China
[4] School of Transportation and Logistics Engineering, Wuhan University of Technology, Wuhan 430062, China
[5] Department of Engineering, University of Cambridge, Cambridge CB2 1TN, UK
[6] Center for Research on Microgrids (CROM), Department of Electronic Engineering, Technical University of Catalonia, 08019 Barcelona, Spain; joz@energy.aau.dk
[7] Catalan Institution for Research and Advanced Studies (ICREA), Pg. Lluís Companys 23, 08010 Barcelona, Spain
[8] Center for Research on Microgrids (CROM), AAU Energy, Aalborg University, 9220 Aalborg, Denmark
* Correspondence: tangdaogui@gmail.com (D.T.); ypyuan@whut.edu.cn (Y.Y.)

Abstract: Hybrid ships offer significant advantages in energy efficiency and environmental sustainability. However, their complex structures present challenges in developing effective energy management strategies to ensure optimal power distribution and stable, efficient operation of the power system. This study establishes a mathematical model of a hybrid system for a specific ship and proposes an energy management strategy based on the deep deterministic policy gradient (DDPG) algorithm, a reinforcement learning technique. The proposed strategy's feasibility and effectiveness are validated through comparisons with alternative energy management strategies and real-world ship data. Simulation results demonstrate that the DDPG-based strategy optimizes the diesel engine's operating conditions and reduces total fuel consumption by 3.6% compared to a strategy based on the deep Q-network (DQN) algorithm.

Keywords: allocation; deep deterministic policy gradient; energy management strategies; hybrid ship; reinforcement learning

check for updates

Received: 14 March 2025
Revised: 30 March 2025
Accepted: 1 April 2025
Published: 3 April 2025

Citation: Guo, X.; Tang, D.; Yuan, Y.; Yuan, C.; Shen, B.; Guerrero, J.M. A New Energy Management Strategy Supported by Reinforcement Learning: A Case Study of a Multi-Energy Cruise Ship. *J. Mar. Sci. Eng.* **2025**, *13*, 720. https://doi.org/10.3390/jmse13040720

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

1.1. Background

Because of the carbon peak and carbon neutrality goals, the shipping industry, which is the major carbon emitter of the transportation industry, must save more energy and reduce its emissions [1]. Therefore, the development of green ships with low energy consumption, low emissions, and high efficiency and the development of new energy power technology for ships have become the top priority of the shipbuilding industry [2].

Against this background, hybrid ships are widely developed and used worldwide. A hybrid ship is defined as a ship whose power system consists of two or more different sources of energy. The use of alternative energy sources is a distinctive feature of hybrid

ships. For example, they run on solar energy, wind energy, and fuel cells. When different energy sources are combined, they complement each other, thus overcoming the limitations of the use of a single energy source [3].

However, due to the differing operating characteristics and electricity generation methods of respective energy systems, a hybrid system that integrates new energy systems with traditional diesel engine systems is inherently more complex. Therefore, efficient cooperation between traditional diesel engine systems and new energy systems is essential [4]. The ship's energy management strategy functions as the brain of its power system. Its role is to optimize the processes of power generation, distribution, and consumption; to intelligently control the power system; to optimize the working conditions of the ship's power system components while meeting power demands; and to enhance the ship's fuel economy [5].

1.2. Literature Review

Classical transport system models, such as those leveraging AIS data to analyze maritime networks [6], have laid the groundwork for understanding ship routing, congestion, and energy demand patterns. However, these studies primarily focus on macroscopic network efficiency rather than real-time, ship-level energy management. Building on these foundations, our work extends classical models by introducing reinforcement learning (RL) to address dynamic energy allocation problems in hybrid ships. As shown in Figure 1, we categorize energy management strategies into three classes: rule-based energy management strategies, optimization-based energy management strategies, and learning-based energy management strategies.

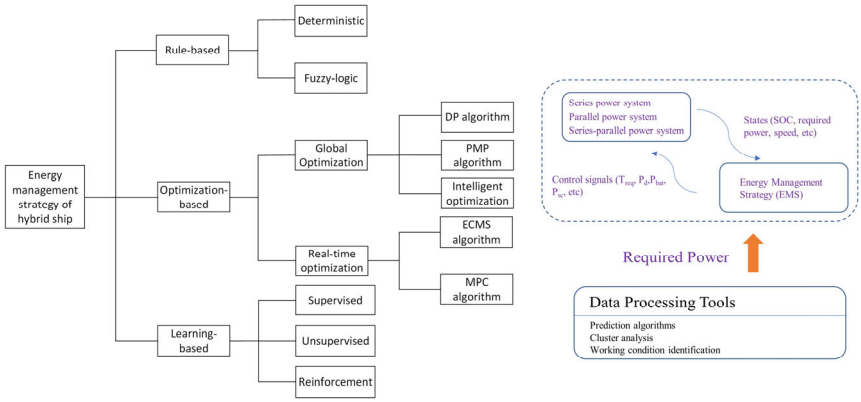

Figure 1. Classification of strategies for ship energy management.

Shen [7] proposed a robust fuzzy control method to design nonlinear control laws for energy management systems. By addressing uncertainties in driver power demand, this approach enables coordinated operation among fuel cell stack health state estimators, energy storage system schedulers, and optimization objectives. In a complementary study, Seng [8] developed a hybrid grid system integrating photovoltaics, wind turbines, and battery storage, accompanied by a fuzzy logic-based energy management framework. While this framework demonstrates effective load–demand balancing and ensures reliable power supply, its broader implementation in Malaysia requires further investigation of regional adaptability and scalability. Compared to conventional rule-based methods, fuzzy rule-based strategies exhibit enhanced intelligence through fuzzified operational criteria. These strategies dynamically classify power system operating modes using human-like decision-making mechanisms, thereby bridging the gap between automated control and expert judgment. However, the development of fuzzy rules relies heavily on engineering expertise and experimental data, limiting their adaptability and scalability. Overall, both

rule-based and fuzzy rule-based energy management strategies often depend on empirical data and engineering experience, leading to suboptimal solutions that lack global optimization potential [9].

Optimization-based energy management strategies are founded on optimization theories and methods, designing corresponding algorithms according to various objective functions and constraints. Global optimization-based energy management strategies utilize algorithms such as dynamic programming (DP) [10] and Pontryagin's minimum principle (PMP) [11]. These strategies aim to optimize single or multiple objectives, including ship economy, emissions, and power performance. They do so by setting the energy supply–demand balance and safe operation of equipment as constraints, obtaining the optimal power distribution solution under the premise of having prior partial knowledge of operating conditions.

Kanellos [12] combined DP and particle swarm optimization to manage energy consumption with the goal of minimizing the consumption of a ship and greenhouse gas emissions. The research results showed that the method ensures the stable operation of the ship's power system and improves its energy efficiency. However, the DP-based energy management strategy is computationally heavy and time-consuming, which limits its use in practice. To overcome these problems, Zhang [13] used the stochastic DP algorithm, which treats the demand as a Markov chain with state transfer probabilities and then uses dynamic programming methods to determine the optimal control strategy. This way, it shortens the computation time and ensures optimization. However, an energy management strategy must be able to operate offline, anticipate all operating conditions, and achieve real-time responses. Ou [14] proposed an energy management strategy based on the adaptive PMP algorithm for a hybrid power system consisting of a fuel cell and battery. The researchers conducted simulation experiments and built a physical platform to verify the proposed strategy. The experimental results showed that the energy management strategy can force the battery to work within a reasonable charge state and improve the system economy. However, the implementation of an energy management strategy based on global optimization requires knowledge about all the operating conditions in advance. This makes the strategy only applicable for simulation analysis because it is too difficult for real systems.

A real-time optimization strategy can optimize and control the real-time conditions faced by hybrid ships. It provides strong real-time performance, requires a smaller computational load than global optimization, and is easy to implement. Although it does not need complete information on the operating conditions beforehand, its optimization results are not guaranteed to be globally optimal. Common real-time optimal control strategies include the equivalent consumption minimization strategy (ECMS) and model predictive control (MPC). Kalikatzarakis [15] utilized a ship's historical operational data to train an adaptive equivalent factor based on predicted power demand. Experimental results showed that the adaptive ECMS (A-ECMS) strategy reduced fuel consumption by approximately 4% compared to the fixed-factor ECMS, while maintaining a more stable state of charge (SOC) in the energy storage system. Despite its real-time applicability, ECMS is constrained by limited optimization performance and weak control robustness. MPC addresses these limitations through model prediction, rolling optimization, and feedback correction, with performance influenced by model accuracy, sampling intervals, and prediction horizons. Zhang [16] proposed a two-level MPC framework, where the high-level MPC optimizes fuel consumption trends, and the low-level MPC manages supercapacitor power distribution. Simulation results demonstrate that this strategy effectively reduces fuel consumption of the diesel generator and maintains the battery state of charge within a reasonable range, outperforming traditional MPC approaches. However, high-frequency power adjustments

may increase computational complexity, and the robustness and real-time performance in complex scenarios require further enhancement. Kofler [17] proposed an EMS for fuel cell electric vehicles that integrates both long-term and short-term predictions to optimize the power distribution between the fuel cell system and the battery. This approach effectively combines long-term static forecasts with real-time short-term data, enhancing performance, particularly in fuel efficiency and load cycles. However, the method's limitations include the rough discretization used in the DP algorithm, which may reduce accuracy in complex driving conditions, and the reliance of MPC on the accuracy of short-term predictions, which may be influenced by real-time data fluctuations. In conclusion, the development of effective ship energy management strategies necessitates the construction of an accurate system model. Nevertheless, the growing intricacy of ship models presents a challenge in accurately modeling the system. Consequently, it is of paramount importance to investigate methods that do not necessitate the use of accurate models or a priori knowledge [18].

As a non-model-based intelligent optimization algorithm, reinforcement learning is extraordinarily suitable for the design of energy management strategies. It does not rely on expert experience and information about the working conditions of the complete segment. The optimized model can be trained according to the current information of the ship [19]. The deep learning method identifies and extracts the characteristics of the system with the help of neural networks to realize the perception of the system state, which is suitable for the classification of working conditions and load forecasting of the ship's power demand.

Huotari [20] developed an energy management strategy based on Q-learning for generator startup control and load sharing. However, the method only allocates load based on the rated capacity ratio of the generator, ignoring energy storage devices and operational constraints. Xiao [21] developed an enhanced DQN algorithm for hybrid-powered vessel energy management optimization. Experimental validation demonstrates that the proposed strategy achieves a 4.11% reduction in fuel consumption, 24.4% improvement in energy storage system efficiency, and 31.3% shorter exploration time during agent training compared to conventional DQN-based approaches. Fu [22] proposed RL-C-DMPC, a reinforcement learning-compensated distributed model predictive control strategy, to mitigate power imbalances in shipboard systems caused by prediction uncertainties. It employs value-decomposition networks for training and reduces imbalances by 90% yet lacks validation under diverse environmental dynamics and long-term operational scenarios. Chen [23] proposed an optimized power management framework for hybrid fuel cell vessels employing MPC and reinforcement learning coordination. Their methodology systematically minimizes operational costs through multi-objective optimization, considering fuel consumption, emissions, and component degradation. Comparative analysis revealed RL-based solutions achieve superior fuel economy under ideal conditions, while noting potential stability concerns in real-world operational scenarios.

Rule-based control strategies offer simplicity and fast computation, often using lookup tables or state machines [24]. However, their optimization capacity is limited due to heuristic rule dependency, although their practicality maintains wide engineering applications. Optimization-based strategies generally outperform rule-based methods but face inherent trade-offs; global optimization lacks real-time feasibility, while real-time optimization sacrifices global optimality. Unlike other machine learning methods, reinforcement learning focuses more on decision-making and does not rely on models. Instead, it obtains corresponding rewards through the interaction between the agent and the environment, thereby learning and optimizing strategies. RL provides a model-free alternative that bypasses precise digital modeling requirements. Once trained, RL agents execute energy management through pre-mapped state–action policies, eliminating online computations [25]. This

unique advantage positions RL as a transformative paradigm for energy management systems, driving its rapid adoption in recent research.

1.3. Motivation and Contributions

In conclusion, the majority of hybrid ships achieve the power allocation of the hybrid power system through rule-based and optimization-based energy management strategies. The formulation of the strategies either relies on engineering experience or cannot be applied in real time. Although reinforcement learning is a widely used technique in the field of distributed energy microgrids and hybrid vehicles, its application in the field of hybrid ships is less common [26]. Consequently, the specific challenges currently encountered by reinforcement learning-based energy management strategies for ships are as follows:

(1) The action space of reinforcement learning algorithms is predominantly discrete, with action values that will inevitably result in discretization errors. One traditional solution is to enhance the precision of discretization. However, this improvement in precision may result in a significant increase in dimensional complexity, which could negatively impact the calculation speed. This limitation significantly constrains the applicability of reinforcement learning algorithms in the energy management of hybrid ships.
(2) Existing energy management strategies are less likely to utilize the output power of multiple diesel generator sets as an action space, thereby preventing the coordinated operation of multiple generator sets and energy storage units to achieve the objective of optimal fuel efficiency.
(3) The most prevalent approach to addressing intricate constraints during maritime navigation is the penalty function methodology. However, when there are multiple constraints, it is of the utmost importance to select the most appropriate penalty coefficients.

Inspired by the above research and analyzing the literature, the motivation of this paper can be divided into three aspects:

(1) Combining the structure of deep learning and the idea of reinforcement learning and synthesizing the advantages of both, cutting-edge deep reinforcement learning algorithms are introduced in the energy management problem of hybrid ships. Energy management strategies such as DDPG based on the actor–critic framework are proposed for the control problem on continuous state–action space.
(2) Multiple energy sources increase the complexity of the system, and it is a key problem for hybrid ships to rationally distribute the power output of each energy unit and improve the fuel economy of the whole ship without jeopardizing the healthy life of the components. Therefore, the strategy proposed in this study considers the cooperative work of multiple diesel generator sets and supercapacitors.
(3) Through comparative analysis, the superiority of the DDPG algorithm to other reinforcement learning algorithms and rule-based strategies, as well as real fuel consumption data of real ships, is concluded.

2. Mathematical Model of Hybrid Power System

2.1. Object Ship

In this study, the five-star Yangtze River cruise ship "Victoria Sabrina" was chosen as the research object (Figure 2). The ship adopts the electric propulsion mode of a DC network, which exhibits a low noise level, low vibrations, and good comfort [27].



Figure 2. Victoria Sabrina hybrid cruise ship.

The Victoria Sabrina uses a series hybrid power system, the structure of which is shown in Figure 3. The main components and parameters of the power system are listed in Table 1.

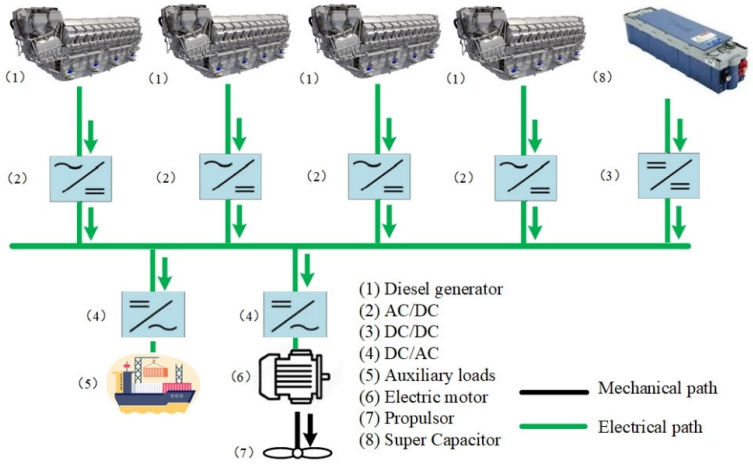

Figure 3. Structure of Victoria Sabrina hybrid power system.

Table 1. Main component parameters of Victoria Sabrina hybrid power system.

Component	Parameter	Value
Prime motor	Maximum Power/kW	1935
	Rated Power/kW	1320
	Rated Speed (r/min)	1000
	Quantity	4
Generator	Rated Speed/kW	1250
	Rated Torque/Nm	11,937.5
	Rated Speed (r/min)	1000
	Quantity	4
Super capacitor	Capacity /F	129
	Maximum/Minimum Voltage/V	820/604
	Rated Power/kW	1000
	Quantity	2

Table 1. *Cont.*

Component	Parameter	Value
Main propulsion	Rated Voltage/V	690
	Rated Power/kW	1680
	Quantity	2
Side propulsion	Rated Power/kW	450
	Rated Voltage/V	690
	Rated Speed (r/min)	1500
	Quantity	1

2.2. Modeling of Diesel Generator

The model built in this study mainly serves as a simulation object for the energy management strategy of the hybrid power system. It focuses on fuel consumption and power output and ignores the internal workings of the engine, temperature, and heat transfer. Instead of using the principle modeling method, the data modeling approach was used to construct the model for the diesel generator. The data modeling method uses the test data from an engine bench test such that a data cloud map of the engine fuel consumption can be generated. By querying the cloud diagram of the diesel generator, the fuel consumption data for the corresponding power level of the engine can be obtained via interpolation. This method reduces the complexity of the engine modeling process and ensures model accuracy. The experimentally obtained cloud diagram of the oil consumption data is shown in Figure 4.

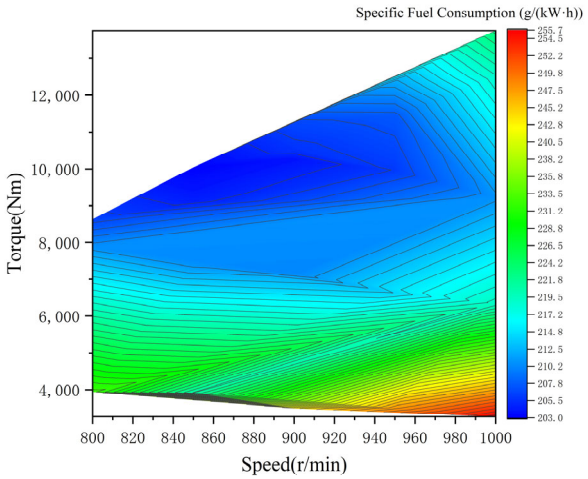

Figure 4. Fuel consumption cloud map of diesel generator.

The specific fuel consumption b_e of a diesel engine is determined by the speed and torque.

$$b_e = f(N_e, T_e) \tag{1}$$

where b_e is the specific fuel consumption, g/(kW·h); N_e is the speed of diesel engine, r/min; and T_e is the torque of the diesel engine, Nm.

The fuel consumption of the diesel engine per unit time is $\dot{m}_e$.

$$\dot{m}_e = \frac{P_e b_e}{3600} \tag{2}$$

where $\dot{m}_e$ is the fuel consumption of the diesel engine per unit time, g/s; P_e is the output power of the diesel engine, kW.

Thus, the total fuel consumption over a certain period is as follows:

$$m_e = \int_{t_0}^{t_f} \dot{m}_e dt \tag{3}$$

where t_0 is the start, s; t_f is the end, s; and m_e is the fuel consumption, g. According to Equations (2) and (3),

$$m_e = \int_{t_0}^{t_f} \frac{P_e b_e}{3600} dt \tag{4}$$

To ensure that the diesel engine operates normally, its output power should not exceed its own maximum output power.

$$0 \leq P_e \leq P_{e_max} \tag{5}$$

where P_{e_max} is the maximum output power of the diesel engine (1320 kW according to the data provided by the ship manufacturer).

The generator model is the efficiency model of the generator. In this study, the generator efficiency η_{gen} is set to 0.955, and the output power of the generator is as follows:

$$P_{gen} = P_e \cdot \eta_{gen} \tag{6}$$

where P_{gen} is the output power of the diesel engine, kW.

2.3. Modeling of Supercapacitor

The super capacitor adopts a simplified equivalent circuit model, as shown in Figure 5. The supercapacitor in this study is primarily used to enhance the hybrid system's response to load fluctuations. Even without other renewable energy sources, the supercapacitor effectively smooths load variations, reducing the frequent start–stop of the diesel generator and improving system efficiency and stability [28].

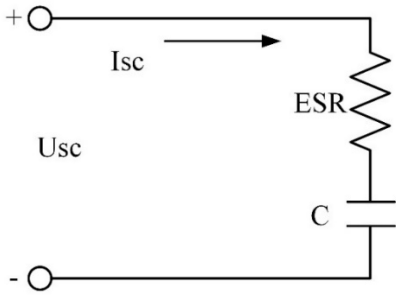

Figure 5. Super capacitor theoretical model.

The output power and SOC of the supercapacitor can be calculated with Equations (7) and (8):

$$P_{sc} = U_{sc} I_{sc} - I_{sc}^2 R_{sc} \tag{7}$$

where P_{sc} is the output power of the supercapacitor, kW; U_{sc} is the terminal voltage of the supercapacitor, V; I_{sc} is the current of the supercapacitor, A; and R_{sc} is the resistance of the supercapacitor, Ω.

$$SOC = \frac{U_{sc} - U_{sc_min}}{U_{sc_max} - U_{sc_min}} \tag{8}$$

where U_{sc_max} and U_{sc_min} are the maximum and minimum voltages of the supercapacitor V, respectively. According to the nomenclature plate of the supercapacitor, $U_{sc_max} = 820V$, and $U_{sc_min} = 604V$.

According to Equation (8), the SOC of the supercapacitor is linearly proportional to its terminal voltage. For convenience, the denominator in Equation (8) was set to U_{scm}, and the molecule was set to U_{sc}:

$$SOC = \frac{U_{sc}}{U_{scm}} \tag{9}$$

To ensure that the supercapacitor operates normally and to avoid overcharging and over-discharging, the supercapacitor SOC is within [0.3, 0.8] based on the data and engineering experience of the manufacturer. That is, $0.3 = SOC_{min} \leq SOC \leq SOC_{max} = 0.8$.

The supercapacitor current can be determined with Equations (7) and (9).

$$I_{sc} = \frac{U_{scm}SOC - \sqrt{(U_{scm}SOC)^2 - 4P_{sc}R_{sc}}}{2R_{sc}} \tag{10}$$

To ensure that the power system operates safely, the supercapacitor should work within its allowable power range to prevent excessive power caused by a plate breakdown.

$$P_{sc_min} \leq P_{sc} \leq P_{sc_max} \tag{11}$$

where P_{sc_min} is the minimum output power of the supercapacitor, kW; P_{sc_max} is the maximum output power of the supercapacitor, kW.

2.4. Modeling of Load

In this study, the real-time power load data measured by a target vessel on a section of the Chongqing–Yichang route were applied as the power demanded of the vessel (Figure 6).

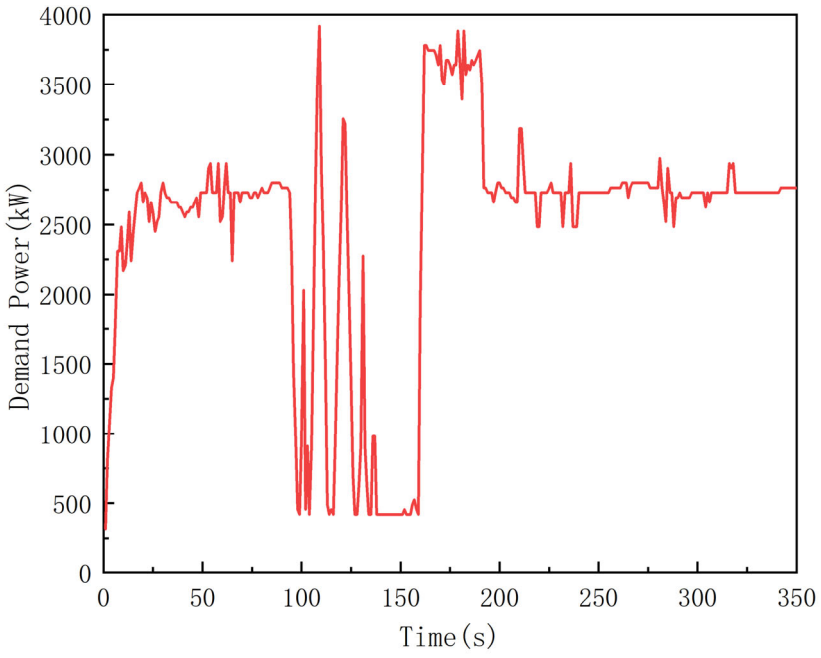

Figure 6. Ship demanded power.

To simulate the working state of the ship's power system under actual working conditions, to ensure that the ship's power system operates reliably and safely, and to simplify the simulation process, four diesel generator sets and a supercapacitor system were simultaneously run in the network to validate the method; the equipment jointly

supplied the ship with power. By considering the structure of a hybrid power system, a full ship power balance model was obtained.

$$P_{req} = (P_1 + P_2 + P_3 + P_4)\eta_{gen} + P_{sc} \cdot \eta_{DC} \tag{12}$$

where $P_{sc} > 0$ indicates that the supercapacitor is discharged; $P_{sc} < 0$ indicates that the supercapacitor is being charged; and η_{DC} is the DC/DC efficiency of the current converter, which is set to 0.98 according to the data provided by the manufacturer.

3. Proposed Energy Management Strategy

3.1. DDPG Algorithm

The DDPG algorithm is a policy-based reinforcement learning algorithm that has evolved from the policy gradient (PG) algorithm. In engineering practice, the output power P_e of the prime mover of the four diesel generators is a continuous variable because a continuous action output is closer to actual engineering conditions. While PPO and SAC can operate in continuous domains, they often require additional discretization layers or policy parameterization to achieve comparable control precision. Furthermore, DDPG's deterministic policy during deployment significantly reduces stochastic fluctuations in power allocation—a critical safety advantage for vessel operations where unstable power distribution could compromise navigation reliability. The PG algorithm is a classical algorithm for learning continuous actions in reinforcement learning. The optimal policy for each step is represented by a probability distribution $\pi(s_t|\theta^\pi)$. Subsequently, the action is sampled according to this probability distribution to obtain the best current action value:

$$a = \pi(s|\theta^\pi) \tag{13}$$

This stochastic gradient algorithm must sample the probability distribution of the optimal strategy to obtain the specific value of the current action while acting in each step. Therefore, the computational complexity of the algorithm is high when the system is a high-dimensional action space [29]. By contrast, in the DDPG algorithm, the action of each step is directly determined by the function μ, rather than by probability resampling the output action. This deterministic PG algorithm accelerates the calculation and improves convergence. The determination of the action is expressed as follows:

$$a = \mu(s|\theta^\mu) \tag{14}$$

The DDPG algorithm is based on the actor–critic (AC) network framework (Figure 7). It incorporates the previously presented deterministic PG algorithm into the AC network framework.

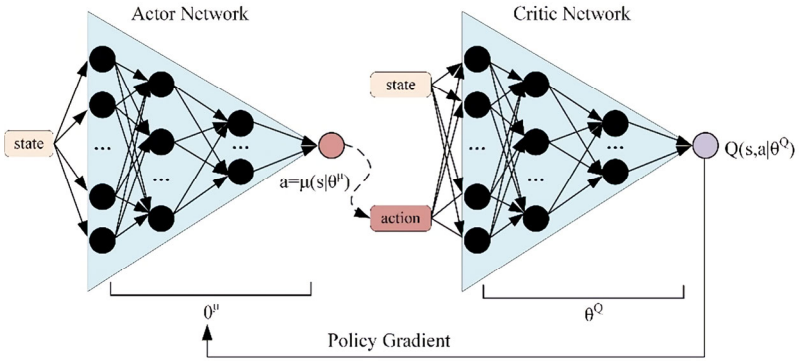

Figure 7. Actor–critic network framework structure.

The deep neural network actor with the parameter θ^{μ} is used to fit the strategy function $a = \mu(s|\theta^{\mu})$. It interacts with the environment according to the continuous action determined by the input state vector output. The deep neural network critic with the parameter θ^{Q} is used to fit the state–action value function $Q(s,a) = Q(s,a|\theta^{Q})$, evaluate the corresponding actions of the state and output, and update the actor network parameters according to the gradient ascent algorithm [30]. For the actor and critic networks, an estimation network and target network are the actor and critic to improve the stability of the algorithm, respectively. The Q estimation network and target Q network in the critic network are expressed as follows, respectively:

$$Q(s,a) = Q\left(s,a\middle|\theta^{Q}\right) \tag{15}$$

$$Q_{target} = r + \gamma maxQ\left(s',a'\middle|\theta^{Q'}\right) \tag{16}$$

The parameters of the Q estimation network are updated by minimizing the error between the target Q network and Q estimation network, namely the loss function L.

$$L\left(\theta^{Q}\right) = \left(Q\left(s,a\,|\,\theta^{Q}\right) - \left(r + \gamma maxQ\left(s',a'\middle|\theta^{Q'}\right)\right)\right)^{2} \tag{17}$$

The action a' for maximizing the Q value is obtained by the deterministic strategy and not by traversing the Q value in the DQN algorithm. That is, $a' = \mu(s|\theta^{\mu})$. Thus, the loss function can be rewritten as follows:

$$L\left(\theta^{Q}\right) = \left(Q\left(s,a\,|\,\theta^{Q}\right) - \left(r + \gamma maxQ\left(s',\mu\left(s'\middle|\theta^{\mu}\right)\middle|\theta^{Q'}\right)\right)\right)^{2} \tag{18}$$

By minimizing the loss function, the parameter θ^{Q} in the critic network is updated. Its specific update process consists of the following two steps:

(1) The loss function $L(\theta^{Q})$ takes the partial derivative $\nabla_{\theta^{Q}} L(\theta^{Q})$ with respect to the network parameters θ^{Q};

(2) $\theta^{Q} = \theta^{Q} + \alpha \nabla_{\theta^{Q}} L(\theta^{Q})$;

where α is the learning rate in the gradient descent.

While the critic network is constantly being updated to better estimate the state–action pair Q values, the actor network must be constantly updated to obtain better value functions. Therefore, the parameters of the actor network must be updated in the direction of the increasing Q. It is updated along the direction of the partial derivative of $Q(s,a|\theta^{Q})$ with respect to a, $a = a + \nabla_{a} Q(s,a|\theta^{Q})$. The action a is determined by the state s and actor network parameter θ^{μ}. That is, $a = \mu(s|\theta^{\mu})$. According to the chain rule, the update of action a can be expressed as the update of the actor network parameter θ^{μ}. The specific update rule is as follows:

$$\theta^{\mu} = \theta^{\mu} + \nabla_{a} Q\left(s,a\middle|\theta^{Q}\right) \nabla_{\theta^{\mu}} \mu(s|\theta^{\mu}) \tag{19}$$

In addition, the DDPG algorithm introduces an empirical replay mechanism and a parameter freezing mechanism. However, instead of directly pasting the parameters of the estimated network into the target network every C round, such as in the DQN algorithm, its parameter freezing mechanism gradually approximates the estimated network parameters to the target network by a tiny amount.

$$\theta^{Q'} = \tau\theta^{Q} + (1-\tau)\theta^{Q'} \tag{20}$$

$$\theta^{\mu'} = \tau\theta^{\mu} + (1-\tau)\theta^{\mu'} \tag{21}$$

where τ is the approximation coefficient, $0 < \tau \ll 1$.

3.2. Energy Management Strategy Based on DDPG Algorithm

Because of the basic principles of reinforcement learning, it can solve the following problems:

(1) The state of the system at the next instance depends only on the state of the system at the current moment and on the action taken, which means that the system has Markovian property.
(2) The optimization objective of the system is not the optimum at a specific moment in time. It is the cumulative target optimum of the entire process.
(3) The system does not necessarily need a specific model in which the system controller only needs to be aware of the current system state and the reward for the action to be taken.

Correspondingly, the energy management problem of diesel–electric series hybrid ships studied in this study also has the previously presented characteristics:

(1) The power demand of the system at the next instance and the SOC of the supercapacitor are constantly changing. In addition, the system state at the next instance is only related to the system controller action at the current moment.
(2) The optimization objective of the energy management problem is to minimize the fuel consumption of the ship during the entire journey rather than only at a certain time.
(3) The controller cannot predict the power demand of the ship during its journey. The currently consumed fuel volume can be determined, and the system enters the next state after the current action has been executed.

According to the previously presented analysis, the reinforcement learning algorithm is suitable for energy management of hybrid ships. The energy management strategy based on the DDPG algorithm for the series hybrid system of the target ship is designed as follows:

(1) State: The selection of the state space S directly affects the optimization effect of the reinforcement learning strategy and the computation speed of the reinforcement learning algorithm. In this study, the power demand P_{req}, SOC of the supercapacitor, and running time t of the hybrid ship during the entire operation were considered the state variables of reinforcement learning. The SOC range is [0.3, 0.8] according to the data provided by the manufacturer.

$$S = \left\{ P_{req}, SOC, t \right\} \tag{22}$$

(2) Action: For the series-connected ship hybrid system investigated in this study, the system state SOC and power demand P_{req} must be able to make a state transfer after an action. The hybrid system is rewarded according to this action. Therefore, the action space in the presented hybrid ship energy management strategy based on the DDPG algorithm is set to the output power P_i ($i = 1, 2, 3, 4$ of each of the four diesel generator prime movers:

$$A = \{continuousP_1,\ continuousP_2,\ continuousP_3,\ continuousP_4\} \tag{23}$$

where $P_i \in [0, 1320\text{kW}]$

(3) Reward Function: The objective of the reinforcement learning algorithm is to determine the optimal set of strategies that maximize the cumulative discount reward. For the energy management strategy of the series ship hybrid system, the optimization goal is to minimize the fuel consumption of the entire ship. Therefore, the reward

function is set to the quantity associated with the sum of the specific fuel consumption and power consumption of the supercapacitor at a given moment.

$$r = \dot{m}_e(P_1) + \dot{m}_e(P_2) + \dot{m}_e(P_3) + \dot{m}_e(P_4) + \beta\left(SOC - SOC_{ref}\right)^2 \quad (24)$$

where β represents the weight of the power retention of the supercapacitor, which is similar to the weight factor of the equivalent fuel consumption. The SOC_{ref} is the capacity expected to be maintained by the supercapacitor (it was set to 0.6). The SOC reference level of 0.6 was chosen to optimally balance energy availability for abrupt load fluctuations, component lifespan preservation (avoiding overcharging > 0.8 or deep discharge < 0.3 per manufacturer guidelines), and system efficiency gain. In engineering practice, to reduce the fuel consumption of a ship during navigation, the previously presented sum is expected to be as small as possible. In addition, the supercapacitor is expected to work within its reasonable operating range. Because the aim of reinforcement learning is to obtain the highest possible reward, the tanh function was introduced to process the sum mathematically. The processed reward function is expressed as follows:

$$R = \begin{cases} -1, \ SOC \leq 0 \ or \ SOC \geq 1 \\ tanh\left(\frac{1}{r}\right), 0 < SOC < 1 \end{cases} \quad (25)$$

The reward function expects the charge state of the supercapacitor to remain within the allowed operation interval $0 < SOC < 1$. If the supercapacitor charge state exceeds this interval, the agent gains a negative reward. Otherwise, the reward is positive and related to the sum of the specific fuel consumption and supercapacitor charge consumption.

In the neural network structure, the actor network comprises an input layer, a hidden layer, and an output layer. The input layer consists of three nodes: the ship's demanded power P_{req}, the ship's operating time t, and the SOC of the supercapacitor. There are 30 hidden layer nodes. The hidden layer uses the Relu activation function, and the output layer outputs the action $\mu(s|\theta^{\mu})$. The input layer has the three states Preq, t, and SOC and action α. The number of neural network nodes in the hidden layer is set to 30, the neural network nodes use the Relu activation function, and the output layer outputs the Q values of the state–action pairs. The other hyperparameters of the algorithm are shown in Table 2.

Table 2. The hyperparameter setting of the DDPG algorithm.

Hyperparameter	Value
Episode	2000
Experience pool capacity	500
The number of samples of the minimum sample set	32
Episode of replication network parameters	300
Learning rate α	0.01
Discount rate γ	0.9
Weight factor β	0.1
ε in ε-greedy	0.9
The reduction of ε	0.05
Approximation coefficient τ	0.01

The framework of the hybrid ship energy management strategy based on the DDPG algorithm is shown in Figure 8. The DDPG algorithm consists of four neural networks. Each neural network is trained by the collected power demand data to force the algorithm

to converge. The trained neural network parameters are pasted into the ship energy management system, which directly outputs continuous actions according to the current power demand of the ship and SOC of the supercapacitor. Simultaneously, the real-time working condition data obtained by the algorithm are continuously recorded and stored to update the training set data and improve the adaptability of the algorithm.

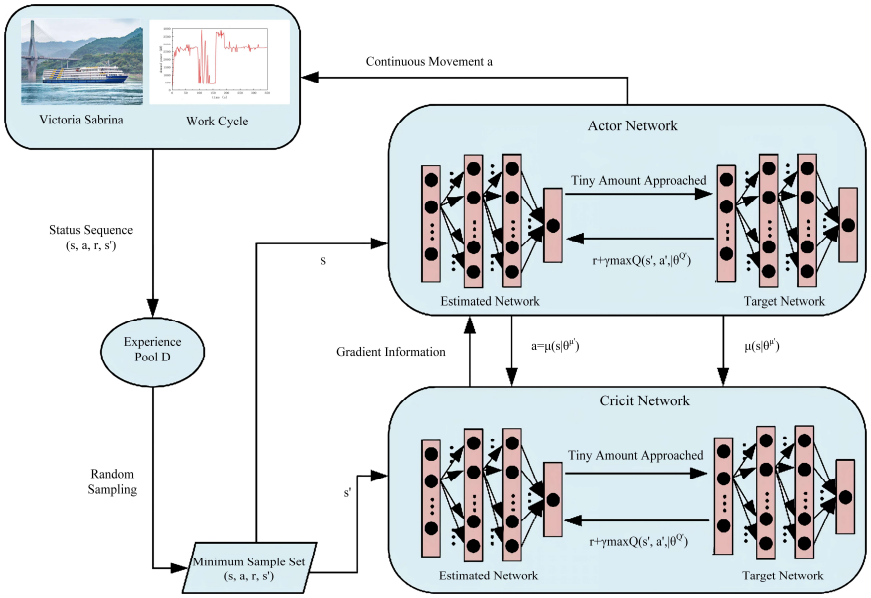

Figure 8. Energy management strategy framework on the DDPG algorithm.

3.3. Comparative Energy Management Strategy

To verify the advancement and reliability of the DDPG algorithm, its total fuel consumption was compared with that of the energy management strategy based on the DQN algorithm. The DQN algorithm is a value-based reinforcement learning algorithm based on the classical Q-learning algorithm and a neural network. If a DQN algorithm is used to manage the energy consumption of a ship, the action space must be discretized within its allowable working range. However, because the output power P_e of the prime mover of the four action diesel generators is a continuous variable, the DQN algorithm cannot traverse all action spaces.

The deep neural network in the DQN algorithm is a supervised learning model. Thus, it is generally required that the front and back samples of the input neural network are independent and identically distributed. However, the front and back states of the reinforcement learning problem have evident probability relations. To eliminate the correlation among the input states, experience replay and parameter freezing are added to the DQN algorithm [31]. The resulting working mechanism is shown in Figure 9.

The experience replay mechanism stores the data pair samples (s, a, r, s'), generated via agent interaction with the environment in the experience pool D. The samples from the experience pool D are used as the input of the current Q estimation network via uniform random sampling when the intelligent agent is trained, and the network weights θ are updated by the gradient descent algorithm. The parameter freezing mechanism, whereby the parameters of the Q estimation network are pasted into the target Q network every C round, reduces the correlation between the estimated and target Q values and improves the algorithm stability to a certain extent.

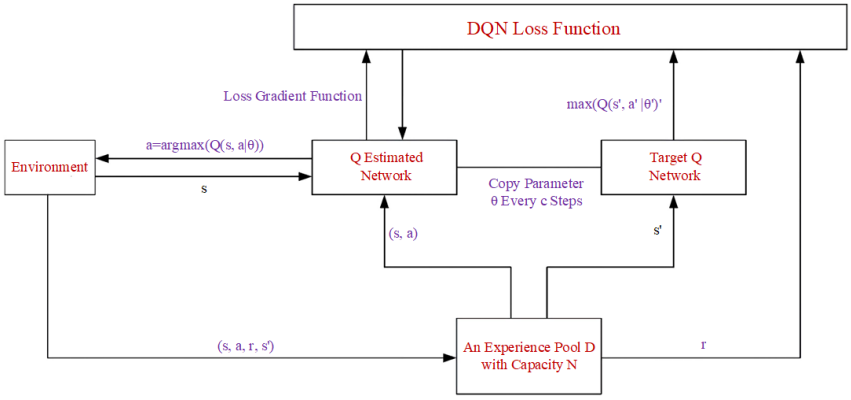

Figure 9. Working process of DQN algorithm.

4. Simulation Results and Analysis

4.1. Analysis of DDPG Algorithm

The code for the energy management strategy based on the DDPG algorithm was written in Python 3.8 and run to perform simulations. To ensure that the algorithm converges to manage the energy consumption, the convergence and stabilizing effect of the algorithm were verified. By accumulating the rewards of each round and then calculating the average reward according to the period of the whole navigation condition, the average reward curve of the DDPG algorithm can be obtained (Figure 10).

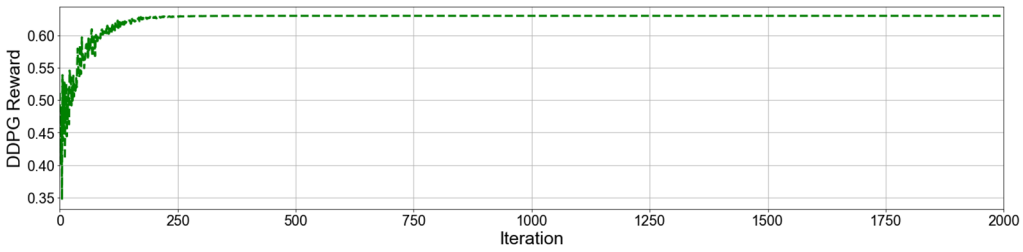

Figure 10. Average reward of the DDPG algorithm.

As shown in Figure 10, the average reward of the DDPG algorithm continues to increase as the algorithm continues to learn. While the algorithm iterates until approximately the 250th round, the average reward of a single round becomes stable, and the algorithm converges. Under the energy management strategy based on the DDPG algorithm, the SOC curve and the output power curve of the supercapacitor for the whole navigation condition are shown in Figure 11 and Figure 12, respectively.

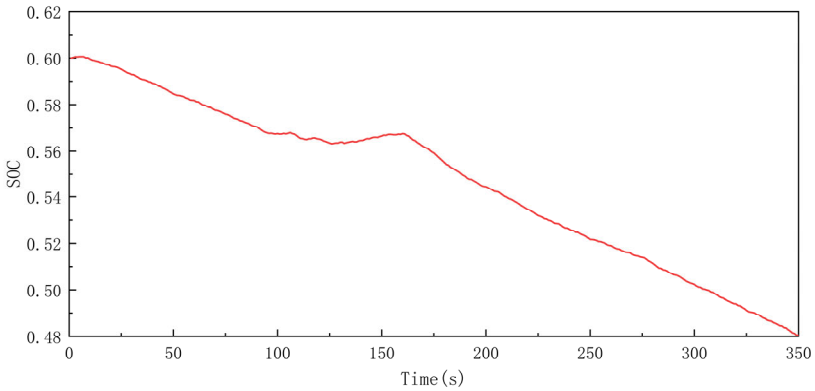

Figure 11. SOC of supercapacitor-based DDPG algorithm.

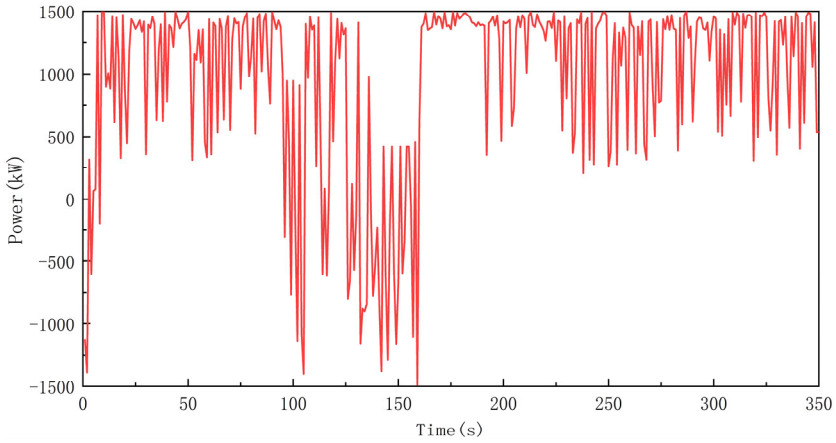

Figure 12. Output power of supercapacitor based on the DDPG algorithm.

As shown in Figures 11 and 12, the SOC of the supercapacitor fluctuates between 0.6 and 0.48, and the output power of the supercapacitor fluctuates frequently around 0. The supercapacitor is charged and discharged. Thereby, it functions as a high-power density energy storage device that "peak-cutting and valley-filling". During the first 10 s of navigation, the ship's power demand is maintained at a low level. When the ship leaves the port, its power demand increases rapidly. The trend of the output power of the supercapacitor is consistent with the trend of the ship's power demand (from early charging to discharging). Moreover, the SOC of the supercapacitor increases slightly in the first 5 s, thereby preparing for the subsequent discharge output power. Between 10 and 100 s, the sudden increase in the ship's power demand ends. Nevertheless, the power demand remains high for a long time, which increases the load of the diesel generator. Thus, the supercapacitor is discharged and shares the ship's power demand with the diesel generator. Consequently, the SOC decreases. Between 100 and 200 s, the ship's power demand fluctuates sharply again, and the supercapacitor is still being discharged. Thus, the DC bus voltage can be kept stable even when the ship's power demand fluctuates sharply. Between the 200th second until the end of the operation, the ship's power demand gradually converges to a high power level. At this point, the SOC of the supercapacitor is still sufficient and much higher than the preset lower limit (0.2). The DDPG algorithm takes the minimum fuel consumption as a reward such that the supercapacitor can be discharged based on the algorithm to minimize fuel consumption globally. The supercapacitor will cooperate with the diesel generator to bear the power demand of the ship. Consequently, the diesel generator set can work in the working state with the low fuel consumption rate. The DDPG algorithm dynamically optimizes power-sharing between the supercapacitor and diesel generators during transient conditions. For sudden load spikes (Figure 12 at 160 s), the supercapacitor provides 60–80% of the transient power within 10–50 ms, while the diesel generators adjust smoothly to supply the remaining demand. When a 500 kW load is abruptly disconnected (simulated at 180 s in Figure 12), the supercapacitor absorbs 400 kW of surplus power within 20 ms, limiting DC bus voltage spikes to +3%. This prioritization minimizes fuel consumption and maintains DC bus voltage stability.

Figure 13 shows the output power curves of the four diesel generators under whole navigation conditions. The output power of the diesel generators fluctuates frequently owing to the frequent charging and discharging of the supercapacitors under sailing conditions and the large power fluctuations. The entire ship's energy demand is provided by the diesel generator set. Thus, it must balance the charging and discharging power of the supercapacitor. However, the calculated reward function is "low" throughout the dynamic process, which effectively reduces the fuel consumption of the diesel generator

set. Quantitatively, the standard deviation of generator power under DDPG is 32% lower than that of the DQN-based strategy, indicating smoother load transitions. This stability stems from the DDPG algorithm's continuous action space, which enables fine-grained power adjustments compared to DQN's discrete steps. Consequently, the supercapacitor's SOC remains within the optimal range (0.48–0.6), avoiding abrupt charging/discharging events that degrade fuel efficiency. Figure 14 further quantifies the load distribution among generators. Generator 1 bears 28% of the total load on average, while Generators 2–4 contribute 24%, 23%, and 25%, respectively. This balanced allocation contrasts sharply with the DQN strategy, where Generator 1 disproportionately handles 34% of the load due to discrete action limitations. The DDPG algorithm's ability to coordinate multiple generators reduces individual component stress, prolonging operational lifespan.

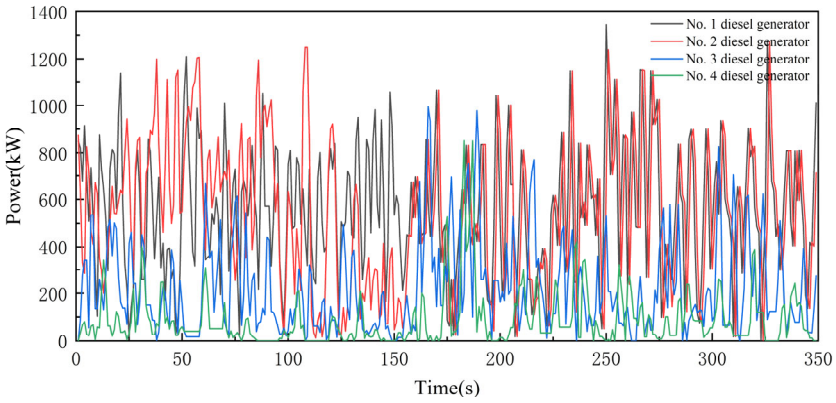

Figure 13. Output power of diesel generator based on DDPG algorithm.

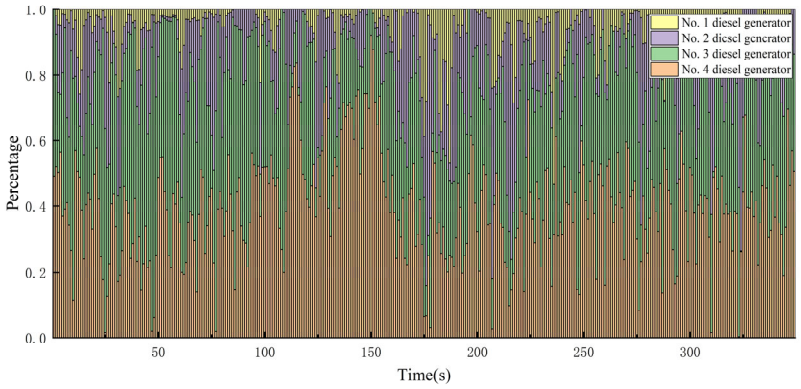

Figure 14. Percentage of output power for each diesel generator based on DDPG algorithm.

This algorithm also illustrates that a high power demand does not mean that the supercapacitor must be discharged and that a low power demand allows the diesel generator to charge the supercapacitor. This is because the DDPG algorithm automatically learns the optimum output power for the diesel generator at different power loads and allows the diesel generator to operate in this state as long as possible.

To investigate further the operating state of the prime mover of the hybrid diesel generator set under the energy management strategy based on the DDPG algorithm, the operating points of the prime mover of the previously presented simulation experiments were plotted (Figure 15). The operating points of the diesel engine are mostly distributed in the upper region of the cloud [in the range of 203 to 213 g/(kW-h) specific fuel consumption]. Some operating points are distributed in the high specific fuel consumption range [218 to 242 g/(kW-h)]. According to Equations (1) and (4), the average specific fuel consumption is 212.65 g/(kW-h), and the total consumed fuel amount is 48.54 kg for the entire journey.

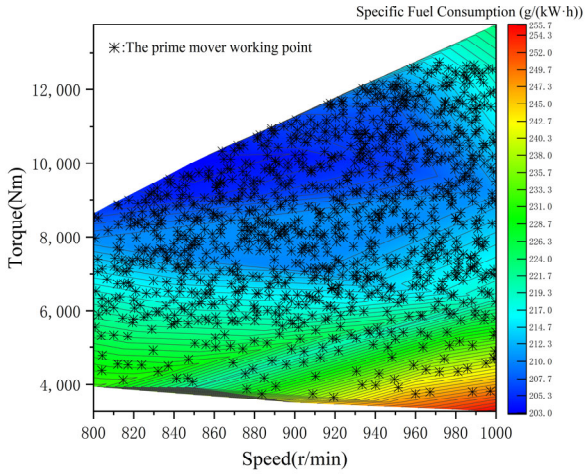

Figure 15. The prime mover working point based on the DDPG algorithm.

4.2. Comparative Analysis

The code for the energy management strategy based on the DQN algorithm was written in Python 3.8 language and run to perform simulations. To ensure that the algorithm converges to manage the energy consumption, the convergence and stabilizing effect of the algorithm were verified. By accumulating the rewards of each round and then calculating the average reward according to the period of the whole navigation condition, the average reward curve of the DQN algorithm was obtained (Figure 16).

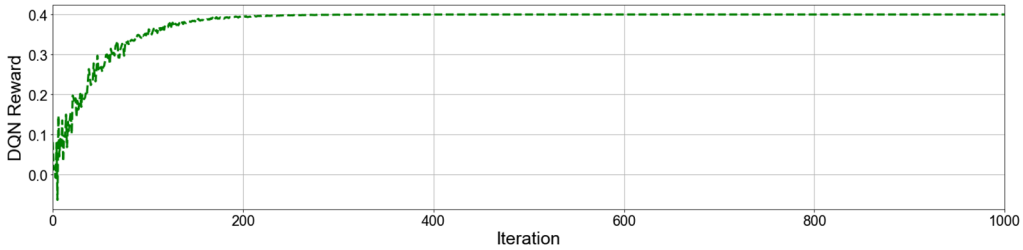

Figure 16. Average reward of DQN algorithm.

Figure 16 shows that the average reward of the DQN algorithm keeps increasing. The average reward of a single round becomes stable, and the algorithm converges at approximately the 200th round of iteration. By contrast, as shown in Figure 10, the DDPG algorithm iterates until approximately the 250th round. This is when the average reward for a single round becomes stable, and the algorithm converges. The comparison highlights key differences in computational efficiency and control performance. While DQN converges faster (9.8 h vs. DDPG's 12.3 h), DDPG achieves higher fuel savings and stability. This trade-off arises because DDPG's actor–critic framework requires simultaneous optimization of policy and value functions, whereas DQN's single-network structure simplifies training but limits adaptability. Compared to the DQN algorithm, although the DDQP algorithm converges more slowly, the average reward of the energy management strategy based on the DDPG algorithm is larger than that of the DQN algorithm. This is because the action space in the DDPG algorithm is continuous, and the intelligent body can traverse all the actions within the allowed action space during the selection of actions to interact with the environment. This enables it to perform actions with higher rewards, which increases the average reward.

The SOC curve for the entire trip with the energy management strategy based on the DQN algorithm is shown in Figure 17.

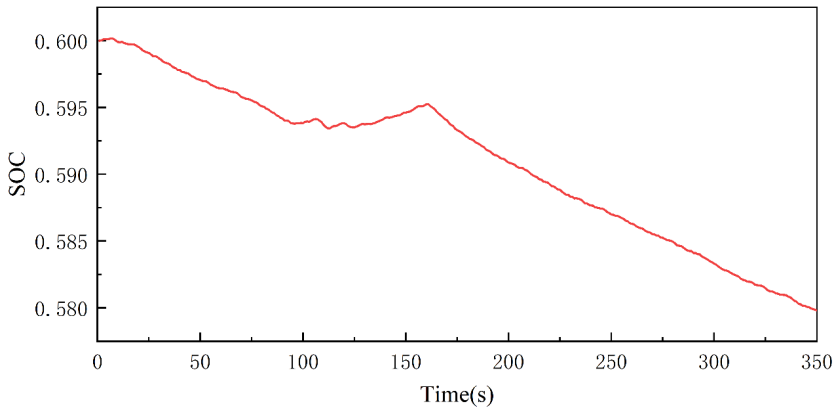

Figure 17. SOC of supercapacitor-based DQN algorithm.

Accordingly, the SOC of the supercapacitor fluctuates between 0.6 and 0.57. The power demand of the ship is low during the first 5 s when the SOC of the supercapacitor slightly increases in preparation for the subsequent discharge of output power. After the 10th s, the ship's power demand increases, the supercapacitor is discharged and shares the ship's power demand with the diesel generator. Thus, the SOC of the supercapacitor decreases. The ship's power demand drops suddenly at 100 s when the SOC of the supercapacitor increases and the supercapacitor is being charged, thereby absorbing the surplus power from the diesel generator. After 150 s, as the SOC of the supercapacitor is still at a stable intermediate value, the difference to the desired SOC value is very small. Therefore, the supercapacitor is still being discharged to keep the diesel power generation at a lower fuel consumption rate. Figure 18 presents the output power curve of the supercapacitor. Throughout the trip, the supercapacitor's output power fluctuates frequently around 0 kW, thereby performing charging and discharging operations to provide power deficits caused by sudden increases in the power demand or to absorb power overflows caused by sudden decreases in the power demand.

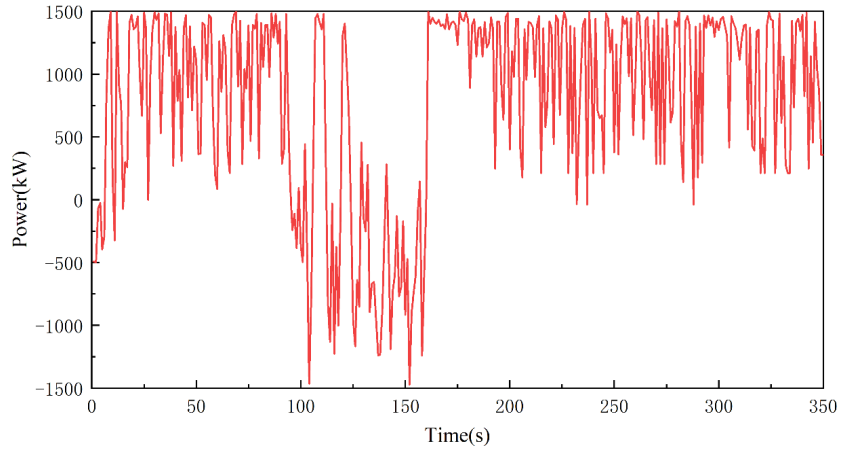

Figure 18. Output power of supercapacitor based on DQN algorithm.

As shown in Figure 19, the DQN-based strategy exhibits significant output power fluctuations, resulting in an increased frequency of generator startup/shutdown. This is because the action space in the DDPG algorithm is continuous, and the agent can traverse all actions within the allowed action space when selecting actions to interact with the environment. Therefore, it eliminates some of the effects of fluctuations. This inefficiency stems from the discrete action space of DQN, which forces the agent to choose suboptimal power levels. For example, during the low-demand period (100–150 s), the tight control of DQN

causes the supercapacitors to be underutilized, resulting in a 12% higher fuel consumption during these intervals than DDPG. Figure 20 illustrates the percentage distribution of total output power among the four diesel generators under the DQN-based energy management strategy during the ship's voyage. Unlike the balanced allocation observed in the DDPG strategy (Figure 14), the DQN algorithm exhibits significant load imbalance, primarily due to its discrete action space. Key observations include the following: Generator 1 handles 34% of the total load on average, significantly higher than Generators 2–4 (22%, 21%, and 23%, respectively). This imbalance stems from DQN's discrete action choices, which limit fine-grained adjustments and force over-reliance on specific generators during abrupt load changes.

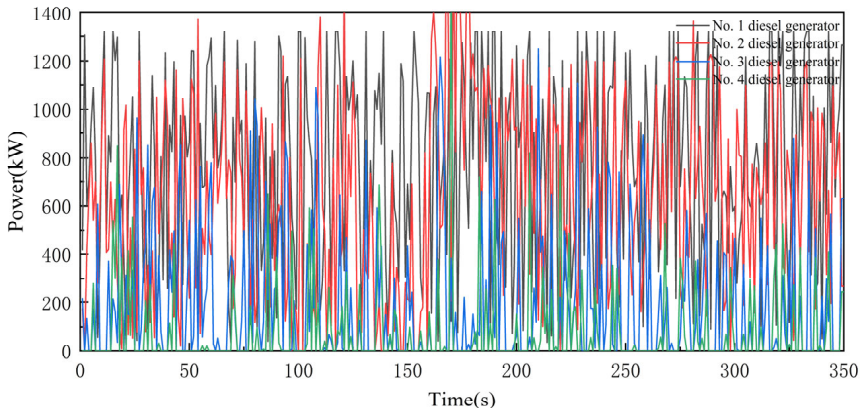

Figure 19. Output power of diesel generator based on DQN algorithm.

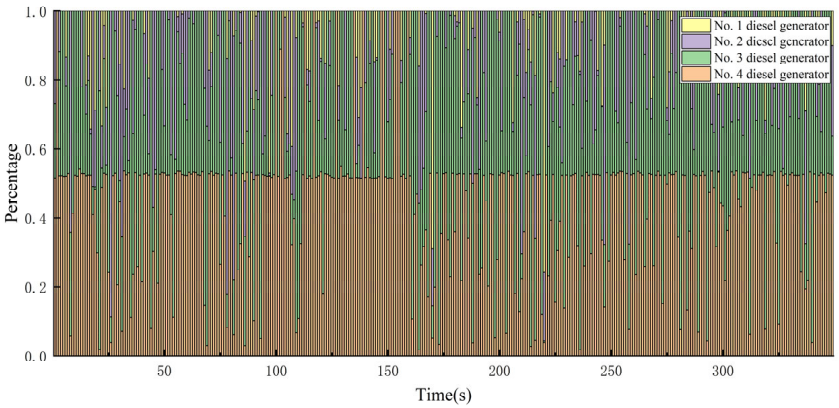

Figure 20. Percentage of output power for each diesel generator on DQN algorithm.

To further investigate the operating state of the prime mover of the hybrid marine diesel generator set under the energy management strategy based on the DQN algorithm, the operating points of the prime mover from the previously presented simulation experiments were plotted (Figure 21). The operating points of the diesel engine are mostly distributed in the upper region of the cloud [in the 203 to 211 g/(kW-h) specific fuel consumption range]. Some operating points are distributed in the high specific fuel consumption range [225 to 248 g/(kW-h)]. According to Equations (1) and (4), the average specific fuel consumption is 219.37 g/(kW-h), and the total consumed fuel amount is 50.35 kg for the entire trip.

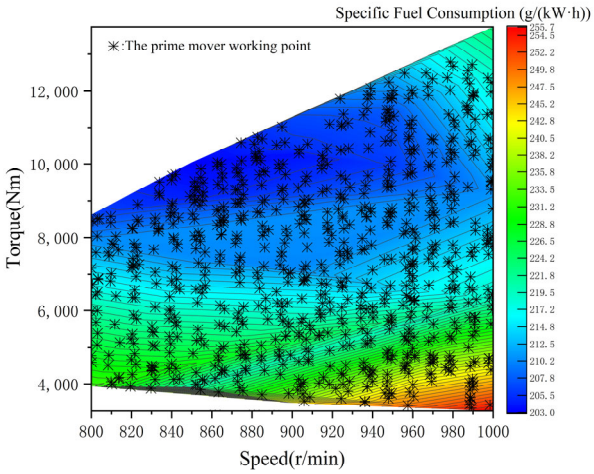

Figure 21. The prime mover working point based on DQN algorithm.

The comparison reveals that the action space based on the DQN energy management strategy is discrete. Hence, the respective operating points of the prime mover are distributed in clusters in the discrete space. Those based on the DDPG algorithm are more dispersed. Most points are concentrated in the range of 203–213 g/(kW-h), and a few are located in the range of 218–231 g/(kW-h). This is because the action output under the DDPG algorithm is continuous, and there is a small probability that the agent will output random actions for the reinforcement learning algorithm. Consequently, the resulting action points are spread over its allowed working interval.

In addition to the distribution of the operating points, the average specific fuel consumption and total fuel consumption during the journey of the ship under the three strategies are compared in Table 3. The rule-based energy management strategy performs worst. It results in the highest average specific fuel consumption and total fuel consumption. However, it is still more fuel-efficient than the real ship data, with a reduction in the total fuel consumption of approximately 5.5%. The energy management strategy based on the DDPG algorithm results in the lowest average specific fuel consumption and lowest total fuel consumption. The strategy consumes 15.8% less fuel in total compared to the real ship. This is followed by the energy management strategy based on the DQN algorithm, which reduces the total fuel consumption by 12.6% compared to that of the real ship.

Table 3. Fuel consumption data under different strategies.

Data Strategy	Average Specific Fuel Consumption/g/(kW-h)	Total Fuel Consumption/kg	Relative Reduction/%
Real ship	230.46	57.63	0
Ruled-based	227.48	54.44	5.5%
DQN-based	219.37	50.35	12.6%
DDPG-based	212.65	48.54	15.8%

The proposed energy management system delivers practical benefits across maritime applications. For commercial vessels, the framework significantly reduces fuel consumption, directly lowering operational costs while minimizing environmental impacts. In passenger ships, the adaptive control mechanism enhances stability during dynamic load transitions, improving onboard comfort through optimized power distribution. Furthermore, the system's modular architecture allows for seamless integration of supplementary energy sources without necessitating major algorithmic redesigns. This scalability en-

ables retrofitting existing cargo ships and ferries with hybrid power systems, offering a transitional strategy to support global decarbonization efforts in the shipping industry.

5. Conclusions

In this study, the fuel consumption of the Victoria Sabrina diesel–electric hybrid ship was the study object. In accordance with the characteristics of a hybrid power system, reinforcement learning was applied to solve the energy management problem of hybrid ships. An energy management strategy based on the DDPG algorithm was established and compared with an energy management strategy based on the DQN algorithm. The conclusions are as follows:

1. Both DDPG and DQN algorithms converge stably. Although the convergence speed of the DDPG algorithm is slightly lower than that of the DQN algorithm, the energy management strategy based on the DDPG algorithm obtains higher rewards. This indicates that the DDPG algorithm is more suitable for the control of continuous systems. The output power fluctuations of the diesel generator under the DQN algorithm are more pronounced.
2. In terms of fuel economy, the rule-based energy management strategy performed worst. The average specific fuel consumption and total fuel consumption were the highest. However, the strategy was still more fuel-efficient than the real ship data, with a reduction in the total fuel consumption of approximately 5.5% compared to that of the real ship. The energy management strategy based on the DDPG algorithm had the lowest average specific fuel consumption and total fuel consumption, with a 15.8% reduction in the total fuel consumption compared to that of the real ship. This is followed by the energy management strategy based on the DQN algorithm, which reduced the total fuel consumption by 12.6% compared to that of the real ship. This result is consistent with the characteristics of the different energy management strategies.

The proposed strategy demonstrates broad applicability across diverse maritime applications due to its inherent compatibility with hybrid power configurations. The framework supports seamless integration of additional energy sources (e.g., batteries and fuel cells) and multiple diesel generators, making it adaptable to cargo ships, ferries, and offshore support vessels. Its capability to manage abrupt load fluctuations—a critical requirement for ships with highly variable power demands such as icebreakers or tugboats—stems from the algorithm's dynamic load adaptability. Furthermore, the methodology ensures scalability through parameter adjustments; ship-specific parameters (e.g., generator capacities, propulsion power ratings) can be customized within the state vector without modifying the core algorithmic architecture, enabling tailored implementations across vessel types while maintaining operational consistency and fuel efficiency.

The proposed strategy's real-world validation will advance through three phases: Hardware-in-the-Loop (HIL) testing using RT-Lab simulators to validate multi-energy system performance under extreme conditions, followed by a 6-month onboard trial on a Yangtze River cargo vessel targeting $\geq$3% fuel savings compared to conventional methods. Finally, collaboration with the China Classification Society (CCS) and IMO will establish certification protocols for AI-driven energy management systems, addressing safety, cybersecurity, and interoperability to enable global maritime adoption.

In summary, as the reinforcement learning-based energy management strategy has a certain self-learning ability, it better balances the optimization effect and the applicability of the strategy in practice. It is also not commonly used in the marine sector. Thus, it has greater research value and prospect for engineering applications. The research results

in this paper can provide technical support for the stable, safe, and efficient operation of diesel–electric hybrid ship power systems.

Author Contributions: Conceptualization, Y.Y. and X.G.; methodology, D.T.; software, X.G.; validation, Y.Y. and C.Y.; formal analysis, D.T. and B.S.; investigation, X.G.; resources, Y.Y.; data curation, X.G. and D.T.; writing—original draft preparation, X.G.; writing—review and editing, C.Y. and J.M.G.; visualization, B.S.; supervision, Y.Y. and C.Y.; project administration, Y.Y. and X.G.; funding acquisition, Y.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This work is Supported by the National Natural Science Foundation of China (Grants U23A20680), Jiangsu Province Carbon Peaking and Carbon Neutrality Science, Technology Innovation Special Fund (Industry Foresight and Key Technology Core Research) Project (BE2023091-2), and Provincial Science and Technology Innovation Strategy Special Project of Shaoguan City (230317166277914).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors would like to thank the anonymous reviewers for their constructive comments, which helped us considerably improve the quality of the paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Uyan, E.; Atlar, M.; Gürsoy, O. Energy Use and Carbon Footprint Assessment in Retrofitting a Novel Energy Saving Device to a Ship. *J. Mar. Sci. Eng.* **2024**, *12*, 1879. [CrossRef]
2. Du, Z.; Chen, Q.; Guan, C.; Chen, H. Improvement and Optimization Configuration of Inland Ship Power and Propulsion System. *J. Mar. Sci. Eng.* **2023**, *11*, 135. [CrossRef]
3. Tang, D.; Tang, H.; Yuan, C.; Dong, M.; Diaz-Londono, C.; Tinaiero, G.D.A.; Guerrero, J.M.; Zio, E. Economic and resilience-oriented operation of coupled hydrogen-electricity energy systems at ports. *Appl. Energy* **2025**, *390C*, 125825. [CrossRef]
4. Guo, X.; Lang, X.; Yuan, Y.; Tong, L.; Shen, B.; Long, T.; Mao, W. Energy Management System for Hybrid Ship: Status and Perspectives. *Ocean Eng.* **2024**, *310*, 118638. [CrossRef]
5. Gao, Y.; Tan, Y.; Jiang, D.; Sang, P.; Zhang, Y.; Zhang, J. An Adaptive Prediction Framework of Ship Fuel Consumption for Dynamic Maritime Energy Management. *J. Mar. Sci. Eng.* **2025**, *13*, 409. [CrossRef]
6. Rindone, C. AIS Data for Building a Transport Maritime Network: A Pilot Study in the Strait of Messina (Italy). In *Computational Science and Its Applications—ICCSA 2024 Workshops*; Springer Nature: Cham, Switzerland, 2024; pp. 213–226. [CrossRef]
7. Shen, D.; Lim, C.-C.; Shi, P. Fuzzy Model Based Control for Energy Management and Optimization in Fuel Cell Vehicles. *IEEE Trans. Veh. Technol.* **2020**, *69*, 14674–14688. [CrossRef]
8. Seng, U.K.; Malik, H.; García Márquez, F.P.; Alotaibi, M.A.; Afthanorhan, A. Fuzzy Logic-Based Intelligent Energy Management Framework for Hybrid PV-Wind-Battery System: A Case Study of Commercial Building in Malaysia. *J. Energy Storage* **2024**, *102*, 114109. [CrossRef]
9. Nivolianiti, E.; Karnavas, Y.L.; Charpentier, J.-F. Energy Management of Shipboard Microgrids Integrating Energy Storage Systems: A Review. *Renew. Sustain. Energy Rev.* **2024**, *189*, 114012. [CrossRef]
10. Li, L.; Yang, C.; Zhang, Y.; Zhang, L.; Song, J. Correctional DP-Based Energy Management Strategy of Plug-In Hybrid Electric Bus for City-Bus Route. *IEEE Trans. Veh. Technol.* **2015**, *64*, 2792–2803. [CrossRef]
11. Upadhyaya, A.; Mahanta, C. Optimal Online Energy Management System for Battery-Supercapacitor Electric Vehicles Using Velocity Prediction and Pontryagin's Minimum Principle. *IEEE Trans. Veh. Technol.* **2025**, *74*, 2652–2666. [CrossRef]
12. Kanellos, F.D. Optimal Power Management With GHG Emissions Limitation in All-Electric Ship Power Systems Comprising Energy Storage Systems. *IEEE Trans. Power Syst.* **2014**, *29*, 330–339. [CrossRef]
13. Zhang, H.; Qin, Y.; Li, X.; Liu, X.; Yan, J. Power Management Optimization in Plug-in Hybrid Electric Vehicles Subject to Uncertain Driving Cycles. *eTransportation* **2020**, *3*, 100029. [CrossRef]
14. Ou, K.; Yuan, W.-W.; Choi, M.; Yang, S.; Jung, S.; Kim, Y.-B. Optimized Power Management Based on Adaptive-PMP Algorithm for a Stationary PEM Fuel Cell/Battery Hybrid System. *Int. J. Hydrogen Energy* **2018**, *43*, 15433–15444. [CrossRef]
15. Kalikatzarakis, M.; Geertsma, R.D.; Boonen, E.J.; Visser, K.; Negenborn, R.R. Ship Energy Management for Hybrid Propulsion and Power Supply with Shore Charging. *Control Eng. Pract.* **2018**, *76*, 133–154. [CrossRef]

16. Zhang, Y.; Xue, Q.; Gao, D.; Shi, W.; Yu, W. Two-Level Model Predictive Control Energy Management Strategy for Hybrid Power Ships with Hybrid Energy Storage System. *J. Energy Storage* **2022**, *52*, 104763. [CrossRef]
17. Kofler, S.; Du, Z.P.; Jakubek, S.; Hametner, C. Predictive Energy Management Strategy for Fuel Cell Vehicles Combining Long-Term and Short-Term Forecasts. *IEEE Trans. Veh. Technol.* **2024**, *73*, 16364–16374. [CrossRef]
18. 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]
19. Jung, W.; Chang, D. Deep Reinforcement Learning-Based Energy Management for Liquid Hydrogen-Fueled Hybrid Electric Ship Propulsion System. *J. Mar. Sci. Eng.* **2023**, *11*, 2007. [CrossRef]
20. Huotari, J.; Ritari, A.; Ojala, R.; Vepsäläinen, J.; Tammi, K. Q-Learning Based Autonomous Control of the Auxiliary Power Network of a Ship. *IEEE Access* **2019**, *7*, 152879–152890. [CrossRef]
21. Xiao, H.; Fu, L.; Shang, C.; Bao, X.; Xu, X.; Guo, W. Ship Energy Scheduling with DQN-CE Algorithm Combining Bi-Directional LSTM and Attention Mechanism. *Appl. Energy* **2023**, *347*, 121378. [CrossRef]
22. Fu, J.; Sun, D.; Peyghami, S.; Blaabjerg, F. A Novel Reinforcement-Learning-Based Compensation Strategy for DMPC-Based Day-Ahead Energy Management of Shipboard Power Systems. *IEEE Trans. Smart Grid* **2024**, *15*, 4349–4363. [CrossRef]
23. Chen, W.; Tai, K.; Lau, M.W.S.; Abdelhakim, A.; Chan, R.R.; Kåre Ådnanes, A.; Tjahjowidodo, T. Optimal Power and Energy Management Control for Hybrid Fuel Cell-Fed Shipboard DC Microgrid. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 14133–14150. [CrossRef]
24. Rudolf, T.; Schürmann, T.; Schwab, S.; Hohmann, S. Toward Holistic Energy Management Strategies for Fuel Cell Hybrid Electric Vehicles in Heavy-Duty Applications. *Proc. IEEE* **2021**, *109*, 1094–1114. [CrossRef]
25. Hu, X.; Liu, T.; Qi, X.; Barth, M. Reinforcement Learning for Hybrid and Plug-In Hybrid Electric Vehicle Energy Management: Recent Advances and Prospects. *IEEE Ind. Electron. Mag.* **2019**, *13*, 16–25. [CrossRef]
26. Li, Y.; Tang, D.; Yuan, C.; Diaz-Londono, C.; Agundis-Tinajero, G.D.; Guerrero, J.M. The Roles of Hydrogen Energy in Ports: Comparative Life-Cycle Analysis Based on Hydrogen Utilization Strategies. *Int. J. Hydrogen Energy* **2025**, *106*, 1356–1372. [CrossRef]
27. Guo, X.; Yuan, Y.; Tong, L. Research on Online Identification Method of Ship Condition Based on Improved DBN Algorithm. In Proceedings of the 2023 7th International Conference on Transportation Information and Safety (ICTIS), Xi'an, China, 4–6 August 2023; pp. 349–354. [CrossRef]
28. Rekioua, D.; Mokrani, Z.; Kakouche, K.; Rekioua, T.; Oubelaid, A.; Logerais, P.O.; Ali, E.; Bajaj, M.; Berhanu, M.; Ghoneim, S.S.M. Optimization and Intelligent Power Management Control for an Autonomous Hybrid Wind Turbine Photovoltaic Diesel Generator with Batteries. *Sci. Rep.* **2023**, *13*, 21830. [CrossRef]
29. Zhou, C.; Wang, Y.; Wang, L.; He, H. Obstacle Avoidance Strategy for an Autonomous Surface Vessel Based on Modified Deep Deterministic Policy Gradient. *Ocean Eng.* **2022**, *243*, 110166. [CrossRef]
30. Wu, Y.; Tan, H.; Peng, J.; Zhang, H.; He, H. Deep Reinforcement Learning of Energy Management with Continuous Control Strategy and Traffic Information for a Series-Parallel Plug-in Hybrid Electric Bus. *Appl. Energy* **2019**, *247*, 454–466. [CrossRef]
31. Du, G.; Zou, Y.; Zhang, X.; Liu, T.; Wu, J.; He, D. Deep Reinforcement Learning Based Energy Management for a Hybrid Electric Vehicle. *Energy* **2020**, *201*, 117591. [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.