A Hierarchical Multi-Agent Deep Reinforcement Learning Framework for Coordinated Logistics-Energy Scheduling in Electrified Seaports

1st Wentao Lv
School of Electrical Engineering
Southeast University
Nanjing 210096, China
E-mail: 230249170@seu.edu.cn

2nd Yujian Ye
Department of Electrical and Electronic Engineering
Imperial College London
London, SW7 2AZ, UK
E-mail: yujian.ye11@imperial.ac.uk

3rd Tianxiang Cui
School of Computer Science
University of Nottingham
Nottingham, NG7 2RD, UK
E-mail: tianxiang.cui@nottingham.ac.uk

4th Goran Strbac
Department of Electrical and Electronic Engineering
Imperial College London
London, SW7 2AZ, UK
E-mail: g.strbac@imperial.ac.uk

***Abstract*—Port electrification is a key strategy for maritime decarbonization, which introduces tight coupling between logistics and energy systems. Conventional optimization methods are computationally expensive and struggle to meet the dynamic, real-time scheduling demands of ports. To address these challenges, this paper proposes a logistics-energy coordinated scheduling method for electrified ports. This method formulates the complex co-optimization problem as a multi-agent markov decision process. To handle the the mixed action space inherent in port scheduling, a two-layer Deep Reinforcement Learning (DRL) framework is designed based on the "Centralized Training with Decentralized Execution" (CTDE) paradigm. The upper (logistics) layer is responsible for coordinating the discrete actions of heterogeneous agents and the lower (energy) layer manages the continuous power decisions of all energy agents. The case study based on data from the Port of Ningbo demonstrates that, compared to the separated strategy, the proposed coordinated method achieves peak shaving and valley filling by proactively avoiding power loads during high-price periods. Numerical results show that the coordinated method reduces the total operational cost by 17.42%, including a 100% reduction in RES curtailment cost and a 74.34% reduction in carbon emission cost.**

***Index Terms*—Electrified Seaports, Coordinated Logistics-Energy Scheduling, Multiple Agent Reinforcement Learning, Hierarchical Framework**

I. INTRODUCTION

Maritime transport carries over 80% of global freight volume and accounts for 70% of trade value, processed through ports worldwide. As critical nodes linking maritime and terrestrial transport, ports form a vital component of the global supply chain [1]. However, the shipping industry and associated port operations entail significant energy consumption and are major sources of pollution, accounting for approximately 3%–5% of global greenhouse gas (GHG) emissions. Pollution within port areas is particularly acute, especially concerning GHG emissions and air pollution, primarily stemming from ship activities within and near the port vicinity.

Port electrification is a key decarbonization strategy, replacing traditional diesel equipment with electric alternatives like shore power, quay cranes (QC), and automated guided vehicles (AGV) to eliminate local emissions [2]. This transition, however, creates tight coupling between the logistics and energy systems. Consequently, port management evolves from a set of singular scheduling problems into a complex logistics-energy co-optimization challenge.

In the field of logistics scheduling, the Berth Allocation Problem (BAP) aims to optimize the berthing time and specific berth assignment for ships. Focusing on dynamic ship arrivals and continuous QC scenarios, [3] optimizes berth scheduling to minimize the total makespan for serving all ships. [4] investigates the berth allocation and QC assignment problem (BAQCAP) across multiple heterogeneous terminals, determining the assigned terminal and the number of allocated QCs for ships to minimize the comprehensive service cost. Furthermore, [5] extends the BAQCAP by considering QC operators as a scheduling object. Regarding internal transport within the port, [6] focuses on the scheduling of AGVs aiming to achieve dynamic and conflict-free AGV operations through integrated task assignment and path planning. For yard management, [7] establishes a stochastic optimization model to jointly optimize yard space allocation, container assignment, and yard crane (YC) deployment.

Within the context of port electrification, the co-optimization of energy management and logistics operations has emerged as a critical research focus. [8] investigates the coordinated management of hybrid Energy Storage Systems

This work was supported by the National Natural Science Foundation of China (52477083, 52207082, 52507139).

979-8-3315-8138-1

(ESS) and microgrids to address the high peak power demands generated by QCs. Focusing on the energy efficiency, [9] employs a dynamic programming approach to minimize the hydrogen fuel consumption of fuel cell hybrid AGVs under specific operating conditions. Recent research highlights the coupled nature between logistics and energy systems. [10] analyzes the energy consumption patterns of logistics elements including ship and QCs to optimize the cost of the integrated energy system. Building upon this coupled relationship, [11] designs a nested bi-level energy management method for ESS, which considers both carbon emissions and storage degradation costs, striving to coordinate the supply-demand balance between hydrogen and electricity.

Traditional port logistics scheduling methods are mainly divided into exact algorithms (e.g., Mixed-Integer Programming, often combined with decomposition [12]) and heuristic approaches, ranging from simple dispatching rules [13] to meta-heuristic algorithms [14]. However, these conventional methods face two major challenges: 1) Port scheduling problems are typically large-scale and NP-hard, making exact algorithms computationally intractable for practical use; 2) Traditional offline optimization requires recalculation for each operating day, which is ill-suited for the rapid, real-time decision-making required in port operations. Given these limitations, Reinforcement Learning (RL), a paradigm based on agent-environment interaction, offers a more effective and adaptive pathway for solving such complex sequential decision-making problems. [15] addresses the joint berth and yard scheduling problem in bulk terminals, proposing a Deep Reinforcement Learning (DRL) method based on prioritized experience replay to minimize ship turnaround time. [16] puts forward a bi-level DRL strategy for energy management, integrating berth allocation, energy scheduling, and battery swapping scheduling within a port microgrid.

The novel contributions of this paper are outlined as follows:
- A novel two-layer DRL architecture is designed, integrating DDQN and PPO. This framework effectively addresses the modeling and solution challenges in the logistics-energy coordinated scheduling of electrified ports, which are characterized by the mixed action space and the vast state space.

- The case study based on data from the Port of Ningbo validates that the proposed coordinated method, compared to the traditional separated scheduling, reduces the total operational cost by 17.42%, while simultaneously achieving a 100% reduction in RES curtailment and a 74.34% reduction in carbon emission costs.

II. Problem Formulation

In the day-ahead stage, the port operator establishes a baseline berth and QC allocation schedule based on the shipping lines provided expected arrival/departure times and container workload for electric ships (ES). During the intraday stage, uncertainties in voyage and cargo demand may cause delays or mismatches, rendering the baseline schedule inapplicable. Consequently, the port operator must reschedule the berth and QC allocation based on practical conditions, while E-AGVs, YCs, and yards are dynamically scheduled according to sequentially observed uncertainties.

A. Day-ahead Scheduling

Day-ahead planning comprises berth allocation and QC allocation. The former connects the port to the maritime transportation system and is crucial for generating the baseline schedule for intraday rescheduling. The latter is closely coordinated with berth allocation, as the number of QCs assigned to an ES directly influences its docking end time. The detailed mathematical modeling of the port day-ahead planning can be found in [17].

Day-ahead planning focuses on logistics scheduling, where the port operator aims to minimize the total logistics costs of ESs in the baseline.

$$F^{da} = c^{w} \sum_{s} (t_s^b - t_s^a) + c^{d} \sum_{s} (t_s^e - t_s^b) \tag{1}$$

Objective function (1) includes two components: 1) the waiting cost, defined as the difference between docking begin times and arrival times, and 2) the docking cost, defined as the difference between docking end times and begin times.

B. Intraday Scheduling

Day-ahead planning comprises 1) Berth and QC Rescheduling, 2) YC and Yard Allocation, 3) ES Charging Scheduling, 4) E-AGV Charging Scheduling and 5) Port Energy System Scheduling. The detailed mathematical modeling of the port intraday scheduling can be found in [17].

The objective function minimizes the total operational cost, which consists of the logistics-related schedule deviation cost and the total energy system cost. The latter is the sum of three components: the energy procurement cost, the carbon emission cost , and the RES curtailment cost.

$$F_t^{id} = Cost_t^{re} + Cost_t^{en} + Cost_t^{\mathrm{co}_2} + Cost_t^{cu} \tag{2}$$

III. The proposed DRL approach

Conventional solution methods are often computationally expensive when applied to real-time port scheduling demands, making them ill-suited for developing adaptive policies that require rapid responses. In contrast, RL offers an adaptive strategy capable of reacting swiftly to uncertainties and making online adjustments. Applying RL first requires reformulating the original co-optimization problem as a Markov Decision Process (MDP).

A. MDP reformulation

To model the complex, coupled logistics-energy system of the port, the problem is formulated as a Multi-Agent Reinforcement Learning (MARL) system. In this framework, the decision-making units are defined as cooperative agents: ES agents, QC agents, YC agents, E-AGV agents, ESS agent, GT agent, and grid agent.

979-8-3315-8138-1

1) State space: The state encompasses the core operational status of the environment, ESs, yards, E-AGVs, and ESS.

$$S_t^{env} = \{t, \overline{P_t^{WT}}, \overline{P_t^{PV}}, \lambda_t^{grid}, \lambda^{gas}\}$$
$$S_{s,t}^{ES} = \{f_{s,t}^{state}, B_{s,t}^{es}, E_{s,t}^{es}, TEU_{s,t}^{left}, t_s^a, t_s^d\}$$
$$S_{y,t}^{YC} = \{c_{y,t}\},\ S_{k,t}^{EV} = \{E_{k,t}^{ev}\},\ S_t^{ESS} = \{E_t^{ess}\}$$

Where $\overline{P_t^{WT}}$ and $\overline{P_t^{PV}}$ denote the maximum output power of wind and PV generators; λ_t^{grid} and λ^{gas} denote the grid electricity and natural gas price; $f_{s,t}^{state}$ denotes the operational state of ES s at time t (including unarrived, waiting, docking and lefted); $B_{s,t}^{es}$ denotes the berth b at which ES s is docked at time t; $E_{s,t}^{es}$ denotes the energy content of ES s; $TEU_{s,t}^{left}$ denotes the remaining containers to be handled for ES s; t_s^a and t_s^d denote the expected arrival time and latest leaving time of ES s; $c_{y,t}$ denotes the number of containers stored in yard y at time t; $E_{k,t}^{ev}$ and E_t^{ess} denote the energy content of E-AGV k and ESS at time t.

2) Action space: The joint actionencompasses all decisions made by the logistics (ES, QC, YC, E-AGV) and energy (ESS, GT) agents at time t to control the system's operation.

$$A_{s,t}^{ES} = \{\zeta_{s,b,t}, P_{s,t}^{esc}, P_{s,t}^{esd}\}, A_{q,t}^{QC} = \{\delta_{s,q,t}\}$$
$$A_{y,t}^{YC} = \{\mu_{s,y,t}\}, A_{k,t}^{EV} = \{e_{k,t}, r_{k,t}, z_{k,t}, \varphi_{s,k,t}, P_{k,t}^{evc}, P_{k,t}^{evd}\}$$
$$A_t^{ESS} = \{P_t^{essc}, P_t^{essd}\},\quad A_t^{GT} = \{P_t^{gt,g}\}$$

Where $\zeta_{s,b,t}$ indicates whether ES s is docked at berth b at time t; $P_{s,t}^{esc}$ and $P_{s,t}^{esd}$ denote the charging/discharging power for ES s at time t; $\delta_{s,q,t}$/$\mu_{s,y,t}$/$\varphi_{s,k,t}$ denotes whether QC q/YC y/E-AGV k served ES s at time t; $e_{k,t}$ /$o_{k,t}$/$z_{k,t}$ indicate whether E-AGV k enters the charging/working/idling mode at time t; $P_{k,t}^{evc}$ and $P_{k,t}^{evd}$ denote the charging/discharging power for E-AGV k at time t; P_t^{essc} and P_t^{essd} denote the charging/discharging power for the ESS at time t; and $P_t^{gt,g}$ denotes the gas consumption power decision for the GT.

3) Reward function: The goal of the MARL agents is to maximize the cumulative reward. To align with the original problem's objective of minimizing total operational costs, the instantaneous reward is defined as the negative of intraday operational costs (F_t^{id}) and penalty costs (F_t^{pe}).

$$R_t = -\sum\nolimits_t (F_t^{id} + F_t^{pe}) \tag{3}$$
$$F_t^{pe} = C_t^{im} + C_t^{un} + C_t^{end} \tag{4}$$
$$C_t^{im} = \max(0, P_t^{load} - P_t^{supply}) \tag{5}$$
$$C_t^{un} = \sum_{s\in S} \mathbb{I}(t > t_s^d \text{ and } TEU_{s,t}^{left} > 0) \tag{6}$$
$$C_t^{end} = \begin{cases} |E_t^{ess} - E_0^{ess}| & \text{, if } t = T \\ 0 & \text{, otherwise} \end{cases} \tag{7}$$

Where F_t^{pe} denotes the total penalty cost incurred at time t; C_t^{im} denotes the power imbalance penalty, which is positive when the total power supply P_t^{supply} fails to meet the total power load P_t^{load}; C_t^{un} denotes the task non-fulfillment penalty, which is applied at every time step t for each ES s that has exceeded its latest departure time but still has unprocessed containers; $\mathbb{I}(\cdot)$ denotes the indicator function, which equals 1 if the condition is true, and 0 otherwise; C_t^{end} denotes the end-state penalty, which is applied only at the final time step if the ESS charge level does not match its initial state (E_0^{ess}).

4) State transition function: In this model, the transition process is a combination of deterministic components (driven by agent actions) and stochastic components (driven by environmental uncertainty).

$$TEU_{s,t+1}^{left} = TEU_{s,t}^{left} - \eta^q \sum\nolimits_q \delta_{s,q,t} \Delta t \tag{8}$$

The deterministic transitions follow physical and operational rules, representing the direct consequences of the joint actions A_t. On the logistics side, the remaining container volume of ES s decreases from its state at time t based on the handling volume of the QC actions, as shown in Eq. 8. The detailed state transition formulas for all other variables, including the container inventory in the yard and the energy levels of the ESs, E-AGVs, and ESS, are modeled according to their physical operational characteristics as detailed in [17].

The stochastic transitions represent external events beyond control: the actual RES generation at time $t+1$ is a random variable, sampled by the environment from the generation state at time t according to the Markov process; meanwhile, the operational state of ESs also stochastically transition from unarrived to waiting based on its arrival probability.

B. Implementation of MARL

To address these challenges, we have designed a two-layer DRL framework based on the centralized training with decentralized execution paradigm. This framework decouples the problem into two hierarchical levels: an upper logistics layer responsible for handling the discrete and variable logistics decisions, and a lower energy layer responsible for managing the continuous energy balancing decisions.

1) Logistics agent: The upper layer is composed of all logistics agents responsible for physical operations, including ES agents, QC agents, YC agents, E-AGV agents. The primary task of this layer is to coordinate the discrete actions of these heterogeneous agents. To achieve this objective, we employ a MARL algorithm based on the "Centralized Training with Decentralized Execution" (CTDE) paradigm. In this framework, each logistics agent i has its own individual Q-network (with parameters θ_i), which estimates the value of its individual action based on its local observation.

During the training phase, a centralized mixing network (with parameters θ_m) non-linearly combines the individual values from all agents to estimate the joint action total value. The principles of DDQN are used to update this centralized Q_{to} to mitigate the overestimation problem.

The target Q-value calculation is decoupled: the current policy networks are responsible for selecting the optimal joint action in the next state, while the target networks are responsible for evaluating that action's value:

$$Y_t^{to} = R_{t+1} + \gamma Q'_{to}(S_{t+1}, \arg\max_{\mathbf{a}} Q_{to}(S_{t+1}, \mathbf{a}; \theta_{all}); \theta'_{all}) \tag{9}$$

where R_{t+1} is the shared global reward; γ is the discount factor; Q_{to} and Q'_{to} are the centralized policy and target networks, respectively; $\mathbf{a} = \{a_{i,t}\}_{\forall i}$ is the joint action of all logistics agents. The centralized loss function $L(\theta_{all})$ is the Mean Squared Error (MSE) between the target Y_t^{to} and the currently predicted total value:

$$L(\theta_{all}) = \mathbb{E}\left[\left(Y_t^{to} - Q_{to}(S_t, \mathbf{a}_t; \theta_{all})\right)^2\right] \tag{10}$$

2) Energy agent: The lower layer consists of all agents responsible for real-time energy balancing, which includes not only the ESS agent, GT agent and grid agent, but also must include the ES agents and E-AGV agents (their respective continuous power decision components). To this end, the Proximal Policy Optimization (PPO) algorithm is employed for training. PPO is an algorithm based on the Actor-Critic framework, which introduces a clipping mechanism during the actor (policy) network update to ensure the stability of policy updates. The critic network (with parameters θ) is responsible for evaluating the state value $V(S_t)$. It is updated by minimizing the MSE loss $L^V(\theta)$.

$$L^V(\theta) = \hat{\mathbb{E}}_t\left[\left(V(S_t; \theta) - G_t^{target}\right)^2\right] \tag{11}$$

Where $V(S_t; \theta)$ is the critic's value prediction for S_t, and G_t^{target} is the actual return sampled from the environment.

The actor network (with parameters ϕ) is responsible for outputting the continuous action. The actor's update relies on the advantage function $\hat{A}_t$, which represents how much better the current action A_t is compared to the average. The advantage function is typically calculated using generalized advantage estimation (GAE), which relies on the temporal difference error δ_t:

$$\delta_t = R_{t+1} + \gamma V(S_{t+1}) - V(S_t) \tag{12}$$

$$\hat{A}_t = \sum\nolimits_{l=0} (\gamma\lambda)^l \delta_{t+l} \tag{13}$$

Where λ is the balancing parameter. The actor network is updated by maximizing a clipped surrogate objective function:

$$L^{clip}(\phi) = \hat{\mathbb{E}}_t\left[\min\left(r_t(\phi)\hat{A}_t, \text{clip}(r_t(\phi), 1-\epsilon, 1+\epsilon)\hat{A}_t\right)\right] \tag{14}$$

where $r_t(\phi)$ is the probability ratio between the new policy π_ϕ and the old policy $\pi_{\phi_{old}}$; and ϵ is the clipping coefficient.

IV. Case Study

The case study in this paper uses actual data from the Port of Ningbo. The logistics system comprises 5 berths, 25 QCs, 60 E-AGVs, and 30 YCs. The operational parameters for the ESs and E-AGVs are set according to current practices. Key techno-economic parameters, vessel arrival and cargo demand data, RES generation profiles, and the Time-of-Use (ToU) electricity tariffs are provided in [17]. To validate its performance, the proposed MARL method is compared against two benchmark algorithms: a Random Dispatch strategy and a Manual Heuristic algorithm. The Random Dispatch strategy simulates non-optimized, random assignments and serves as the lower performance bound. The Manual Heuristic, in contrast, is a sophisticated scheduling method based on practical port supply-demand rules, representing an advanced and mature solution. As shown in Figure 1, the "Random Dispatch" and "Manual Heuristic" algorithms provide constant performance baselines. The proposed MARL method, after an initial exploration phase ($\leq$ 250 iterations) where it underperforms, rapidly learns and surpasses the heuristic benchmark. Throughout the middle and late stages of training, the MARL method consistently outperforms both benchmarks and converges to a significantly higher reward level, validating its superior optimization capability.

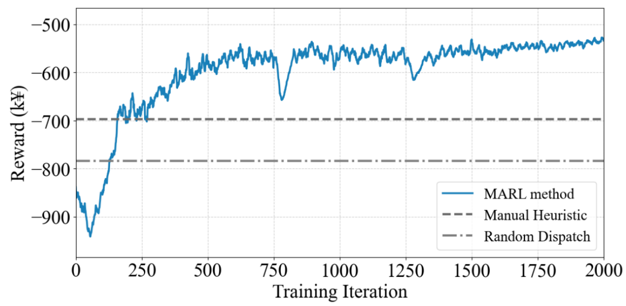
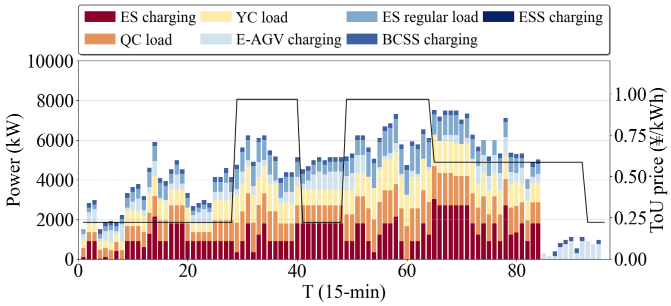


Fig. 1: Convergence performance comparison.

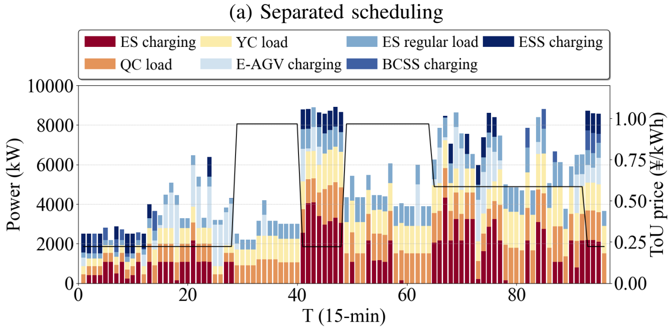
To validate the effectiveness of the proposed coordinated scheduling method, this paper compares two scheduling mechanisms: 1) **Separated Scheduling**, where the logistics system is first optimized independently, and the resulting logistics schedule is then used as an input for the energy system optimization; and 2) **Coordinated Scheduling**, where the logistics and energy systems are co-optimized using the DRL method proposed in this paper. Figures 2 illustrates the power load dispatch results under the different modes, while Table I compares the total costs between the two strategies.

As shown in Figure 2a, the separated scheduling mechanism generates a plan that prioritizes logistical efficiency. This means the operations of the ESs and QCs are determined first, and ESs tend to dock upon arrival. The resulting schedule demonstrates that load timing is completely decoupled from electricity prices, leading to high power consumption even during peak price periods. In this mode, the energy system can only passively accept the fixed load profile, making effective load shifting impossible. In contrast, Figure 2b shows the results of the integrated logistics-energy optimization. This mechanism allows the model to optimize the trade-off between logistical efficiency and energy cost. The model identifies high-price periods and proactively shifts energy-intensive operations away from these intervals, forming a distinct load valley. These loads are transferred to low-price periods, thereby achieving peak shaving and valley filling.

As shown in Table I, the "Separated" scheduling approach primarily minimizes logistics and rescheduling costs but lacks coordination with the energy system. This strategy results in substantial energy-related expenditures, including energy procurement (543.2 k¥), carbon emissions (26.5 k¥), and RES curtailment (20.6 k¥). In contrast, the "Coordinated" scheduling method, through joint logistics-energy optimization, achieves a significant 17.42% reduction in total operational cost, despite a slight compromise on logistics and rescheduling costs. This saving is primarily attributed to effective control of energy-related costs, most notably the complete elimination of RES curtailment (a 100% reduction) and a 74.34% reduction in carbon emission costs.



(a) Separated scheduling



(b) Coordinated scheduling

Fig. 2: Comparison of electric power load dispatch scheduling

TABLE I: Comparison of total operation costs.

	Logistics cost (k¥)	Reschedule cost (k¥)	Energy cost (k¥)	Carbon emission cost (k¥)	RES curtailment cost (k¥)	Total cost (k¥)
Separated	20.0	16.0	543.2	26.5	20.6	**626.3**
Coordinated	58.0	65.0	387.4	6.8	0.0	**517.2**

V. Conclusion

This paper addresses the tight coupling of logistics-energy systems in electrified seaports by proposing a two-layer DRL framework. This method aims to overcome the high costs, high emissions, and RES curtailment caused by traditional ”separated scheduling,” which ignores coupled dynamics and uncertainties. To tackle the challenges of a vast state space and a mixed (discrete-continuous) action space in coordinated scheduling, the framework employs DDQN for upper-layer logistics decisions and PPO for lower-layer energy decisions. A case study based on Port of Ningbo data demonstrates that the proposed coordinated method achieves a 17.42% reduction in total operational cost. By enabling effective ”load shifting” and optimizing internal asset dispatch , this approach also eliminates 100% of RES curtailment and reduces carbon emission costs by 74.34%. The DRL method provides an effective tool for ports to achieve economic, low-carbon, and adaptive operations.

References

[1] Z. Zhang, C. Song, J. Zhang, Z. Chen, M. Liu, F. Aziz, T. A. Kurniawan, and P.-S. Yap, “Digitalization and innovation in green ports: A review of current issues, contributions and the way forward in promoting sustainable ports and maritime logistics,” *Science of the Total Environment*, vol. 912, p. 169075, 2024.

[2] Y. Shu, B. Han, L. Song, T. Yan, L. Gan, Y. Zhu, and C. Zheng, “Analyzing the spatio-temporal correlation between tide and shipping behavior at estuarine port for energy-saving purposes,” *Applied Energy*, vol. 367, p. 123382, 2024.

[3] S.-C. Chang, M.-H. Lin, and J.-F. Tsai, “An optimization approach to berth allocation problems,” *Mathematics*, vol. 12, no. 5, p. 753, 2024.

[4] S. Aslam, M. P. Michaelides, and H. Herodotou, “Optimizing multi-quay combined berth and quay crane allocation using computational intelligence,” *Journal of Marine Science and Engineering*, vol. 12, no. 9, p. 1567, 2024.

[5] J. He, Y. Wang, C. Tan, and H. Yu, “Modeling berth allocation and quay crane assignment considering qc driver cost and operating efficiency,” *Advanced Engineering Informatics*, vol. 47, p. 101252, 2021.

[6] S. Li, L. Fan, and S. Jia, “A hierarchical solution framework for dynamic and conflict-free agv scheduling in an automated container terminal,” *Transportation Research Part C: Emerging Technologies*, vol. 165, p. 104724, 2024.

[7] X. Yang, H. Hu, J. Jin, and N. Luo, “Joint optimization of space allocation and yard crane deployment in container terminal under uncertain demand,” *Computers & Industrial Engineering*, vol. 172, p. 108425, 2022.

[8] M. Kermani, E. Shirdare, G. Parise, M. Bongiorno, and L. Martirano, “A comprehensive technoeconomic solution for demand control in ports: Energy storage systems integration,” *IEEE Transactions on Industry Applications*, vol. 58, no. 2, pp. 1592–1601, 2022.

[9] P. Wang, J. Dou, W. Su, Z. Jiang, and Y. Shi, “An energy management strategy based on dynamic programming for fuel cell hybrid trucks in ports,” *International Journal of Hydrogen Energy*, vol. 82, pp. 123–133, 2024.

[10] A. Mo, Y. Zhang, Y. Xiong, F. Ma, and L. Sun, “Energy–logistics cooperative optimization for a port-integrated energy system,” *Mathematics*, vol. 12, no. 12, p. 1917, 2024.

[11] J. Yin, Q. Huo, J. He, X. Peng, and T. Wei, “Energy management and capacity allocation method of hybrid energy storage system based on port transportation-energy coupling characteristics,” *Journal of Energy Storage*, vol. 83, p. 110745, 2024.

[12] J. He, M. Huang, Y. Wang, Y. Zhang, and H. Yu, “A holistic approach to resilient port management: Synergizing space template and crane deployment via benders decomposition algorithm,” *Ocean & Coastal Management*, vol. 269, p. 107806, 2025.

[13] E. Lalla-Ruiz, X. Shi, and S. Voß, “The waterway ship scheduling problem,” *Transportation Research Part D: Transport and Environment*, vol. 60, pp. 191–209, 2018.

[14] G. Wu, Y. Li, C. Jiang, C. Wang, J. Guo, and R. Cheng, “Multi-vessels collision avoidance strategy for autonomous surface vehicles based on genetic algorithm in congested port environment,” *Brodogradnja: An International Journal of Naval Architecture and Ocean Engineering for Research and Development*, vol. 73, no. 3, pp. 69–91, 2022.

[15] T. Ai, L. Huang, R. Song, H. Huang, F. Jiao, and W. Ma, “An improved deep reinforcement learning approach: A case study for optimisation of berth and yard scheduling for bulk cargo terminal.” *Advances in Production Engineering & Management*, vol. 18, no. 3, 2023.

[16] T. Song, L. Fu, L. Zhong, Y. Fan, and Q. Shang, “Two-layer deep reinforcement learning based port energy management strategy considering transportation-energy coupling characteristics,” *Energy*, vol. 318, p. 134659, 2025.

[17] W. Lv, Y. Ye, T. Cui, S. Chen, D. Xu, W. Yu, D. Huang, Z. Liu, J. Zhu, T. Li *et al.*, “Sustainable electrified seaports: A coordinated energy and logistics scheduling approach for future maritime hubs,” *Applied Energy*, vol. 401, p. 126645, 2025.

979-8-3315-8138-1