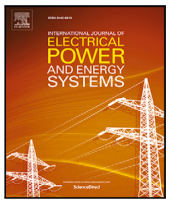
Electrical Power and Energy Systems 173 (2025) 111437

Contents lists available at ScienceDirect

International Journal of Electrical Power and Energy Systems

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



Optimizing green port operations with reinforcement learning and model predictive control

Francesco Conte [a], Gianluca Natrella [b,*], Daniele Sasso [a], Federico Silvestro [b], Fabio D'Agostino [b], Federico D'Antoni [a,c], Mario Merone [a]

[a] Faculty of Engineering, Campus Bio-Medico University of Rome, Via Alvaro del Portillo 21, Rome, 00128, Italy
[b] DITEN, University of Genoa, Via all'Opera Pia 11a, Genoa, 16145, Italy
[c] Fondazione Policlinico Universitario Campus Bio-Medico, Via Alvaro del Portillo 200, Rome, 00128, Italy

ARTICLE INFO

Keywords:
Smart port
Cold ironing
Neural networks
Soft actor critic
Model predictive control

ABSTRACT

While playing a crucial role in global trade, seaports are polluting hubs. To mitigate greenhouse gas emissions from ships at berth, the present study incorporates a combination of solutions to meet ships' demand, including shore-side electricity (Cold ironing), multi-energy storage, renewable generation, and artificial intelligence techniques for forecasting and optimizing renewable energy usage.

Key contributions include the sizing of a port based on real ship traffic data and area from the ports of Genova and Trapani, the comparison of reinforcement learning and model predictive control for effective energy management, and the development of artificial intelligence-driven forecasting models for building demand, and photovoltaic and wind generation. Long Short-Term Memory and Time Delay Neural Networks are used to predict renewable generation and demand, while sensitivity analyses demonstrate the flexibility and efficiency of the Soft Actor Critic algorithm in handling forecast uncertainties. The hyperparameters of the artificial intelligence models are optimized on the data of the port of Genova, and tested on both the ports.

Results highlight the strength of the developed control algorithms to achieve diesel usage reduction at berth up to 100%. Numerical findings prove the model predictive control ability to extract knowledge from input data and demonstrate the advantages of reinforcement learning in optimizing energy management under limited and uncertain information. The model predictive control and the Soft Actor Critic algorithms were able to achieve 95.9% and 94.4%, respectively, of the maximum achievable cash flow in the port of Genova, and 91.3% and 89.3%, respectively, in the port of Trapani.

1. Introduction

The maritime sector accounts for more than 90% of global trade [1], representing the backbone of global economy [2] while contributing to around 3%–5% of global greenhouse gas emissions [3]. However, pollution from ships is concentrated near the ports and inside the ports, due to maneuvering and hoteling [4], contributing to the pollution of coastal cities. Indeed, port regions have been proven to pollute more than non-port regions, especially in terms of greenhouse gas emissions [5]. To this extent, through the regulation 2023/1804/EU [6], the European Union requires the use of shore-side electricity in ports, known as Cold Ironing (CI), to ensure, by the end of 2029, the electric connection of the ships to the port electricity distribution network, thus turning off ships generation units at berth. CI can provide significant environmental benefits, indeed, [7] demonstrates that CO_2 emissions can be reduced by CI in the range of 48%–70%.

1.1. Background

The development of Smart Ports (SPs), characterized by advanced computational processes and integration of cutting-edge technologies, such as CI, AI and RESs, has increasingly contributed to the planning and construction of environmentally friendly and low-carbon ports [2].

Over the last few years, CI has been studied at various levels to overcome some technical and economic issues. In [8], the authors present a procedure for short-circuit currents computation in the case of low-voltage shore power connection systems where multiple ships are allowed to operate simultaneously on one MV/LV transformer. An analysis of the behaviour of the high-voltage shore connection system in the case of single phase-to-ground fault and three-phase short circuit is presented in [9]. The aspects related to reliability of CI have been addressed in [10], which investigates the power system reliability

* Corresponding author.
E-mail address: gianluca.natrella@unige.it (G. Natrella).

https://doi.org/10.1016/j.ijepes.2025.111437
Received 17 June 2025; Received in revised form 27 November 2025; Accepted 30 November 2025
Available online 10 December 2025

considering the forced outage rate of ship electric generators and shore connection. In [11] the study presents a CI control scheme capable to maintain power quality. In [12] a high power charging network and its control mechanism are developed for CI.

Currently, the shipping sector is exploring new alternatives to fossil fuels other than batteries. Among several possibilities, such as ammonia and methanol, hydrogen is the only one that is completely zero-emission when produced with renewable energy [13]. Several hydrogen ships, powered by fuel cells [14], have already been developed and are operating worldwide, such as hydrogen-fuelled passenger ferries [15], and tanker and container ships [16]. Moreover, beyond its use as fuel for ships, hydrogen is investigated as an effective energy vector for storing energy in ports. Indeed, an increasing number of worldwide-spread projects addressing the development of infrastructures dedicated to hydrogen production, storage, distribution, and use for ports decarbonization is reported. Some virtuous examples are the Port of Rotterdam, the Port of Valencia, and the Port of Hamburg [15].

The integration of RESs in ports enhances the environmental benefits of CI, contributing to abate greenhouse gas emissions [17] and shaping the port as a smart grid. A distributed power-sharing control based on adaptive virtual impedance in a port, with RES performing CI, is presented in [18]. An adaptive Direct Load Control approach, integrated with an IoT-based architecture is presented in [19] to optimize day-ahead scheduling and real-time load shedding to enhance consumer comfort while reducing electricity costs. An optimal power scheduling framework for port microgrids, integrating various utility loads, thermal loads and CI is proposed in [20]. A mixed-integer linear programming model is proposed in [21] to address integrated operations planning and energy management in ports with smart grid infrastructure, considering uncertain renewable generation. An optimal stochastic control strategy of Battery Energy Storage System (BESS) and hydrogen storage to manage CI and ZE-Ship hoteling operations in a SP with RESs is presented in [22]. A similar framework is considered in [23] with the objective of optimally sizing the distributed energy resources and the infrastructure of a port.

Recent studies have highlighted the potential of AI, and RL in particular, in modernizing port operations exploiting AI's ability to extract knowledge from large sets of data. Adopting AI enables ports to optimize energy management and respond effectively to climate change by reducing their reliance on fossil fuels and transitioning towards zero-emission operations, thereby enhancing sustainability and efficiency [24]. The study in [25] investigates the use of various data-driven models, including artificial neural networks, for accurate prediction of ship berthing times. These models support the optimization of berth allocation and energy scheduling, enhancing the operational efficiency of ports and reducing environmental impact. In [26], the importance of machine learning in developing efficient Ship Energy Efficiency Management Plans is highlighted. In particular, machine learning techniques are employed to analyse ship activities at container ports, identifying key variables influencing CO_2 emissions. Regarding the utilization of RL for port management, [27] investigates the application of multi-objective deep RL in power scheduling for emission-free ships, not only ensuring zero emissions but also improving the reliability of power systems in all-electric ships, demonstrating the practical application of AI in managing complex and dynamic systems in maritime settings. In [28], an energy management system that uses Double Q RL for shipboard plug-in hybrid fuel cell and battery propulsion systems is proposed, where the policy generated by the RL agent is used directly by ships as a guide strategy to achieve long-term cost performance without prior knowledge of future power loads or demands.

1.2. Motivations and contributions

Recent advances in researches comparing the performance of MPC and RL controllers have been made mainly in the field of thermal energy management. In [29], an MPC algorithm and several RL control algorithms are developed to control the heat pump in a residential building. The analysis focuses on the ability of the control algorithms in reducing the electricity bill and the thermal discomfort, and the results revealed the superiority of the MPC. In [30], RL control algorithms and MPC are compared for energy management of a cold-water buffer tank and the best performances were obtained with a RL controller. Recent contributions have investigated hybrid control schemes that combine RL with predictive control in power electronics. In [31] an architecture has been proposed that combines RL and predictive control to reduce switching frequency and improve robustness in power converters. Similarly, an online learning integrating MPC has been developed in [32] for real-time control of power converters.

However, important issues remain unsolved, *i.e.*, existing comparative analyses on MPC and RL control algorithms have not investigated complex energy management problems, and the analyses are limited to the performance of the algorithms and do not encompass other key features, such as inputs uncertainty and extent of knowledge. Recent studies have not addressed a comparison between the energy management in SP operations achieved with conventional control techniques such as MPC, and with RL. Moreover, a detailed analysis of the impact of using AI-generated predictions for future energy loads and generation is not present. For these reasons, this study aims to answer the following Research Questions (RQs):

RQ1: what are the economical and operational impacts of using AI predictions versus not using predictions?
RQ2: how do MPC-based and RL algorithms perform in optimizing the energy management of an SP?

Specifically, a detailed case study is conducted adopting RL and MPC under various scenarios to determine the potential gain in economic performance when utilizing AI-based predictions. Unlike previous AI/RL-based approaches that optimize ship-to-grid connection with a focus on CO_2 reduction, our work addresses the broader port system, defining optimal operations from an economic perspective. Furthermore, while previous studies primarily evaluated performance metrics, our study explicitly examined the role of uncertainty in inputs and the extent of knowledge required by each approach.

To summarize, this paper puts forward the following contributions:

- Scope expansion: the proposed framework addresses SP operations as an integrated system, rather than focusing on isolated subsystems, it incorporates RESs, and dynamic demands from buildings and ships;
- AI-prediction integration: the analysis considers the economic and operational impacts of integrating AI-based forecasts for buildings energy demands and PV and wind generation, and evaluates the interaction between predictive accuracy and control performance;
- Data uncertainty and availability: the effect of input uncertainty and the extent of model knowledge on the relative performance of MPC and RL controllers is examined, the performances of several RL algorithms are evaluated and compared with two different MPC algorithms that are suitably developed to handle different amount of data;
- Sensitivity and scenarios analysis: robustness and sensitivity analyses are conducted to assess the influence of forecast errors, control parameters, and extent of available information;
- Real-world grounding: the presented case studies are based on actual ship data and port area characteristics from two ports with different seasonality and purposes, ensuring high relevance and practical implementation.

The remainder of this paper is organized as follows: Section 2 describes the framework of the study, detailing the case study and the data, Section 3 illustrates the adopted methodology to manage the port operations, Section 4 presents the results, and Section 5 summarizes the findings of this paper.

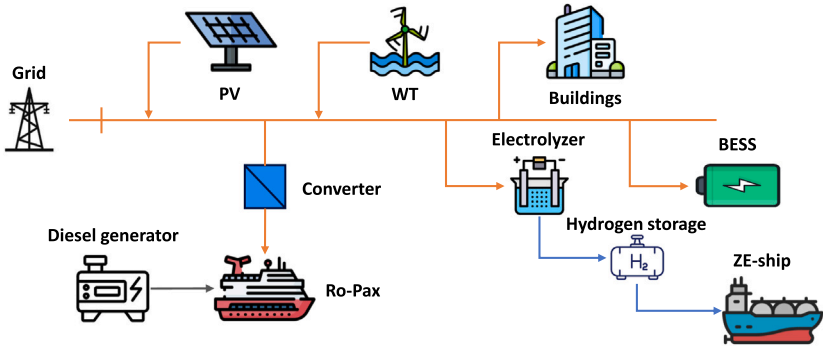


Fig. 1. Smart port architecture.

Table 1
Energy prices.

$c_t^i = 158.4$ €/MWh	$c_t^c = 160$ €/MWh	$c_t^{ci} = 165$ €/MWh	$c_t^h = 165$ €/MWh	$c_t^{dg} = 1.85$ €/L
		c_t^e		
Day/h	0–7	7–8	8–19	19–23	23–midnight
Mon–Fri	51.97 €/MWh	62.23 €/MWh	75.83 €/MWh	62.23 €/MWh	51.97 €/MWh
Sat	51.97 €/MWh	62.23 €/MWh	62.23 €/MWh	62.23 €/MWh	51.97 €/MWh
Sun	51.97 €/MWh	51.97 €/MWh	51.97 €/MWh	51.97 €/MWh	51.97 €/MWh

2. System framework

In this section, the SP set-up is outlined, and the data are presented. The case of the port of Genova (Italy) and the case of the port of Trapani (Italy) are analysed, however the overall configuration of the SP is the same for the two cases. Fig. 1 shows the architecture of the SP. The SP includes multi-energy storage, *i.e.*, a BESS, and an Electrolyzer (Ely) loading a tank and serving ZE-Ships for berth operations. RES generation is present as PV rooftop panels and Wind Turbines (WTs).

The SP is meant to operate CI, allowing the ships at berth to connect to the utility grid through a converter. However, while ZE-Ships must be fed by the Ely, the ships operating CI have their on-board Diesel Generators (DGs) to possibly satisfy their demand.

The yearly electrical demand at the shore-connection was defined based on [33], which provides an estimation of the total electrical load of Ro-Pax ships moored at a quay in the port of Genova and in the port of Trapani, in 2019, and of the yearly demand of hydrogen of ZE-Ships, identified in the categories of chemical and oil tankers. To compute the hydrogen consumption from the energy demand, the equivalence $1\,\mathrm{MW\,h} = 50\,\mathrm{kg}_{H_2}$ was used, which is based on hydrogen lower heating value [34] and considers that ZE-Ships are equipped with on-board fuel cells with an efficiency of 60% [35]. It is assumed that only one Ro-Pax ship and one ZE-Ship at a time can be moored at the corresponding quay.

The architecture of the SP also considers the presence of office buildings, simulating this load using historic data from a university office building located outside the port of Genova and Trapani.

The SP exchanges electrical energy with the utility grid and with ships operating CI. Moreover, it supplies hydrogen to ZE-Ships. Therefore, the study assumes that energy export is remunerated at c^e, energy import is paid at c^i, hydrogen sold to ZE-Ships is remunerated at c^h, and the energy consumed in CI operations is remunerated at c^{ci}. The use of DGs is supposed to cost c^{dg} to Ro-Paxs. Finally, to promote RES utilization, the port is penalized by c^c for the curtailment of RES generation.

2.1. Data

The yearly demand of Ro-Paxs and ZE-Ships in the ports of Genova and Trapani and the corresponding load duration curves are shown in Fig. 2. It is possible to observe how only one ship at a time is moored at each berth, as stated in Section 2, shaping the energy demand profiles as a series of steps. The building demand, shown in Fig. 3, is lower than the ZE-Ships and Ro-Paxs demands, in both the considered ports. When comparing the demands of the ships in the two considered ports, the Ro-Paxs demand in the port of Genova is consistently larger, both in maximum and mean value, while the ZE-Ships demand is larger in terms of maximum value but smaller in terms of mean value. The differences in the load duration curves of the ships in the two ports highlight the different seasonality and characteristics of the two considered ports. Specifically, as a port on an island and far from the mainland coast, the port of Trapani is characterized by low Ro-Paxs traffic, with peaks in the summer, and by a stable tanker ships traffic, *i.e.*, ZE-Ships.

To reproduce PV and wind generation, the software Hybrid Optimization of Multiple Energy Resources (HOMER) was fed with historical weather data from the location of the port of Genova and of the port of Trapani taken from [36].

The values of energy prices are shown in Table 1. The diesel cost c^{dg} is set to 1.85 €/L, which corresponds to 470 €/MWh considering an energy efficiency of the on-board DGs of 40%, a fuel density of 0.820 kg/L, and a lower heating value of 43.2 MJ/kg. Note that it is required $c^{ci} < c^{dg}$ for CI to be profitable for the Ro-Paxs, and also $c^{ci} > c^i$ for CI to be profitable for the SP.

2.2. System sizing

For both the port of Genova and the port of Trapani, the sizing of the SP devices was carried out using the software HOMER [37], which simulates the system for every possible combination of devices in a microgrid and calculates the LCOE in each simulation. The possible sizes of every device in the system and the selected ones are shown in Table 2.

For the case of the port of Genova, to establish the maximum size of the PV power plant, an available area of $123\,800\,\mathrm{m}^2$ to install the panels was considered, estimated by Authority of the port system in [38], which, with an expected efficiency of $0.1\,\mathrm{kW/m}^2$, returned a potential peak power of 12.38 MW. Given the smaller area of the port of Trapani, a maximum installable power of 3 MW was considered.

Similarly, a maximum of 20 Enercon E-48 [39] WTs, each one with a rated power of 800 kW and a maximum power of 810 kW, and a

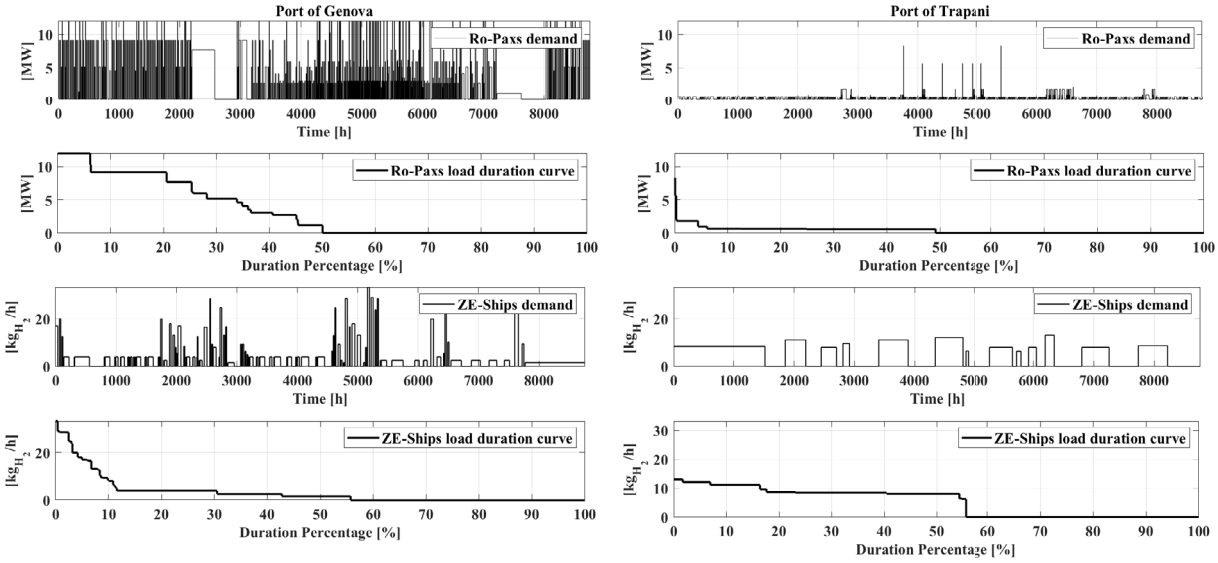


Fig. 2. Left column refers to the case of the port of Genova, right column refers to the case of the port of Trapani. From the top: Ro-Paxs energy demand, Ro-Paxs load duration curve in terms of energy demand, ZE-Ships hydrogen demand, ZE-Ships load duration curve in terms of hydrogen demand.

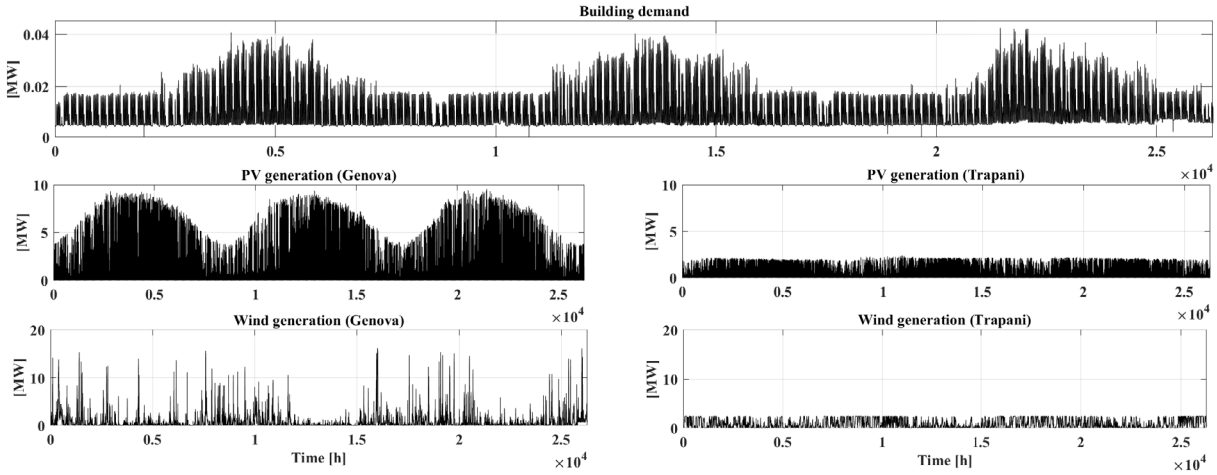


Fig. 3. Building energy demand (used for the ports of Genova and Trapani), PV generation in the port of Genova, PV generation in the port of Trapani, wind generation in the port of Genova, wind generation in the port of Trapani.

maximum of 35 WES 18/80 [40] WTs, each one with a rated power of 80 kW and a maximum power of 83.3 kW, were considered as upper limits for the wind power generation in the port of Genova. For the case of the port of Trapani, the upper limits for the E-48 and the WES-18 were set to 5 and 10, respectively. Additionally, the constraint that only one type of WT can be installed was considered when sizing the SPs.

The VRFB was considered as the BESS technology, because of its low investment cost and outstanding performance in stationary applications [41]. VRFBs are particularly useful in energy applications, when they are employed to provide energy arbitrage and support islanded operations [42]. Specifically, a 0.25C (fully charged/discharged in 4 h at nominal current) VRFB was considered.

The selected technology for the Ely was the proton exchange membrane, having the advantages of being a commercialized technology, operating at low temperatures, and showing better performance than the alkaline (another already commercialized technology), against a higher investment cost [43].

Given the input data, *i.e.*, annual profiles of Ro-Paxs, ZE-Ships and building demand, solar irradiation, and wind speed, HOMER computes the LCOE for every possible combination of devices sizes in the two considered ports, applying the load following dispatch strategy. Under the strategy adopted, whenever a generator is operating, it only produces enough power to meet the primary load, ensuring that the on-board DGs would only be used to meet Ro-Paxs demand. The final system configuration, reported in Table 2, is the one returning the lowest LCOE. Note that, in simulating the system, HOMER has complete and full knowledge of the inputs from the beginning along the whole simulation.

In optimizing the sizes of the devices, a discount rate of 6%, an inflation rate of 2%, and a project lifetime of 25 years were considered as optimization parameters in the HOMER software. The results show the lowest LCOE when 20 E-48 turbines, 0 WES-18 turbines, and 12.38 MW of PV are installed for the case of the port of Genova, and when 3 E-48 turbines, 0 WES-18 turbines, and 2.5 MW of PV are installed for the case of the port of Trapani. Given the sizes of the RESs, the wind and PV generations are shown in Fig. 3. The optimal sizes of the remaining devices are shown in Table 2.

Moreover, for the case of the port of Genova, the optimal solution contemplates the installation of the maximum RES generation and of the converter to operate CI, while for the port of Trapani, the total

Table 2
Smart ports devices characteristics and sizes. Top section: the first column lists the devices, the second column lists the search space for the sizes of each device in the port of Genova, the third column lists the search space for the sizes of each device in the port of Trapani, the fourth column lists the device installation cost, and the fifth column lists the device expected lifetime. Bottom section: the first column lists the devices, the second column lists the sizes of the devices minimizing the LCOE in the port of Genova, the third column lists the sizes of the devices minimizing the LCOE in the port of Trapani.

Device	Size (Genova)	Size (Trapani)	Installation cost	Replacement cost	Lifetime (years)
Grid Connection	[12, 14, 16] MW	[7 – 9] MW	2000 €/kW	–	–
Converter	[0 – 12] MW	[0 – 8.5] MW	200 €/MW	200 €/MW	15
PV	[0 – 12.38] MW	[0 – 3] MW	1000 €/MW	750 €/MW	25
E-48	[0 – 20] units	[0 – 5] units	800 000 €/unit	600 000 €/unit	20
WES-18	[0 – 35] units	[0 – 10] units	96 000 €/unit	72 000 €/unit	20
VRFB	[4, 6, 8, 10, 12] MWh	[0, 0.4, 0.8, …, 4] MWh	36 530 €/MWh	27500 €/MWh	25
Ely	[0.5 – 1] MW	[0.5 – 1] MW	1000 €/kW	750 €/kW	15
Hydrogen Tank	[2.5, 3, 3.5] ton	[2.5, 3, 3.5] ton	205 €/kg	205 €/kg	25

Results of the sizing

Device	Genova	Trapani
Grid Connection	14 MW	7 MW
Converter	12 MW	8.5 MW
PV	2.38 MW	2.5 MW
E-48	20 units	3 units
WES-18	0 units	0 units
VRFB	4 MWh	1.6 MWh
Ely	0.75 MW	0.825 MW
Hydrogen Tank	3.5 ton	3 ton

installed RES generation is lower than the maximum allowable, but the size of the installed converter is the largest possible.

Finally, in the port of Genova, the installed battery is a 1 MW/4 MWh VRFB, the smallest of the possible sizes, while the installed Ely power is 0.75 MW, coupled with the largest hydrogen tank size, *i.e.*, 3.5 ton. Differently, in the port of Trapani, the installed battery is a 0.4 MW/1.6 MWh VRFB, the rated power of the installed Ely is 0.825 MW, and the hydrogen tank size is 3 ton.

3. Methodology

This section presents the equations modelling the SP described in Section 2, the energy forecasting algorithms, and the MPC-based and RL control algorithms to manage the SP operations. Fig. 4 displays the block diagram of the proposed methodology.

3.1. Port model

In this subsection, the models adopted for each system component are presented. In all the following equations, t indicates the discrete-time with a sampling time Δ=1 h and the reported powers are considered as mean values within the sampling time interval.

3.1.1. Connection with the grid

During the timestamp t, the SP can import power P_t^i or export power P_t^e. Therefore, it results that:

$$0 \leq P_t^i \leq \delta_t^g P_{max}^g, \tag{1}$$

$$0 \leq P_t^e \leq \left(1 - \delta_t^g\right) P_{max}^g, \tag{2}$$

where P_{max}^g is the rated power of the connection, *i.e.*, 14 MW, as reported in Table 2, and δ_t^g is a binary variable.

3.1.2. Ro-Paxs

The Ro-Paxs demand is indicated with P_t^s. The power generated by on-board DGs is indicated with P_t^{dg} and it must satisfy the following constraint:

$$0 \leq P_t^{dg} \leq P_t^s. \tag{3}$$

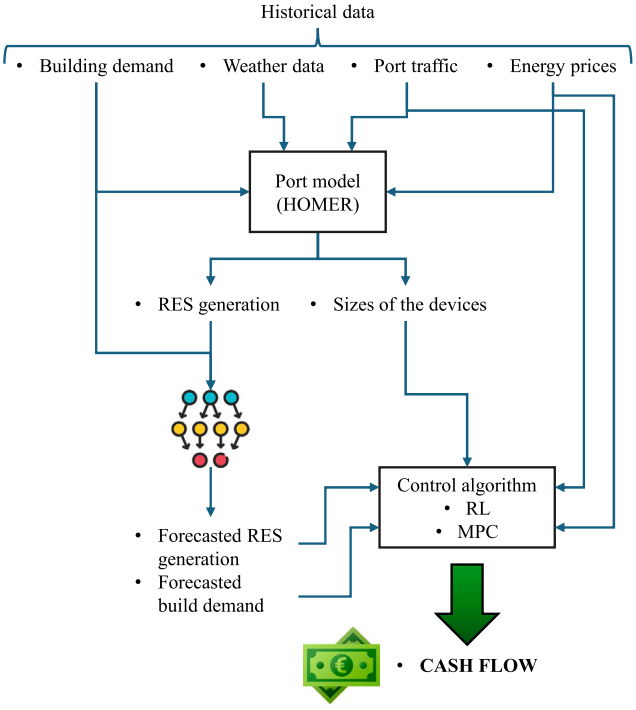


Fig. 4. Block diagram of the proposed methodology. Historical weather, port traffic and energy prices data are fed to the port model in HOMER which returns the sizes of the devices in the port and the profile of renewable generation. Forecasted renewable generation is obtained from the profile of renewable generation, and used as input to the control algorithms, together with the sizes of the devices and the historical port traffic and energy prices, which maximize the port cash flow.

3.1.3. RESs

The total power generated by RESs is indicated with P_t^{res}, whereas P_t^c indicates the curtailment, which must be such that:

$$0 \leq P_t^c \leq P_t^{res}. \tag{4}$$

3.1.4. VRFB

The VRFB can exchange power with the system by charging and discharging, according to its State of Charge (SoC). The following equations model the VRFB power exchange and the dynamic of its SoC:

$$0 \leq P_t^{ch} \leq P_{max}^b \delta_t^b, \tag{5}$$

$$0 \leq P_t^{ds} \leq P_{max}^b \left(1 - \delta_t^b\right), \tag{6}$$

$$S_{t+1} = S_t + \frac{\Delta}{E_t^b}\left(\eta^{ch} P_t^{ch} - \frac{P_t^{ds}}{\eta^{ds}}\right), \tag{7}$$

$$0 \leq S_t \leq 1, \tag{8}$$

$$E_{t+1}^b = E_{nom}^b - 0.8 E_{nom}^b \frac{N_{t+1}^c}{N_{ref}^c}, \tag{9}$$

$$N_{t+1}^c = N_t^c + \mid S_{t+1} - S_t \mid E_t^b, \tag{10}$$

where: P_t^{ch} and P_t^{ds} are the VRFB charging and discharging powers at timestamp t, respectively; $\eta^{ch} = 0.837$ and $\eta^{ds} = 0.837$ are the VRFB charging and discharging efficiencies, respectively, for a round trip efficiency $\eta^b = \eta^{ch} \cdot \eta^{ds} = 0.70$; P_{max}^b is the VRFB nominal power, *i.e.*, 1 MW; S_t [p.u.] is the battery SoC at timestamp t; E_t^b is the battery capacity at timestamp t; E_{nom}^b is the battery nominal capacity, *i.e.*, 4 MWh; N_t^c is the equivalent number of cycles computed by the battery from the beginning of its life to timestamp t; N_{ref}^c is the number of complete cycles that leads to battery end of life, thus $E_t^b = 0.8 E_{nom}^b$; δ_t^b is a binary variable ensuring that the VRFB does not charge and discharge during the same timestamp t.

3.1.5. Hydrogen generation and storage system

The hydrogen tank is fed by the Ely and discharged by the demand of ZE-Ship. The Ely has a technical operation minimum when it is producing hydrogen, equal to 10% of its rated power [44]. Therefore, the dynamic of the hydrogen storage system can be represented by the following equations:

$$P_{min}^{el} \delta_t^{el} \leq P_t^{el} \leq P_{max}^{el} \delta_t^{el}, \tag{11}$$

$$H_{t+1} = H_t + \frac{\Delta}{E^h}\left(\eta^{el} P_t^{el} - P_t^h\right), \tag{12}$$

$$0 \leq H_t \leq 1, \tag{13}$$

where: P_t^{el} is the power consumed by the Ely; P_t^h is the hydrogen required by the ZE-Ship; P_{max}^{el}, P_{min}^{el} are the Ely maximum and minimum power limits, *i.e.*, 0.75 MW and 0.075 MW, respectively; H_t [p.u.] is the LoH; E^h is the hydrogen tank capacity, *i.e.*, 3.5 ton; η^{el} is the efficiency of the Ely, *i.e.*, 0.65; δ_t^{el} is a binary variable indicating the on/off status of the Ely.

3.1.6. Power balance and operational costs

At every timestamp t, the following power balance needs to be matched:

$$P_t^{res} + P_t^i + P_t^{dg} + P_t^{ds} = P_t^s + P_t^e + P_t^{ch} + P_t^{el} + P_t^l + P_t^c, \tag{14}$$

where P_t^l is the building demand. The SP economic return at every timestamp t is:

$$J_t = \Delta\left(c_t^e P_t^e + c_t^{ci}(P_t^l - P_t^{dg}) + c_t^h P_t^h - c_t^i P_t^i - c_t^c P_t^c\right), \tag{15}$$

where at every timestamp t: c_t^e and c_t^i are the cost of the energy exported and imported to the grid, respectively; c_t^{ci} and c_t^h are the rates applied from the port to the Ro-Paxs for the CI service and to the ZE-Ships for the hydrogen bunkering, respectively; c_t^c is a penalty applied on the curtailed renewable energy.

3.2. Energy forecasting algorithms

In order for the MPC and the RL algorithms to work in an optimal manner, an accurate forecasting of some of the key variables is of paramount importance. To this end, three separate regression algorithms were implemented and optimized to forecast the future values of PV generation, wind generation, and power demand from an office building. Each forecasting algorithm adopted a time-series approach, where a sequence of input features was utilized to predict the future value of the respective variable.

In addition to past variable values, temporal information was integrated into each model. To forecast PV and wind energy generations, the models were enriched with hourly and yearly temporal features, while the building demand forecasting model incorporated temporal data related to the hour of the day and the day of the week. To accurately reflect temporal cycles, these temporal features were represented using sine and cosine functions of each temporal value, multiplied by 2π and divided by the appropriate number of instances (7, 24, or 365).

In predicting PV energy generation and building demand, the forecasting model structures identified in prior work [45] were leveraged, based on Time Delay Neural Network (TDNN) with a single hidden layer. These models were optimized from scratch in the current study, because the sampling rate for the features is different. Conversely, for the prediction of WTs energy generation, the performance of a Long Short-Term Memory (LSTM) neural network was explored in addition to the TDNN. TDNNs are renowned for their ability to capture the temporal context within sequential data while maintaining the training time short [46], while LSTMs excel in handling long-term dependencies and are particularly suited for catching short-term information for accurate forecasting [47].

Each model underwent training using an autoregressive approach, *i.e.*, during the training phase, it was trained solely to predict the next value of the variable. However, during the testing phase, each prediction was recursively fed back as an input, allowing the model to produce predictions with a variable prediction horizon, eliminating the need for training multiple separate models. This strategy allowed us to concentrate on fine-tuning a select set of hyperparameters, namely the input sequence size, number of hidden units, and learning rate.

Across all regression models, input sequences ranging from 1 to 36 h, a spectrum of hidden units from 20 to 260, and L2 regularization values spanning from 10^{-9} to 1 were explored. To optimize each model, a comprehensive grid search was conducted utilizing the first two years of available data, with the third year reserved for evaluating unbiased performance during the final testing phase. Specifically, the initial 80% of the two years of data served as a discovery set, while the remaining 20% functioned as a test set for selecting the optimal configuration. Furthermore, to mitigate overfitting, the discovery set was randomly divided in an 80%/20% ratio for training and validation, enabling early stopping measures.

3.3. Model predictive control-based algorithm

To optimally manage the SP operations, two MPC-based algorithms were developed. The first one is an MPC algorithm optimizing the port operations, coupled with a method to optimally schedule the minimum LoH needed to supply ZE-Ships looking at a time horizon longer than the one of the MPC. The second one consists of a slightly modified version of the same MPC algorithm, without the minimum LoH scheduling.

The MPC optimal algorithm is described in Section 3.3, introducing the issues calling for either the changes in its formulation, presented in Section 3.3.2, or the adoption of the minimum LoH scheduling problem, detailed in Section 3.3.3.

3.3.1. Basic MPC algorithm

The MPC technique is a well-renowned control technique that, for a discrete-time system, consists of solving at a given timestamp a suitably defined optimal control problem over a finite prediction horizon, employing inputs forecasts. Once the optimal control problem has been solved, an optimal control trajectory, spanning the whole prediction horizon length, is obtained, and the first element (the one related to the current timestamp) is applied to the system. Then, the same procedure is repeated at the following timestamp [48].

The objective of the MPC is to maximize the SP economic return, assuming that at the hour t, given a prediction horizon of length T, the following data are available:

- a forecast profile of the RES generation $\{\hat{P}^{res}_{t+k}\}^{T-1}_{k=0}$;
- a forecast profile of the building demand $\{\hat{P}^{l}_{t+k}\}^{T-1}_{k=0}$;
- a scheduled profile of ZE-Ships hydrogen demand $\{P^{h}_{t+k}\}^{T-1}_{k=0}$;
- a scheduled profile of Ro-Paxs energy demand $\{P^{s}_{t+k}\}^{T-1}_{k=0}$;
- the current SoC S_t and LoH H_t;
- all energy prices from time t to time $t+T-1$.

Thus, the optimization problem, at every hour t, is formulated as follows:

$$\max_{\{X_k\}} \sum_{k=0}^{T-1} J_{k+t} \qquad X_k = \left[P^i_k, P^e_k, P^{dg}_k, P^c_k, P^{ch}_k, P^{ds}_k, P^{el}_k\right]^\top \tag{16}$$

subject to: (1)–(8), (11)–(14).

To implement the MPC strategy, it is then required to know at every timestamp t the forecasts of the following $T-1$ timestamps of RES generation and building demand. Given the features to forecast, typical values of T are lower than 24 h. Much longer prediction horizons are unreliable and unrealistic. However, under these circumstances, problem (16) cannot be solved due to insufficient LoH. Indeed, given the step-shaped profile of the ZE-Ships demand, where no hydrogen is demanded for periods much longer than T, the MPC tends to empty the hydrogen tank, since no hydrogen demand is foreseen in the next T hours, and since operating the Ely is an energy consumption, thus a cost. Once the tank is empty and a ZE-Ship arrives, the Ely cannot meet the ship hydrogen demand, not even working at full load, and the constraint (13) is violated. Note that the violation of (13) does not mean that the system is poorly sized. Indeed, under the assumption of having, at the beginning of the simulation, the perfect knowledge of all inputs for the whole simulation period (HOMER sizing condition), no constraint would be violated.

3.3.2. Modified cost function

One way to solve the above-mentioned issue consists of rewarding high values of LoH, modifying problem (16) as follows:

$$\max_{\{X_k\}} \sum_{k=0}^{T-1} \left(J_{k+t} + wH_{k+t}\right) \qquad X_k = \left[P^i_k, P^e_k, P^{dg}_k, P^c_k, P^{ch}_k, P^{ds}_k, P^{el}_k\right]^\top \tag{17}$$

subject to: (1)–(8), (11)–(14); where w is a positive weight helping to raise the LoH. Note that, this solution is only effective if the controller prioritizes the hydrogen tank loading over the expense of importing energy from the grid. Therefore, it must be $w > c^i_t P^g_{max}$.

The change introduced in (17) ensures compliance with (13), but it compromises the purely economic nature of the cost function, which now differs from the SP operating profit. In the following, this algorithm will be addressed as MPC_m.

3.3.3. Minimum LoH scheduling

Another way to solve the issue on the violation of (13), while maintaining pure economic cost function, consists of scheduling at a given timestamp τ a profile of minimum LoH $\{\underline{H}_{\tau+k}\}^{HT-1}_{k=0}$ over a prediction horizon with length HT, ensuring that (13) would not be violated, and change (13) with:

$$\underline{H}_t \le H_t \le 1, \quad t \in [\tau, \tau+T-1]. \tag{18}$$

Note that the necessary condition to avoid constraint violation is $HT > T$. The profile of minimum LoH $\{\underline{H}_{\tau+k}\}^{HT-1}_{k=0}$ depends on the expected ZE-Ships hydrogen demand $\{\hat{P}^h_{t+k}\}^{HT-1}_{k=0}$, whose knowledge can be more or less accurate, depending on the length of the prediction horizon HT. Assuming a normally distributed zero-mean error $\varepsilon_{\tau+k} = P^h_{\tau+k} - \hat{P}^h_{\tau+k}$ with standard deviation σ^h_τ, the constraints (12)–(13) are probabilistic, and the standard deviation relative to the LoH is found as:

$$\sigma^h_{\tau+k} = \frac{\Delta}{E^h}\sqrt{\sum_{j=0}^{k-1}\left(\sigma^h_{\tau+j}\right)^2}. \tag{19}$$

For a generic normally distributed random variable, $x \sim \mathcal{N}\left(\hat{x}, \sigma^2\right)$, the chance constraint $\mathbf{P}\left(\underline{x} \le x \le \overline{x}\right) \ge 1-\beta$ can be written as:

$$\underline{x} + \theta\sigma \le \hat{x} \le \overline{x} - \theta\sigma, \tag{20}$$

where $\theta = \sqrt{2}\text{erf}^{-1}(1-2\beta)$ [49]. Therefore, the probabilistic version of (12)–(13), applied to the case of minimum LoH, are $\forall t \in [\tau, \tau+HT-1]$:

$$\hat{H}_{t+1} = \hat{H}_t + \frac{\Delta}{E^h}\left(\eta^{el}P^{el}_t - \hat{P}^h_t\right), \tag{21}$$

$$\theta\sigma^H_t \le \hat{H}_t \le 1 - \theta\sigma^H_t. \tag{22}$$

Then, to find the scheduled $\{\underline{H}_{\tau+k}\}^{HT-1}_{k=0}$, the following optimization problem is solved :

$$\{\underline{H}_{\tau+k}\}^{HT-1}_{k=0} = \min_{\{P^{el}_k\}} \{\hat{H}_{\tau+k}\}^{HT-1}_{k=0} \tag{23}$$

subject to: (11), (21)–(22). Once $\{\underline{H}_{\tau+k}\}^{HT-1}_{k=0}$ has been found, problem (16) can be solved, subject to (1)–(8), (11), (12), (14), (18). In the following, this algorithm will be addressed as MPC_LoH.

To summarize, the overall MPC_LoH algorithm consists of the following steps to be replicated after a fixed time interval D:

1. At timestamp $t = \tau$, get $\{\underline{H}_{\tau+k}\}^{HT-1}_{k=0}$ by solving (23) over the finite prediction horizon $t \in [\tau, \tau+HT-1]$ subject to (11), (21)–(22);
2. $\forall t \in [\tau, \tau+D-1]$, solve (16) over the finite prediction horizon $t \in [t, t+T-1]$ applying the MPC technique;
3. Set $\tau = \tau + D$ and repeat.

3.4. Reinforcement learning algorithm

The RL algorithm used in this study is Soft Actor Critic (SAC), an off-policy actor-critic deep RL method incorporating entropy maximization into the RL objective to enhance exploration and training stability [50]. SAC's ability to efficiently handle complex tasks and its robustness in high-dimensional action spaces make it particularly suitable for applications in robotics, autonomous driving, and video games, where it has demonstrated considerable success. The incorporation of entropy into the optimization process not only facilitates exploration but also significantly stabilizes the training process, establishing SAC as a versatile and powerful tool in the RL landscape. The SAC aims to balance exploration and reward maximization by augmenting the RL objective with an entropy term, thus encouraging the policy to explore more diverse actions. The importance and difficulty of shaping a reward function in deep RL-based energy management is well known [51], especially when it comes to hydrogen storage optimization [52]. Recent studies

highlighted the effectiveness of including entropy maximization in RL-based optimizers, culminating with the review [53] who connected entropy-regularized SAC to energy system control, justifying its use for complex storage incentives.

The state-value function $J(\pi)$ estimates the expected return (total discounted reward) over all future timestamps from state s following policy π according to the equation:

$$J(\pi) = \mathbb{E}_{\tau\sim\pi}\left[\sum_{t=0}^{\infty}\gamma^t\Big(R(s_t,a_t,s_{t+1}) + \alpha H(\pi(\cdot|s_t))\Big)\right], \tag{24}$$

where:

- $\pi(\cdot|s_t)$ is the policy, *i.e.*, a probability distribution over all possible actions $(\cdot)$ given the state s_t;
- $\mathbb{E}_{\tau\sim\pi}$ is the expectation taken over the trajectory τ generated by following policy π;
- γ^t is the discount factor raised to the power of the time step t, quantifying the diminishing importance of rewards received further in the future, and it is a value between 0 and 1;
- $R(s_t,a_t,s_{t+1})$ is the reward function, giving the immediate reward received when transitioning from state s_t to state s_{t+1} due to action a_t;
- α is a scaling factor for the entropy term, used to control the degree of exploration in policy optimization;
- $H(\pi(\cdot|s_t))$ is the entropy of the policy π at state s_t, measuring the randomness of the policy, with larger entropy implying more exploration.

The action-value function under policy π, $Q^\pi(s,a)$, that incorporates the entropy term is given by the equation:

$$Q^\pi(s_t,a_t) = \mathbb{E}\Big[R(s_t,a_t,s_{t+1}) + \gamma\Big(Q^\pi(s_{t+1},a_{t+1}) + \alpha H(\pi(\cdot|s_{t+1}))\Big)\Big], \tag{25}$$

where:

- $\pi(\cdot|s_{t+1})$ is the policy, a probability distribution over actions given the state s_{t+1}.
- $Q^\pi(s_t,a_t)$ is the action-value function under policy π, estimating the expected return (total discounted reward) from state s, taking action a, and thereafter following policy π;
- $\mathbb{E}$ is the expected value operator, computed taking the expectation over the random variables involved, specifically over the next state s_{t+1} and next action a_{t+1};
- γ is the discount factor, quantifying the difference in importance between future rewards and immediate rewards, and it is a value between 0 and 1;
- $Q^\pi(s_{t+1},a_{t+1})$ is the action-value function evaluated at the new state s_{t+1} and new action a_{t+1} chosen according to policy π;
- $H(\pi(\cdot|s_{t+1}))$ is the entropy of the policy π at state s_{t+1}, measuring the randomness of action selection, with larger entropy encouraging more exploration.

Actor updates are performed using gradient ascent on the policy network's parameters to maximize (24).

Establishing an environment is a fundamental step for the implementation of a RL algorithm, delineating both an action space and an observation space to determine the agent's interaction with the environment. The action space encompasses all possible actions that the agent can execute, either discrete or continuous. Similarly, the observation space encompasses all relevant information available to the agent, providing insights into the current state of the environment, which is essential for informed decision-making. Typically, an action generates a variation of some observations, which allows the RL algorithm to explore how its actions modify the observed quantities and learn from its environment effectively.

The environment is designed to mimic the SP architecture depicted in Fig. 1. Both the action and observation spaces are defined as continuous, requiring only the specification of the minimum and maximum values for each quantity. The range of optimal parameters was identified by preliminary tests, and their optimal combination was selected via a grid search evaluating performance on the last 20% of the first two years of available data (i.e., last 20% of the training set). We investigated the variable number of hidden neurons of the actor (n_a) and the critic (n_c) neural networks in the range between 4 and 100, the learning rate in the range between 10^{-3} and 10^{-1}, and the entropy value between 1 and 5.

The action space consists of three actions: "(dis)charge the VRFB", "charge the Ely", and "use DG". After having optimized the configuration through a grid search, the observation space consists of 20 observations, which are:

- the current values, at timestamp t, of RES energy generation, ship loads, hydrogen loads, SoC, LoH, energy price (Table 1), power from/to the grid, diesel utilized, economic gain/loss, timestamp of the year, battery capacity and number of equivalent cycles, and a dummy variable acting as an internal counter for data storage;
- the predictions over the next 6 h performed by the algorithms described in Section 3.2 regarding the total RES generation (sum of the predictions concerning PV and WTs) and total office building load;
- the future values, assumed as known in advance, of the total ship load and hydrogen request over the next 24 h, and of the single hydrogen requests for timestamps $t+1, t+3$, and $t+5$.

The upper and lower limits of the actions are determined by the technical constraints and parameters described in Section 2, while those for the observations are based on the values observed in the training set. At each timestamp, the quantities in the observation space vary depending on the actions selected according to the current policy following the equations described in Section 3.1. Therefore, the reward function utilized to update the RL model at each timestamp is defined as follows:

$$R_t(s_t,a_t,s_{t+1}) = J_t - a_1\cdot P_t^{dg} - a_2\cdot \Delta_A \tag{26}$$

where J_t is the net economic gain, a_1 and a_2 are two constants empirically set to 150 and 10, respectively, and Δ_A represents the total difference between the actions suggested by the RL controller and the actions actually implemented, which may have been adjusted according to the post-optimization procedures outlined in Section 3.5. Here, P_t^{dg} refers to the generated power at timestamp t, which is associated with an operational cost that the controller seeks to minimize. The RL algorithm is trained on 500 episodes, each of length 17520 h (2 years), whereas it is tested on the last year of available data in order to provide a fair comparison with the performance of the MPC. To minimize the likelihood of system failure and ensure the controller's ability to manage all load situations, an episode was interrupted and a negative reward of 10^{-6} was assigned whenever the controller was unable to satisfy all loads despite the post-optimization operations. The training curves of the SAC algorithm are presented as supplementary materials.

3.5. Post-optimization operations

The MPC-based and RL algorithms described above use forecasts of RESs generation and building demand to decide the optimal control to apply. However, forecasts may be different from the actual RES generation and building demand. Thus, the scheduled optimal control needs to be modified accordingly, since all the constraints must hold true at every timestamp t. The connection with the main grid is called to account for any excess of energy generation or demand so that (14) holds true, by modifying the programmed values of imported and exported powers, in the limits imposed by (1) and (2). Thus, the new values of P_t^i and P_t^e are computed from the scheduled ones $\hat{P}_t^i$ and $\hat{P}_t^e$ as follows:

$$P_t^i = \max\left(\hat{P}_t^i, \hat{P}_t^i + \left(\hat{P}_t^{res} - \hat{P}_t^l\right) - \left(P_t^{res} - P_t^l\right)\right), \tag{27}$$

$$P_t^e = \max\left(\hat{P}_t^e, \hat{P}_t^e - \left(\hat{P}_t^{res} - \hat{P}_t^l\right) + \left(P_t^{res} - P_t^l\right)\right). \quad (28)$$

From (27) and (28), it results that if the expected overall generation, $\hat{P}_t^{res} - \hat{P}_t^l$, is larger than the actual one $P_t^{res} - P_t^l$, meaning there has been a deficit of generation, such deficit is imported from the grid. On the other hand, if the expected overall generation, $\hat{P}_t^{res} - \hat{P}_t^l$, is lower than the actual one $P_t^{res} - P_t^l$, meaning there has been an excess of generation, such excess is exported to the grid.

However, by modifying the power exchanged with the grid, constraints (1) and (2) can be violated. To account for this issue, if the new value of P_t^i is such that $P_t^i > P_{max}^g$, the following procedure is adopted. At the end of every step, the new modified value of P_t^i ensuring (14) is found, and if it is in the limits set by (1) the procedure stops, otherwise it goes on to the next step:

- the RES curtailment is decreased in the limited amount set by (4);
- the VRFB charge is decreased, or eventually, the VRFB discharge is increased, in the limited amount set by (5), (6), and (8);
- the DG power is increased in the limited amount set by (3).

Instead, if the new value of P_t^e is such that $P_t^e > P_{max}^g$, the following procedure is adopted. At the end of every step, the new modified value of P_t^e ensuring (14) is found, and if it is in the limits set by (2) the procedure stops, otherwise it goes on to the next step:

- the DG power is reduced in the limited amount set by (3);
- the Ely power is increased in the limited amount set by (11), and (13);
- the VRFB charge is increased, or eventually, the VRFB discharge is decreased, in the limited amount set by (5), (6), and (8);
- the RES curtailment is increased in the limited amount set by (4).

After having balanced the forecasting error, the new VRFB SoC, capacity and number of equivalent cycles, and the new LoH are computed with (7), (9)–(10), (12).

Finally, in real-world applications, any control algorithm must be complemented with backup strategies to ensure operational safety in case of failures. This applies universally to RL, MPC, and traditional control methods.

Critical infrastructure management always requires fail-safe mechanisms and redundancy to mitigate risks. However, since this is a general requirement for all control strategies, a detailed discussion on risk mitigation is beyond the scope of this paper.

4. Simulations and results

The hyperparameters of the RL and energy forecasting models were optimized for the port of Genova, while the port of Trapani is considered as a validation case study. Straightforwardly, Section 4.1–4.5 presents the detailed models and the results obtained for the case study of Genova, while the results related to Trapani are reported in Section 4.6.

The control algorithms were tested on the data of the third year; the first and second years of the available data were used to train the neural networks and obtain the RESs and building demand forecasts.

This section presents the results referred to the port of from the developed AI forecasting models, which provide predictions used as inputs for the controllers. Additionally, the section includes the outcomes of the MPC algorithms with prediction sequences of varying lengths as input, as well as the results of the SAC algorithm with different observation spaces. A comparative analysis of the results achieved with the two approaches is also presented.

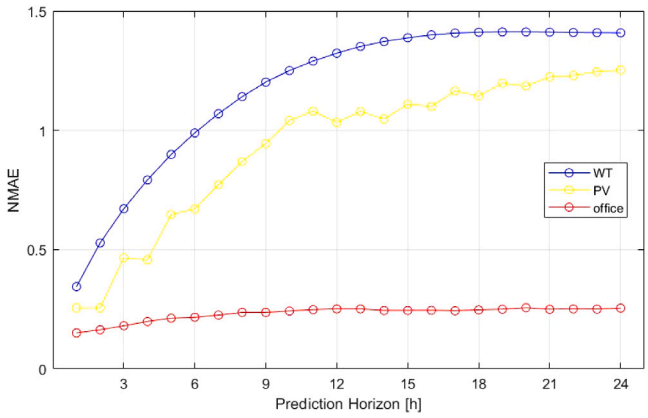


Fig. 5. NMAE of the forecasting algorithms over a prediction horizon of 24 h for the port of Genova.

4.1. Sensitivity analyses

In order to test the developed algorithms under different levels of information knowledge, two sensitivity analyses were conducted, one on the length of the prediction horizon of the MPC_m and MPC_{LoH} algorithms, and one on the type of predictions passed to the SAC.

The sensitivity analysis on the length of the prediction horizons of the MPC_m and MPC_{LoH} was carried out to assess the ability of the developed algorithms to optimize the SP operations at various levels of information and risk. Indeed, the MPC technique requires the knowledge of the inputs along the whole prediction horizon, thus the longer the prediction horizon, the more extended the information on the inputs. However, longer prediction horizons lead to less risky decisions. In fact, once the first step of the control trajectory is applied, the MPC will repeat the optimization with the additional knowledge of just the last timestamp of the prediction horizon. Since the state of the system is already defined for every timestamp of the optimization but the last one, the last timestamp will be less likely to create disturbances on the state of the system if it represents a small percentage of the prediction horizon.

Specifically, the MPC_m and the MPC_{LoH} were tested at the following values of T: 6, 12, 18, and 24 h. Concerning the MPC_{LoH}, HT was kept constant to 1 week, assuming that the expected ZE-Ships hydrogen demand $\{\hat{P}_{t+k}^h\}_{k=0}^{HT-1}$ was exact only for the first $T+$ 24 h. At midnight of the current day, and every 24 h, thus τ is hour 00:00 and D is 24 h, the actual ZE-Ships demand was perturbed starting from the timestamp $t = T+$ 25 by randomly increasing and decreasing the vessel demand and randomly shifting forward and backward in time the vessel arrival and departure time. The choice of not perturbing the ZE-Ships demand for the first $T+$ 24 h lies in the assumption that the MPC control sees an exact scheduled profile of both Ro-Paxs and ZE-Ships demands, possible only if during the first $T+$ 24 h the ZE-Ships demand is not perturbed. In fact, at 11 p.m. on any simulation day, for which the minimum LoH profile to be met was scheduled 23 h earlier, the MPC algorithm looks ahead until hour $T-$ 2 of the following day, a timestamp $T+$ 23 h away from when the minimum hydrogen profile was last defined.

Regarding the RL algorithm, the performance of the SAC algorithm was examined when:

1. it has complete knowledge of future inputs;
2. it has zero knowledge of the future inputs;
3. predictions of the inputs are available.

To accomplish the analysis, the SAC was tested using (a) real observations, (b) no observations, and (c) AI predictions as observations of the next 6 h of total RES generation and office building demand. The

Table 3
Sensitivity analyses.

Algorithm	Sensitivity variable
	Prediction horizon length T
MPC_m	[6, 12, 18, 24] h
MPC_{LoH}	[6, 12, 18, 24] h
	RES generation and building demand
SAC w/ AI predictions	AI predictions
SAC w/ ground truth	Real values hours
SAC w/o predictions	–

Table 4
Results over one year of simulations for the case of the port of Genova.

		Port of Genova			
Algorithm		Cash flow [M€]	Final S [p.u.]	Final H [p.u.]	Overall cash flow [M€]
	T [h]				
MPC_m	6	2.403	0.212	0.994	2.423
	12	2.348	0.584	1	2.367
	18	2.332	0.639	1	2.351
	24	2.326	0.426	0.990	2.346
	T [h]				
MPC_{LoH}	6	2.542	0.213	0.036	2.543
	12	2.553	0.587	0.043	2.544
	18	2.560	0.589	0.049	2.561
	24	2.564	0.589	0.035	2.564
	n_A/n_C				
SAC w/ AI predictions	50/100	2.513	0	0.525	2.523
SAC w/ ground truth	50/100	2.506	0	0.562	2.517
SAC w/o predictions	10/50	2.491	0	1	2.510
Oracle					2.674

last case represents a real application. The first, though unrealistic, and the second cases serve as comparisons to assess the ability of the SAC to extract patterns from the input data and provide flexibility against error-corrupted or absent future observations.

Table 3 resumes the conducted sensitivity analyses.

4.2. Performance analysis of AI forecasting models

The grid searches regarding the predictive neural networks yielded the following optimized configurations: for PV energy generation forecasting, a TDNN with 64 hidden neurons, taking as input 12-hour (12 timestamps) input sequences, and a regularization coefficient of 10^{-6}; for building demand forecasting, a TDNN with 144 hidden neurons, taking as input 36-hour input sequences, and a regularization coefficient of 0.3; for WTs energy generation forecasting, an LSTM with 70 hidden units, taking as input 11-hour input sequences, and a regularization coefficient of 10^{-9}.

The predictions were performed in an autoregressive manner, meaning the network predicts the next value in the sequence, which is then fed back into the network to predict subsequent values. In this study, autoregressive predictions were extended up to 24 h into the future. The Normalized Mean Absolute Error (NMAE) on the test set, *i.e.*, the last year of data with respect to the average value of the models are reported in Fig. 5. NMAE is the Mean Absolute Error normalized by the mean of the ground truth values for a given sequence of true values y_i^{true} and corresponding predictions y_i^{pred}. It is defined as:

$$NMAE = \frac{1}{N}\frac{1}{|\bar{y}^{true}|}\sum_{i=1}^{N}|y_i^{true} - y_i^{pred}| \tag{29}$$

where N is the total length of the simulation, *i.e.*, the number of predicted values, and $\bar{y}^{true}$ represents the mean of the ground truth values over the entire sequence.

The results for PV generation and building load align with those presented in our previous study [45]. The NMAE of building energy forecasting ranges from 0.15 for 1-hour-ahead predictions to 0.25 for 24-hour-ahead predictions. For PV energy forecasting, the NMAE ranges from 0.25 to 1.25. The NMAE for WT energy forecasting ranges from 0.34 to 1.41, with WT forecasting proving to be the most challenging. This difficulty is expected given the limited number of features considered in this study, as the best-performing studies on this topic use additional meteorological features for medium-term predictions [54]. Based on these results, only short-term predictions up to 6 h were used as input for the SAC regarding RES forecasting (PV+WT), with NMAEs of 0.67 and 0.99, respectively.

4.3. MPC Results

Table 4 shows the results in terms of annual cash flows for all the developed algorithms. Specifically:

- the column "Algorithm" refers to the technique adopted in the simulation;
- the column "Cash Flow" refers to the sum of the SP economic return J_t (15) over a simulation of one year;
- the column "Final S" refers to the VRFB SoC at the end of the simulation;
- the column "Final H" refers to the tank LoH at the end of the simulation;
- the column "Overall Cash Flow" refers to the sum of the cash flow and the profits coming from fully discharging the VRFB and the hydrogen tank to meet ships' demand.

The algorithm denominated "Oracle" represents the maximum achievable cash flow and is used as a benchmark, and consists of an optimization procedure on the simulated year maximizing the SP economic return, assuming to have perfect knowledge of all the inputs from the beginning, *i.e.*, HOMER sizing conditions.

From Table 4 it can be seen how the MPC_{LoH}, described in Section 3.3.3, outperforms the MPC_m, described in Section 3.3.2, for every prediction horizon length T.

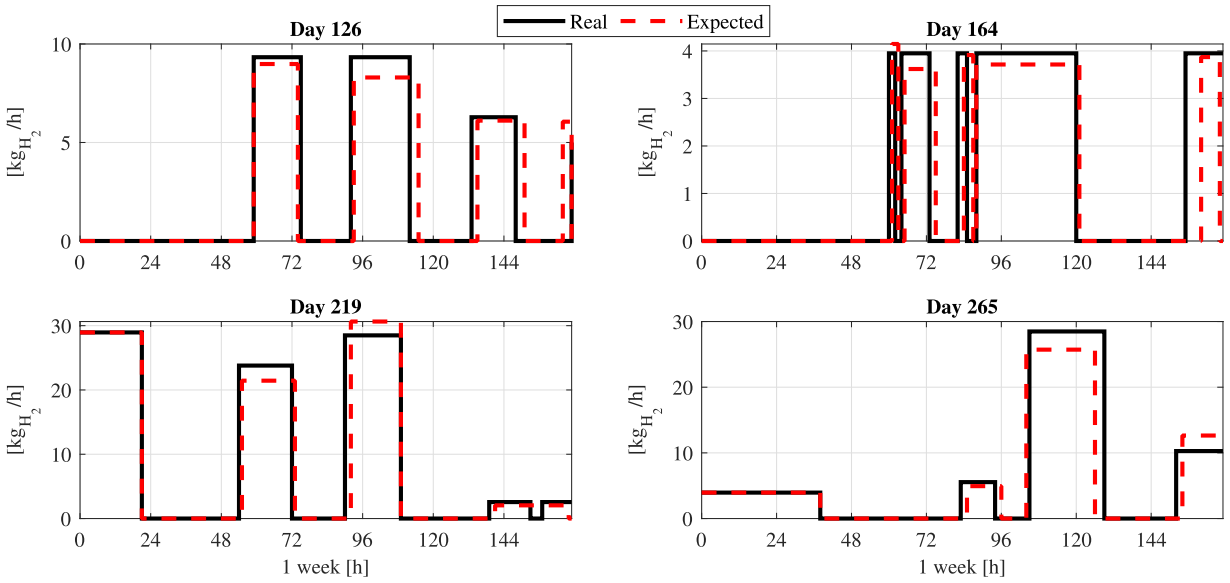


Fig. 6. Real and expected ZE-Ships hydrogen demand.

Table 5
Diesel usage reduction at berth for the case of the port of Genova.

Algorithm		Diesel usage reduction [%]
	T [h]	
MPC_m	6	77.1
	12	68.4
	18	62.7
	24	58.4
	T [h]	
MPC_{LoH}	6	100
	12	100
	18	100
	24	100
	n_A/n_C	
SAC w/ AI predictions	50/100	98.1
SAC w/ ground truth	50/100	98.6
SAC w/o predictions	10/50	75.6
Oracle		100

Regarding the MPC_{LoH}, results show that larger cash flows are obtained at larger prediction horizon lengths. This trend reflects the ability of the LoH scheduling algorithm to make the most of the input information. For the sake of completeness, Fig. 6 shows four examples of the expected ZE-Ships hydrogen demand profile $\{\hat{P}^h_{t+k}\}_{k=0}^{HT-1}$, compared with the actual demand profile.

On the other hand, the MPC_m with the modified cost function shows better performances at smaller values of T. Short prediction horizons cause the MPC to take riskier actions, however the LoH is always kept in safe intervals by (17) and thus the riskier actions cause no harm and provide better results.

Finally, as reported in Table 5, it is worth remarking that MPC_{LoH} does not use DGs, for every prediction horizon length of T, thus contributing to reduce by 100% diesel usage at berth. While MPC_m occasionally uses on-board DGs, especially at larger prediction horizon lengths T. In this case, CI contributes to reduce diesel usage at berth by 77.1%, 68.4%, 62.7%, and 58.4%, for T equal to 6, 12, 18, and 24 h, respectively.

4.4. Reinforcement learning results

All RL models were trained using Adam as optimizer and with a discount factor of 0.99. The actor network architecture was designed to generate the policy for the agent, consisting of the following layers: an input layer, a fully-connected hidden layer with n_A neurons followed by another fully connected layer with $n_A/2$ neurons, followed by a single neuron with ReLU transfer function; then, the network splits into two branches: a mean branch, consisting of a fully connected layer with 3 units, responsible for predicting the mean of the action distribution, and a standard deviation branch, which starts with a fully connected layer with 3 units and then incorporates a standard deviation layer from the original actor network. These layers and branches are connected to form the complete actor network. The critic network evaluates the value of state–action pairs, guiding the agent by providing feedback on the actions taken. The critic net mimics the structure of the actor net, with the exception that the number of hidden neurons n_C is optimized separately and a fully connected layer with one unit produces the critic's output, which represents the value estimate.

Beside SAC that takes as input the predictions from the developed AI algorithms, two additional SAC configurations were examined: one that takes the future ground truth as a prediction (simulating perfect predictions), and one that does not receive information on the predicted features. The results of the RL algorithms are reported in Table 4. The optimal values of n_A and n_C were identified through a grid search on the discovery set. Optimization also returned a best learning rate value of 0.03 and an entropy weight of 3, and a target entropy value of −3 for each model.

The results reported for the RL are the average results achieved over 100 runs, each initialized with the same experimental setup as the

Table 6
Best results for different RL models with AI predictions for the port of Genoa.

Model	Cash flow [M€]	Final S [p.u.]	Final H [p.u.]	Overall cash flows [M€]
SAC	**2.513**	**0**	**0.525**	**2.523**
TD3	1.895	1	1	1.936
DDPG	0.639	0.12	0	0.639
AC	1.992	1	0	1.993
PPO	Training failed (agent never completed an episode)			
TRPO	Training failed (agent never completed an episode)			
Oracle				2.674

MPC simulations. The SAC algorithm fed with AI predictions achieves a cash flow of 2.523 M€, which is slightly higher than the cash flow achieved by the SAC algorithm with ground truth (2.517 M€) and the SAC without predictions (2.510 M€). This result demonstrates that the integration of AI predictions can slightly improve the cash flow performance over no predictions. Interestingly, the model that receives AI predictions performs slightly better than the model that receives the ground truth, *i.e.*, predictions without errors. This could be explained by the nature of RL: the AI predictions might introduce a beneficial bias that can result in better performance despite individual prediction errors. By incorporating predictions with a degree of error, the model might explore a wider range of strategies, potentially discovering more profitable approaches over time [55]. This could also lead the model to be able to generalize in a wider range of situations [56].

A similar behaviour is observed for the overall cash flow. Indeed, the final SoC is zero for all the SAC models; conversely, the SAC without predictions shows a different outcome with a final value of one for the LoH, which is due to the system's tendency to charge the Ely as much as possible. Overall, the results indicate that incorporating AI predictions into the SAC framework can enhance the economic performance of the system. However, this improvement is limited compared to the performance of the model without predictions, indicating that knowledge of future hydrogen and ship loads (assumed to be known in advance in both cases) is more crucial for the system's performance. In addition to the SAC algorithms, the performance under different configurations of other state-of-the-art RL models was investigated, including Actor Critic (AC), Deep Deterministic Policy Gradient (DDPG), Twin Delayed Deep Deterministic Policy Gradient (TD3), and Proximal Policy Optimization (PPO). However, these algorithms achieved considerably worse results compared to SAC, as reported in Table 6.

Unlike the MPC_{LoH}, the RL controller sporadically uses DGs in some situations, as reported in Table 5. However, combining CI with SAC and AI predictions reduces diesel usage by 98.1%, consistently higher than what is achieved by the MPC_m. This significant reduction is notable compared to the 75.6% reduction achieved by SAC without predictions, which, however, is close to the 77.1% reduction achieved by the MPC_m. As a comparison, using SAC with ground truth as a prediction reduces diesel use by 98.6%.

The Oracle method achieves the highest overall cash flow of 2.674 M€, serving as a benchmark for the maximum achievable performance. The SAC models, while not reaching the Oracle's performance, show competitive results, particularly when enhanced with AI predictions. These results highlight the effectiveness of using AI predictions in conjunction with SAC models to optimize cash flow, with an improvement over models using no predictions.

4.5. Comparison between MPC and RL performance

The MPC strategy itself has the limitations discussed in Section 3.3.1 given by the difficulties of handling step-shaped loads. To overcome these limitations, two variations were made to the classic MPC problem formulation, one with a modified cost function rewarding high LoH values, MPC_m, and the other scheduling a minimum optimal LoH profile to be observed, MPC_{LoH}.

The MPC_{LoH} algorithm tends to perform better when it has access to more detailed and accurate information. This trend is evident from the results where the resulting control algorithm with long prediction horizon yields outcomes closer to the oracle, which represents the maximum achievable performance. Conversely, the MPC_m shows better performance at lower prediction horizon length, *i.e.*, less information.

On the other hand, SAC demonstrates superior robustness and flexibility in conditions with limited information, yielding outcomes close to the MPC_{LoH}. RL can better leverage shorter-term predictions and more effectively manage and mitigate inevitable prediction errors. This result translates into consistent financial performance even with less accurate forecasts, making the use of SAC with AI predictions a competitive choice.

While MPC excels with highly realistic predictions, RL may be more advantageous in scenarios where predictions are uncertain or prone to errors. This difference suggests that the choice between MPC and RL heavily depends on the specific operational context: if very realistic long-term predictions can be obtained, MPC might deliver optimal results; conversely, in situations where predictions are less reliable, RL might offer a better trade-off due to its adaptability and error management capabilities.

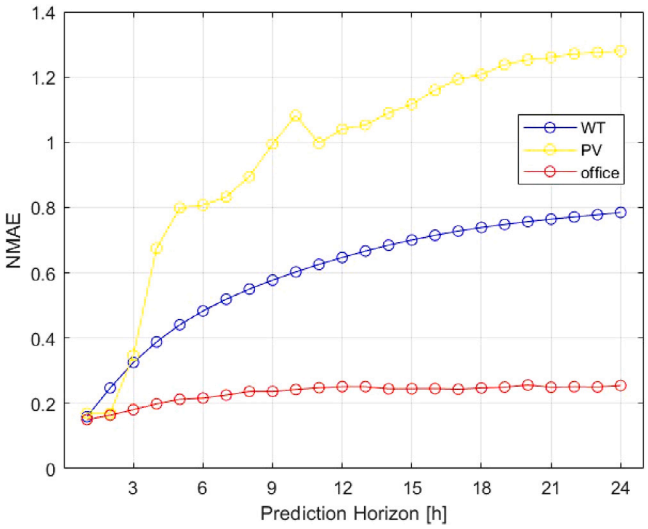


Fig. 7. NMAE of the forecasting algorithms over a prediction horizon of 24 h for the port of Trapani.

4.6. External validation case study: Port of Trapani

This section presents the results obtained in the case study of the port of Trapani. The aim of this Section is to evaluate the flexibility of the proposed approach dealing with a port with different seasonality; straightforwardly, the best-performing configuration, identified on the data from the port of Genova, was applied to the new dataset after retraining the models on the new data, without modifying their structure or re-optimizing hyperparameters. This approach allows for testing the adaptability of the framework to different port configurations and seasonal patterns with minimal additional computational cost.

Table 7
Results over one year of simulations for the case of the port of Trapani.

Port of Trapani					
Algorithm		Cash flow [M€]	Final S [p.u.]	Final H [p.u.]	Overall cash flow [M€]
	T [h]				
MPC_m	6	1.087	0.002	0.816	1.100
	12	1.083	0.196	0.995	1.100
	18	1.073	0.197	0.991	1.089
	24	1.066	0.039	0.966	1.082
	T [h]				
MPC_{LoH}	6	1.102	0.296	0.999	1.119
	12	1.107	0.968	1	1.124
	18	1.109	0.796	1	1.126
	24	1.109	1	1	1.126
	n_A/n_C				
SAC w/ AI predictions	50/100	1.089	0.228	1.000	1.102
SAC w/ ground truth	50/100	0.975	0.418	0.924	0.999
SAC w/o predictions	10/50	0.984	0.010	0.500	0.992
Oracle					1.234

Table 8
Diesel usage reduction at berth for the case of the port of Trapani.

Algorithm		Diesel usage reduction [%]
	T [h]	
MPC_m	6	81.5
	12	82.7
	18	85.8
	24	83.6
	T [h]	
MPC_{LoH}	6	99.9
	12	99.9
	18	99.9
	24	99.9
	n_A/n_C	
SAC w/ AI predictions	50/100	95.0
SAC w/ ground truth	50/100	100
SAC w/o predictions	10/50	99.0
Oracle		100

The same sensitivity analyses utilized for the Genova case study were performed. The neural networks utilized for the predictions maintain the same configurations described in Section 4.2, achieving the results on the test set reported in Fig. 7. The prediction NMAEs over 24 h of autoregressive forecasts are similar to those observed for the port of Genoa, with the error in wind prediction being significantly lower, likely due to greater wind regularity in Trapani.

Table 7 reports the results of 1 year of simulations for the port of Trapani. The results of the MPC algorithms completely respect the trends shown for the results of the port of Genova. Indeed, MPC_{LoH} outperforms the MPC_m for every prediction horizon length T, at larger prediction horizon length the MPC_{LoH} shows larger cash flows while MPC_m shows smaller cash flows. Specifically, the MPC_{LoH} and the MPC_m at $T = 24$ are able to achieve 91.3% and 87.7% of the maximum achievable cash flow, *i.e.*, "Oracle" cash flow.

Moreover, as reported in Table 8, it is worth remarking that MPC_{LoH} does not use DGs but to cover the peaks of Ro-Paxs demand higher than 7 MW (nominal grid connection), for every prediction horizon length of T, thus contributing to reduce by 99.9% diesel usage at berth. While MPC_m occasionally uses on-board DGs, especially at larger prediction horizon lengths T. In this case, CI contributes to reduce diesel usage at berth by 81.5%, 82.7%, 85.8%, and 83.6%, for T equal to 6, 12, 18, and 24 h, respectively.

The SAC algorithm, retrained using the same configuration described in Section 4.4, achieved a cash flow of 1.102 M€, 0.999 M€, and 0.992 M€ with a corresponding reduction in diesel utilization of 95.0%, 100%, and 99.0% when fed with AI predictions, the ground truth, and without predictions, respectively.

The results of the case of the port of Trapani proved again the ability of the MPC_{LoH} to extract knowledge from the input and maximize the cash flow and the ability of the SAC to deal with uncertain information. Indeed, the SAC fed with AI predictions is still able to outperform the MPC_m for every T.

4.7. Robustness analysis

The performance of the proposed algorithms can depend on the values of the parameters of the controlled devices, *i.e.* the VRFB and the Ely. Parameters such as the efficiencies of the VRFB and Ely, may change in real operations, and they can be different from the ones in the models used during the algorithms design. Therefore, to test the robustness of the proposed methods, two mismatch analyses were conducted: maintaining the Ely efficiency η^{el} at 0.65 and the VRFB round trip efficiency η^{b} at 0.70 during the control design, operations were simulated in the following cases:

- **mismatch up**: $\eta^{el} = 0.70$ and $\eta^{b} = 0.75$;
- **mismatch down**: $\eta^{el} = 0.60$ and $\eta^{b} = 0.65$.

In this way, the controllers take action considering efficiencies that are lower, in the case **mismatch up**, and larger, in the case of **mismatch down**, than the real values. Such analyses were conducted on all the tests listed in Tables 4 and 7, but the case of SAC with ground truth and the Oracle, as these cases are not applicable in real applications. The results are shown in Table 9 and present a situation in which the percentage variation of the overall cash flow, with respect to the results listed in Table 4 and in Table 7, is particularly limited for the

Table 9
Results of mismatch analyses on model parameters. Percentage variations of the overall cash flow in the case of increased electrolyzer and battery efficiencies (mismatch up) decreased electrolyzer and battery efficiencies (mismatch down).

Algorithm		Port of Genova [%]		Port of Trapani [%]	
	T [h]	mismatch up	mismatch down	mismatch up	mismatch down
MPC_m	6	1.0	−1.0	1.3	−1.5
	12	1.3	−1.0	1.3	−1.7
	18	1.0	−1.0	1.6	−1.7
	24	0.7	−0.9	1.5	−1.6
	T [h]				
MPC_{LoH}	6	0.6	−0.9	1.5	−2.0
	12	1.2	−0.6	1.3	−2.2
	18	0.8	−1.0	1.2	−2.2
	24	0.7	−1.0	1.2	−2.3
	n_A/n_C				
SAC w/ AI predictions	50/100	0.3	−0.7	0.1	0.1
SAC w/o predictions	10/50	0.2	−0.1	0.3	0.5

SAC algorithms. The maximum and minimum differences pertain to the MPC algorithms with longer time horizons; however, they are limited to 1.6% and −2.3%. In general, the proposed algorithms appear to be robust to parameter mismatch, particularly in the case of the SAC with AI predictions.

4.8. Implementation aspects and computational complexity

The software General Algebraic Modeling System (GAMS) [57] and the solver Complex Linear Programming Expert (CPLEX) [58] were used to solve the problem (17) in the algorithm MPC_m and the problems (16) and (23) in the algorithm MPC_{LoH}. The computational time of MPC algorithms depends on the number of variables to be optimized and the type of programming. Both MPC_m and MPC_{LoH} are mixed-integer linear programming algorithms, and their number of variables increases with the length of the prediction horizon. However, even at the longest prediction horizon length, *i.e.*, 24 h, both the algorithms take less than 1 s to run. Such running time is largely short considering the 1-hour sampling time interval.

The computational complexity of the considered neural networks was analysed to evaluate their scalability and computational feasibility. The complexity of a TDNN is $O(n_h \times I)$, while the complexity of the LSTM neural network is given by $O(4n_h^2 + 2n_h \times I)$ where n_h is the number of hidden units and I is the length of the input sequence. The computational complexity of the SAC during training depends primarily on the number of model parameters and the number of training steps. The actor network consists of two fully connected layers with n_a and $n_a/2$ neurons, while the critic network comprises two independent Q-networks, each with layers of n_c neurons. Given an observation space of size s and an action space of size a, the forward pass complexity for each step scales as

$$O(sn_a + n_a \times \frac{n_a}{2} + 2(sn_c + an_c + n_c^2)) \tag{30}$$

where the dominant term is $O(n_c^2)$ since the number of hidden units is significantly larger than the input and output dimensions. During training, the overall complexity can be approximated as $O(ENn_c^2)$), where E is the number of episodes and T is the length of each episode. In contrast, the testing phase scales as $O(Nn_c^2)$, with a contribution of $O(n_c^2)$ for each estimation, corresponding to a testing time of less than 1 s, which is largely feasible in real-life applications given the 1-hour sampling of the data. All tests were performed using Matlab R2023B on a system running Windows 11 Pro (version 23H2), equipped with a 13th Gen Intel Core i7-13700K processor (3.40 GHz), 32 GB of RAM, and an NVIDIA GeForce RTX 3060 GPU.

5. Conclusions

The presented paper compares the performance of two advanced control strategies, MPC and SAC, in optimizing cash flow, and contemporarily reducing fossil fuel usage, in SP operations. The study highlights the strengths and limitations of each approach under varying levels of prediction accuracy.

Answer to RQ1: The MPC_{LoH} algorithm performs optimally with longer prediction horizons, making effective use of detailed and accurate information. It closely approaches the benchmark "Oracle" method, which assumes perfect knowledge of future inputs, achieving 95.9% and 91.2% of the maximum possible cash flow in the use cases of Genova and Trapani, respectively. The MPC_m performs better with shorter prediction horizons, it occasionally uses DGs, and it shows the poorest performances among the investigated algorithms. The SAC algorithm demonstrates superior robustness and flexibility, performing well even with less accurate and shorter-term predictions, achieving 94.4% and 89.3% of the maximum possible cash flow in the use cases of Genova and Trapani, respectively. This approach shows a consistent performance and a significant reduction in diesel usage, especially when enhanced with AI predictions. Moreover, SAC shows better performances than other state-of-the-art RL models like DDPG, TD3, and PPO.

Answer to RQ2: Both MPC and SAC approaches contribute to a reduction in fossil fuel usage and greenhouse gas emissions up to 100% in both use cases. The choice of the best technique should be based on the operational context and the availability of accurate prediction information. The MPC_{LoH} proves to be better in scenarios with highly realistic and long-term predictions, whereas the SAC algorithm offers better performance in environments with uncertain and less reliable predictions.

In summary, both MPC and SAC approaches are promising techniques for enhancing the efficiency and sustainability of shore power operations, with SAC providing a more adaptable and error-tolerant solution in scenarios with less certain predictions. Regarding computational complexity, the computational cost of SAC is primarily driven by the number of hidden units in the critic network. Consequently, reducing the size of the actor network can lower overall complexity while maintaining a good balance between performance and efficiency, albeit with potentially suboptimal results. Similarly, although re-optimizing the hyperparameters of the neural network might lead to improved results, the case study of the port of Trapani demonstrated that satisfactory performance can be achieved also without resorting to extensive model retraining. Moreover, dealing with larger ports could mean dealing with multiple batteries, electrolyzers and/or hydrogen tanks, thus the only variables that can increase in their number are those related to additional energy storage systems. However, given the very restrained inference time of the developed control algorithms, larger case studies would not be problematic to handle.

Finally, possible future developments could extend the framework to integrate EU ETS trends in the analysis and explore other adaptive control algorithms, such as Multi-Agent RL, or hybrid RL-MPC algorithms that can hopefully integrate the ability of the MPC of producing optimal results and the robustness and flexibility of the RL.

CRediT authorship contribution statement

Francesco Conte: Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization. **Gianluca Natrella:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis. **Daniele Sasso:** Writing – original draft, Software, Methodology. **Federico Silvestro:** Visualization, Resources, Funding acquisition. **Fabio D'Agostino:** Visualization, Resources, Funding acquisition. **Federico D'Antoni:** Writing – original draft, Validation, Supervision, Software, Methodology, Investigation, Formal analysis. **Mario Merone:** Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work has been developed in the framework of the project "Network 4 Energy Sustainable Transition - NEST", funded under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.3 - NextGenerationEU. All authors have read and agreed to the published version of the manuscript

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ijepes.2025.111437.

Data availability

The data that has been used is confidential.

References

[1] Čampara L, Hasanspahić N, Vujičić S. Overview of MARPOL ANNEX VI regulations for prevention of air pollution from marine diesel engines. SHS Web Conf 2018;58:01004. http://dx.doi.org/10.1051/shsconf/20185801004.

[2] Zhang C, Yang Y, Wang N. Port governance and sustainable development: The impact of port smartization on port carbon emission efficiency. Ocean & Coastal Management 2024;259:107485. http://dx.doi.org/10.1016/j.ocecoaman.2024.107485.

[3] Stolz B, Held M, Georges G, Boulouchos K. The CO2 reduction potential of shoreside electricity in Europe. Appl Energy 2021;285:116425. http://dx.doi.org/10.1016/j.apenergy.2020.116425.

[4] Jahangiri S, Nikolova N, Tenekedjiev K. An improved emission inventory method for estimating engine exhaust emissions from ships. Sustain Environ Res 2018;28(6):374–81. http://dx.doi.org/10.1016/j.serj.2018.08.005.

[5] Ducruet C, Martin BP, Sene MA, Prete ML, Sun L, Itoh H, Pigné Y. Ports and their influence on local air pollution and public health: A global analysis. Sci Total Environ 2024;915:170099. http://dx.doi.org/10.1016/j.scitotenv.2024.170099.

[6] The European Parliament and the Council of the European Union. Regulation (EU) 2023/1804 of the European parliament and of the council of 13 September 2023 on the deployment of alter- native fuels infrastructure, and repealing Directive 2014/94//eu (Text with EEA relevance). Off J of Eur Union 2023. http://data.europa.eu/eli/reg/2023/1804/oj.

[7] D'Agostino F, Fidigatti A, Ragaini E, Silvestro F. Integration of shipboard microgrids within land distribution networks: Employing a ship microgrid to meet critical needs. IEEE Electrif. Mag 2019;7(4):69–80. http://dx.doi.org/10.1109/MELE.2019.2943979.

[8] Parise G, Su C-L, Marchetti FA, Francucci J, Liao C-H. Short-circuit calculations in LV cold ironing systems: Characteristic currents method CCM and IEC method. IEEE Trans Ind Appl 2022;58(4):4394–400. http://dx.doi.org/10.1109/TIA.2022.3174185.

[9] D'Agostino F, Grillo S, Infantino R, Pons E. High-voltage shore connection systems: Grounding resistance selection and short-circuit currents evaluation. IEEE Trans Transp Electrification 2022;8(2):2608–17. http://dx.doi.org/10.1109/TTE.2021.3137717.

[10] Tsekouras GJ, Kanellos FD. Ship to shore connection - Reliability analysis of ship power system. In: 2016 XXII international conference on electrical machines. 2016, p. 2955–61. http://dx.doi.org/10.1109/ICELMACH.2016.7732944.

[11] Kumar N, Panda SK. A multipurpose and power quality improved electric vessels charging station for the seaports. IEEE Trans Ind Inform 2023;19(3):3254–61. http://dx.doi.org/10.1109/TII.2022.3170424.

[12] Kumar N, Panda SK. Smart high power charging networks and optimal control mechanism for electric ships. IEEE Trans Ind Inform 2023;19(2):1476–83. http://dx.doi.org/10.1109/TII.2022.3170484.

[13] McKinlay CJ, Turnock SR, Hudson DA. Route to zero emission shipping: Hydrogen, ammonia or methanol? Int J Hydrog Energy 2021;46(55):28282–97. http://dx.doi.org/10.1016/j.ijhydene.2021.06.066.

[14] van Biert L, Godjevac M, Visser K, Aravind P. A review of fuel cell systems for maritime applications. J Power Sources 2016;327:345–64. http://dx.doi.org/10.1016/j.jpowsour.2016.07.007.

[15] Pivetta D, Dall'Armi C, Sandrin P, Bogar M, Taccani R. The role of hydrogen as enabler of industrial port area decarbonization. Renew Sustain Energy Rev 2024;189:113912. http://dx.doi.org/10.1016/j.rser.2023.113912.

[16] Dall'Armi C, Pivetta D, Taccani R. Hybrid PEM fuel cell power plants fuelled by hydrogen for improving sustainability in shipping: State of the art and review on active projects. Energies 2023;16(4). http://dx.doi.org/10.3390/en16042022.

[17] Kotrikla AM, Lilas T, Nikitakos N. Abatement of air pollution at an aegean island port utilizing shore side electricity and renewable energy. Mar Policy 2017;75:238–48. http://dx.doi.org/10.1016/j.marpol.2016.01.026.

[18] Zhao P, Ma K, Yang J, Yang B, Guerrero JM, Dou C, Guan X. Distributed power sharing control based on adaptive virtual impedance in seaport microgrids with cold ironing. IEEE Trans Transp Electrification 2023;9(2):2472–85. http://dx.doi.org/10.1109/TTE.2022.3211204.

[19] Oprea S-V, Bâra A. Edge and fog computing using IoT for direct load optimization and control with flexibility services for citizen energy communities. Knowl-Based Syst 2021;228:107293. http://dx.doi.org/10.1016/j.knosys.2021.107293.

[20] Fang S, Wang C, Liao R, Zhao C. Optimal power scheduling of seaport microgrids with flexible logistic loads. IET Renew Power Gener 2022;16(12):2711–20. http://dx.doi.org/10.1049/rpg2.12401.

[21] Iris Ç, Lam JSL. Optimal energy management and operations planning in seaports with smart grid while harnessing renewable energy under uncertainty. Omega 2021;103:102445. http://dx.doi.org/10.1016/j.omega.2021.102445.

[22] Conte F, D'Agostino F, Kaza D, Massucco S, Natrella G, Silvestro F. Optimal management of a smart port with shore-connection and hydrogen supplying by stochastic model predictive control. In: 2022 IEEE power & energy society general meeting. IEEE; 2022, p. 1–5. http://dx.doi.org/10.1109/PESGM48719.2022.9916817.

[23] D'Agostino F, Kaza D, Silvestro F, Conte F, Rrukaj R, Zadeh M. Green smart port energy system design: Optimal sizing. In: 2023 IEEE power & energy society general meeting. IEEE; 2023, p. 1–5. http://dx.doi.org/10.1109/PESGM52003.2023.10253013.

[24] Kovalishin P, Nikitakos N, Svilicic B, Zhang J, Nikishin A, Dalaklis D, Kharitonov M, Stefanakou A-A. Using Artificial Intelligence (AI) methods for effectively responding to climate change at marine ports. J Int Marit Saf Environ Aff Shipp 2023;7(1):2186589. http://dx.doi.org/10.1080/25725084.2023.2186589.

[25] Bakar NNA, Bazmohammadi N, Çimen H, Uyanik T, Vasquez JC, Guerrero JM. Data-driven ship berthing forecasting for cold ironing in maritime transportation. Appl Energy 2022;326:119947. http://dx.doi.org/10.1016/j.apenergy.2022.119947.

[26] Dawangi ID, Budiyanto MA. Ship energy efficiency management plan development using machine learning: Case study of CO 2 emissions of ship activities at container port. Int J Technol 2021;12(5):1048–57. http://dx.doi.org/10.14716/ijtech.v12i5.518.

[27] Hasanvand S, Rafiei M, Gheisarnejad M, Khooban M-H. Reliable power scheduling of an emission-free ship: Multiobjective deep reinforcement learning. IEEE Trans Transp Electrification 2020;6(2):832–43. http://dx.doi.org/10.1109/TTE.2020.2983247.

[28] Wu P, Partridge J, Bucknall R. Cost-effective reinforcement learning energy management for plug-in hybrid fuel cell and battery ships. Appl Energy 2020;275:115258. http://dx.doi.org/10.1016/j.apenergy.2020.115258.

[29] Wang D, Zheng W, Wang Z, Wang Y, Pang X, Wang W. Comparison of reinforcement learning and model predictive control for building energy system optimization. Appl Therm Eng 2023;228:120430. http://dx.doi.org/10.1016/j.applthermaleng.2023.120430.

[30] Brandi S, Fiorentini M, Capozzoli A. Comparison of online and offline deep reinforcement learning with model predictive control for thermal energy management. Autom Constr 2022;135:104128. http://dx.doi.org/10.1016/j.autcon.2022.104128.

[31] Liu X, Qiu L, Fang Y, Wang K, Li Y, Rodríguez J. Event-driven based reinforcement learning predictive controller design for three-phase NPC converters using online approximators. IEEE Trans Power Electron 2024.

[32] Liu X, Qiu L, Fang Y, Wang K, Li Y, Rodríguez J. Combining data-driven and event-driven for online learning predictive control in power converters. IEEE Trans Power Electron 2024.

[33] D'Agostino F, Kaza D, Schiapparelli P, Federico S, Bossi C, Colzi F. Assessment of the potential shore to ship load demand: the Italian scenario. IEEE PES Gen Meet 2021. http://dx.doi.org/10.1109/PESGM46819.2021.9638000.
[34] Perry JH, Green DW. Perry's chemical engineers' handbook. 8th ed. McGraw-Hill Education; 2008, https://www.aiche.org/sites/default/files/cep/20080670.pdf.
[35] Fan L, Tu Z, Chan SH. Recent development of hydrogen and fuel cell technologies: A review. Energy Rep 2021;7:8421–46. http://dx.doi.org/10.1016/j.egyr.2021.08.003.
[36] NASA. NASA Prediction of Worldwide Energy Resources. https://power.larc.nasa.gov/data-access-viewer.
[37] Hybrid Optimization of Multiple Energy Resources. https://www.homerenergy.com.
[38] Autorita' di Sistema Portuale del Mar Ligure Occidentale. Documento di pianificazione energetico ambientale del sistema portuale del mar ligure occidentale [in Italian]. 2019, https://www.portsofgenoa.com/it.
[39] Enercon E-48. https://en.wind-turbine-models.com/turbines/529-enercon-e-48.
[40] WES 18/80. https://windenergysolutions.nl/wes/windturbine-wes-80.
[41] Baxter R. 2020 energy storage pricing survey. 2021, http://dx.doi.org/10.2172/1866526.
[42] Mohiti M, Mazidi M, Rezaei N, Khooban M-H. Role of vanadium redox flow batteries in the energy management system of isolated microgrids. J Energy Storage 2021;40:102673. http://dx.doi.org/10.1016/j.est.2021.102673.
[43] Kumar SS, Lim H. An overview of water electrolysis technologies for green hydrogen production. Energy Rep 2022;8:13793–813. http://dx.doi.org/10.1016/j.egyr.2022.10.127.
[44] NEL PEM electrolyser - M series. https://nelhydrogen.com/product/m-series-electrolyser.
[45] Conte F, D'Antoni F, Natrella G, Merone M. A new hybrid AI optimal management method for renewable energy communities. Energy AI 2022;10:100197. http://dx.doi.org/10.1016/j.egyai.2022.100197.
[46] Peddinti V, Povey D, Khudanpur S. A time delay neural network architecture for efficient modeling of long temporal contexts. In: Interspeech. 2015, p. 3214–8. http://dx.doi.org/10.21437/Interspeech.2015-647.
[47] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9(8):1735–80. http://dx.doi.org/10.1162/neco.1997.9.8.1735.
[48] Borrelli F, Bemporad A, Morari M. Predictive Control for Linear and Hybrid Systems. Cambridge University Press; 2017, http://dx.doi.org/10.1017/9781139061759.
[49] Conte F, Massucco S, Schiapparelli G-P, Silvestro F. Day-ahead and intra-day planning of integrated BESS-PV systems providing frequency regulation. IEEE Trans Sustain Energy 2020;11(3):1797–806. http://dx.doi.org/10.1109/TSTE.2019.2941369.
[50] Haarnoja T, Zhou A, Abbeel P, Levine S. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In: International conference on machine learning. PMLR; 2018, p. 1861–70. http://dx.doi.org/10.48550/arXiv.1801.01290.
[51] Desportes L, Fijalkow I, Andry P. Deep reinforcement learning for hybrid energy storage systems: Balancing lead and hydrogen storage. Energies 2021;14(15). http://dx.doi.org/10.3390/en14154706, URL https://www.mdpi.com/1996-1073/14/15/4706.
[52] Dreher A, Bexten T, Sieker T, Lehna M, Schütt J, Scholz C, Wirsum M. AI agents envisioning the future: Forecast-based operation of renewable energy storage systems using hydrogen with Deep Reinforcement Learning. Energy Convers Manage 2022;258:115401. http://dx.doi.org/10.1016/j.enconman.2022.115401, URL https://www.sciencedirect.com/science/article/pii/S0196890422001972.
[53] Perera A, Kamalaruban P. Applications of reinforcement learning in energy systems. Renew Sustain Energy Rev 2021;137:110618. http://dx.doi.org/10.1016/j.rser.2020.110618, URL https://www.sciencedirect.com/science/article/pii/S1364032120309023.
[54] Simankov V, Buchatskiy P, Teploukhov S, Onishchenko S, Kazak A, Chetyrbok P. Review of estimating and predicting models of the wind energy amount. Energies 2023;16(16):5926. http://dx.doi.org/10.3390/en16165926.
[55] Schäfer L, Christianos F, Hanna J, Albrecht SV. Decoupling exploration and exploitation in reinforcement learning. In: ICML 2021 workshop on unsupervised reinforcement learning. 2021, https://openreview.net/forum?id=NwuIIOcznYt.
[56] Wang H, Zariphopoulou T, Zhou XY. Reinforcement learning in continuous time and space: A stochastic control approach. J Mach Learn Res 2020;21(198):1–34, http://jmlr.org/papers/v21/19-144.html.
[57] GAMS Development Corporation. General algebraic modeling system (GAMS) release 36.1.0. 2021, https://www.gams.com/download/. Fairfax, VA, USA.
[58] Cplex, IBM ILOG. V12. 1: User's manual for CPLEX. Int Bus Mach Corp 2009;46(53):157.