

Perspective

Deep Reinforcement Learning-Based Energy and Power Management for Ships: A Perspective Review of Methods and Applications

Yujeong Kang, Dita Puspita and Il-Yop Chung * 

School of Electrical Engineering, Kookmin University, 77 Jeongneung-ro, Seongbuk-gu, Seoul 02707, Republic of Korea; kangyj17@kookmin.ac.kr (Y.K.); ditapuspita@kookmin.ac.kr (D.P.)

* Correspondence: chung@kookmin.ac.kr; Tel.: +82-2-910-4702

Abstract

Energy and power management systems (EMS/PMS) are essential for electric-propulsion ships, affecting propulsion performance, fuel consumption, emissions, and component lifetime. As shipboard power systems integrate heterogeneous energy resources and face nonlinearity, uncertain load demand, and multi-source interactions, deep reinforcement learning (DRL) has emerged as a promising adaptive, sequential decision-making tool in shipboard EMS/PMS. This perspective reviews DRL studies through a hierarchical decision-making framework comprising power dispatch, energy coordination, and operational strategy. Most research focuses on real-time power dispatch, while emerging research addresses energy coordination via multi-source cooperation, multi-objective operation, degradation awareness, and uncertainty handling. However, operational strategy remains underexplored, despite its role in speed control, route-aware planning, predictive operation, and voyage scheduling. This paper argues that future shipboard EMS/PMS adopt integrated hierarchical DRL frameworks across all three decision layers, leveraging DRL's strengths in sequential policy learning in dynamic environments and supporting multi-time-scale decision-making. This paper clarifies current research trends, identifies gaps, and outlines future directions toward adaptive, reliable, and autonomous shipboard EMS/PMS in next-generation electric-propulsion ships.

Keywords: electric propulsion ship; energy-management system; power management system; deep reinforcement learning; hierarchical decision-making



Academic Editor: Panos Kotsampopoulos

Received: 29 May 2026

Revised: 3 July 2026

Accepted: 15 July 2026

Published: 16 July 2026

Copyright: © 2026 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The maritime industry is rapidly transitioning toward low-carbon, high-efficiency propulsion systems due to stringent environmental regulations and the International Maritime Organization (IMO)'s greenhouse-gas reduction strategy and energy-efficiency regulations. This has accelerated the shift from conventional fossil-fuel-based propulsion toward electrified and hybrid onboard power systems. Inal et al. [1] reviewed marine power generation, energy storage, energy and power management systems (EMS/PMS), and hybrid propulsion topologies, confirming hybridization as a key pathway toward low-emission and eventually zero-carbon shipping. Accordingly, next-generation shipboard power systems in electric and hybrid ships increasingly integrate diesel generators (DGs), gas turbine generators (GTGs), fuel cells (FCs), energy storage systems (ESSs), renewable sources, and shore power.

Such systems act as isolated or mobile microgrids with tightly coupled energy sources, propulsion/service loads, ESSs, and power electronic converters. Kumar and Zare [2] reviewed maritime microgrid architectures and emphasized that electric propulsion, integrated power systems, and power electronic converters introduce challenges of efficiency, power quality, reliability, and regulation. EMS/PMS serve as supervisory layers coordinating generation, storage, propulsion, and service loads but face difficulty from time-varying loads, nonlinear efficiency, operational constraints, uncertain voyage conditions, and conflicting objectives. Hou et al. [3] showed propulsion-load fluctuations in all-electric ships include low- and high-frequency components from operating conditions such as propeller in-and-out-of-water effects. Therefore, modern shipboard EMS/PMS problems are high-dimensional, nonlinear, and multi-objective optimization problems.

To address these challenges, conventional approaches include rule-based control, heuristic scheduling, mixed-integer optimization, model predictive control (MPC), dynamic programming (DP), and equivalent consumption minimization strategy (ECMS). These studies provide foundations for shipboard scheduling and energy optimization. Kalikatzarakis et al. [4] applied ECMS to a hybrid ship propulsion and power-supply system with shore charging, while Hein et al. [5] coordinated voyage planning and energy management for an all-electric ship using DP-based route planning and multi-objective energy dispatch. However, as shipboard power systems grow larger and more heterogeneous, rule-based strategies remain limited by predefined logic, whereas optimization-based methods require accurate mathematical models and risk high online computational burdens.

These limitations have motivated artificial intelligence (AI)-based approaches for adaptability under uncertainty and nonlinearity. Liu et al. [6] proposed a prior-knowledge-based two-layer EMS for an FC hybrid ship, using offline optimization to train neural-network- and support-vector-machine-based models for real-time energy distribution. Consequently, deep reinforcement learning (DRL) has gained attention for learning sequential control policies through operating environment interaction. Wu et al. [7] proposed a double deep Q-network (Double DQN)-based EMS for a plug-in hybrid FC/battery coastal ferry, demonstrating DRL's applicability for FC power control, battery management, and real-time energy management under stochastic loads.

However, deploying AI-based EMS/PMS in safety-critical shipboard power systems requires governance, risk management, and secure lifecycle design. Recent maritime cybersecurity, AI and risk management standards, and industrial-control security standards indicate that such systems need trustworthy, risk-managed, and cyber-secure frameworks [8–12]. Therefore, DRL-based shipboard EMS/PMS should be reviewed for algorithmic performance and also for reliable, secure, and operationally acceptable decision-making structures.

Despite these developments, existing AI- and DRL-based shipboard EMS/PMS studies remain fragmented across individual applications, such as generator scheduling, ESS charge/discharge control, FC/battery power allocation, emission-aware operation, and voyage optimization. Existing reviews covered sustainable ship energy systems [13], maritime microgrids, and shipboard power-system architectures. Zohrabi et al. [14] summarized design specifications and requirements for medium-voltage direct-current (MVDC) shipboard power systems. Nevertheless, these studies lack an integrated perspective that interprets DRL studies by their decision-making roles.

Motivated by this gap, this paper reinterprets shipboard EMS/PMS as a hierarchical decision-making structure, categorizing DRL studies into three decision layers: power dispatch, energy coordination, and operational strategy. Based on this classification, this paper analyzes DRL's role at each layer, discusses research imbalances and limitations,

and proposes future directions for advancing shipboard EMS/PMS toward hierarchical, integrated, and intelligent operational frameworks.

This Perspective review was informed by literature searches using combinations of keywords related to “shipboard EMS,” “shipboard PMS,” “ship energy management,” “ship power management,” “electric ship,” “hybrid electric ship,” “electric propulsion,” “reinforcement learning,” and “deep reinforcement learning.” The retrieved studies were screened for relevance to reinforcement-learning-based energy and power management in electric and hybrid-electric shipboard systems. The selected studies were subsequently analyzed in terms of energy-system configuration, decision layer, DRL methodology, reward design, and practical implementation considerations.

The main contributions of this paper are summarized as follows:

- This study defines a hierarchical classification framework for DRL-based shipboard EMS/PMS by reorganizing fragmented optimization studies into three decision layers: power dispatch, energy coordination, and operational strategy. This framework clarifies their decision scope and operational role.
- This paper methodologically analyzes DRL-based shipboard EMS/PMS by focusing on Markov decision process (MDP) formulations, DRL algorithm structures, objective and reward designs, and validation approaches.
- This paper identifies key research gaps and future directions toward integrated hierarchical DRL frameworks that unify power dispatch, energy coordination, and operational strategy as separate problems within a single intelligent decision-making architecture.

The remainder of this paper is organized as follows. Section 2 describes the architecture and key characteristics of EMS/PMS in electric-propulsion ships. Section 3 reviews reinforcement learning (RL)-based studies and their limitations. Section 4 analyzes DRL-based methodologies and hierarchical decision-making structures. Section 5 presents future research directions and practical considerations. Finally, Section 6 concludes the paper.

2. Shipboard Energy and Power Management System

2.1. Shipboard Power-System Configuration

Modern ships are evolving from conventional single-main-engine propulsion to integrated electric and hybrid architectures combining multiple power sources, ESSs, and power electronic converters. In electric-propulsion ships, propulsion and service loads, generators, ESSs, converters, propulsion motors, and service loads interconnect through a unified power network. Wang et al. [15] and Si et al. [16] showed that power-source sizing, configuration optimization, and EMS/PMS design are closely coupled in hybrid ship systems, indicating that power-system configuration strongly shapes shipboard EMS/PMS problems.

Representative onboard power sources include DGs, GTGs, FCs, and ESSs. DGs are widely used as primary or backup sources due to their maturity, reliability, and stable output, while GTGs provide high power density and fast response but raise fuel-consumption and emission concerns. Recent FC/battery and FC/ESS hybrid ship studies reveal that FCs and ESSs improve efficiency, reduce emissions, manage degradation, and support stable operation under dynamic loads [17–19]. Therefore, FCs and ESSs are key controllable resources in shipboard EMS/PMS together with conventional DG and GTG sources.

Advances in power electronics and DC-based energy sources have accelerated the adoption of DC-bus and hybrid AC/DC architectures. FCs, ESSs, and PV systems output DC, so DC-bus architectures reduce conversion stages and improve power-flow flexibility. Aziz et al. [20] modeled a hybrid multi-PEMFC/BESS system using a common DC bus, confirming that DC architectures suit FC, ESS, and DC resource integration. Auxiliary resources such as PV systems, shore power, wind-assisted sources, and waste-heat recov-

ery also reduce fuel consumption and emissions. Vahabzad et al. [21] showed that PV generation, ESSs, and shore power can reduce onboard DG operation during berthing, although these resources remain supplementary due to intermittency and installation or infrastructure constraints.

Accordingly, this paper focuses on a representative hybrid electric-propulsion ship power system with DGs, GTGs, FCs, and ESSs as core onboard resources and PV and shore power as supplementary resources (Figure 1). The heterogeneous characteristics of these resources require shipboard EMS/PMS to determine which sources should operate, when, and at what power time-varying navigation conditions and load demands. This configuration-level complexity provides the physical basis for formulating shipboard EMS/PMS as high-dimensional, nonlinear, and multi-objective optimization problems.

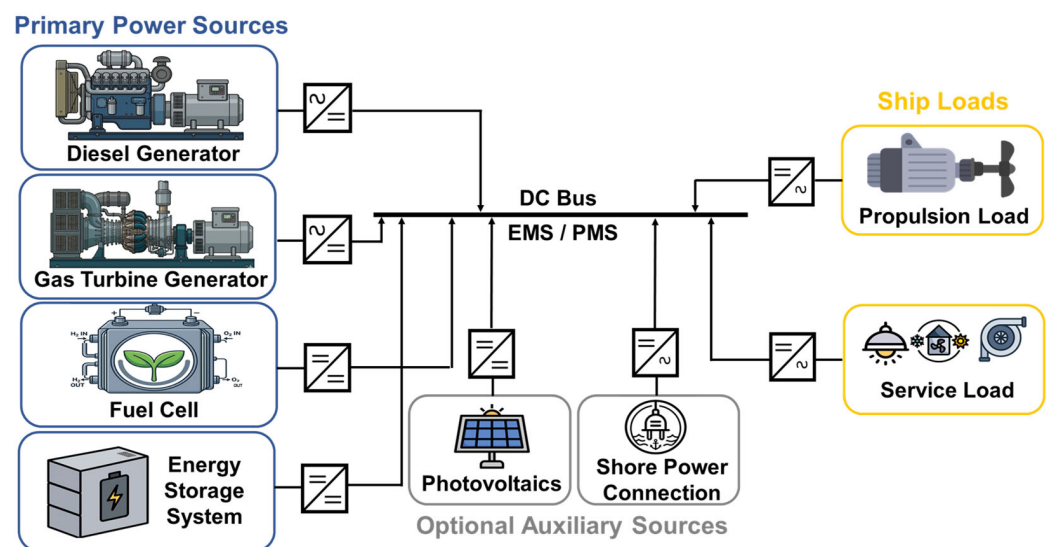


Figure 1. Representative power-system configuration of a hybrid electric-propulsion ship.

2.2. Roles of EMS/PMS

Shipboard EMSs and PMSs are supervisory control layers ensuring stable, efficient, and reliable ship operation. PMS is generally the lower-level layer for real-time power balance, system stability, and power quality, whereas EMS serves as the upper-level layer for fuel economy, energy efficiency, emission reduction, and component-lifetime management. Table 1 summarizes the representative differences between PMS and EMS.

Table 1. Comparison of EMS/PMS in ship power systems.

Category	PMS	EMS
Primary objective	Power balance, system stability	Cost, efficiency, emission optimization
Time scale	Short-term, real-time	Medium-/long-term
Control focus	Power balance, source output control	Energy scheduling, source coordination
Decision type	Reactive, immediate	Strategic, predictive
Typical functions	Generator start/stop, load sharing, voltage/frequency control, protection	Economic dispatch, ESS scheduling, fuel consumption and emission reduction
Decision level	Power dispatch	Energy coordination, operational strategy
Input information	Real-time electrical states	Load forecast, route, operating condition, fuel or emission factors

PMS maintains power balance and operational security under rapid load variations. Propulsion, hotel, and pulse loads may change abruptly depending on sailing speed, sea state, maneuvering, and onboard demand. Accordingly, PMS performs real-time generator start-up/shutdown, load sharing, output adjustment, voltage/frequency regulation, protection, fault handling, and blackout prevention. Xie et al. [22] proposed a real-time two-layer PMS for hybrid electrical ships under fluctuating propulsion loads, showing that PMS should prioritize power tracking, stability, and supply continuity.

In contrast, EMS optimizes medium- to long-term performance by considering fuel consumption, operating cost, emissions, degradation, and storage-state conditions. Its functions include generator scheduling, power allocation, ESS charge/discharge planning, FC utilization, renewable-energy prioritization, and shore-power coordination. Yang et al. [23] developed a two-stage EMS combining mixed-integer nonlinear programming (MINLP) and MPC for a hybrid diesel–electric tugboat, providing reference trajectories for real-time power allocation with battery dynamics. Therefore, EMS supports economic, low-emission, lifetime-aware, and predictive operation.

In practice, PMS and EMS integrate into a hierarchical control structure. EMS determines upper-level objectives and reference commands, while PMS executes real-time actions for power balance, stability, and power quality. PMS feedback (generator conditions, load variations, fault status, and ESS state) informs EMS for higher-level optimization, linking real-time control with long-term energy optimization [24]. This multi-time-scale, multi-objective, and constraint-driven optimization problem motivates interest in DRL-based methods.

2.3. Key Challenges and Characteristics of Shipboard EMS/PMS Problems

Shipboard EMS/PMS problems are complex due to heterogeneous energy sources, operational constraints, uncertain environments, and multiple time scale interactions. Therefore, they are not simple power-allocation tasks, but coupled optimization problems requiring coordination across power sources, storage devices, loads, and operational objectives.

The first challenge is heterogeneous multi-energy integration. Modern electric and hybrid ships may use DGs, GTGs, FCs, ESSs, renewable energy sources, and shore power, each with different efficiency, response speeds, emissions, degradation behaviors, and operating limits. Yuan et al. [25] showed that multi-energy hybrid ship power systems improve economic and environmental performance, but increase system complexity in reliability, control, and efficiency. Therefore, EMS/PMS must coordinate heterogeneous resources, not optimize each source independently.

The second challenge is multi-time-scale decision-making. Power dispatch and power-quality control need near-real-time responses, whereas generator scheduling, ESS charge/discharge planning, and FC utilization are optimized over longer horizons. Voyage planning, speed scheduling, routing, and port-operation decisions operate at voyage scale. Xie et al. [26] distinguished PMS (instantaneous power flow) from EMS (longer-term energy utilization and planning). Therefore, shipboard EMS/PMS cannot be captured by a single-time-scale optimization.

The third challenge is uncertainty in loads, operating conditions, and external environments. Propulsion/hotel loads vary with sailing speed, maneuvering conditions, sea state, and onboard demand, while renewable generation and shore-side operation depend on weather and port infrastructure. Pang et al. [27] found propulsion-load prediction difficult due to highly dynamic and environment-dependent ship operating data. Park et al. [28] showed berth dwell-time uncertainty and time-of-arrival constraints affect fuel-efficient routing. Therefore, uncertainty impacts both operational and strategic decision-making.

The fourth challenge is the presence of conflicting objectives. Fuel consumption, operating cost, emissions, component degradation, system reliability, power quality, and voyage schedule adherence must be considered simultaneously. Tsoumpris and Theotokatos [29] demonstrated that reliability improvements can increase fuel consumption. Consequently, shipboard EMS/PMS is a multi-objective optimization problem involving economic, environmental, reliability, and safety-related criteria.

Figure 2 summarizes how multi-energy integration, system uncertainty, and operational constraints lead to hierarchical decision layers (power dispatch, energy coordination, and operational strategy), making shipboard EMS/PMS a multi-objective, high-dimensional, and coupled optimization problem. These characteristics also carry important implications for DRL modeling. In particular, heterogeneous energy integration, uncertainty, multi-time-scale interactions, and conflicting objectives increase the complexity of state representation, action selection, and reward design, while also requiring adaptive sequential decision-making under dynamic operating conditions. Consequently, DRL represents a promising approach to addressing these challenges, as it enables adaptive policies to be learned through interaction with complex environments rather than relying solely on predefined rules or accurate mathematical models. Detailed DRL methodologies and their implications are discussed in Sections 3 and 4.



Figure 2. Key challenges and hierarchical structure of shipboard EMS/PMS problems.

3. Reinforcement Learning for Shipboard EMS/PMS

In shipboard EMS/PMS research, RL formulates power-system operation as a sequential decision-making problem. Unlike conventional optimization methods that require system models, predefined rules, or future load profiles, RL allows an agent to improve its policy through operating environment interaction. Early RL-based studies examined whether shipboard power-system operation could be modeled as an MDP and solved through experience-based policy learning.

Several early studies illustrate this direction. Huotari et al. [30] formulated a passenger ship's auxiliary power network as an MDP and applied Q-learning to determine generator-set operating policies using operational data. Das et al. [31] implemented Q-learning to post-fault shipboard power-system reconfiguration, learning both the final switch configuration and switching sequences required for restoration. Oo and Kong [32] proposed a Markov-approximation-based real-time EMS for uncertain vessel power demand. These studies reveal that RL can support generator scheduling, auxiliary power-network control, real-time energy management, and power-system reconfiguration, while reducing dependence on predefined operating rules under uncertain or time-varying operating conditions.

However, early RL-based approaches have structural limitations. First, they use discrete action spaces [30,31], such as generator-set operating policies and switching sequences, rendering them less suitable for continuous-control variables such as generator output, ESS charge/discharge power, FC operating level, and power-sharing ratio. Second, these studies address subproblems: auxiliary generator operation [30], post-fault reconfiguration [31], and real-time energy management [32] under uncertain vessel power demand. These studies overlook the hierarchical EMS/PMS structure involving power dispatch, energy

coordination, and operational strategy. Third, as the number of states, actions, energy sources, and constraints increases, tabular or simple value-based RL suffers scalability issues, including slow learning, unstable convergence, and degraded policy quality.

These limitations motivate the transition from conventional RL to DRL. By incorporating RL with deep neural networks (DNNs), DRL approximates complex value functions or policies, handling large state spaces, nonlinear system behavior, and diverse decision variables. Accordingly, DRL is promising for adaptive decision-making under uncertainty, sequential control across multiple time scales, and coordination among heterogeneous energy resources.

4. Deep Reinforcement Learning for Ship EMS/PMS

4.1. DRL for Ship EMS/PMS

Shipboard EMS/PMS's complex optimization problems involve multiple energy sources, hierarchical decision-making, and uncertain operating environments. Conventional RL (Section 3) struggles with large state spaces, continuous/hybrid actions, and coupled multi-objective decisions. DRL addresses these limitations via DNNs for value/policy approximation, extending learning-based control from small-scale discrete decisions to continuous power control, hybrid scheduling, and coordinated operation among heterogeneous energy resources. Jung and Chang [33] applied DRL to a PEMFC/battery hybrid electric ship by considering hydrogen consumption and component degradation. Hasanvand et al. [34] proposed a multi-objective DQN-based power scheduling method for an emission-free ferry, balancing operating cost and reliability. Song et al. [35] formulated all-electric ship EMS as a constrained MDP with hybrid discrete–continuous decision variables using a hybrid penalized proximal policy optimization (HP3O). Together, these studies demonstrate that DRL extends shipboard EMS/PMS from discrete, subproblem-specific control to multi-objective, constrained, and hybrid decision-making.

4.2. DRL Methodology for Ship EMS/PMS

4.2.1. MDP Formulation

Most DRL-based studies formulate the shipboard EMS/PMS control problem as an MDP, where an agent observes the system state, selects an action, receives a reward, and transitions to the next state through interaction with the shipboard power-system environment. The state transition captures the system-level response of the shipboard power system after an action is applied. MDP formulation varies by ship architecture, energy-source configuration, control objective, and operational constraints. Figure 3 illustrates the DRL–ship interaction, where the ship power system serves as the MDP environment and the DRL agent derives control policies by interacting with system states, objective functions, and constraints. Table 2 summarizes representative MDP components used in DRL-based shipboard EMS/PMS studies.

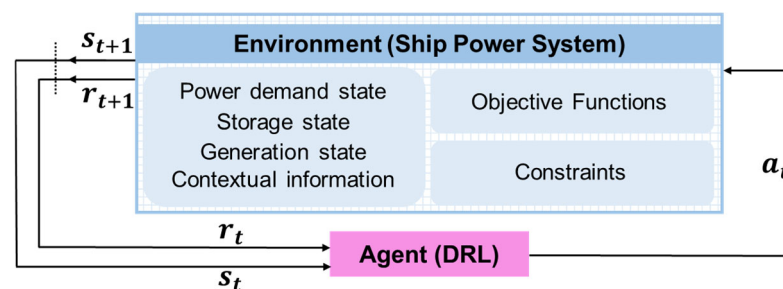


Figure 3. Agent–environment interaction framework for DRL-based ship EMS/PMS.

Table 2. MDP components in DRL-based ship EMS/PMS studies.

Component	Category	Variables
State	Power demand	Load demand, propulsion load, service load
	Storage	Battery SOC, health/degradation state
	Generation	Generator status, FC output, hydrogen level
	Voyage/environment	Speed, route progress, weather, draft/loading condition
	Temporal information	Forecast load, historical sequence, time step
Action	Discrete control	Generator ON/OFF, switching, operating-mode selection
	Continuous control	Power dispatch, charge/discharge power
	Hybrid control	UC + dispatch + ESS control, flexible-load control
	Operational decision	Speed command, voyage scheduling, route-related decision
Reward	Reward signal	Cost, emission, SOC deviation, constraint violation, degradation

The state represents agent-observed operating conditions. Common state variables include load demand, generator output and status, ESS state of charge (SOC), FC output, hydrogen storage level, power balance, and time. Xiao et al. [36] defined the state of an all-electric ship energy-scheduling environment using scheduling time, load power, ESS SOC, and generator power. Fang et al. [37] included ship speed, service load, PV output, ESS SOC, and time under PV uncertainty. Therefore, state design has expanded from compact power-system variables to incorporate renewable uncertainty, temporal information, and operational context. In practical shipboard operation, however, several variables relevant to EMS/PMS, such as battery state of health, fuel-cell degradation state, and spatially or temporally varying sea-state conditions, may be incompletely observed or unavailable as direct control inputs, instead being inferred through state-estimation, health-monitoring, onboard sensing, or environmental forecasting procedures. Under these conditions, the decision problem is more appropriately treated as partially observable, and recurrent policy architectures such as LSTM-based networks can exploit observation histories to encode latent temporal information and improve robustness to incomplete observations [38,39].

Actions are control decisions selected by the agent. Typical actions include generator on/off and output adjustment, ESS charge/discharge power, FC operating level, power-splitting ratio, cruising speed, and flexible-load control. Wu et al. [40] extended the action space from one-dimensional discrete to multi-dimensional continuous control for PEMFC cluster management. Ahmadi et al. [41] defined continuous power flows among FCs, batteries, cold ironing, and a recuperative organic Rankine cycle. Shang et al. [42] incorporated generator adjustment, battery power, and ship speed into the action design. Therefore, action design is moving from discrete switching or limited power-level selection toward continuous and hybrid control formulations.

The reward defines the DRL agent's learning objective, typical cost, emissions, SOC deviation, constraint violations, component degradation, and reliability loss. Reward construction directly shapes the learned policy (Section 4.2.3). Consequently, MDP design determines the decision scope of the DRL framework, from low-level power dispatch to broader energy coordination and operational strategy.

4.2.2. DRL Algorithms

Various DRL algorithms have been applied to shipboard EMS/PMS problems, from value-based methods for discrete decisions to actor–critic algorithms for continuous control and hybrid/hierarchical structures. This progression reflects the requirements for high-dimensional state representation, continuous power control, operational constraints, multi-objective optimization, and coordination across decision layers. Figure 4 illustrates

the methodological evolution of DRL algorithms, while Table 3 summarizes their key characteristics and representative applications in shipboard EMS/PMS.

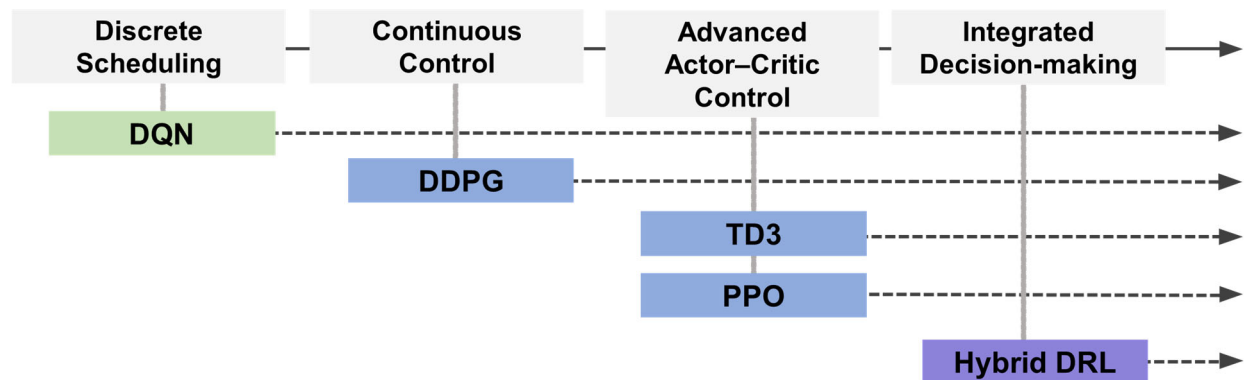


Figure 4. Methodological evolution of DRL algorithms for ship EMS/PMS.

Table 3. Comparative characteristics of DRL algorithms for ship EMS/PMS.

Action Space	Type	Algorithm	Reference	Strengths	Limitations
Discrete	Value-based	DQN	[34]	Simple implementation, stable learning for discrete decisions	Not suitable for continuous control
Continuous	Actor–critic	DDPG	[41]	Fine-grained continuous control, direct policy learning	Training instability
		TD3	[40]	Improved training stability, reduced overestimation	Higher computational complexity
Continuous/hybrid	Policy gradient/Actor–critic	PPO	[35]	Stable policy updates, robust constrained learning	Lower sample efficiency
Hybrid	Hybrid architecture	Hybrid DRL	[42]	Prediction integration, hierarchical decision support	Increased model complexity

In early DRL-based ship EMS/PMS studies, DQN was widely adopted for discrete action-space problems (generator commitment, limited output-level selection, energy scheduling, and switching control) by approximating the Q-function with a DNN. Hasanvand et al. [34] applied DQN-based multi-objective power scheduling to an emission-free ferry, considering cost-effective operation and reliability-oriented scheduling. This study exhibits that value-based DRL supports discrete or discretized scheduling decisions.

As ship EMS/PMS problems involve continuous power control, actor–critic algorithms have gained attention. Deep Deterministic Policy Gradient (DDPG) learns a deterministic policy via actor–critic, suitable for continuous actions. Ahmadi et al. [41] applied DDPG to a zero-emission ferry, managing continuous power-flow decisions among PEM FCs, batteries, cold ironing, and a recuperative organic Rankine cycle system. This study marks a methodological shift from value-based scheduling to continuous power-control-oriented DRL.

To improve the stability of deterministic actor–critic methods, Twin Delayed Deep Deterministic Policy Gradient (TD3) has also been adopted in ship EMS studies. TD3 extends DDPG with twin critics, delayed policy updates, and target policy smoothing to reduce Q-value overestimation. Wu et al. [40] proposed a TD3-based intelligent EMS for a hybrid-electric coastal ferry, using real-ship load profiles to coordinate multiple PEMFC clusters and a battery system in continuous state and action spaces. Their study

demonstrates TD3's suitability for multi-source continuous power control under stochastic operating conditions.

Proximal policy optimization (PPO)-based methods have been applied for stable policy updates and simple implementation. In 2024, Song et al. [35] proposed a Hybrid Penalty proximal policy optimization (HP3O) for all-electric ship EMS, extending PPO to a hybrid discrete–continuous action problem. Their framework used separate continuous and discrete actor networks to control generator power, ESS power, cruising speed, flexible load, and generator on/off status within a constrained MDP. This indicates PPO's usefulness for ship EMS problems involving operational constraints and hybrid actions.

Beyond single algorithms, hybrid and extended DRL frameworks have emerged for higher-dimensional, integrated ship decision-making. In 2024, Shang et al. [42] proposed a bi-level multi-step dueling double DQN (MSD3QN) for joint optimization of power generation and voyage scheduling in all-electric ships under uncertain navigation conditions and variable load demands. The method integrates dueling network architecture, double Q-learning, multi-step bootstrap technology, and an optimization solver integrated through action classification. This framework extends DRL beyond low-level power dispatch toward coordinated power generation, energy storage, and voyage-related operation.

Overall, DRL algorithms for ship EMS/PMS have evolved from value-based DQN (discrete scheduling) to actor–critic methods (e.g., DDPG and TD3) for continuous power control, and further to PPO-based and hybrid frameworks for constrained, hybrid, and hierarchical decision-making. This methodological evolution reflects the need for high-dimensional states, continuous controls, multi-objectives, uncertain environments, and integrated decision structures in modern ship power systems.

4.2.3. Objective Functions and Reward Design

Objective-function and reward-design strategies determine the DRL agent's learning direction, influencing convergence behavior, policy quality, and operational performance. In shipboard EMS/PMS problems, reward functions transform operational objectives and constraints into scalar feedback, including fuel or hydrogen consumption, operating cost, emissions, ESS SOC maintenance, component degradation, power-balance satisfaction, reliability, and voyage-level performance. Therefore, reward design is not merely a technical element of MDP formulation, but a key modeling step shaping the DRL agent's learned operational behavior. Table 4 summarizes representative reward components in DRL-based shipboard EMS/PMS studies, illustrating their expansion toward multi-objective formulations.

The fundamental objective is economic operation, representing fuel consumption, hydrogen consumption, electricity cost, generator operating cost, or total voyage cost. Zhu et al. [43] designed a multi-objective reward for FC hybrid ships combining voyage cost, FC lifespan, and battery SOC maintenance. Their DQL-based EMS achieved 92.6% of DP's voyage economy while maintaining FC durability and battery SOC. This finding validates that economic reward design can extend beyond short-term fuel-cost minimization to lifecycle-aware operation.

Environmental objectives have grown in importance. In 2025, Abdalla et al. [44] designed DRL-based EMSs for hybrid-powered LNG vessels, minimizing cumulative greenhouse-gas emissions (methane slip and CO₂ emissions). Their DRL-based EMS reduced GHG emissions by up to 14.6% versus peak-shaving and load-leveling, showing that reward design can shift from cost-oriented power allocation to emission-oriented operation.

Another objective concerns storage-state management and component health. In hybrid ship power systems, ESSs and FCs directly affect efficiency, lifetime, and reliability. Zhu et al. [43] incorporated FC lifespan and battery SOC into the reward, while

Fang et al. [37] designed a segmented reward to maintain ESS SOC within 0.27–0.80 under PV uncertainty, reducing fuel consumption to 90.32% and 94.57% of the non-PV and robust-optimization cases, respectively. Guo et al. [45] formulated a DDPG-based reward using fuel consumption and supercapacitor SOC deviation, indicating that storage-state regulation is often embedded directly into the learning objective.

Table 4. Reward components used in DRL-based ship EMS/PMS studies.

Objective Category	Representative Reward Components	Main Purpose
Economic operation	Fuel, hydrogen, start-up/shutdown, operating, and energy costs, efficiency reward	Cost reduction, efficiency improvement
Constraint handling	Power-balance, reserve-margin, demand-mismatch, feasibility, and output-limit penalties	Feasible and secure operation
Storage management	SOC tracking, SOC-bound and battery-power penalties, end-of-trip SOC requirement, charge/discharge regulation	Storage availability, energy balancing
Sustainability and asset health	Emission, battery degradation, fuel-cell aging, smoothness, and switching penalties	Emission reduction, lifetime extension
Voyage-level operation	Mode- and schedule-related rewards/penalties, mission-efficiency and voyage-time objectives	Voyage-aware operation
Reward design strategy	Normalization/scaling, fixed and adaptive weighting, reward shaping, constraint-aware design	Balanced and context-aware prioritization

Reward functions also serve as implicit constraint-handling mechanisms. To satisfy power balance, output limits, SOC bounds, and safety constraints in shipboard EMS/PMS, many studies introduce penalty terms for constraint violations. Zhao et al. [46] designed a reward function for a green ship EMS combining economic cost, power-balance penalty, and SOC safety penalty under PV uncertainty. This study reveals that reward design can guide DRL agents toward physically feasible and operationally safe policies without optimization constraints.

Some studies extend reward design toward voyage- and environment-aware operation. Wang et al. [47] developed a DDPG-based EMS for a diesel–electric hybrid ship considering sailing-route division and navigation environment. Their framework reduced fuel consumption by 2.3%, 5.5%, 18.1%, and 20.1% compared with DQN, Q-learning, HAAR, and rule-based algorithms, respectively. This case shows linking of reward and objective design to onboard power allocation and route-dependent energy management.

However, the central challenge of multi-objective reward design is not merely selecting several reward components, but converting heterogeneous objectives into a coherent learning signal. Operating cost, emissions, SOC deviation, degradation proxies, and power-balance errors differ in units, numerical ranges, and temporal consequences. When such objectives are combined through a fixed weighted sum without appropriate normalization, differences in numerical scale can disproportionately influence the learning signal, while a single fixed preference vector may impose trade-offs that are unsuitable across all operating conditions. Reward components should therefore be normalized or scaled using physically meaningful reference ranges, and the suitability of fixed scalarization should be carefully assessed for the intended operating context [48].

Moreover, the preferred objective hierarchy is operating-mode dependent. During port entry/departure, adequate power availability, maneuvering reliability, and local-emission compliance may require greater emphasis; during anchorage, low-emission hotel-load supply and sufficient SOC reserve for contingencies become more important; during full-

speed navigation, propulsion-power tracking, fuel efficiency, and component thermal and degradation stress become increasingly consequential. Fixed scalarization is therefore unlikely to represent the full range of maritime operating priorities. Context-conditioned or adaptive reward weighting based on operating mode, mission phase, and system condition provides a practical direction for reflecting these changing objective priorities [49].

Overall, objective-function and reward-design strategies in DRL-based shipboard EMS/PMS have evolved from cost minimization toward multi-objective and context-aware structures incorporating economic, environmental, storage, constraint, and voyage-aware considerations. This evolution reflects the hierarchical nature of shipboard EMS/PMS, where power dispatch, energy coordination, and operational strategy emphasize different optimization priorities. Therefore, reward design helps interpret each layer's decision scope and operational objective.

4.3. Hierarchical Classification of DRL-Based Ship EMS/PMS Studies

Shipboard EMS/PMS is a hierarchical decision-making structure involving real-time power control, multi-source energy coordination, and voyage-level operational planning. Therefore, this paper reinterprets DRL-based shipboard EMS/PMS studies by their decision-making roles, not just algorithms or system architectures.

Existing studies are categorized into three layers: power dispatch, energy coordination, and operational strategy. Power dispatch focuses on real-time source-level output, including generator or FC output, ESS charge/discharge power, and power-balance control at the execution level [50]. Energy coordination addresses cooperative operation among heterogeneous resources, including DGs, GTGs, FCs, ESSs, renewable sources, and shore power [51]. Operational strategy involves voyage-level decisions incorporating sailing speed, route conditions, future load demand, mission schedules, and port-related information [38]. Figure 5 illustrates this inclusive hierarchy, where the decision scope and time horizon expand from lower-level dispatch to higher-level operational planning. This perspective reveals where DRL research is concentrated and where integrated shipboard EMS/PMS development is needed.

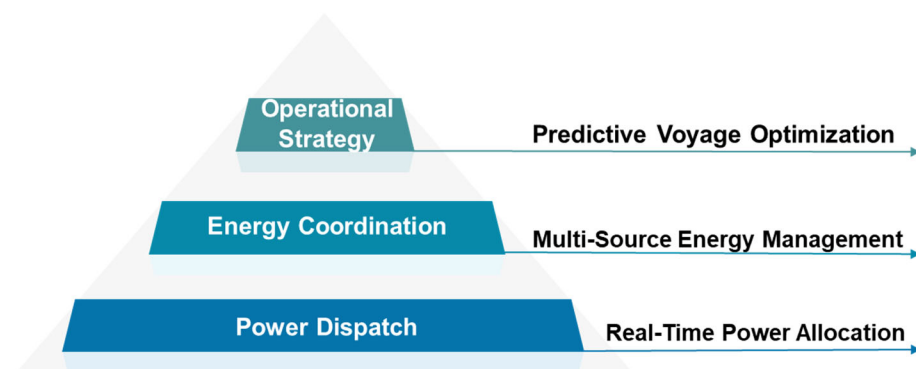


Figure 5. Hierarchical classification of DRL-based studies in ship EMS/PMS.

4.3.1. Power Dispatch

Power dispatch is the most fundamental execution layer in DRL-based shipboard EMS/PMS research. It reflects short-term control of source-level power outputs, such as generator power, FC output, ESS charge/discharge power, and power-sharing ratios, under time-varying load demand and operational constraints. In this review, studies are classified as power dispatch when primary DRL decisions are real-time or short-term power-allocation commands, even with multiple onboard energy sources. The main objec-

tive is power balance with improved short-term economic efficiency, load-following, and operational feasibility.

Most reviewed DRL-based shipboard EMS/PMS studies fit within this layer. Shang et al. [50] formulated all-electric ship dispatch as an MDP and applied DQN for generator–ESS under uncertain propulsion and pulse loads. This study represents value-based DRL for real-time dispatch. Guo et al. [45] proposed a DDPG-based EMS for diesel–electric series hybrid ships, with actions as continuous output power of four diesel-generator prime movers and rewards for fuel consumption and supercapacitor SOC regulation. Ahmadi et al. [41] also applied DDPG to a zero-emission ferry, optimizing hourly power flows among PEM FCs, batteries, cold ironing, and a recuperative organic Rankine cycle system. These studies demonstrate that continuous-action DRL can represent source-level power-flow optimization. In 2025, Wu et al. [40] proposed a TD3-based EMS for a hybrid-electric coastal ferry controlling multiple PEMFC clusters and a battery system under stochastic loads. Although this study involves multiple-source coordination, it is classified as power dispatch because the primary decision variables are short-term source-level power references.

Overall, power dispatch is the most accessible and explored layer because states, actions, and rewards are clearly definable around load demand, source output, SOC, fuel consumption, and power-balance constraints. However, it addresses only short-term execution control, not higher-level decisions such as source operating priorities, long-term degradation management, voyage scheduling, or route-aware energy planning. These limitations motivate the expansion from power dispatch toward energy coordination and operational strategy.

4.3.2. Energy Coordination

Energy coordination represents the intermediate decision layer between real-time power dispatch and voyage-level operational strategy. Unlike power dispatch (short-term source-level output commands), energy coordination focuses on jointly utilizing, prioritizing, and managing multiple onboard energy resources over operating intervals. Since DGs, GTGs, FCs, batteries, supercapacitors, renewable sources, and shore power differ in efficiency characteristics, response capabilities, degradation behaviors, emission profiles, and operating constraints, the key issue is coordinating their roles to improve system-level performance.

In this layer, DRL learns cooperative operating policies among multiple energy sources, not just following instantaneous load demand. Reddy et al. [51] developed a health-aware energy and emissions management strategy for zero-emission ships powered by FCs and batteries. Their framework implemented DDQL, SAC, and PPO with rewards for operating expenses, fuel consumption, and component degradation. This study highlights energy coordination in FC–battery source prioritization and degradation-aware multi-source management. In 2026, Li et al. [52] proposed a TD3-based real-time EMS for a methanol range-extended series hybrid port tugboat. The TD3 agent determined power distribution between the engine and battery while balancing fuel economy, emission reduction, SOC maintenance, and dynamic response under transient port-duty cycles. This study reveals energy coordination extends beyond power allocation by balancing economy, emissions, storage management, and real-time implementability. Fan et al. [53] further extended energy coordination to adaptive operation under uncertain navigation conditions using an LSTM-augmented DRL framework for an FC/battery hydrogen-hybrid ship. Compared with the original DDPG strategy, LSTM-DDPG reduced FC operating pressure and hydrogen consumption under unknown operating conditions, demonstrating robust, adaptive multi-source management.

Overall, energy coordination extends power dispatch to include source prioritization, storage-state management, degradation-aware operation, emission reduction, and robustness under uncertainty. However, it still does not fully address higher-level decisions such as speed planning, route selection, port scheduling, or mission-level operation. Therefore, energy coordination intermediately bridges real-time power dispatch and operational strategy.

4.3.3. Operational Strategy

Operational strategy represents the highest decision layer in DRL-based shipboard EMS/PMS. It determines vessel-level operating directions via voyage conditions, external environments, mission requirements, and operational information (e.g., sailing speed, route conditions, weather, port electricity price, arrival schedule, regulatory requirements, and future load demand). These factors directly affect propulsion load, fuel consumption, emissions, and energy-storage utilization. Therefore, operational strategy is a long-horizon sequential decision-making problem linking onboard power management with voyage-level operation.

Shang et al. [38] addressed this upper-layer problem by proposing an IMPALA-RND-based-multitask DRL framework that jointly optimized power generation and voyage scheduling in all-electric ship power systems in real time under uncertain sailing environments, stochastic load demand, and seasonal port electricity-price fluctuations. This study illustrates operational strategy by using DRL for onboard power management and voyage-aware planning under heterogeneous operating scenes. Xiao et al. [54] proposed a DQN knowledge-distillation compression algorithm for coordinated ship speed and energy scheduling, addressing operational strategy with deployable voyage-level scheduling. Their target problem jointly considered generator scheduling, ESS operation, and navigation speed. The compressed model reduced parameters, FLOPs, and memory by 87.7%, 92.61%, and 88.3%, respectively, with only 0.33% higher economic consumption compared with the teacher model. This study's operational strategy includes not only high-level planning performance, but also deployability under practical computational constraints. Alshareef and Alghanmi [55] implemented DRL to regulate-aware vessel operation across multiple vessel types and operational scenarios by considering cargo load, vessel speed, weather conditions, fuel type, and engine performance. Their approach improved fuel efficiency by up to 10%, reduced EEXI values by 8–15%, and improved CII ratings by 10–30%, extending DRL to vessel-level operational optimization.

Overall, operational strategy is the layer most aligned with future intelligent and autonomous ship operation. It encompasses lower-layer functions such as power allocation and energy coordination, while incorporating voyage planning, operating-scene uncertainty, deployability, and regulatory compliance. However, this layer introduces challenges of long-horizon reward design, sparse feedback, real-voyage data requirements, high-dimensional environmental uncertainty, and simulation-to-validation gaps. Therefore, operational strategy bridges DRL-based EMS/PMS with ship-wide operational intelligence.

Based on this classification, Figure 6 summarizes hierarchical DRL architectures for shipboard EMS/PMS. Figure 6a depicts the power-dispatch framework for generating dispatch actions from system-state information through real-time allocation and source output control. Figure 6b presents the energy coordination architecture, where multi-source energy inputs are processed through state assessment, source prioritization, and energy scheduling to produce decision outputs including energy allocation, source priority, charge/discharge scheduling, and reserve planning. Figure 6c illustrates the operational strategy framework designed for predictive planning and high-level decision-making. In this framework, operational inputs and contextual feedback are processed through

forecasting, context analysis, and policy generation, supporting voyage planning, mission scheduling, and strategic energy planning. Figure 6d integrates the three layer-specific architectures, illustrating the top-down propagation of mission targets, commands, and constraints together with bottom-up feedback on system status across the hierarchy, and distinguishing operating time scales from voyage-level planning to millisecond-to-second local control.

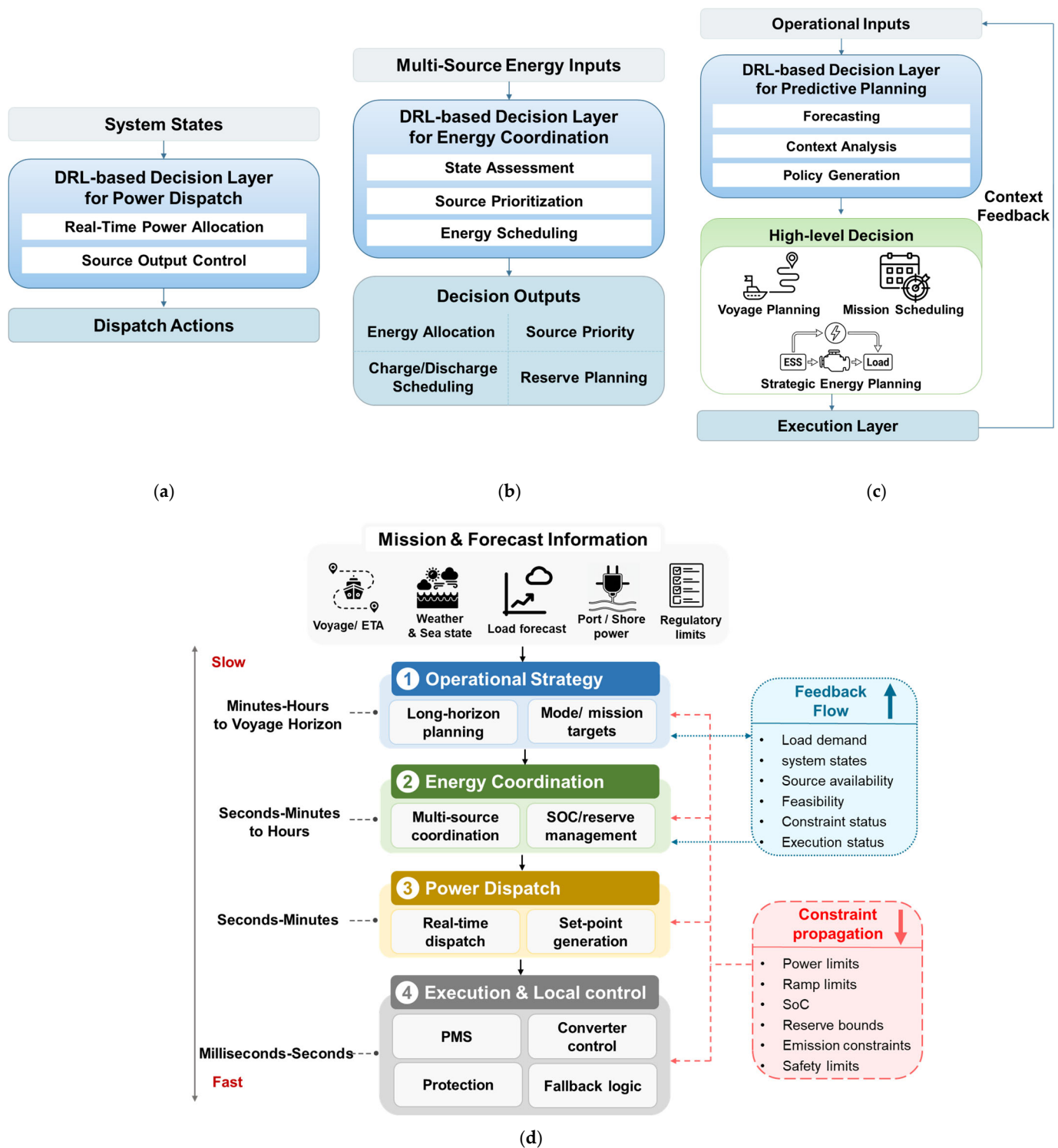


Figure 6. Hierarchical DRL architectures for shipboard EMS/PMS: (a) power dispatch, (b) energy coordination, (c) operational strategy, and (d) integrated hierarchical DRL architectures.

Table 5 compares the three decision layers in terms of decision scope, control targets, time scale, objectives, state, action, and key limitations, underscoring the need for integrated hierarchical frameworks connecting all layers. Table 6 summarizes representative quantitative outcomes reported in DRL-based shipboard EMS/PMS studies across the three decision layers. Taken together, the reported outcomes illustrate a consistent correspondence between the decision scope of each layer and the performance measures used to assess it. Power-dispatch studies emphasize source-level fuel use and allocation performance, whereas energy-coordination studies assess trade-offs among operating economy, storage management, component durability, and emissions. Operational-strategy studies extend the assessment to voyage-level efficiency, regulatory performance, and computational deployability. This layer-dependent pattern provides literature-based quantitative support for the proposed hierarchical perspective.

Table 5. Comparison of hierarchical decision layers in DRL-based shipboard EMS/PMS.

Aspect	Power Dispatch	Energy Coordination	Operational Strategy
Decision scope	Real-time allocation	Multi-source coordination	Voyage-level planning
Primary target	Generator output, ESS power, power balance	Source priority, energy flow, FC–ESS coordination	Route, speed, schedule, mode
Time scale	Supervisory: seconds–minutes, Local control: milliseconds–seconds	Seconds–minutes to hours	Minutes–hours to voyage horizon
Inter-layer interface	Source priority, fault status, power-limit envelopes, power-balance residuals	Mission, SOC/reserve targets, source priorities, power-limit envelopes	Route, speed, mission targets, feasibility, reserve, health feedback
Objectives	Fuel cost, power balance, load tracking	Cost–emission trade-off, efficiency, SOC/degradation	Voyage efficiency, prediction, schedule adherence
State	Load, SOC, generator status	Source states, SOC/SOH, reserve, forecast load	Weather, route, ETA, forecasts
Action	Dispatch, ON/OFF, power sharing	Resource allocation, source-priority control	Route/speed planning, voyage scheduling
Key limitations	Short-term focus	Reward tuning, source coupling	High uncertainty, sparse rewards, sim-to-real gap

Table 6. Representative quantitative outcomes across the proposed decision layers in DRL-based shipboard EMS/PMS.

	Reference	Evaluation Basis	Metrics	Result
Power dispatch	[47]	DDPG vs. DQN, Q-learning, HAAR, rule-based EMS	Fuel consumption	Fuel-consumption reduction of 2.3–20.1% across methods
Energy coordination	[43]	DQL vs. dynamic programming (DP)	Voyage economy, fuel-cell durability, battery SOC	92.6% of DP-based voyage economy, with maintained fuel-cell durability and battery SOC
	[44]	TD3 vs. peak-shaving and load-leveling (PS-LL) EMS	GHG-reduction efficiency	GHG-reduction efficiency of 79.3% (TD3) vs. 66.2% (PS-LL)
Operational strategy	[53]	DQN-KD student model vs. teacher DQN model	Parameters, FLOPs, memory, economic consumption	87.7%/92.61%/88.3% reduction in parameters/FLOPs/memory; 0.33% increase in economic consumption
	[54]	DRL-based optimization vs. traditional optimization methods	Fuel efficiency, EEXI, CII	Fuel-efficiency improvement of up to 10%, EEXI reduction of 8–15%, CII-rating improvement of 10–30%

5. Future Perspective and Research Directions

5.1. Research Trends and Limitations

Research on DRL-based shipboard EMS/PMS has evolved from RL with discrete sequential decision-making toward high-dimensional, continuous, multi-objective frameworks. Early studies by Huotari et al. [30], Das et al. [31], and Oo and Kong [32] framed shipboard power-system operation as sequential decision-making problems involving generator operation, power-system reconfiguration, and real-time energy management. However, their discrete or sampled decision structures limited scalability to realistic EMS/PMS problems with continuous power variables, multiple energy sources, and complex operational constraints, motivating the transition to DRL.

The first dominant trend is DRL research concentration at the power-dispatch layer. Studies such as Wu et al. [40], Ahmadi et al. [41], Guo et al. [45], and Shang et al. [50] formulate dispatch problems around short-term measurable variables, including load demand, generator output, FC output, ESS SOC, fuel consumption, and power-balance error. These variables allow clear state, action, and reward definitions and quantifiable objectives such as fuel-cost reduction, load tracking, SOC regulation, and power-balance maintenance. Therefore, power dispatch is the most accessible and actively explored layer.

The second trend is the expansion from source-level dispatch toward energy coordination. With DGs, GTGs, FCs, batteries, renewable sources, and shore power, EMS/PMS must coordinate source priorities, storage utilization, degradation behavior, emissions, and operating efficiency over longer horizons. Reddy et al. [51] addressed health-aware energy and emissions management for FC/battery ships. Li et al. [52] used TD3 for methanol engine-generator set and battery pack coordination in a range-extended port tugboat. Fan et al. [38] combined LSTM augmentation with DRL for hydrogen-hybrid ships under navigation uncertainty. Therefore, DRL is moving beyond instantaneous output allocation toward multi-source coordination, degradation/emission awareness, and robustness under uncertainty.

The third, less mature, trend is operational strategy-level DRL, connecting shipboard EMS/PMS with voyage-level decisions such as sailing speed, route conditions, mission scheduling, regulatory compliance, and future load demand. Shang et al. [38] optimized power generation and voyage scheduling using multitask DRL, Xiao et al. [54] connected ship speed and energy scheduling with model compression for deployable DQN-based control. Alshareef and Alghanmi [55] applied DRL to regulation-aware vessel operation by optimizing energy efficiency, fuel use, and EEXI/CII-related performance. These studies indicate a shift from reactive onboard energy management to predictive, voyage-aware operational planning. However, this layer remains underexplored due to challenges in long-horizon learning, real voyage data, environmental prediction, sparse-reward handling, and coordination with lower EMS/PMS execution layers.

Algorithmically, a clear transition is evident. Value-based methods (e.g., DQN) dominate discrete scheduling and unit-commitment decisions, whereas actor-critic algorithms (e.g., DDPG, TD3, PPO, and SAC) handle continuous or hybrid action problems. This transition reflects the physical reality that generator output, FC power, ESS charge/discharge power, and sailing speed are often continuous decision variables. The emergence of LSTM-augmented DRL, hybrid action frameworks, multitask DRL, and hierarchical DRL signals a shift from single-algorithm comparison toward learning architectures for practical operating environments [38,42,53].

Despite progress, research gaps remain. First, real-world validation is insufficient. Many studies rely on simulation with assumed load profiles, or scenario-based evaluation in dispatch- and strategy-oriented studies [38,50,55]. Although useful for benchmarking, simulation cannot fully capture maritime uncertainties (e.g., weather variation, sensor noise,

unexpected load changes, equipment aging, schedule changes, and faults). Li et al. [51] advanced toward hardware-in-the-loop (HIL) testing, but onboard transferability of learned policies remains insufficiently demonstrated.

Second, generalization and scalability remain open challenges. Many DRL policies are trained for specific ships, energy configurations, routes, or operating conditions. Shang et al. [38] and Alshareef and Alghanmi [55] considered multiple but predefined operating scenarios, while Fan et al. [53] employed LSTM-based data augmentation to improve adaption to navigation uncertainty. However, DRL policies that perform reliably across different vessels, unseen routes, and changing environments remain underdeveloped.

Third, safety and reliability requirements remain inadequately addressed. Inappropriate EMS/PMS actions risk power imbalance, voltage/frequency instability, blackout risk, propulsion loss, or excessive battery and FC stress. Some studies incorporated reliability or constraint-related terms, such as loss-of-load expectation [34] and constrained MDP [35]. However, strict safety guarantees, fault-tolerant operation, and robust behavior under abnormal conditions remain rarely demonstrated, making safety-aware and risk-sensitive DRL a major unresolved issue.

Fourth, explainability and operator acceptance remain important barriers. Most DRL-based EMS/PMS frameworks rely on DNNs with uninterpretable decision logic. Song et al. [35], Wu et al. [40], and Shang et al. [42] focus on control performance, optimization quality, or decision-making capability, not policy explainability. This limits confidence in deploying autonomous or semi-autonomous EMS/PMS in safety-critical shipboard environments.

Finally, integrated multi-time-scale decision-making remains underdeveloped in DRL-based shipboard EMS/PMS. Power dispatch, energy coordination, and operational strategy are functionally connected (Section 4.3), but many studies treat them separately. Dispatch studies focus on short-term output allocation [7,40,41,45], coordination studies manage energy-source interactions over longer intervals [51–53], and strategy studies address voyage-level planning [38,54,55]. Practical ship operation, however, requires continuous interaction across these layers. Therefore, future DRL-based shipboard EMS/PMS should move beyond isolated control tasks to integrated frameworks connecting real-time power allocation, multi-source energy management, predictive planning, and safety-aware operation across decision layers. Table 7 summarizes these key research gaps and their practical implications for DRL-based shipboard EMS/PMS, highlighting the need for real-world validity, safety, scalability, interpretability, and deployability.

Table 7. Key research gaps in DRL-based shipboard EMS/PMS.

Research Gap	Limitations in the Reviewed Studies	Practical Implication
Real-world validation	Reliance on simulation and scenario-based validation	Uncertain onboard performance under real operating conditions
Generalization and scalability	Policies trained for specific vessels, routes, energy configurations, operating conditions	Limited transferability to new vessels, routes, environments
Safety and reliability	Limited treatment of safety constraints and abnormal conditions	Safety-critical barriers to onboard deployment
Explainability and operator acceptance	Black-box policies with limited interpretability	Low operator trust, difficult certification
Integrated multi-time-scale decision-making and physical deployment	Separate decision layers with limited treatment of timing, command validity, computing placement	Inconsistent inter-layer commands, compromising coordination, fallback safety, hierarchical deployability

These limitations in current DRL-based shipboard EMS/PMS research should be distinguished from the broader ship energy-system optimization literature. Recent model-based approaches have already applied double-layer or nested optimization to jointly address propulsion-system sizing and operational energy management, coordinate multiple on-board energy sources, and evaluate economic and environmental trade-offs across mission profiles and operating conditions [15,56]. Wang et al. [15] integrated hybrid propulsion-system sizing with MILP-based energy management for diesel engines, batteries, and fuel cells, whereas Maloberti and Zaccone [56] combined outer-level component sizing with inner-level power-split optimization and assessed solution robustness under power-demand variations. Future integrated DRL frameworks should therefore be evaluated not only against simple rule-based strategies but also against advanced model-based optimization benchmarks, and should demonstrate added value under model uncertainty, changing operating conditions, unmodeled dynamics, or online decision-making requirements.

5.2. Future Research Directions

Future research on DRL-based shipboard EMS/PMS should move beyond algorithmic improvements toward integrated, reliable, explainable, and deployable intelligent decision-making systems. Key challenges (Section 5.1) include real-world validation, generalization, safety, explainability, multi-time-scale coordination, and practical deployment. Therefore, future studies should develop DRL frameworks that operate under real maritime uncertainty and integrate power dispatch, energy coordination, and operational strategy into a unified hierarchy. Figure 7 presents a maritime-oriented roadmap for DRL-based shipboard EMS/PMS, linking research stages to ship-operating contexts and a deployment pathway from simulation to supervised pilot deployment. To ensure maritime relevance, the following directions are framed around recurrent ship operating contexts, including berth operation with shore-power connection or disconnection, harbor maneuvering, coastal and open-sea transit, port approach, and abnormal operating conditions.

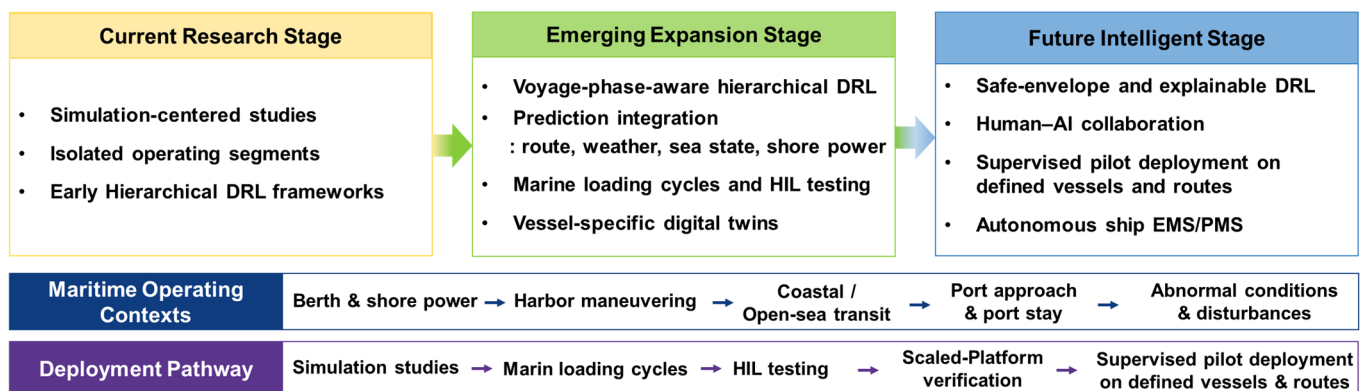


Figure 7. Roadmap for future development of DRL-based shipboard EMS/PMS.

First, future research should prioritize integrated hierarchical DRL frameworks connecting power dispatch, energy coordination, and operational strategy as interdependent decision functions across time scales. An upper-level module determines voyage strategy, operating mode, speed profile, or energy-source priority based on mission objectives and forecasts. A middle-level module coordinates roles and energy flows among DGs, GTGs, FCs, batteries, renewable sources, and shore power. A lower-level module executes real-time dispatch commands under power-balance and equipment constraints. Shang et al. [42] linked power generation with voyage scheduling through bi-level DRL. Lai and Illindala [57] exhibited distributed EMS structures for resilience and scalability.

Furthermore, Fan et al. [53] combined prediction with DRL under unknown navigation conditions. Future work should develop cross-layer DRL frameworks unifying power dispatch, energy coordination, and operational strategy. Building on this, an implementable pathway evaluates cross-layer policies over complete voyage sequences, translating voyage-level objectives into phase-specific source priority, energy-reserve, and power-limit references.

However, physically deployable cross-layer policies require a multi-rate, asynchronous execution architecture rather than a single controller operating at one update rate. The operational strategy layer updates route-, speed-, mission-, and mode-related targets over minutes to hours, the energy coordination layer translates these into source priority, SOC/reserve targets, and power-limit envelopes over seconds to minutes, and the DRL-based power-dispatch supervisor updates feasible source-level set-points at an appropriate interval, while millisecond-to-second voltage/current regulation, load sharing, protection, and fault isolation remain the responsibility of existing PMS, converter controllers, and protection devices [22,24].

Future studies should explicitly define inter-layer message semantics and failure behavior, including time stamps, command versions, validity intervals, operating-mode identifiers, acknowledgements, health/feasibility feedback, and event-triggered updates after faults or major load deviations, and should quantitatively evaluate how communication delay, packet loss, and asynchronous updates affect stability, reserve preservation, and fallback activation. A practical architecture should run latency-sensitive inference, constraint checking, and fallback logic onboard, while shore-side infrastructure supports offline training, digital-twin updating, fleet-level analytics, and controlled policy retraining, with message interfaces compatible with maritime ethernet and safety/security network practices [58,59].

Second, future studies should strengthen real-world validation and experimental verification. Several DRL-based EMS/PMS studies rely on numerical simulation, which is useful for algorithm development but insufficient for onboard feasibility. Fan et al. [60] called for progression from virtual simulation to HIL and physical testing. Liu et al. [61] developed a scaled hydrogen-electric ship EMS platform considering equipment response, time delay, environmental noise, transient operation, and steady-state behavior. Ünlübayir et al. [62] validated machine-learning-based EMS using propulsion-component measurements and test-bench experiments. Future DRL-based EMS/PMS should utilize HIL testing, scaled platforms, pilot deployment, and real voyage data. Specifically, an implementable validation pathway should progress from representative marine loading cycles to HIL testing with propulsion and power-electronic dynamics, scaled-platform verification, and supervised pilot deployment on a defined vessel and route.

Third, future research should develop prediction-integrated and data-driven DRL frameworks. Shipboard EMS/PMS decisions depend on future load demand, weather, sea state, voyage schedule, port operation, and equipment condition. Therefore, DRL should not be purely reactive based only on the current system state. Ünlübayir et al. [63] developed a data-driven ship EMS using automated tracking system data and weather information to predict load demand for steadier FC operation. Jiang et al. [64] proposed a two-level real-time EMS with multi-step load forecasting for hydrogen FC-powered ships, verified by HIL experiments. Future DRL should integrate forecasting models, time-series learning, data fusion, and uncertainty-aware state representations for proactive energy management. Practically, prediction-integrated DRL should fuse onboard power-system data with automatic identification system and route information, weather or sea-state forecasts [61], port-call information [28], and shore-power availability [21], translating the resulting uncertainty into source priority, energy-reserve, and fuel or hydrogen management decisions for scenarios such as delayed berthing or unavailable shore power.

Fourth, future research should develop safety-aware, robust, and health-aware DRL. Inappropriate EMS/PMS decisions risk power imbalance, blackout risk, propulsion loss, excessive component stress, or accelerated degradation. Kim et al. [65] proposed robust power scheduling for PEMFC/BESS electric ships under environmental and load uncertainties. Tsoumpris and Theotokatos [29] developed a health-aware EMS combining reliability estimation with energy management. Han et al. [66] highlighted availability, redundancy, and maintenance for autonomous maritime machinery systems. Future DRL-based EMS/PMS should incorporate constrained RL, robust DRL, risk-sensitive rewards, health-state estimation, degradation-aware control, and supervisory safety mechanisms. Before deployment, policies should accordingly be constrained by vessel- and architecture-specific safe operating envelopes, including source-output and ramp limits, energy reserves, equipment-health limits, and propulsion reserve, with independent supervisory override of infeasible actions.

Fifth, future research should enhance explainability and human-AI collaboration. DRL's dependence on DNNs renders its black-box nature limiting to expert trust in energy and power systems [67]. Nguyen and Ahn [68] emphasized that human-in-the-loop AI can improve transparency, adaptability, and reliability in uncertain energy environments. For shipboard EMS/PMS, operators must understand, verify, and override AI-based decisions when necessary. Future systems should therefore provide interpretable policy analysis, action-level explanations, uncertainty indicators, and HIL interfaces to support operator trust and regulatory acceptance. Building on this, an implementable research pathway evaluates role-specific decision displays, override procedures, and audit records through operator-in-the-loop bridge or engine-room simulator studies during port approach, berth operation, loss of shore power, or weather-driven load changes, presenting the recommended action, active constraints, and safe alternatives. Such development should be aligned with class-society guidance on autonomous and remote-control functions [69,70].

Sixth, future studies should expand digital-twin-based validation and lifecycle-aware EMS/PMS. Digital twins bridge simulation-based DRL training and real ship operation via real-time monitoring, virtual testing, model updating, and performance prediction. Kim [71] developed a digital-twin-based lifecycle analysis model for electric/hybrid propulsion ships, while Madusanka et al. [72] proposed a four-layer digital-twin framework for marine vessels. Future DRL-based EMS/PMS should use digital twins not only as simulation environments, but also as continuously updated operational platforms connecting onboard sensor data, shore monitoring, lifecycle assessment, and predictive control. Vessel-specific digital twins should be calibrated during commissioning and updated to represent route and seasonal changes, hull and propeller fouling, battery or FC aging, and component replacement before policy updates are deployed. For practical deployment, real-time onboard digital-twin functions should focus on state estimation, condition monitoring, and scenario evaluation using current operational data, while computationally intensive model calibration, policy retraining, and extensive validation should generally be performed onshore or during scheduled offline maintenance periods. Only verified policy updates that satisfy predefined safety and performance criteria should subsequently be deployed onboard.

Finally, future research should develop benchmark datasets and evaluation protocols. Divergent ship models, load profiles, energy-source configurations, objective functions, and performance metrics currently complicate fair comparison. The field necessitates open datasets, representative voyage scenarios, standardized uncertainty profiles, and common metrics covering fuel cost, emissions, safety constraints, degradation, computational burden, real-time feasibility, robustness, and interpretability. Such standards would support fair comparison and accelerate the transition from academic studies to reliable industrial applications. Rather than a single generic benchmark, the field requires tiered maritime

scenario libraries covering vessel classes, propulsion architectures, berth and shore-power operation, maneuvering, transit, schedule disruption, and source-fault or degradation events. Shared maritime benchmark datasets should provide time-synchronized voyage and route information, propulsion and service-load profiles, source and ESS operating states, and component performance maps, including fuel-consumption and efficiency characteristics, together with operating constraints such as power, ramp-rate, and SOC limits. The datasets should further specify energy-system configurations, control actions, operating scenarios, fault or degradation events where available, and standardized training, validation, and test splits with common performance metrics to support reproducible controller evaluation.

Overall, future DRL-based shipboard EMS/PMS research should progress from isolated simulation toward integrated, prediction-informed, safety-aware, explainable, and experimentally validated energy-management frameworks. The objective is not just better DRL algorithms, but coherent unification of real-time dispatch, multi-source energy coordination, and operational strategy into a coherent decision-making system. Such development will evolve DRL from a promising optimization tool into a practical enabler for autonomous, low-emission, and resilient ship operation.

5.3. Practical Considerations

DRL-based shipboard EMS/PMS faces practical barriers beyond algorithmic performance as an intelligent optimization tool. In real ship operation, EMS/PMS is embedded in a safety-critical cyber-physical system connected to propulsion equipment, converters, sensors, communication networks, automation systems, operators, and regulations. Therefore, deployment requires real-time execution, system integration, safety supervision, certification, data management, cybersecurity, and economic feasibility.

The first consideration is real-time onboard execution. DRL policies must meet shipboard PMS/EMS response times under rapid propulsion transients, hotel-load variations, maneuvering, and environmental disturbances. Planakis et al. [73] emphasized EMS evaluation using realistic marine loading cycles over simplified load profiles. Therefore, practical DRL-based EMS/PMS requires evaluation not only by offline training performance, but also by inference latency, deterministic timing, embedded compatibility, and stability under realistic load transients.

Second, integration with existing shipboard control systems is essential. Vessels already include PMS, generator controllers, converter controllers, protection systems, automation systems, and supervisory monitoring layers. Therefore, DRL-based EMS/PMS should act as an upper-level optimization or decision-support layer, not replace existing control and protection logic. This direction aligns with class-society guidance on autonomous and remotely operated ship functions, which emphasizes verifiable decision-making, functional boundaries, and risk-based approval procedures [69,70]. Lai and Illindala [57] showed that distributed EMS structures improve resilience and scalability, while Madusanka et al. [72] highlighted the need for connected data infrastructures in maritime digitalization and digital twins. Accordingly, practical deployment of shipboard control logic requires standardized interfaces, clear authority boundaries, conflict-prevention mechanisms, and explicit communication timing and command-validity rules.

At the implementation level, latency-sensitive functions should be separated from longer-horizon learning and analytics functions. DRL inference at the dispatch layer, constraint checking, and fallback activation should run onboard so that safe operation does not depend on uninterrupted shore-side connectivity, while shore-side infrastructure supports offline training, digital-twin updating, fleet-level analytics, and controlled policy retraining. Inter-layer messages should include a timestamp, command version, validity

interval, operating-mode identifier, acknowledgement, and feasibility/health status, and a receiving layer should reject stale or inconsistent commands and, if a valid higher-layer command is absent, transition to a pre-validated deterministic PMS or fallback controller. Communication interfaces should follow maritime ethernet and network safety/security requirements [58,59], with control-function placement consistent with class-society guidance for onboard, off-ship, or hybrid control arrangements [69].

Third, operational safety and backup logic must be embedded in the implementation architecture. Under abnormal conditions (sensor faults, communication delays, sudden load surges, component degradation, and actuator malfunctions), safe behavior outweighs average cost reduction. Therefore, DRL-based EMS/PMS should include independent safety monitors, hard constraint-checking layers, fallback controllers, and manual or automatic override functions. Implementation-oriented testing under realistic marine conditions should be required before deployment.

For example, if an SOC or propulsion-load sensor produces unavailable or implausible data, the corresponding observation should be treated as unreliable, and the DRL agent should not directly execute actions dependent on that input. The candidate action should first be checked against predefined power, ramp-rate, SOC, reserve, and safety limits, and if no feasible action can be identified, the controller should switch to a pre-validated rule-based PMS or fallback dispatch mode that preserves essential power balance and reserve margins. Similarly, if a generating unit fails to follow its commanded output, source availability should be updated and the dispatch recalculated or replaced by fallback control [74].

Onboard explainability should translate this process into concise operational information for the ship master and engineering personnel. Rather than presenting internal DRL-model details, the interface should indicate the selected operating mode, the detected anomaly, the active safety constraint, and the required response. For example, conservative dispatch may be activated due to an unreliable SOC measurement, ESS discharge may be limited to preserve reserve capacity, and sensor verification may be required [75].

Fourth, regulatory compliance, certification, and accountability are deployment barriers. Learning-based controllers are difficult to certify due to dependence on training data, neural-network structure, and operating context. Kaiser and Chowdhury [76] noted similar barriers for hydrogen-powered marine vessels (regulatory, standardization, infrastructure, and cost barriers) as DRL-based EMS/PMS. Even if a policy performs well in simulation, deployment needs verifiable performance, reproducible testing, traceable decision records, and clear responsibility allocation, including audit trails, version control, input records, safety-case documentation, and operator approval procedures.

Fifth, data management and model maintenance support long-term reliability. A DRL policy trained on a fixed dataset may lose validity with aged ships, changed routes, drifted sensors, replaced components, or shifted operating profiles. Karatuğ et al. [77] demonstrated that onboard operational data can support cost-effective energy-efficiency monitoring without new sensing infrastructure. Therefore, DRL-based EMS/PMS deployment requires data monitoring, anomaly detection, periodic evaluation, retraining criteria, and model update procedures.

Sixth, cybersecurity and software resilience are implementation requirements. DRL-based EMS/PMS may depend on sensor data, communication networks, software updates, remote monitoring, and digital interfaces with other ship systems. Manipulated input signals, communication channels, or model parameters can cause the controller to generate unsafe or inefficient decisions. Lai and Illindala [57] addressed privacy-preserving distributed EMS operations, and Madusanka et al. [70] revealed that maritime digital-twin applications rely on continuous data exchange. Ship cyber-resilience standards require

secure access, integrity checks, protected updates, anomaly detection, and safe-mode transition strategies [78,79].

Finally, economic feasibility determines industrial adoption of DRL-based EMS/PMS from research prototypes. Even with improved DRL fuel economy, emissions, or component lifetime, adoption may remain limited if sensing, computing, software development, validation, certification, maintenance, and crew-training costs are excessively high. Therefore, DRL-based EMS/PMS should be assessed from a total-cost-of-ownership perspective, including capital cost, operational savings, maintenance reduction, certification cost, downtime avoidance, component lifetime extension, and training requirements.

Table 8 summarizes these practical deployment requirements, emphasizing the shift from simulation-oriented algorithms to deployable intelligent control systems.

Table 8. Practical deployment requirements for DRL-based shipboard EMS/PMS.

Practical Consideration	Deployment Requirement	Industrial Challenge
Real-time onboard execution	Low-latency inference, deterministic timing, embedded compatibility	Limited computing and fast load transients
System integration	Supervisory decision support compatible with PMS, converters, protection, automation	Legacy interfaces and control-logic conflicts
Safety and backup logic	Safety monitoring, constraint checking, fallback control, override function	Sensor faults, abnormal loads, delays, actuator failures
Certification and accountability	Audit trail, version control, reproducible testing, responsibility allocation	Limited certification for learning-based controllers
Data and model maintenance	Data monitoring, anomaly detection, periodic evaluation, model updates	Sensor drift, route changes, missing data, equipment replacement
Cybersecurity	Secure access, integrity checks, protected updates, safe-mode transition	Data manipulation, intrusion, model tampering
Economic feasibility	Total-cost-of-ownership assessment	High validation, certification, hardware, software, training costs

6. Conclusions

This paper systematically reviewed DRL-based shipboard EMS/PMS studies and restructured them into three decision layers: power dispatch, energy coordination, and operational strategy. Moving beyond conventional algorithm-centered classifications, this hierarchical perspective clarifies how DRL has been applied to shipboard energy and power management and how the field is evolving.

The review reveals that most studies concentrate on power dispatch, with recent expansion toward energy coordination. In contrast, operational strategy remains underexplored, despite its importance for linking EMS/PMS with speed control, route-aware planning, predictive operation, and voyage-level scheduling. The central implication is that these three layers should not be treated independently. Future DRL-based shipboard EMS/PMS should be an integrated hierarchical decision-making framework where operational strategy provides voyage-level planning and predictive guidance, energy coordination manages multi-source energy flows and operating priorities, and power dispatch executes real-time power allocation under system constraints. This structure is well-suited to DRL due to its sequential policy learning in dynamic environments and support multi-time-scale decision-making across short-term, medium-term, and long-horizon operations. To realize this transition, future research should move beyond isolated algorithm comparisons to address real-world validation, generalization, safety assurance, explainability, regulatory compliance, and industrial deployment requirements. Therefore, DRL should evolve from

a standalone control method into a practical intelligence layer cooperating with existing shipboard EMS/PMS, protection systems, operators, and digital infrastructure. From a methodological perspective, this review highlights that the effectiveness of DRL-based shipboard EMS/PMS depends not only on the choice of learning algorithm but also on the formulation of the MDP and reward functions. MDP formulations are evolving to incorporate richer operational context and predictive information, while DRL algorithms are progressing from value-based approaches toward continuous-control and hierarchical architectures capable of addressing increasingly complex decision-making problems. Reward design, in turn, has shifted from purely economic objectives toward multi-objective formulations that jointly consider operational cost, emissions, component degradation, operational constraints, and voyage-level performance.

Overall, DRL has strong potential to support adaptive, predictive, and autonomous shipboard EMS/PMS. However, its practical value lies in integrating real-time dispatch, multi-source energy coordination, and voyage-level operational strategy into a coherent, reliable ship-wide decision-making system. The proposed hierarchical perspective provides a foundation for advancing next-generation green and autonomous ship energy-management technologies.

Author Contributions: Conceptualization, Y.K., D.P. and I.-Y.C.; validation, Y.K., D.P. and I.-Y.C.; formal analysis, Y.K.; investigation, Y.K.; data curation, Y.K.; writing—original draft preparation, Y.K.; writing—review and editing, Y.K. and D.P.; visualization, Y.K.; supervision, I.-Y.C.; project administration, I.-Y.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Korea Planning & Evaluation Institute of Industrial Technology (KEIT) funded by the Korea Ministry of Trade, Industry and Energy (MOTIE) (Grant No. RS-2025-02633139), and by the Institute for Information & Communications Technology Planning & Evaluation (IITP) under the Information Technology Research Center (ITRC) support program (Grant No. RS-2023-00259004) funded by the Ministry of Science and ICT (MSIT), Korea.

Data Availability Statement: No new data were created or analyzed in this study.

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

References

1. 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. [\[CrossRef\]](#)
2. Kumar, D.; Zare, F. A comprehensive review of maritime microgrids: System architectures, energy efficiency, power quality, and regulations. *IEEE Access* **2019**, *7*, 67249–67277. [\[CrossRef\]](#)
3. Hou, J.; Sun, J.; Hofmann, H. Adaptive model predictive control with propulsion load estimation and prediction for all-electric ship energy management. *Energy* **2018**, *150*, 877–889. [\[CrossRef\]](#)
4. 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. [\[CrossRef\]](#)
5. Hein, K.; Xu, Y.; Wilson, G.; Gupta, A.K. Coordinated optimal voyage planning and energy management of all-electric ship with hybrid energy storage system. *IEEE Trans. Power Syst.* **2021**, *36*, 2355–2365. [\[CrossRef\]](#)
6. Liu, L.; Yang, X.; Li, X.; Zhou, X.; Wang, Y.; Tang, T.; Song, Q.; Liu, Y. Prior knowledge-based two-layer energy management strategy for fuel cell ship hybrid power system. *J. Mar. Sci. Eng.* **2025**, *13*, 94. [\[CrossRef\]](#)
7. Wu, P.; Partridge, J.; Anderlini, E.; Liu, Y.; Bucknall, R. Near-optimal energy management for plug-in hybrid fuel cell and battery propulsion using deep reinforcement learning. *Int. J. Hydrogen Energy* **2021**, *46*, 40022–40040. [\[CrossRef\]](#)
8. International Maritime Organization. *Guidelines on Maritime Cyber Risk Management*, MSC-FAL.1/Circ.3/Rev.3; IMO: London, UK, 2025. Available online: <https://wwwcdn.imo.org/localresources/en/OurWork/Security/Documents/MS-C-FAL.1-Circ.3-Rev.3.pdf> (accessed on 1 July 2026).
9. ISO/IEC 42001:2023; Information Technology—Artificial Intelligence—Management System. International Organization for Standardization: Geneva, Switzerland; International Electrotechnical Commission: Geneva, Switzerland, 2023. Available online: <https://www.iso.org/standard/42001> (accessed on 1 July 2026).

10. ISO/IEC 23894:2023; Information Technology—Artificial Intelligence—Guidance on Risk Management. International Organization for Standardization: Geneva, Switzerland; International Electrotechnical Commission: Geneva, Switzerland, 2023. Available online: <https://www.iso.org/standard/77304.html> (accessed on 1 July 2026).
11. IEC 62443-2-1:2024; Security for Industrial Automation and Control Systems—Part 2-1: Security Program Requirements for IACS Asset Owners. International Electrotechnical Commission (IEC): Geneva, Switzerland, 2024. Available online: <https://webstore.iec.ch/en/publication/62883> (accessed on 1 July 2026).
12. IEC 62443-4-1:2018; Security for Industrial Automation and Control Systems—Part 4-1: Secure Product Development Lifecycle Requirements. International Electrotechnical Commission (IEC): Geneva, Switzerland, 2018. Available online: <https://webstore.iec.ch/en/publication/33615> (accessed on 1 July 2026).
13. Triviza, N.L.; Rentizelas, A.; Theotokatos, G.; Boulougouris, E. Decision support methods for sustainable ship energy systems: A state-of-the-art review. *Energy* **2022**, *239*, 122288. [\[CrossRef\]](#)
14. Zohrabi, N.; Shi, J.; Abdelwahed, S. An overview of design specifications and requirements for the MVDC shipboard power system. *Int. J. Electr. Power Energy Syst.* **2019**, *104*, 680–693. [\[CrossRef\]](#)
15. Wang, X.; Shipurkar, U.; Haseltalab, A.; Polinder, H.; Claeys, F.; Negenborn, R.R. Sizing and control of a hybrid ship propulsion system using multi-objective double-layer optimization. *IEEE Access* **2021**, *9*, 72587–72601. [\[CrossRef\]](#)
16. Si, Y.; Wang, R.; Zhang, S.; Zhou, W.; Lin, A.; Zeng, G. Configuration optimization and energy management of hybrid energy system for marine using quantum computing. *Energy* **2022**, *253*, 124131. [\[CrossRef\]](#)
17. Wu, P.; Bucknall, R. Hybrid fuel cell and battery propulsion system modelling and multi-objective optimisation for a coastal ferry. *Int. J. Hydrogen Energy* **2020**, *45*, 3193–3208. [\[CrossRef\]](#)
18. Zhang, Z.; Guan, C.; Liu, Z. Real-time optimization energy management strategy for fuel cell hybrid ships considering power sources degradation. *IEEE Access* **2020**, *8*, 87046–87059. [\[CrossRef\]](#)
19. Liu, H.; Fan, A.; Li, Y.; Bucknall, R.; Vladimir, N. Multi-objective hierarchical energy management strategy for fuel cell/battery hybrid power ships. *Appl. Energy* **2025**, *379*, 124981. [\[CrossRef\]](#)
20. Aziz, M.; Trinh, P.-H.; Hudaya, C.; Chung, I.-Y. Coordinated control strategy for hybrid multi-PEMFC/BESS in a shipboard power system. *Electr. Power Syst. Res.* **2025**, *243*, 111495. [\[CrossRef\]](#)
21. Vahabzad, N.; Mohammadi-Ivatloo, B.; Anvari-Moghaddam, A. Optimal energy scheduling of a solar-based hybrid ship considering cold-ironing facilities. *IET Renew. Power Gener.* **2021**, *15*, 532–547. [\[CrossRef\]](#)
22. Xie, P.; Tan, S.; Bazmohammadi, N.; Guerrero, J.M.; Vasquez, J.C. A real-time power management strategy for hybrid electrical ships under highly fluctuated propulsion loads. *IEEE Syst. J.* **2023**, *17*, 395–406. [\[CrossRef\]](#)
23. Yang, Q.; Zhang, S.; Zhao, W.; Tao, J.; Lu, X.; Meng, C.; Zhao, Y.; Zhang, H.; Zhao, G.; Jiang, C.; et al. A two-stage energy management strategy for hybrid ship power systems considering the dynamic output characteristics of batteries. *Appl. Energy* **2025**, *396*, 126232. [\[CrossRef\]](#)
24. Xie, P.; Asgharian, H.; Guerrero, J.M.; Vasquez, J.C.; Araya, S.S.; Liso, V. A two-layer energy management system for a hybrid electrical passenger ship with multi-PEM fuel cell stack. *Int. J. Hydrogen Energy* **2024**, *50*, 1005–1019. [\[CrossRef\]](#)
25. Yuan, Y.; Wang, J.; Yan, X.; Shen, B.; Long, T. A review of multi-energy hybrid power system for ships. *Renew. Sustain. Energy Rev.* **2020**, *132*, 110081. [\[CrossRef\]](#)
26. Xie, P.; Guerrero, J.M.; Tan, S.; Bazmohammadi, N.; Vasquez, J.C.; Mehrzadi, M.; Al-Turki, Y. Optimization-based power and energy management system in shipboard microgrid: A review. *IEEE Syst. J.* **2022**, *16*, 578–590. [\[CrossRef\]](#)
27. Pang, S.; Zou, L.; Zhang, L.; Wang, H.; Wang, Y.; Liu, X.; Jiang, J. A hybrid TCN-BiLSTM short-term load forecasting model for ship electric propulsion systems combined with multi-step feature processing. *Ocean Eng.* **2025**, *316*, 119808. [\[CrossRef\]](#)
28. Park, Y.; Kim, S.; Eom, J.; Kim, S. Development of a reinforcement learning-based ship voyage planning optimization method applying machine learning-based berth dwell-time prediction as a time constraint. *J. Mar. Sci. Eng.* **2026**, *14*, 43.
29. Tsoumpris, C.; Theotokatos, G. A decision-making approach for the health-aware energy management of ship hybrid power plants. *Reliab. Eng. Syst. Saf.* **2023**, *235*, 109263.
30. 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. [\[CrossRef\]](#)
31. Das, S.; Bose, S.; Pal, S.; Schulz, N.N.; Scoglio, C.M.; Natarajan, B. Dynamic reconfiguration of shipboard power systems using reinforcement learning. *IEEE Trans. Power Syst.* **2013**, *28*, 669–676. [\[CrossRef\]](#)
32. Oo, T.Z.; Kong, A.W.-K. Real-time energy management for marine applications using Markov approximation. *IEEE Trans. Power Syst.* **2023**, *38*, 4341–4354. [\[CrossRef\]](#)
33. 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. [\[CrossRef\]](#)
34. Hasanvand, S.; Rafiei, M.; Gheisarnejad, M.; Khooban, M.-H. Reliable power scheduling of an emission-free ship: Multiobjective deep reinforcement learning. *IEEE Trans. Transp. Electr.* **2020**, *6*, 832–843. [\[CrossRef\]](#)

35. 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. [\[CrossRef\]](#)
36. 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. [\[CrossRef\]](#)
37. Fang, H.; Zhang, H.; Wen, S.; Li, Z.; Zeng, Z.; Zhu, M.; Lin, P. A reinforcement learning based real-time energy management method for mobile microgrid considering photovoltaic uncertainty. *Int. J. Electr. Power Energy Syst.* **2025**, *170*, 110844. [\[CrossRef\]](#)
38. Shang, C.; Fu, L.; Xiao, H.; Bao, X.; Xu, X. Joint optimization of power generation and voyage scheduling in ship power system based on operating scene clustering and multitask deep reinforcement learning. *IEEE Trans. Transp. Electr.* **2024**, *10*, 8432–8442. [\[CrossRef\]](#)
39. Hausknecht, M.; Stone, P. Deep Recurrent Q-Learning for Partially Observable MDPs. In *Proceedings of the 2015 AAAI Fall Symposium Series*; Association for the Advancement of Artificial Intelligence: Washington, DC, USA, 2015; pp. 29–31.
40. Wu, P.; Partridge, J.; Anderlini, E.; Liu, Y.; Bucknall, R. An intelligent energy management framework for hybrid-electric propulsion systems using deep reinforcement learning. *Int. J. Hydrogen Energy* **2025**, *106*, 282–294. [\[CrossRef\]](#)
41. Ahmadi, H.; Rafiei, M.; Igder, M.A.; Gheisarnejad, M.; Khooban, M.-H. An energy efficient solution for fuel cell heat recovery in zero-emission ferry boats: Deep deterministic policy gradient. *IEEE Trans. Veh. Technol.* **2021**, *70*, 7571–7581. [\[CrossRef\]](#)
42. Shang, C.; Fu, L.; Bao, X.; Xiao, H.; Xu, X.; Hu, Q. Dynamic joint optimization of power generation and voyage scheduling in ship power system based on deep reinforcement learning. *Electr. Power Syst. Res.* **2024**, *229*, 110165. [\[CrossRef\]](#)
43. Zhu, L.; Liu, Y.; Zeng, Y.; Guo, H.; Ma, K.; Liu, S.; Zhang, Q. Energy management strategy for fuel cell hybrid ships based on deep reinforcement learning with multi-optimization objectives. *Int. J. Hydrogen Energy* **2024**, *93*, 1258–1267. [\[CrossRef\]](#)
44. Abdalla, A.; Kirchen, P.; Gopaluni, B. Deep reinforcement learning for methane slip reduction in hybrid-powered liquefied natural gas marine vessels. *Sustain. Energy Technol. Assess.* **2025**, *81*, 104404. [\[CrossRef\]](#)
45. 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. [\[CrossRef\]](#)
46. 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. [\[CrossRef\]](#)
47. Wang, X.; Yuan, Y.; Tong, L.; Yuan, C.; Shen, B.; Long, T. Energy management strategy for diesel–electric hybrid ship considering sailing route division based on DDPG. *IEEE Trans. Transp. Electr.* **2024**, *10*, 187–202. [\[CrossRef\]](#)
48. Roijers, D.M.; Vamplew, P.; Whiteson, S.; Dazeley, R. A Survey of Multi-Objective Sequential Decision-Making. *J. Artif. Intell. Res.* **2013**, *48*, 67–113. [\[CrossRef\]](#)
49. Abels, A.; Roijers, D.M.; Lenaerts, T.; Nowé, A.; Steckelmacher, D. Dynamic Weights in Multi-Objective Deep Reinforcement Learning. In *Proceedings of the 36th International Conference on Machine Learning, Long Beach, CA, USA, 9–15 June 2019*; PMLR: Cambridge, MA, USA, 2019; Volume 97, pp. 11–20.
50. 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. [\[CrossRef\]](#)
51. Reddy, N.P.; Skjetne, R.; Os, O.S.; Papageorgiou, D. A comparison of the state-of-the-art reinforcement learning algorithms for health-aware energy and emissions management in zero-emission ships. *IEEE J. Emerg. Sel. Top. Ind. Electron.* **2024**, *5*, 149–166. [\[CrossRef\]](#)
52. Li, Z.; Long, W.; Tian, H. Research on energy management optimization for hybrid-powered port tugboat systems based on a dual-delay deep deterministic policy gradient algorithm. *Energies* **2026**, *19*, 905. [\[CrossRef\]](#)
53. Fan, A.; Liu, H.; Wu, P.; Yang, L.; Guan, C.; Li, T.; Bucknall, R.; Liu, Y. LSTM-augmented DRL for generalisable energy management of hydrogen-hybrid ship propulsion systems. *eTransportation* **2025**, *25*, 100442. [\[CrossRef\]](#)
54. Xiao, H.; Fu, L.; Shang, C.; Bao, X.; Xu, X. A knowledge distillation compression algorithm for ship speed and energy coordinated optimal scheduling model based on deep reinforcement learning. *IEEE Trans. Transp. Electr.* **2025**, *11*, 945–960. [\[CrossRef\]](#)
55. Alshareef, M.H.; Alghanmi, A.F. Optimizing maritime energy efficiency: A machine learning approach using deep reinforcement learning for EEXI and CII compliance. *Sustainability* **2024**, *16*, 10534. [\[CrossRef\]](#)
56. Maloberti, L.; Zaccone, R. A Nested Optimization Framework for Ship Hybrid Propulsion Systems Considering Equivalent CO₂ Emissions. *Ocean Eng.* **2026**, *355*, 125076. [\[CrossRef\]](#)
57. Lai, K.; Illindala, M.S. A distributed energy management strategy for resilient shipboard power system. *Appl. Energy* **2018**, *228*, 821–832. [\[CrossRef\]](#)
58. IEC 61162-450:2024; Maritime Navigation and Radiocommunication Equipment and Systems—Digital Interfaces—Part 450: Multiple Talkers and Multiple Listeners—Ethernet Interconnection. International Electrotechnical Commission (IEC): Geneva, Switzerland, 2024.
59. IEC 61162-460:2024; Maritime Navigation and Radiocommunication Equipment and Systems—Digital Interfaces—Part 460: Multiple Talkers and Multiple Listeners—Ethernet Interconnection—Safety and Security. International Electrotechnical Commission (IEC): Geneva, Switzerland, 2024.

60. Fan, A.; Li, Y.; Liu, H.; Yang, L.; Tian, Z.; Li, Y.; Vladimir, N. Development trend and hotspot analysis of ship energy management. *J. Clean. Prod.* **2023**, *389*, 135899. [\[CrossRef\]](#)
61. Liu, H.; Fan, A.; Bucknall, R.; Wu, P.; Liu, Y.; Xia, M. Experimental verification of a dual-objective energy management strategy: A case study of hydrogen-electric ship. *Energy Convers. Manag.* **2026**, *348*, 120659. [\[CrossRef\]](#)
62. Ünlübayir, C.; Youssfi, H.; Khan, R.A.; Salas Ventura, S.; Fortunati, D.; Rinner, J.; Börner, M.F.; Quade, K.L.; Ringbeck, F.; Sauer, D.U. Comparative analysis and test bench validation of energy management methods for a hybrid marine propulsion system powered by batteries and solid oxide fuel cells. *Appl. Energy* **2024**, *376*, 124183. [\[CrossRef\]](#)
63. Ünlübayir, C.; Mierendorff, U.H.; Börner, M.F.; Quade, K.L.; Blömeke, A.; Ringbeck, F.; Sauer, D.U. A data-driven approach to ship energy management: Incorporating automated tracking system data and weather information. *J. Mar. Sci. Eng.* **2023**, *11*, 2259. [\[CrossRef\]](#)
64. Jiang, J.; Zou, L.; Liu, X.; Han, Z.; Wang, R. A two-level real-time energy management strategy incorporating load forecasting for hydrogen fuel cell-powered ships. *Energy Convers. Manag.* **2025**, *341*, 120032. [\[CrossRef\]](#)
65. Kim, G.; Lee, M.; Chung, I.-Y. Robust optimal power scheduling for fuel cell electric ships under marine environmental uncertainty. *Energies* **2025**, *18*, 2837. [\[CrossRef\]](#)
66. Han, Z.; Zhang, D.; Fan, L.; Zhang, J.; Zhang, M. A dynamic Bayesian network model to evaluate the availability of machinery systems in Maritime Autonomous Surface Ships. *Accid. Anal. Prev.* **2024**, *194*, 107342. [\[CrossRef\]](#) [\[PubMed\]](#)
67. Machlev, R.; Heistrene, L.; Perl, M.; Levy, K.Y.; Belikov, J.; Mannor, S.; Levron, Y. Explainable artificial intelligence (XAI) techniques for energy and power systems: Review, challenges and opportunities. *Energy AI* **2022**, *9*, 100169. [\[CrossRef\]](#)
68. Nguyen, A.T.; Ahn, Y. Human-in-the-loop artificial intelligence in the energy sector: A systematic review of paradigm shifts toward Industry 5.0. *Adv. Appl. Energy* **2026**, *22*, 100267. [\[CrossRef\]](#)
69. DNV. *Autonomous and Remotely Operated Ships—AROS Class Notations*; DNV: Høvik, Norway, 2025. Available online: <https://www.dnv.com/maritime/autonomous-remotely-operated-ships/aros-class-notation/> (accessed on 1 July 2026).
70. American Bureau of Shipping. *Requirements for Autonomous and Remote-Control Functions*; ABS: Spring, TX, USA, 2024. Available online: <https://ww2.eagle.org/content/dam/eagle/rules-and-guides/current/other/323-requirements-for-autonomous-and-remote-control-functions-2024/323-autonomous-reqts-oct24.pdf> (accessed on 1 July 2026).
71. Kim, H. A real-time lifecycle analysis model with digital twin and novel control method for enhancing the environmental performance of electric/hybrid propulsion ships. *Ocean Eng.* **2025**, *329*, 121195. [\[CrossRef\]](#)
72. Madusanka, N.S.; Fan, Y.; Yang, S.; Xiang, X. Digital twin in the maritime domain: A review and emerging trends. *J. Mar. Sci. Eng.* **2023**, *11*, 1021. [\[CrossRef\]](#)
73. Planakis, N.; Papalambrou, G.; Kyrtatos, N. Ship energy management system development and experimental evaluation utilizing marine loading cycles based on machine learning techniques. *Appl. Energy* **2022**, *307*, 118085. [\[CrossRef\]](#)
74. Wabersich, K.P.; Zeilinger, M.N. A Predictive Safety Filter for Learning-Based Control of Constrained Nonlinear Dynamical Systems. *Automatica* **2021**, *129*, 109597. [\[CrossRef\]](#)
75. Sandberg, A.; Hinojosa, M.; Lekkas, A.M. Deep Reinforcement Learning for Ship Dynamic Positioning: A Live Explainable AI Approach on an Autonomous Ferry Prototype. *IEEE Access* **2026**, *14*, 59064–59094. [\[CrossRef\]](#)
76. Kaiser, R.; Chowdhury, A.M. Hydrogen-powered marine vessels: A rewarding yet challenging route to decarbonization. *Clean. Technol.* **2025**, *7*, 68. [\[CrossRef\]](#)
77. Karatuğ, Ç.; Tadros, M.; Ventura, M.; Guedes Soares, C. Strategy for ship energy efficiency based on optimization model and data-driven approach. *Ocean Eng.* **2023**, *279*, 114397. [\[CrossRef\]](#)
78. UR E26; Cyber Resilience of Ships, Rev.1. International Association of Classification Societies (IACS): London, UK, 2023. Available online: <https://iacs.org.uk/resolutions/unified-requirements/ur-e/> (accessed on 1 July 2026).
79. UR E27; Cyber Resilience of On-Board Systems and Equipment, Rev.1. International Association of Classification Societies (IACS): London, UK, 2023. Available online: <https://iacs.org.uk/resolutions/unified-requirements/ur-e/> (accessed on 1 July 2026).

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.