Decarbonising coastal shipping using fuel cells and batteries

Peng Wu

A thesis submitted in partial fulfilment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Department of Mechanical Engineering

University College London

June 15, 2020

I, Peng Wu, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

Fuel cells and batteries hold the potential to revolutionise emission performance for coastal ships. A critical review of previous research and projects in this field highlighted potential benefits offered through using fuel cell and battery technologies to decarbonise the coastal shipping. Currently, there is a lack of holistic design methods for such systems operating in different regions with variable energy properties.

This research project aims to develop a holistic power sourcing sizing methodology that can optimise the alternative propulsion system designs, and effective energy management strategies for such systems. This research employs energy efficiency oriented propulsion system models which are calibrated and validated by experimental data. The proposed multi-objective sizing methodology explores the optimal power source sizing of a typical coastal ferry with sensitivity studies of different energy properties. The sizing methodology has been validated using the random search method. Reinforcement learning and deep reinforcement learning agents were trained with real-ship power profiles to generate near-optimal energy management strategies. These strategies have been compared and verified with results solved by deterministic dynamic programming.

The conclusions of this research suggest that energy properties and power source characteristics can significantly influence the designs of hybrid fuel cell and battery systems. Ships operating on short routes can potentially benefit from such systems. Additionally, continuously monitored power profiles can be used to train reinforcement learning agents to achieve near-optimal operating costs.

Impact Statement

This research work identifies opportunities and constraints of using fuel cell and battery power to decarbonise coastal shipping. A novel power source sizing methodology for coastal ships focusing on both life-cycle greenhouse gas emissions and costs of the alternative propulsion systems has been developed. The developed design methodology can be used by the shipping industry to achieve a balanced plug-in hybrid fuel cell and battery propulsion system design, considering life-cycle greenhouse gas emissions, costs and shipboard constraints.

Reinforcement learning and deep reinforcement learning based on-line energy management strategies have been developed using extensive continuous monitoring data to achieve near-optimal cost performance in un-predicted voyages. The near-optimal cost performance of the developed energy management strategies has highlighted the potential of deep reinforcement learning algorithms in the control of hybrid propulsion systems. The proposed intelligent hybrid propulsion system control framework can be adopted by the industry to minimise cost impacts from the plug-in hybrid fuel cell and battery systems with multiple power sources.

Though the proposed hybrid propulsion system design methodology and intelligent energy management strategies are initially developed for coastal ships in this study, they are transferable to other applications such as road vehicles. Other researchers should be able to benefit from the output of this work as formulate and solve the hybrid propulsion system optimal energy management problem with extensive continuous monitoring data using suitable reinforcement learning algorithms.

Acknowledgements

First and foremost, a special mention goes to my supervisor, Professor Richard Bucknall, for giving me the opportunity to undertake this study. I would like to thank him for the invaluable assistance and continuous support he has provided to me throughout my PhD. Without his academic guidance and reassurance of my progress, it would not have been possible to complete this thesis.

I would like to thank Dr Julius Partridge and Mr Konrad Yearwood for reviewing my research, contributing ideas and proofreading this thesis.

I would like to thank Professor Alistair Greig and my second supervisor Professor Giles Thomas and for providing constructive and valuable comments on my MPhil/PhD upgrade report.

I would also like to thank Mr Jens Christian Bjeldorf and Molslinje A/S for approving using the case ship data in this study. I am also grateful to Mr Stig Eriksen and his colleagues for collecting the case ship data.

Thanks to Yuanchang, Ting for the conversations on reinforcement learning and my PhD. Thanks to Luke, John and Santiago for the collaborations we had. Thanks to Bin, Rui, Peipei, Kang, Liwei and Xinjie for having years of good time together.

Last but not least, my sincere gratitude to my family, uncles and aunts.

Contents

List of Figures

List of Tables

Nomenclature

Chapter 3

Roman Symbols

f	Objective function
g	Inequality constraint
l	Equality constraint
x	Vector of decision variable

Chapter 4

Roman Symbols

A_m	Active area of membrane electrode assembly, cm^2
a	Water activity
CR	Inlet air compression ratio
C	Battery C-rate
C_1	Battery cell model capacitance, F
c_p	Air specific heat capacity, $\mathrm{kJ\,kg^{-1}\,K^{-1}}$
c_1	Concentration overpotential coefficient
D	Fuel cell degradation rate
F	Faraday constant, 96485 $\mathrm{C\,mol^{-1}}$
$F(Z)$	Sizing objective function vector
f_k	Sizing objective function
G	Gibbs free energy of activation
I	Current
i	Current density, $\mathrm{Acm^{-2}}$

i_0^k	Exchange current density, Acm^{-2}
i_k	Battery cell model current, A
i_L	Limiting current density, Acm^{-2}
I_{fc}	Fuel cell stack current, A
k_1	Fuel cell balancing of plant power coefficient
LHV	Hydrogen lower heating value, $\mathrm{kJ\,kg}^{-1}$
$\dot{m}$	Mass flow rate, $\mathrm{kg\,s}^{-1}$
m_1	Number of fuel cell strings connected in parallel
m_2	Number of fuel cells connected in series
n_1	Number of battery strings connected in parallel
n_2	Number of battery cells connected in series
n	Molar flow rate, $\mathrm{mol\,s}^{-1}$
P	Power in system model, kW
P	Pressure in fuel cell model, Pa
p	Partial pressure
R	Universal gas constant
R_k	Battery model resistance, Ω
spA	Shore power availability
SOC	Battery state of charge
T	Temperature, K
t_m	Membrane thickness, m
V_k	Voltage for fuel cell system, V
V_D	System volume for system model, m^3
V_x	Alternative system volume, m^3
W_x	Alternative system weight, kg
W_D	Original diesel system weight, kg
Z	Sizing decision variable vector

Greek Symbols

β	Concentration overpotential coefficient
γ_a	Air specific heat ratio
γ_m	Electrode surface roughness factor
δ	Degradation function
Δt	Time step, s
η_c	Air compressor efficiency
η_m	Air compressor motor efficiency

η_{fc}	Fuel cell stack net efficiency
η	Efficiency
λ	Degree of humidification
σ_k	Price, USD
σ_m	Fuel cell membrane conductivity, $\mathrm{S\,m^{-1}}$
υ_k	GWP, kg CO_2 $\mathrm{kg^{-1}}$
ψ_{fc}	Fuel cell specific H_2 consumption, $\mathrm{kgkWh^{-1}}$

Chapter 5

Roman Symbols

A	Action space
a	Action
$cost$	Cost incurred in one time step
$\mathbb{E}$	Expectation
N_{max}	Maximum episode
n	Episode
$\mathbb{P}$	Transition probability
P_{dem}	Power, kW
P_{sh}	Shore power, kW
Q	Action-value function
R	Reward function
r	Immediate reward
S	State space
s	State
s_{actual}	Actual state
spA	Shore power availability
SOC	Battery state of charge
T	Time horizon
t	Time step, s
x	Fuel cell power state

Greek Symbols

α	Learning rate
γ	Discount rate

ε	Exploration probability
Δt	Time step, s
δ	Degradation function
σ_k	Price, USD
υ_k	GWP, kg CO_2 kg^{-1}
ψ_{fc}	Fuel cell specific H_2 consumption, $kgkWh^{-1}$

Chapter 6

Roman Symbols

A	Action batch
a	Action
$cost$	Cost incurred in one time step
D	Mini-batch size
j	$j-th$ element of the mini-batch
K	Time step when battery charged to higher state of charge limit
L	Loss function
M	Replay memory size
N_s	Episode number when α and ε start to be constants
P_{dem}	Power demand, kW
p_{dem}	Normalised power demand
Q	Q network
$\hat{Q}$	Q-target network
R	Reward batch
r	Immediate reward
S	Current state batch
S'	Next state batch
s	State
s'	Next state
T	Total time steps of one power profile
t	Time step, s
Y	Target value batch
y	Target value
Z	Training interval in time steps

Greek Symbols

α	Learning rate
β_1	Exponential decay rate for the first moment estimates of Adam optimiser
β_2	Exponential decay rate for the second moment estimates of Adam optimiser
γ	Discount factor
δ	Temporal difference
θ	Q network parameters
θ^-	Q-target network parameters
τ	Q-target network soft-update rate
ϕ	transition sequence

Chapter 7

Roman Symbols

a	Action vector
$\tilde{a}$	Action given by target actor network with noise
a_k	$k-th$ $(k = 1,2,\ldots,m-1,m)$ PEMFC cluster per unit power adjustment
a_M	PEMFC maximum allowed per unit power adjustment amplitude
a_{M-}	PEMFC maximum allowed per unit power decreasing limit
a_{M+}	PEMFC maximum allowed per unit power increasing limit
c	Policy noise clip factor
D	Mini-batch capacity
d	Delayed policy update step
$\mathbb{E}$	Expectation
J	Expected return
M	Experience memory size
m	PEMFC cluster number
$\mathscr{N}$	Gaussian noise
$P_{fc,rated}$	Total installed fuel cell power, kW
P_c	PEMFC cluster power, kW

P_1	Delivered fuel cell power after the power converters, kW
Q_{θ_i}	$i-th$ critic network
$Q_{\theta'_i}$	$i-th$ critic target network
R_t	discounted sum of rewards from time step t
r	Immediate reward
s	State
s'	Next state
t	Time step, s
y	Target value
Z	Training interval in time steps

Greek Symbols

α	Learning rate
β_1	Exponential decay rate for the first moment estimates of Adam optimiser
β_2	Exponential decay rate for the second moment estimates of Adam optimiser
γ	Discount factor
σ	Standard deviation of policy exploration noise
$\tilde{\sigma}$	Standard deviation of policy value smooth noise
$\delta_{i,j}$	Temporal difference of $j-th$ sample in the mini-batch of $i-th$ critic
ε	Action exploration noise
$\tilde{\varepsilon}$	Clipped noise for $\tilde{a}$
$\eta_{1,k}$	Uni-directional power converter efficiency of $k-th$ PEMFC cluster
θ, θ_i	Q network parameters
θ'_i	Q-target network parameters
π	Actor, policy
π_ϕ	Actor network with parameters ϕ
ρ^π	State distribution following policy π
τ	Target network soft-update rate
ϕ	Actor network parameters
ϕ'	Actor target network parameters

Other Symbol

∇	Gradient

Abbreviations

AC	Alternating Current
AFC	Alkaline Fuel Cell
CAPEX	Capital Expenditure
DC	Direct Current
DWT	Deadweight Tonnage
DDP	Deterministic Dynamic Programming
DDPG	Deep Deterministic Policy Gradient
DPG	Deterministic Policy Gradient
DQN	Deep Q-Network
ECA	Emission Controls Areas
EEDI	Energy Efficiency Design Index
EEOI	Energy Efficiency Operational Indicator
EMS	Energy Management System
EU	European Union
EV	Electric Vehicles
GHG	Greenhouse Gas
GWP	Global Warming Potential
HEV	Hybrid Electric Vehicles
HT-PEMFC	High-Temperature PEMFC
IFEP	Integrated Full Electric Propulsion
IMO	International Maritime Organization
LNG	Liquefied Natural Gas
MCFC	Molten Carbonate Fuel Cell
MDO	Marine Diesel Oil
MDP	Markov Decision Process
MOPSO	Multiple Objective Particle Swarm Optimisation

MSE	Mean Squared Error
NSGA	Non-dominated Sorting Genetic Algorithm
NOx	Nitrogen Oxides
OCV	Open Circuit Voltage
OPEX	Operating Expenses
PAFC	Phosphoric Acid Fuel Cell
PEMFC	Proton Exchange Membrane Fuel Cell
PM	Particulate Matter
RL	Reinforcement Learning
SEEMP	Ship Energy Efficiency Management Plan
SFC	Specific Fuel Consumption
SOC	State of Charge
SOFC	Solid Oxide Fuel Cell
SOx	Sulphur Oxides
SSS	Short Sea Shipping
TD3	Twin Delayed Deep Deterministic Policy Gradient
UK	United Kingdom

Chapter 1

Introduction

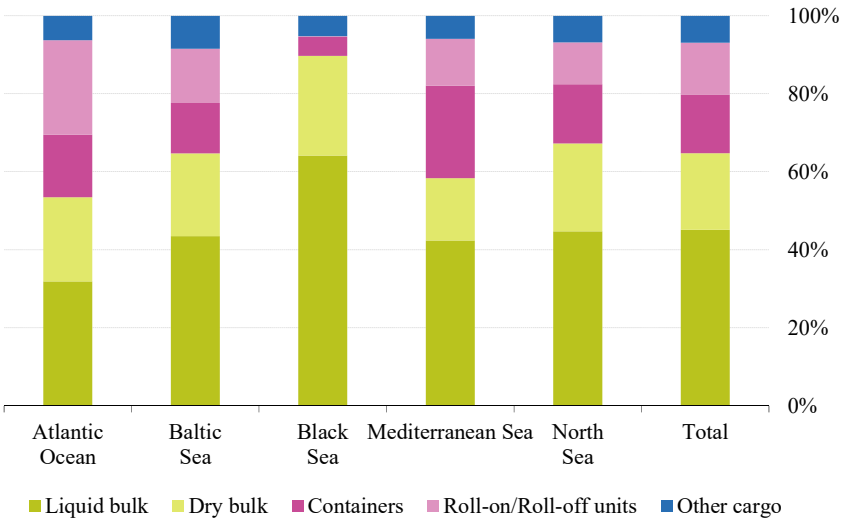
1.1 Coastal shipping

1.1.1 Coastal shipping scope

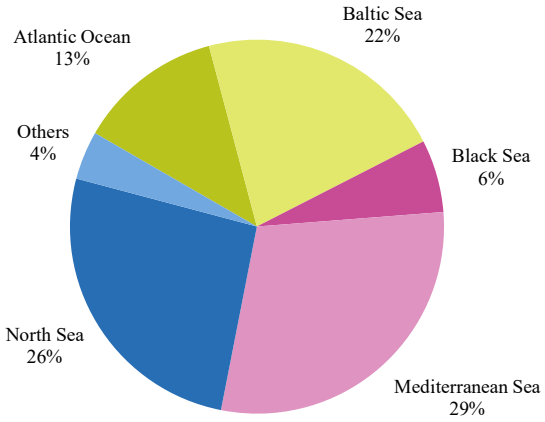
Coastal shipping, also known as Short Sea Shipping (SSS) in the European Union (EU), is the transport of cargo and passengers by sea over relatively short distances that does not involve an ocean crossing (Johnson and Styhre, 2015). It is considered a mode of transport that can alleviate road congestion and provide better energy efficiency by shifting 30% of road freight over 300 km (Douet and Cappuccilli, 2011; Johnson et al., 2014; Eurostat, 2018). Early definition of SSS limits the ship size to a maximum of 5000 gross tonnage (Crilley and Dean, 1993). According to Musso et al. (2002), typical coastal ship sizes can vary from to 1000 to 15 000 Deadweight Tonnage (DWT). However, there is no decisive academic agreement on the definition on coastal shipping (Douet and Cappuccilli, 2011). Nonetheless, the scope of this research project is limited to coastal ships sailing on short routes that do not cross oceans, which could potentially benefit from integrating fuel cells and energy storage technologies.

1.1.2 Coastal shipping overview

Figure 1.1 provides an overview of coastal shipping of the EU in 2016. In 2016, 60% of EU goods transported by sea (19 Gt) were by coastal shipping (Eurostat, 2018). As depicted in Figure 1.1a, liquid and dry bulk are the main cargo types transported via coastal shipping, followed by containers and Roll-on/Roll-off units. The Mediterranean Sea, the North Sea and the Baltic Sea are the main areas where these shipping activities occurred (Figure 1.1b).



(a)



(b)

Figure 1.1: Short sea shipping in the European Union in 2016. (a) percentage of total gross weight of goods transported by sea region and (b) percentage of total gross weight of goods transported for each sea region (source: (Eurostat, 2018)).

1.2 Motivation

1.2.1 Global shipping emissions

The global commercial shipping fleet emitted 2.2% of global anthropogenic Greenhouse Gases (GHGs) in 2012 (International Maritime Organization, 2014a). The International Maritime Organization (IMO) has set a long-term GHG reduction goal for the shipping industry to be in line with the global GHG reduction target to limit the global temperature rise to no more than 2 °C above pre-industrial levels (Vergara et al., 2012; Cames et al., 2015). By 2050, the maritime transport segment will need to reduce its total annual GHG emissions by 50% compared to 2008 levels and achieve zero GHG emissions as soon as is practicable in this century (Figure 1.2) (International Maritime Organization, 2019).

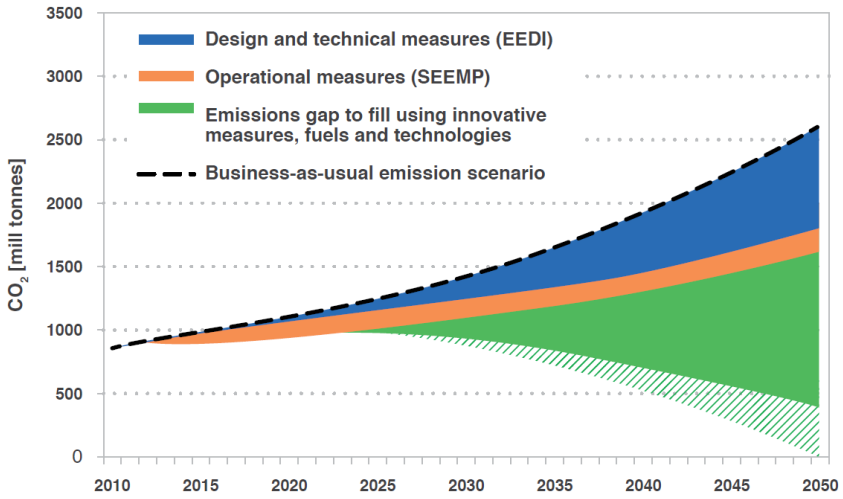


Figure 1.2: IMO strategy on reducing GHG emissions from ships. (source: International Maritime Organization (2019)).

Figure 1.3 shows the evolution of IMO GHG reduction regulations since 1997 (International Maritime Organization, 2015). Two mandatory emission reduction mechanisms, i.e. the Energy Efficiency Design Index (EEDI) and the Ship Energy Efficiency Management Plan (SEEMP) have been introduced by the IMO to guide energy-efficient ship design and operation, respectively.

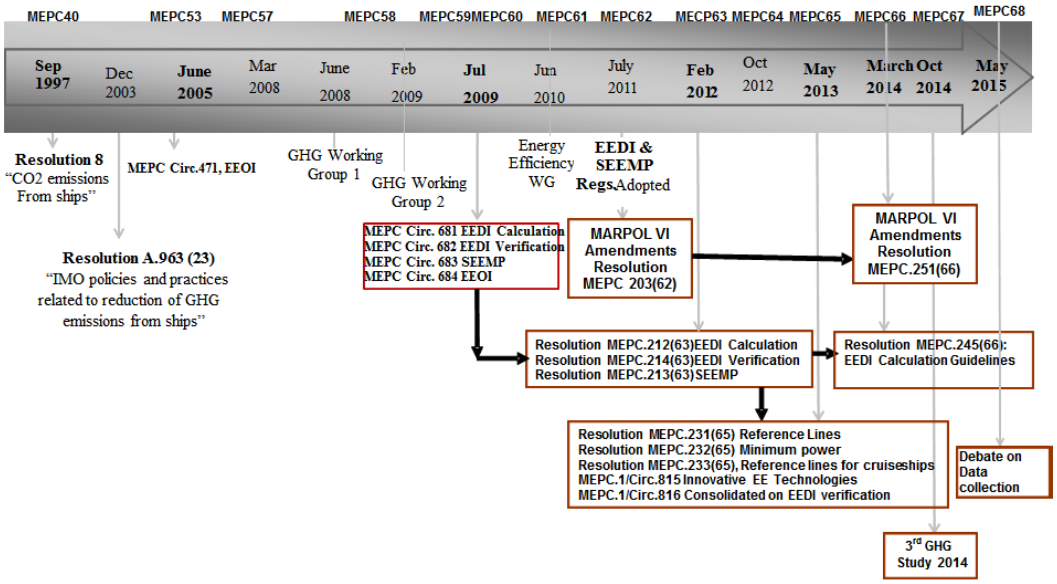


Figure 1.3: IMO energy efficiency regulatory developments. (source: International Maritime Organization (2015)).

In 2009, the EEDI was proposed by the IMO to improve the energy efficiency of new ships by promoting more energy efficiency technologies (International Maritime Organization, 2011). In 2014, the EEDI scope was extended to a broader range of ships (accounting for 85% of global shipping GHG emissions) including Roll-on/Roll-off cargo and passenger ships.

The SEEMP, proposed in 2009, establishes an operational approach to improve the energy efficiency of ships cost-effectively over time using, for example, the Energy Efficiency Operational Indicator (EEOI) as a monitoring tool for setting energy efficiency targets and evaluating energy efficiency levels (International Maritime Organization, 2009, 2016).

1.2.2 Coastal shipping emissions

Although typical coastal ship sizes are much smaller than those vessels used for ocean transits, the emissions from coastal ships make up a significant proportion of the total emissions from the global commercial fleet. It has been estimated that 70% of global shipping emissions are produced within 400 km of coastlines, where

almost half of the world's population resides (Corbett et al., 1999; Viana et al., 2014; Kay and Alder, 2017). Aulinger et al. (2016) indicate that smaller ships (less than 60 000 gross tonnage) consumed 16.4% of the diesel fuel of the whole shipping fleet in the North Sea region. The author's earlier work (Wu et al., 2016) suggests that ships of less than 15 000 DWT mainly operate in coastal waters, and these ships account for approximately 25% of global shipping emissions.

The EEDI and SEEMP mechanisms focus primarily on energy efficiency and GHG emissions, while Emission Control Areas (ECAs) (Figure 1.4) have also been designated to limit SOx, NOx and Particulate Matter (PM) emissions, mainly in coastal waters (International Maritime Organization, 2014b; Zhen et al., 2018; Xia et al., 2019). The PM emissions are only limited within the North American area and the United States Caribbean Sea areas. The Baltic sea and the North sea areas only limit SOx but not NOx. From 2020, the sulphur content in marine fuels is limited to no more than 0.5%.

Figure 1.4: Emission Control Areas (source: (International Maritime Organization, 2014b; Zhen et al., 2018; Xia et al., 2019)).

Almost 15% of total global anthropogenic NOx emissions and 4-9% of SOx emissions are from shipping which is a significant and increasing source of air pollutants (Endresen et al., 2003; Eyring et al., 2010; Viana et al., 2014). Nearly 70%

of these emissions are estimated to occur within 400 km of coastlines, which is a potential influence on air quality in coastal areas where 45% of the world's population resides (Corbett et al., 1999; Viana et al., 2014; Kay and Alder, 2017). SOx emissions can harm human respiratory systems and cause damage to sensitive ecosystems by contributions to acid rain (Halkos and Tsilika, 2019). Long-term exposure to NOx can cause respiratory and lung cancer deaths (Lu et al., 2016). According to Sofiev et al. (2018), even with cleaner marine fuels, ship-related $PM_{2.5}$ will still account for approximately 2.5×10^5 deaths and 6.4×10^6 childhood asthma cases annually. There is a necessity to cut down emissions from ships, especially those operating in coastal waters.

Marine diesel engines, as the primary sources of the pollutants and GHG emissions from coastal shipping, are not a viable long-term powering solution as regards the required reduction in GHG and pollutants (Deniz and Zincir, 2016). Various efforts have been investigated to improve the existing diesel engine based technologies, e.g. using Liquefied Natural Gas (LNG) in dual-fuel and gas engines (Thomson et al., 2015). Though the NOx, SOx and PM emissions can be reduced significantly with LNG, the GHG savings offered by LNG is limited to no more than 21%, and methane slip could potentially cancel out that benefit (Brynolf et al., 2014; Thomson et al., 2015; Ekanem Attah and Bucknall, 2015). Ships operating within Emissions Control Areas have adopted exhaust gas treatment devices which could potentially lead to the GHG emission performance being even worse as a consequence of the additional power requirement and the negative impact on engine efficiency (Verschaeren et al., 2014; Di Natale and Carotenuto, 2015; Boscarato et al., 2015).

1.3 Alternative powering solutions

There is a need for the shipping industry to develop alternative power and propulsion plants to meet these emission reduction goals. Proton Exchange Membrane Fuel Cell (PEMFC) and lithium battery technologies have seen sustained development

and are now commercially available, though improvements are still expected from both (Sharaf and Orhan, 2014; Larcher and Tarascon, 2015). Coastal ships operating on short routes at modest speed can potentially benefit from hybrid propulsion systems with PEMFC and batteries by utilising clean H_2 energy and shore power (McConnell, 2010; de Troya et al., 2016; van Biert et al., 2016). However, factors including both the power and energy density of the power sources, degradation characteristics, energy properties and operational requirements can vary from place to place and influence the design of such a hybrid propulsion system significantly (Wu and Bucknall, 2018). For instance, in 2016, the average GHG emission for grid power in the UK was 281 g CO_2kWh^{-1}, while it was 166 g CO_2kWh^{-1} in Denmark (European Environment Agency, 2019). A holistic design methodology and a suitable Energy Management System (EMS) are required to inform an overall optimum alternative propulsion system design.

PEMFC operating on H_2 has gone through rapid development in recent decades with improving performance and reducing cost and has been adopted by parallel industries such as road transport (Alaswad et al., 2016; Dicks and Rand, 2018). There has been an increasing interest in utilising PEMFC for ship applications (van Biert et al., 2016; Choi et al., 2016; Sasank et al., 2016; de Troya et al., 2016). Compared to other fuel cell types, such as the solid oxide fuel cell or the molten carbonate fuel cell, PEMFC offers better power density and transient performance but lower efficiency and less fuel flexibility (Sharaf and Orhan, 2014). When used in transport applications, PEMFC is typically used with batteries or supercapacitors to provide better overall efficiency and capability for managing power transients.

Lithium battery technology is evolving rapidly and is recognised as having great potential for utilising renewable energy and improving the performance of existing powering solutions (Luo et al., 2015; Hannan et al., 2017). Batteries can be used as stand-alone or in hybrid configurations in ship propulsion systems. When used in a hybrid configuration, batteries can help optimise the loadings of other power sources (e.g. fuel cells or internal combustion engines). Ovrum and Bergh (2015) reported that a 30% fuel reduction could be achieved with regenerative en-

ergy recovered from ship crane operations. In configurations without regeneration, much less fuel saving can be expected, depending on the configuration and power profiles. 15% fuel saving is reported in a case of a non-regenerative hybrid ship (Stefanatos et al., 2015). When only a battery is used, the shipboard system becomes zero-emission, but the life-cycle emission performance depends on the shore-generated electricity. Currently battery technologies are still constrained by limited energy densities, even though the world's largest battery package in ship applications has reached a capacity of 6.1 MWh (Larcher and Tarascon, 2015; Corvus, 2019a).

Although H_2 appears as a clean fuel when the scope is only limited to the propulsion system, the H_2 production process (e.g. natural gas reforming) could be carbon-intensive (Acar and Dincer, 2014). Also, the power generation process that provides the energy for H_2 production and battery charging would have a certain carbon intensity, and this intensity would vary depending upon location. Other factors including power source power capacity, energy densities, degradation characteristics, energy properties and operational requirements can differ from place to place and significantly influence the design of such a hybrid propulsion system.

The hybridisation of PEMFC and Lithium-ion batteries for coastal ship propulsion systems may potentially offer beneficial emission performance. However, the design of such hybrid propulsion systems must consider a series of variables to achieve an overall optimal design. As shown in Figure 1.5, the alternative system has to provide sufficient and reliable power, without occupying too large a volume and mass margins from the propulsion plant. Moreover, the alternative system needs to be more environmentally friendly yet commercially competitive. When multiple power sources are integrated into one propulsion system, an effective EMS is essential to manage power flows.

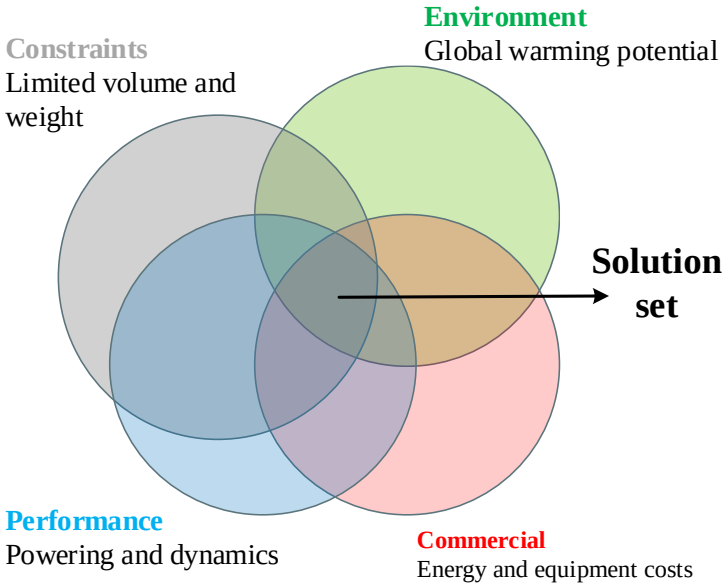


Figure 1.5: The complexity of alternative propulsion system design.

1.4 Research questions

This research aims to answer the following: How can coastal shipping substantially reduce harmful GHG emissions from their propulsion and power systems using fuel cells and batteries and remain attractive commercially? This key research question is split into sub-questions as follows:

1. What are the opportunities and constraints of new technologies and how do they compare to traditional technologies for the design of propulsion and power systems for coastal shipping?

2. How can fuel cells and batteries be best integrated into an electric propulsion system for coastal ships?

3. How do the fuel cell and battery propulsion and power system designs perform and compare with conventional arrangements?

4. For a given hybrid propulsion and power system configuration, how much saving can be achieved from reinforcement learning and deep reinforcement

learning based energy management strategies using continuous monitoring data?

1.5 Thesis outline

This thesis is organised as follows:

Chapter 1 introduces the motivation behind this research project and provides an overview of the project and this thesis. It also describes the research contributions with the author's publications listed.

Chapter 2 details a critical literature review of existing diesel engines and their improvement technologies, alternative fuels, fuel cell technologies, energy storage systems and marine hybrid-electric propulsion system architecture. This chapter aims to identify the opportunities and constraints of alternative power sources in comparison to existing systems, and to further identify the challenges of integrating alternative power sources to coastal ships.

Chapter 3 provides justifications for the research questions via a critical gap analysis and an overview of methodologies adopted in this study. The main methodologies are constrained multi-objective optimisation for hybrid propulsion system design optimisation, reinforcement learning (RL) and deep RL energy management strategies for hybrid propulsion systems.

Chapter 4 deals with the problem of optimal sizing of alternative propulsion and power systems utilising grid power, based upon a proposed plug-in hybrid fuel cell and battery propulsion system model. A two-layer multi-objective optimisation sizing methodology is proposed considering both economic and environmental performance. Such a methodology can be used to guide practical shipboard alternative propulsion system design with variable energy and power source properties.

Chapter 5 formulates and solves the optimal online power split problem for the proposed plug-in hybrid system using RL in discrete state and action spaces. The RL agents are trained with a set of historical power profiles and validated with data over another period.

Chapter 6 extends discrete state spaces in Chapter 5 to continuous state spaces utilising deep neural networks as function approximators to achieve higher resolution.

Chapter 7 further improves Chapter 6 by extending both the discrete state and action spaces to continuous spaces. This chapter also extends the optimal control of fuel cell actions to multiple clusters. In Chapter 5 and 6, the fuel cells are controlled uniformly.

Chapter 8 concludes this project and recommends future work.

1.6 List of achievements

1.6.1 Journal publications

- **Wu, P.**, Partridge, J. & Bucknall, R. (2020). Cost-effective reinforcement learning energy management for plug-in hybrid fuel cell and battery ships. Submitted to *Applied Energy*. Accepted.

- **Wu, P.** & Bucknall, R. (2020). Hybrid fuel cell and battery propulsion system modelling and multi-objective optimisation for a coastal ferry. *International Journal of Hydrogen Energy*. DOI: 10.1016/j.ijhydene.2019.11.152

1.6.2 Conference publications

- Yin, Q., Li, M., Liu G., **Wu, P.** & Bucknall, R. (2019). The design of energy efficiency management system for an electric propulsion passenger ship in inland river. In *5th International Conference on Transportation Information and Safety*. Liverpool, UK.

- **Wu, P.** & Bucknall, R. (2018). On the design of plug-in hybrid fuel cell and lithium battery propulsion systems for coastal ships. In *Marine Design XIII: Proceedings of the 13th International Marine Design Conference (IMDC 2018)*. Helsinki, Finland: CRC Press.

- Farrier, L. A., **Wu, P.** & Bucknall, R. (2017). Opportunities and constraints of electrical energy storage systems in ships. In *Shipping in Changing Climates Conference 2017*. London, UK.

- **Wu, P.** & Bucknall, R. (2016). Marine propulsion using battery power. In *Shipping in Changing Climates Conference 2016*. Newcastle, UK.

1.6.3 Technical reports

- Suarez de la Fuente, S., **Wu, P.**, Partridge, J. & Bucknall, R., (2018) Marine applications of electric turbo compounding for waste energy recovery: identifying opportunities, technical constraints and environmental impact through engine modelling.

- Smith, T., Raucci, C., Haji Hosseinloo, S., Rojon, I., Calleya, J., Suarez de la Fuente, S., **Wu, P.** & Palmer, K. (2016). CO_2 emissions from international shipping: Possible reduction targets and their associated pathways.

- **Wu, P.**, Osses, J., Bhatt, A. & Bucknall, R., (2016). Marine applications of electric turbo compounding for waste heat recovery–market potential for ETC in Shipping.

1.6.4 Conference presentations

- 'On the design of plug-in hybrid fuel cell and lithium battery propulsion systems for coastal ships', presented at *the 13th International Marine Design Conference (IMDC 2018)*, Helsinki, Finland.

- 'Opportunities and constraints of electrical energy storage systems in ships', presented at *Shipping in Changing Climates Conference 2017*, London, UK.

- 'Marine propulsion using battery power', presented at *Shipping in Changing Climates Conference 2016*, Newcastle, UK.

1.6.5 Awards

- Dean's Prize from University College London Faculty of Engineering Sciences for this PhD research.
- WEGEMT–EU Marine University Association subsidy for attending the course of *Green Shipping: Energy, Emissions & Economy* in Denmark, May 2019.

Chapter 2

Literature review

2.1 Overview

This literature review draws on four key themes. These themes are conventional diesel-based powering solutions, fuel cell technologies, energy storage systems and the system architecture integrating the fuel cells and batteries. The scope is limited to coastal ships sailing on short routes, which could benefit from integrating fuel cells and batteries. This chapter presents the literature review in four parts:

- The first part aims to provide an overview of the baseline diesel-based powering solutions of coastal shipping (Section 2.2).
- The aim of the second part is to offer a review of fuel cell technologies and their applications in marine applications (Section 2.3), the intention of which is to identify suitable fuel cell technology with potentials and constraints provided.
- The aim of the third part is to provide a review of energy storage technologies (with a focus on batteries) and their application in marine propulsion (Section 2.4), the intention of which is to identify suitable energy storage technologies with potentials and constraints highlighted.

- The fourth part of this literature review is aimed to identify suitable hybrid-electric system architectures integrating fuel cells and batteries for coastal shipping (Section 2.5).

2.2 Review of diesel engine based powering systems

2.2.1 A historical review of conventional powering solutions

Figure 2.1 presents a timeline with the historical highlights of the marine vessel's power system development from 1900 to 2017 (El-Gohary, 2013; Barnes, 2014; Skjong et al., 2015; Geertsma et al., 2017). From the 1900s to 1960s, steam turbines were the primary powering solutions for commercial ships, notwithstanding the evolving diesel engine technologies had gradually begun to be used in small ships. Since the 1960s, driven by the need of better fuel economy, marine diesel engines have been the dominant power sources for commercial ships (Barnes, 2014). In 1969, Wärtsilä delivered five 16 080 DWT container ships powered by medium-speed diesel engines. In 1987, the cruise ship *Queen Elizabeth 2* was retrofitted with diesel-electric propulsion, replacing its original steam turbines.

However, diesel engines are not a perfect solution. Technologies such as waste heat recovery and exhaust gas recirculation have been introduced to improve efficiency and reduce emissions (Sprouse III and Depcik, 2013; Baldi and Gabrielii, 2015; Pan et al., 2019). Since 1997, the International Maritime Organisation (IMO) has introduced a series of regulations to limit GHG emissions and pollutants (International Maritime Organization, 2014b, 2015). Since 2010, batteries have been applied to commercial ships in hybridisation with diesel engines to achieve better overall efficiency by load levelling. For instance, an 800 kW h battery is fitted to the ferry *MV Hallaig* with its 3.5 MW generating sets (Geertsma et al., 2017). Also, fuel cells have emerged as power sources for marine applications since 2000 (Pratt and Klebanoff, 2016). For example, in 2015, the *SF-BREEZE* was designed with a total installed PEMFC power of 4920 kW to achieve zero emissions.

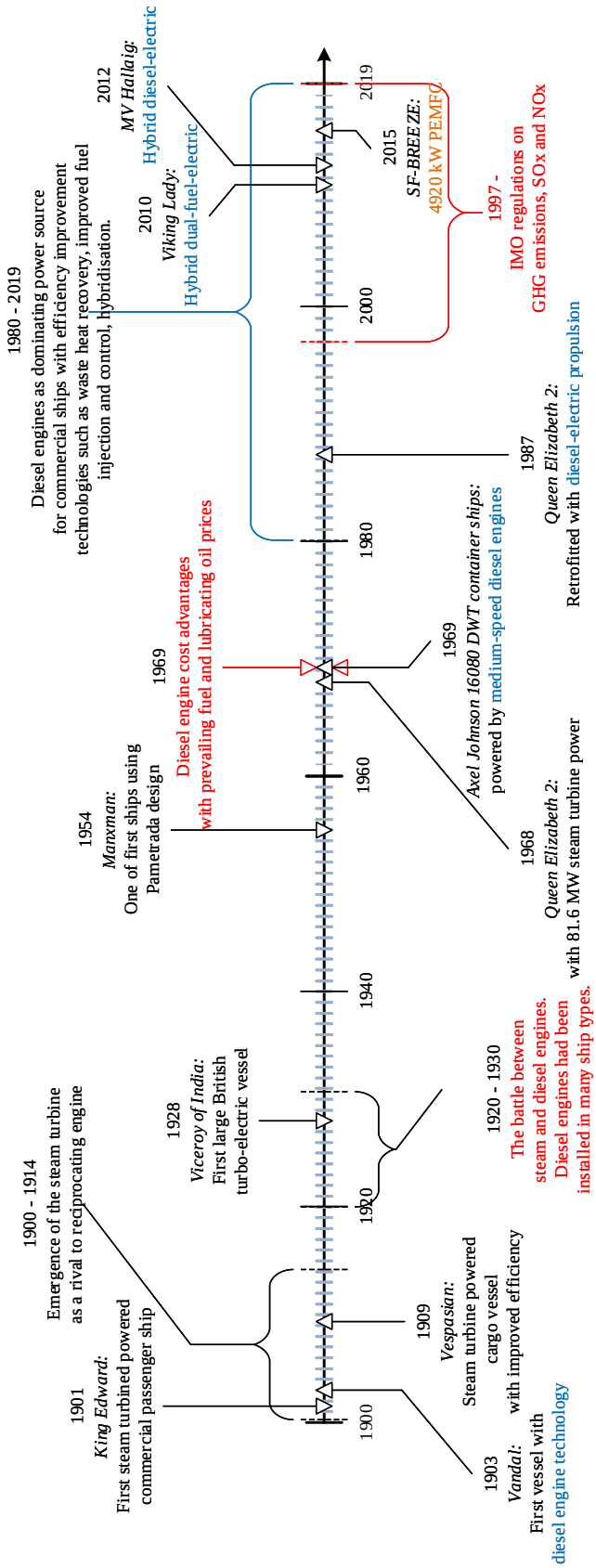
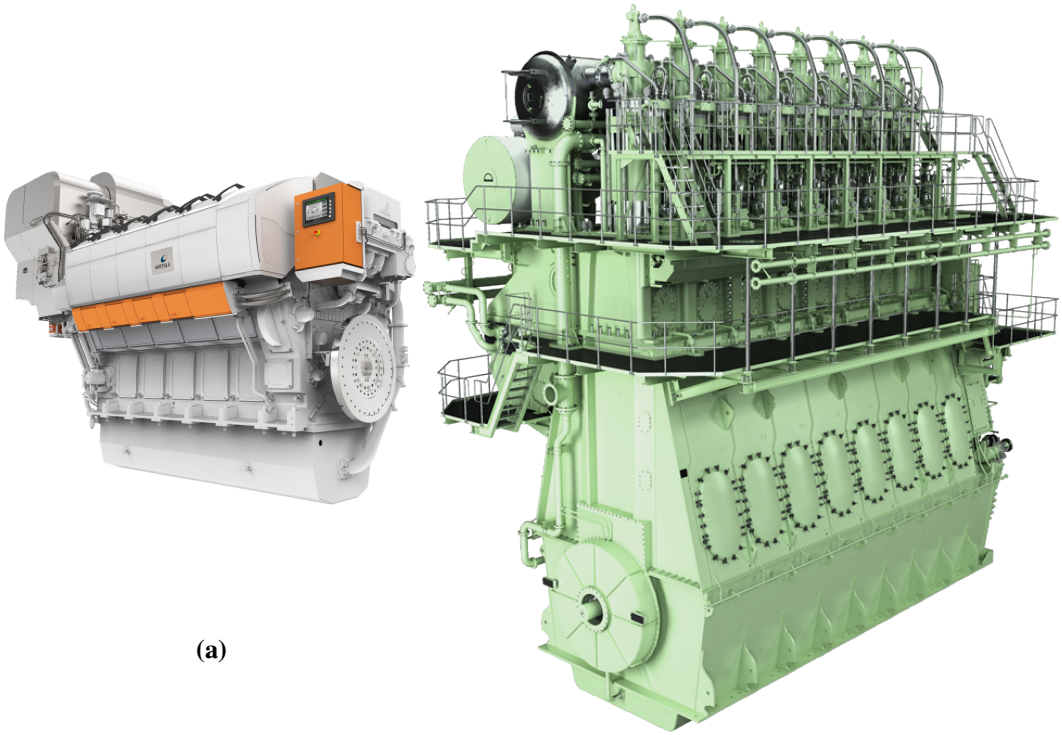


Figure 2.1: Timeline with the historical highlights of the marine vessel's power system development from 1900 (El-Gohary, 2013; Barnes, 2014; Skjong et al., 2015).

2.2.2 The state-of-the-art of diesel engine based systems

Figure 2.2 shows the two types of most widely used diesel engines (4-stroke medium-speed in Figure 2.2a and 2-stroke low-speed in Figure 2.2b). Large 2-stroke low-speed diesel engines (2.2b (MAN, 2019)) with output power of up to 80 MW are mainly used as the main propulsion engines for large ships (Carlton et al., 2013; Barnes, 2014). 4-stroke medium and high-speed engines, which are typically much smaller in size and power output but with higher power density, can either be used as main propulsion engines (through reduction gears) or prime movers of generating sets (Woud and Stapersma, 2002; Mollenhauer et al., 2010; Barnes, 2014). These 4-stroke diesel engines widely used by coastal ships with a typical Specific Fuel Consumption (SFC) of around 195 $gkWh^{-1}$ are approximately 10% less efficient than large 2-stroke engines (Woud and Stapersma, 2002; El-Gohary, 2013). A SFC of 195 $gkWh^{-1}$ corresponds to a specific CO_2 emission of 601 $gkWh^{-1}$.



(a) **(b)**

Figure 2.2: Marine diesel engines. (a) 4-stroke medium-speed and (b) 2-stroke low-speed diesel engines (source: (Wartsila, 2019; MAN, 2019)).

2.2.3 Issues of diesel engine based systems

The global commercial shipping fleet emitted 2.2% of the global anthropogenic Greenhouse Gases (GHGs) in 2012 (International Maritime Organization, 2014a). By 2050, the maritime transport sector needs to reduce its total annual GHG emissions by 50% compared to 2008 to be in line with the global GHG reduction target to limit the global temperature rise to no more than 2 °C above pre-industrial levels (Vergara et al., 2012; Cames et al., 2015). On the other hand, according to Endresen et al. (2003); Eyring et al. (2010); Viana et al. (2014), almost 15% of total global anthropogenic NOx emissions and 4–9% of SOx emissions are from shipping which is a significant and increasing source of air pollutants. Moreover, the studies of Corbett et al. (1999); Viana et al. (2014); Kay and Alder (2017) suggest that nearly 70% of these emissions are estimated to occur within 400 km of coastlines, which has a potential influence on air quality in coastal areas where 45% of the world's population resides.

A timeline of IMO emission regulations on SOx and NOx is shown in Figure 2.3 (Carlton et al., 2013; International Maritime Organization, 2014b). The fuel oil sulphur mass percentage limits are subject to a series of step changes over the years. From 01 January 2020, the sulphur content is limited to no more than 0.5% globally. Moreover, within the Emission Controls Areas (ECAs), Regulation 13 of MARPOL Annex VI controls the engine emissions within the Tier III limits (these vary with engine speed, e.g. $2.5\,\mathrm{gkWh}^{-1}$ for an engine speed of 1000 rpm) (International Maritime Organization, 2014b).

2.2.3.1 SOx

SOx emission, as a significant pollutant from marine diesel engines has been limited in ECAs (Lindstad et al., 2015). SOx emission is a product of the sulphur content of the fuels. To limit SOx emissions, two approaches can be adopted, either by using low sulphur fuel and/or an exhaust gas after-treatment device such as a scrubber, which can be of the wet, dry or hybrid types (Lloyd's Register, 2015). The price

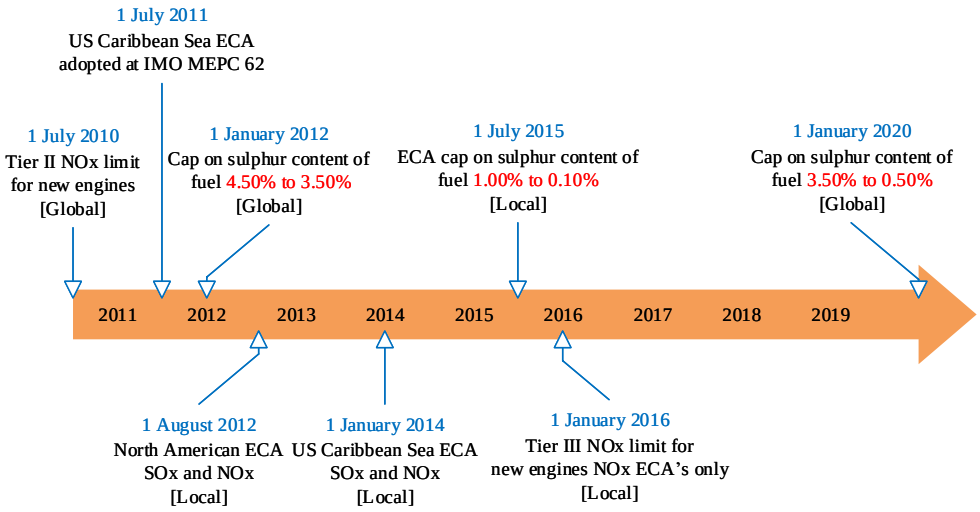


Figure 2.3: Timeline of IMO emission regulations on pollutants (Carlton et al., 2013; International Maritime Organization, 2014b).

of low sulphur fuel with 0.1% sulphur content has been estimated to be 70–80% higher than heavy fuel oil (Holmgren et al., 2014). Instead, sulphur scrubbers have been widely adopted to meet the ECA requirements (Jiang et al., 2014; Panasiuk and Turkina, 2015). However, such devices inevitability will require additional power, which will further impact the overall energy efficiency and GHG emission performance.

2.2.3.2 NOx

NOx emissions are formed during the process of combustion and is highly dependent on the local flame temperature (Zhang et al., 2019). Technologies such as exhaust gas recirculation, selective catalytic reactors, adjusted timing and fuel injection strategies can be used to reduce NOx emissions (Guo et al., 2015). However, these NOx abatement technologies could impact engine efficiency due to reduction in combustion temperature (Pan et al., 2019). Verschaeren et al. (2014) indicate that up to 70% NOx reduction could be achieved with exhaust gas recirculation and Miller timing. However, the drawbacks, including reduced fuel efficiency and other emissions like CO and Particulate Matter (PM) need further investigation.

2.2.3.3 Greenhouse gases

Carlton et al. (2013) argue that although marine diesel engines are reliable and have been primary power sources for the majority of commercial ships for decades, they could not be the long-term powering solution for future ships without step changes to meet the GHG emission reduction goal. Such an argument has been supported by El-Gohary (2013); Barnes (2014); Skjong et al. (2015). Marine diesel engines are typically optimised for a specific load region (e.g. 80% of rated power); operating the engines under non-optimum load would lead to even poorer fuel economy and emission performance, especially under low load conditions (Woud and Stapersma, 2002; Wartsila, 2019). The specific CO_2 emission of 601 $gkWh^{-1}$ is more than double that of UK national grid electricity GHG level (281 $gkWh^{-1}$) in 2016 (Eurostat, 2019). That is, the UK grid electricity GHG emission performance is more than two times better than that of the marine diesel engines.

2.2.3.4 Improvement technologies to diesel engines

There have been various energy efficiency improvement technologies investigated for marine diesel engines (Smith et al., 2010; Calleya, 2014; Scott et al., 2017). Among these technologies, waste heat recovery technologies have been widely applied to marine diesel engines to recover waste heat mainly from exhaust gases (Suárez De La Fuente, 2016; Baldi and Gabrielii, 2015). Singh and Pedersen (2016) indicate that the primary waste heat recovery for marine applications the include Rankine Cycle, Organic Rankine Cycle, Supercritical Rankine Cycle, Kalina Cycle, exhaust gas turbine systems (hybrid turbocharging, mechanical turbo-compound system, hydraulic turbo-compound system and electrical turbo-compound system) and thermoelectric generators. However, the energy efficiency improvement that could be offered by such technologies are limited to 15% (Larsen et al., 2013; Sprouse III and Depcik, 2013; Singh and Pedersen, 2016).

When used in a hybrid configuration, energy storage device (e.g. batteries) can optimise the loadings on the diesel engines. Ovrum and Bergh (2015) reported that 30% fuel reduction could be achieved with regenerative energy recovered from the

ship's crane operations. In configurations without regeneration much less fuel saving can be expected. 15% fuel saving is reported in the case of a non-regenerative hybrid ship (Stefanatos et al., 2015). It should be noted that the achievable saving can vary for different applications, depending on the configuration and power profiles.

2.2.3.5 Alternative fuels for diesel engines

As suggested by Raucci (2017); Balcombe et al. (2019), fuel changes are optional approaches to reduce shipping emissions. Table 2.1 provides a comparison between Marine Diesel Oil (MDO) and potential alternative fuels for coastal shipping (Gilbert et al., 2018; Wartsila, 2019). The alternative fuels included in this review are Liquefied Natural Gas (LNG), H_2 and methanol. Note that biofuels are omitted due to the potential of such fuels being heavily constrained by the complex trade-offs with human essentials such as food and water, and carbon neutrality could be unachievable due to carbon content in the fuels and sacrifice of forests for arable land (Florentinus et al., 2012; Balcombe et al., 2019).

Table 2.1: Potential alternative marine fuels for diesel engines in comparison with MDO (Gilbert et al., 2018; Wartsila, 2019).

Fuel	Net calorific value $MJ\,kg^{-1}$	Volumetric energy density $MJ\,l^{-1}$	SFC $gkWh^{-1}$	Operational fuel emission factor $gkWh^{-1}$					
				CO_2	CH_4	N_2O	SOx	NOx	PM
MDO	42.6	38.3	195	601	0.011	0.030	0.367	16.976	0.184
LNG (a)	48.6	21.9	169	473	3.380	0.018	0.003	1.318	0.030
Liquefied H_2	120.0	8.5	>78 (b)	0	0	(c)	0	(c)	0
Methanol	20.0	15.8	381	522	0	0	0	3.050	0

a: the data is for a typical Wartsila 32 duel-fuel engine; additional 3.5 $gkWh^{-1}$ MDO as pilot fuel
b: SFC not provided as a lack of mature pure H_2 engine;
but the efficiency H_2 engine are not expected to be much higher than diesel engines
c: vary with fraction of H_2 energy in H_2 assisted dual fuel engines

Thomson et al. (2015) indicate that application of LNG as a fuel in the marine industry has been developing rapidly in the past decade mainly due to the IMO emission control regulations. Figure 2.4 shows the trend of LNG powered marine engines installed from 2010 to 2015 (Clarksons, 2016). Although it is a fossil fuel, compared to its liquid fuel equivalents, it is far better in terms of SOx (near zero), NOx (less than 10%) and PM (near zero) emissions when used by marine dual-fuel

or gas engines (Carlton et al., 2013; Balcombe et al., 2019). LNG has a better carbon factor compared to MDO and could provide up to 21% of CO_2 saving without considering methane slip (Brynolf et al., 2014). However, it should be noted that methane is the second most prevalent GHG emitted from human activities, and the GHG effect of methane on the climate is 25 times higher than that of CO_2 over 100 years (Boucher et al., 2009). Methane slip occurs during transporting, bunkering, and combustion process of LNG engines, which could cancel out the GHG saving achieved by its advantage in chemical composition (Anderson et al., 2015). To achieve overall GHG emission savings, methane slip must be controlled throughout the life-cycle of LNG.

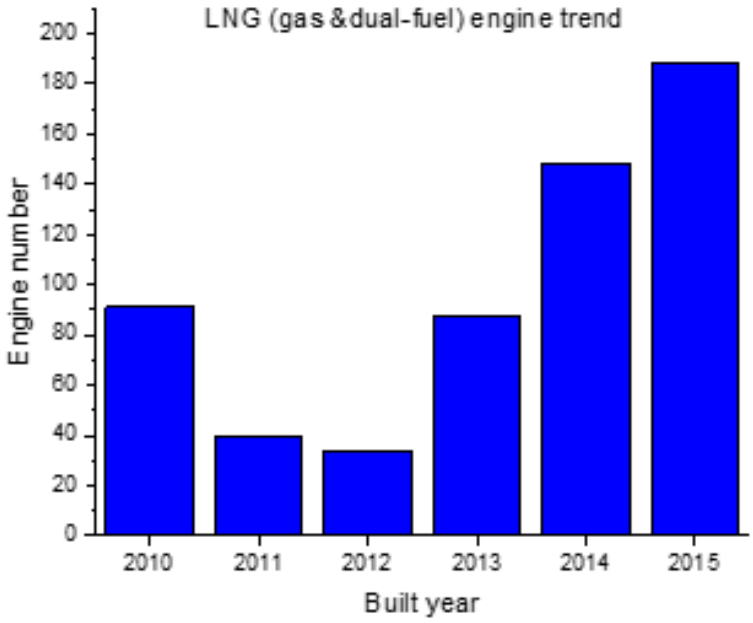


Figure 2.4: LNG fuelled gas and dual-fuel engine number installed on commercial ships built from 2010 to 2015 (data source: (Clarksons, 2016)).

Bicer and Dincer (2018) adopted a life-cycle approach to analyse the potential of utilising H_2 as a marine fuel. Their results indicate that the life-cycle GHG emission could be reduced by about 40% when H_2 is used by dual-fuel engines compared to conventional diesel fuel. Bicer and Dincer (2018) also indicated that the usage of H_2 in marine power and propulsion systems eventually depends on the capability of producing clean and low-cost energy. Dimitriou and Tsujimura (2017)

indicated that developing H_2 fuelled compression ignition engines is still challenging, and most existing H_2 fuelled engine projects focus on using H_2 as an additional fuel to other base fuels such as biofuels. The overall efficiency improvement could be marginal or negative, and NOx is one of the problematic emissions (Zhou et al., 2016; Gurz et al., 2017; Dimitriou et al., 2018). Nevertheless, conversion of H_2 to electricity with fuel cells is typically more efficient than within internal combustion engines, which makes it an effective approach to use H_2 with fuel cells (White et al., 2006; Edwards et al., 2008; van Biert et al., 2016). When used in transportation applications, limited system-level volumetric energy density could be a significant limitation of H_2 (van Biert et al., 2016; de Troya et al., 2016).

It appears that methanol could provide about 13% CO_2 saving compared to MDO (Table 2.1). According to Jenkins (2016), methanol can be produced via two main approaches: natural gas or biomass reforming. However, the natural gas reforming approach requires additional energy, and will inevitably give rise to methane slip during the production process (Hansson et al., 2019; Blumberg et al., 2019). Methanol produced from this approach is not a viable alternative marine fuel from the perspective of GHG reduction. The second approach of producing methanol from biomass, will also inevitably face the problems of biofuels as mentioned at the beginning of this section (Florentinus et al., 2012; Balcombe et al., 2019).

2.2.4 Summary of review on diesel-based systems

As indicated by Barnes (2014); Deniz and Zincir (2016); Balcombe et al. (2019), Marine diesel engines operating with heavy fuel oil or marine diesel oil are not a long-term powering solution for the shipping industry in terms of the required reduction in GHG and pollutants emissions. Various efforts have been made to improve the existing diesel engine based technologies, e.g. using LNG in dual-fuel or gas engines (Thomson et al., 2015; Balcombe et al., 2019). Though the NO_x, SO_x and PM emissions can be reduced significantly with LNG, the GHG savings offered by LNG are limited to no more than 21%, and methane slips could potentially cancel

out the benefit (Brynolf et al., 2014; Thomson et al., 2015; Ekanem Attah and Bucknall, 2015). Ships operating within Emissions Control Areas have adopted exhaust gas treatment devices which could potentially lead to the GHG emission performance being even worse as a consequence of the requirement for additional power and the negative impacts on engine efficiency (Verschaeren et al., 2014; Di Natale and Carotenuto, 2015; Boscarato et al., 2015). Methanol, as an alternative fuel, would only provide very marginal GHG saving. Using H_2 with internal combustion engines remains challenging. For the long-term future, alternative technologies like fuel cells and batteries are promising options.

2.3 Review of fuel cells

A fuel cell is a galvanic device that produces electrical energy from the chemical energy of a fuel (Dicks and Rand, 2018). When H_2 is used as the fuel, fuel cells generate electricity, and the by-product is simply water. Figure 2.5 shows the electrode reactions and charge flow for acid electrolyte (Figure 2.5a) and alkaline electrolyte (Figure 2.5b) fuel cells, respectively (Dicks and Rand, 2018). As in Figure 2.5a, in an acid electrolyte fuel cell, H_2 is oxidised at the anode, creating H^+ by the release of electrons. The H^+ ions then pass through the electrolyte, while the electrons flow from the anode to cathode. By this means electric current is produced, powering external loads (e.g. an electric motor). It should be noted that such an operating principle is identical to the one used by William Grove in 1839 and is still the most widely adopted for commercial fuel cell applications (e.g. PEMFC) (O'Hayre et al., 2016). In a fuel cell with an alkaline electrolyte (Figure 2.5b), OH^- ions produced at the cathode pass through the alkaline electrolyte and react with H_2 fuel at the anode. They then produce H_2O by release of electrons. The electrons flow through an external circuit, powering external loads. This operating principle applies to Alkaline Fuel Cell (AFC). It is worth mentioning that, for other fuel cell types (e.g. Solid Oxide Fuel Cell), the operating principles can vary, depending on the electrolyte used and the reactions at the two electrodes.

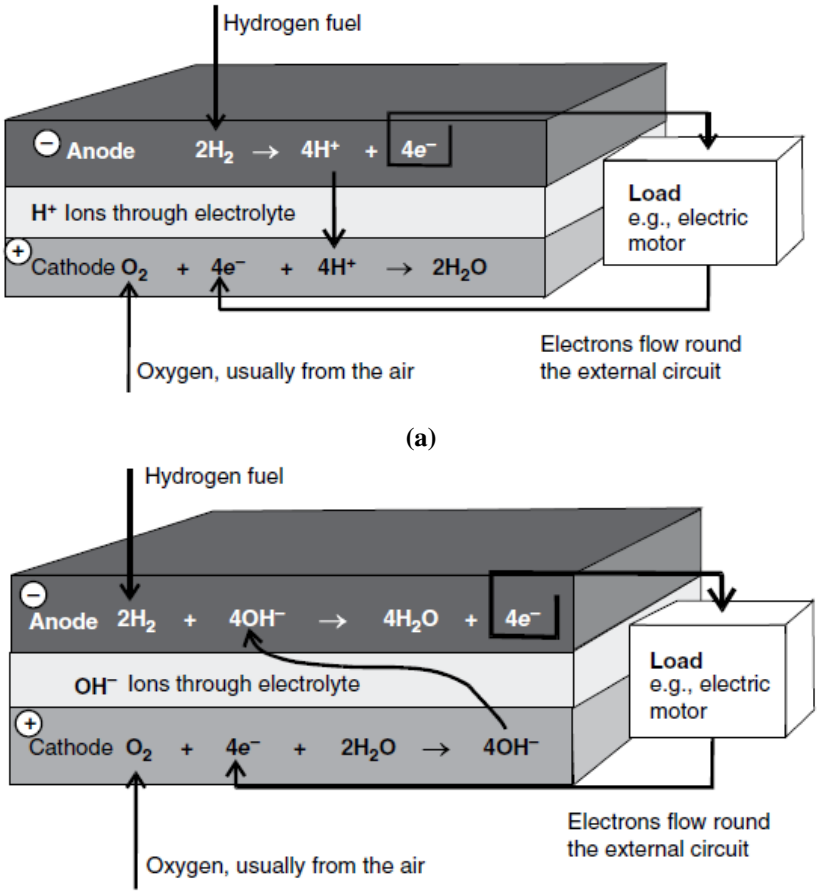


Figure 2.5: Electrode reaction and charge flow for a fuel cell with: (a) an acid electrolyte and (b) an alkaline electrolyte (source: (Dicks and Rand, 2018)).

2.3.1 Fuel cell technologies

There have been a variety of fuel cell technologies developed with specific features for different purposes. For maritime applications, the following aspects should be considered (van Biert et al., 2016): electrical efficiency, power density, energy density (fuel related), load following and system start-up time, reliability, safety and cost. Table 2.2 summarises some relevant features of these major fuel cell technologies (Mekhilef et al., 2012; van Biert et al., 2016; O'Hayre et al., 2016; Wu, 2018; Dicks and Rand, 2018).

Table 2.2: Fuel cell technology comparison (Mekhilef et al., 2012; van Biert et al., 2016; O'Hayre et al., 2016; Wu, 2018; Dicks and Rand, 2018).

Fuel cell type	Electrolyte	Charge carrier	Temperature [°C]	Catalyst	Fuel	Poisonous species	LHV efficiency [%]	Gravimetric density [Wkg^{-1}]	Volumetric density [Wl^{-1}]	Challenges	Commercially available?	Applications	Estimated cost [$\$kW^{-1}$]
LT-PEMFC	Polymer membrane	H^+	65–85	Platinum	H_2	S CO >10 ppm	40-60 -	250–1000	300–1550	Expensive catalysts Sensitive to fuel impurities	Yes	Portable Transport	800–2600
HT-PEMFC	Polymer and phosphoric acid membrane	H^+	100–220	Platinum	H_2	S CO >3%	50-60 -	-	-	Finding suitable electrolyte material	Developing	Portable Transport	-
MCFC	Molten carbonate	CO_3^{2-}	650–700	Molten carbonate	H_2 hydrocarbon fuels	S	40-55	7.75–25	2–20	High temperature corrosion Long start-up time Limited cycling	Yes	Stationary	2600
SOFC	Ceramic	O^{2-}	500–1000	Ceramic	H_2 hydrocarbon fuels	S	45-60	8–80	4–32	High temperature corrosion Long start-up time Limited cycling	Developing	Stationary	1300
AFC	Liquid KOH	OH^-	80	Platinum	Pure H_2	CO_2	60	-	-	Sensitive to CO_2 Electrolyte management and conductivity	Yes	Stationary Transport	6600
PAFC	Liquid H_3PO_4	H^+	150–220	Platinum	H_2	S CO >1.5%	40	16	6.5	Expensive catalysts Long start-up time Sulphur sensitivity	Yes	Stationary	3900

2.3.1.1 Alkaline Fuel Cell

Alkaline Fuel Cell (AFC) operates following the principle detailed in Figure 2.5b. AFCs typically operate within the temperature range of 60–100 °C. Although an AFC can provide LHV efficiency of up to 60%, one of the drawbacks of the AFC is the electrolyte reacts with CO_2, which may be contained within the O_2 or the H_2. Consequently, AFCs cannot even tolerate low concentrations of CO_2 due to the way CO_2 reacts with the electrolyte solution (de Troya et al., 2016). Because of this limitation, the AFC has to use pure H_2 concurrently with O_2 free of CO_2. Such cells have been used in space and transportation applications.

The requirement to use pure O_2 makes the AFC not a practical option for coastal shipping applications due to the additional infrastructures and costs are needed to provide for pure O_2. The high cost of AFCs is another major limitation (Wu, 2018). There is a trend that the AFC is being rapidly replaced the by Proton Exchange Membrane Fuel Cell (PEMFC) in transportation applications (Alaswad et al., 2016). Such cells have never been used in marine applications.

2.3.1.2 Proton Exchange Membrane Fuel Cell

PEMFC also referred to as solid polymer fuel cell, comprises a membrane-electrode assembly, two bipolar plates and two seals (Mehta and Cooper, 2003). PEMFC operates following the principles described in Figure 2.5a. The PEMFC needs to operate with pure H_2 at 60–80 °C and is sensitive to CO as CO can be absorbed by the anode catalyst which consequently leads to blocked active sites which are intended for H_2 oxidation reaction (Stephens et al., 2016; Narayanan and Basu, 2017). Note that traces of CO can be found in H_2 fuel generated via steam reforming of hydrocarbons. Although the PEMFC prices are relatively lower compared to other cell types, due to having to use a precious Pt catalyst, the still high cost and limited life- time are still major limitations of the PEMFC (O'Hayre et al., 2016; van Biert et al., 2016). Nonetheless, as indicated by Thompson and Papageorgopoulos (2019), platinum-group metal-free catalysts and mass production could potentially improve the long-term cost competitiveness of PEMFCs for transportation applications.

With start-up flexibility and relatively better dynamic performance, PEMFCs have been widely used in various applications including submarines, ships and road vehicles, and are commercially available (de Troya et al., 2016; Xie et al., 2019). Another merit offered by the PEMFC is high power density, which is crucial for coastal shipping applications with very limited weight and space margins for their propulsion systems.

The development of High-Temperature PEMFC (HT-PEMFC) could potentially alleviate the poisoning issue inherent to PEMFCs by using polymer and phosphoric acid membranes (Chandan et al., 2013; Rosli et al., 2017). A HT-PEMFC, typically operating in the temperature range of 100–200 °C, is less sensitive to impurities in the fuel and could offer higher overall efficiency, but is still under development (Chandan et al., 2013; Rosli et al., 2017). It should be noted that such cells are not commercially available and could also suffer from the limitation of high cost owing to the use of expensive Pt based catalysts.

2.3.1.3 Molten Carbonate Fuel Cell

The Molten Carbonate Fuel Cell (MCFC) typically operates at high temperatures of 600–700 °C (Dicks and Rand, 2018). Such a high temperature is necessary for the molten carbonate electrolyte to reach an adequate level of conductivity. MCFCs are not sensitive to CO. MCFCs can consume fuels that contain gaseous impurities reformed from natural gas or even coal, which makes it flexible regrading fuels (Dicks, 2004). The high operating temperature of MCFC contributes to an efficiency of approximately 50% (up to 60%). However, high operating temperature requires a long start-up time ($\geq$ 10 h) and leads to high-temperature corrosions, which makes them non ideal for transport applications requiring short-up time (Wu, 2018; Dicks and Rand, 2018).

Although MCFCs are commercially available, high production costs and limited lifetime remain as challenges (van Biert et al., 2016; Wu, 2018). Also, the power density of an MCFC is much lower compared to that of an equivalent PEMFC. It should be noted that although MCFCs can operate with fuels reformed

from hydrocarbons (e.g. natural gas), the GHG savings that could be achieved from the combination of an MCFC with hydrocarbon derived fuels are very limited due to CO_2 inevitably being emitted into the atmosphere without carbon capture.

2.3.1.4 Phosphoric Acid Fuel Cell

The Phosphoric Acid Fuel Cell (PAFC), which uses phosphoric acid (H_3PO_4) as an electrolyte, typically operates in the temperature range of 150–220 °C (Mekhilef et al., 2012). The PAFC is more flexible regarding fuel options compared to the AFC and PEMFC but requires a longer start-up time (1–4 h) (Wu, 2018). Such a long start-up time makes the PAFC impractical for applications in coastal ships which require frequent start-stop of the power sources. The PAFC is commercially available but is only used in stationary applications. Although it is less sensitive to impurities, contamination by carbon monoxide and sulphur compounds, which could poison the anode, still need to be prevented (Mekhilef et al., 2012). The efficiency of a PAFC is 40–50% (Dicks and Rand, 2018).

2.3.1.5 Solid Oxide Fuel Cell

The Solid Oxide Fuel Cell (SOFC) uses a solid ceramic inorganic oxide as the electrolyte and typically has an operating temperature range of 750–1000 °C (Staniforth and Ormerod, 2003). The efficiency can be up to 60% and when combined with gas turbines, even higher efficiencies can be achieved (Ebrahimi and Moradpoor, 2016). A SOFC, being flexible in fuel choices, can generally run on a mixture of H_2 and CO formed by hydrocarbon fuel reforming. The high operating temperature leads to a long start-up time. A limitation in the number of applicable start/stop cycles is one of the issues of the SOFC (Dicks and Rand, 2018). Similar to the MCFC, the system-level power density is much lower compared to that of an equivalent PEMFC due to balancing of plant (van Biert et al., 2016). For these reasons, current SOFCs are not an ideal solution for large scale shipboard applications. Although the SOFC could offer better efficiency compared to low-temperature fuel cells, potential GHG savings that can be achieved from the combination of SOFC and hydrocarbon fuels are limited and could potentially be cancelled out due to issues such as methane

slip.

Based on the comparisons above, the PEMFC is deemed the most promising for application in coastal shipping considering the PEMFC's merits in terms of power density, start-up time, commercial availability, maturity and comparatively lower price driven by the road vehicle applications. Although the HT-PEMFC appears promising, it is still under development, hence not considered in this research. High-temperature fuel cells such as the MCFC and SOFC are not suitable for coastal ships due to long start-up time and low power density. Moreover, any potential GHG savings that could be offered by using high-temperature fuel cells with hydrocarbon fuels are limited due to the life-cycle GHG emissions of hydrocarbon fuels (e.g. LNG).

2.3.2 Fuel cells in marine applications

Figure 2.6 details major fuel cell marine applications from 2003 onwards (Pratt and Klebanoff, 2016; Tronstad et al., 2017). Among the 18 listed projects, the number of ships using PEMFC, HT-PEMFC, SOFC and MCFC are 9, 3, 3 and 3 respectively.

The *FCS Alsterwasser* is a hybrid PEMFC and battery powered 100-passenger boat launched in 2008 (Figure 2.7a) (McConnell, 2010). The *FCS Alsterwasser* is part of the EU funded (€2.4 million) *Zemship* project. Two PEMFC stacks with a total power of 96 kW were fuelled by 12 H_2 tanks (a total of 50 kg H_2 stored at 350 bar). A battery pack was also installed to guarantee operations at peak hours and level the load on the PEMFC while docking and casting-off to prolong PEMFC lifetime (Zemships, 2013). According to (McConnell, 2010), the boat's hybrid PEMFC and battery propulsion system provided double the efficiency of a diesel system. Another PEMFC powered marine example is the *SF-BREEZE* (Figure 2.7e) (Pratt and Klebanoff, 2016). To date, the *SF-BREEZE*, a high speed ferry with a total 4920 kW PEMFC operating on pure H_2, is the largest fuel cell powered concept ship to date .

In the *FellowSHIP* project (2003–2011), a modularised MCFC system was

installed onboard *Viking Lady* to investigate the concept of using a high-temperature fuel cell in marine applications (Figure 2.7b) (Ovrum and Dimopoulos, 2012). The *FellowSHIP* used a prototype 320 kW MCFC with LNG as the fuel. It should be noted that the MCFC output power was kept stable due to the MCFC only being able to handle slow load changes. The measured system-level efficiency of the MCFC at full load is 52%, which is approximately 10% better than the installed diesel engines.

In 2009, the *PA-X-ELL* project was launched to investigate the potential of reducing emissions from cruise ships, yachts and RoPax ferries through the integration of decentralised fuel cell power systems (Tronstad et al., 2017). Three methanol powered HT-PEMFC stacks (total 60 kW) were installed onboard *MS Mariella* (Figure 2.7c). Although it has been claimed that the project was successful, the use of methanol would only provide marginal GHG saving.

In the *SchiBz* project, a 100 kW SOFC was installed onboard *MS Forester*. The installed SOFC stacks operate with low sulphur diesel fuel with an external diesel reformer (Tronstad et al., 2017). Note that the ships shown in Figure 2.7 b to d still operate with the majority of the power being provided from diesel or dual-fuel engines (Tronstad et al., 2017).

It is apparent that PEMFCs with H_2 are the most widely used fuel cell type and logistic fuel combination. The largest marine fuel cell application is the *SF-BREEZE* with a PEMFC power of 4920 kW. HT-PEMFCs, with less stringent requirements on fuels, have been tested in the projects of *E4Ships–MS Mariella*, *RiverCell* and *MF Vagen* with methanol and H_2. Two types of high-temperature fuel cells, i.e. MCFC and SOFC have been tested over 6 projects, operating on carbon or hydrocarbon-based fuels including LNG, diesel and methanol. Such an observation suggests that high-temperature fuel cells are utilised to operate with conventional fuels. It should be noted that the use of conventional fossil fuels does not bring obvious savings in terms of GHG emissions since the efficiency improvements are not hugely significant when compared to diesel engines.

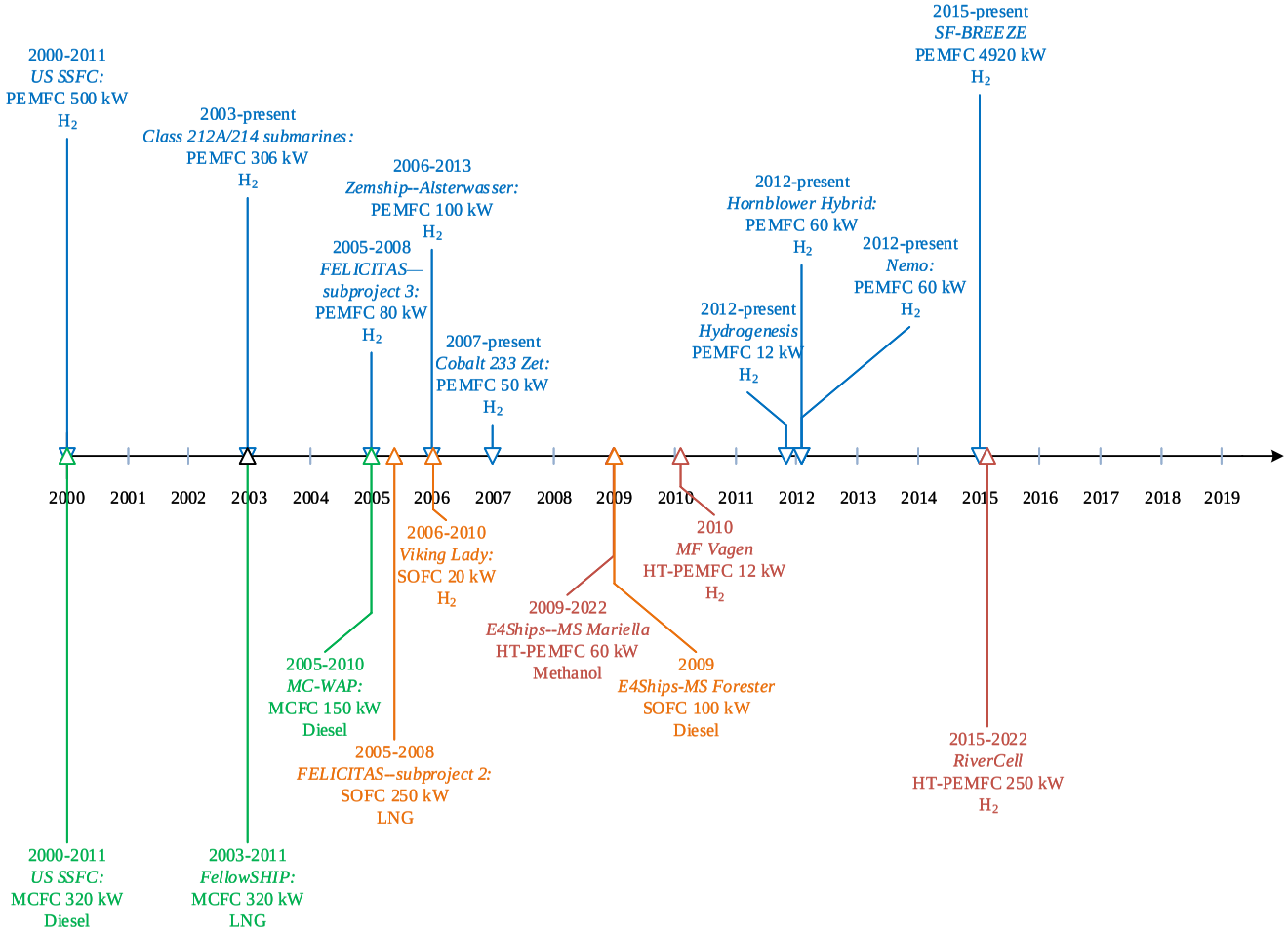


Figure 2.6: Major marine fuel cell applications from 2003 onwards (Pratt and Klebanoff, 2016; Tronstad et al., 2017).

(a) **(b)** **(c)** **(d)** **(e)**

Figure 2.7: Fuel cell powered ships. (a) *FCS Alsterwasser* with 96 kW PEMFC (McConnell, 2010). (b) *FellowSHIP* with 320 kW MCFC (Ovrum and Dimopoulos, 2012). (c) *MS Mariella* with 60 kW HT-PEMFC (Tronstad et al., 2017). (d) *MS Forester* with 100 kW SOFC (Tronstad et al., 2017). (e) *SF-BREEZE* with 4920 kW PEMFC (Pratt and Klebanoff, 2016).

2.3.3 Summary of review on fuel cells

The SOFC and MCFC could offer some efficiency improvement over low-temperature fuel cells (see Table 2.2). However, the life-cycle GHG emission savings that would be realised by using these fuel cells would be limited due to the carbon content in the fuels, without an additional significant efficiency advantage over diesel engines or the low-temperature fuel cells (Stephens et al., 2016). Although either MCFCs or SOFCs could operate with H_2 (Dicks and Rand, 2018), limitations such as long start-up time, low power density, limited cycling number, availability and maturity, make them less appealing for use in coastal ships sailing on short routes that frequent start-stop and manoeuvring.

As depicted by Stephens et al. (2016), low-temperature fuel cells operating with renewable fuels are a more promising option for transportation applications. Although there are still major challenges, such as high production costs, the PEMFC is commercially available, and the cost of production cost is expected to decrease rapidly (Chen et al., 2015; Fletcher et al., 2016; Ehsani et al., 2018). This model of fuel cell can offer desirable life-cycle emission performance provided that the H_2 is produced using clean and low-cost energy (Singh et al., 2015; Ahmed et al., 2016; Gurz et al., 2017; Bicer and Dincer, 2018). The limited system-level energy density of H_2 energy is one of the main obstacles for its maritime applications (van Biert et al., 2016; Raucci, 2017). Nevertheless, for coastal ships operating on short routes, e.g. *SF-BREEZE*, the excessive volumetric impact from H_2 storage could be insignificant, but more frequent replenishment of H_2 may be necessary (Pratt and Klebanoff, 2016).

2.4 Review of energy storage technologies

Table 2.3 compares the main ESS technologies, including Lithium-ion batteries, supercapacitors and flywheels from a range of perspectives, including energy density, power density, lifetime and efficiency. In hybrid-electric propulsion systems, an ESS can be used to store energy from regeneration, to optimise engine loading

conditions and to provide direct propulsion.

Table 2.3: Main ESS technologies (May et al., 2018; Kim et al., 2015a; Hameer and van Niekerk, 2015; González et al., 2016; Farrier et al., 2017; Gür, 2018).

ESS	Energy density [WhL^{-1}]	Power density [WL^{-1}]	Specific energy [$Whkg^{-1}$]	Specific power [Wkg]	Daily Self discharge [% of energy]	Cycle life [cycle]	Lifetime [year]	System efficiency [%]
Lead-acid	60–110	10–400	20-40	75–300	0.2	1200–1800	5–15	50–95
Nickel-Cd	150–300	80–600	50-80	150–300	0.3	1500–2500	10–20	60–83
Lithium-ion	250–675	1500–10000	50-250	500–2000	0.1–0.3	400–9000	15–20	90–99
Flywheel	20–80	1000–2000	10–30	400–1500	$\geq$20% per hour	>1000000	15–20	70–95
Supercapacitor	10–30	>100000	1–10	500–10000	10–20	>1000000	10–20	85–98

2.4.1 Supercapacitors

Supercapacitors are electrochemical energy storage devices. As shown in Figure 2.8, supercapacitors store electric energy by means of an electrolyte solution between two solid conductors which are typically designed to offer large surface areas (Sabri et al., 2016; Farrier et al., 2017) such that large capacitances can be achieved with a large amount of electric charge stored with a small distances between the two electrodes. Supercapacitors can discharge stored electric charge within a short period. As a consequence, supercapacitors are characterised as having high power density and specific power, which makes them ideal for short period high pulse loads. However, supercapacitors offer low energy density, which makes them less suitable for marine propulsion applications (González et al., 2016).

2.4.2 Flywheels

Figure 2.9 illustrates the structure and components of a flywheel (Amiryar and Pullen, 2017). Flywheels store the kinetic energy by using a high-speed rotating flywheel inside a containment vacuum on a set of bearings (Faraji et al., 2017; Amiryar and Pullen, 2017). The flywheel rotor is coupled to a motor/generator which functions as the electromechanical interface between the flywheel and the external system via a bi-directional converter (Luo et al., 2015). High-speed flywheels operate at speeds of up to 1×10^5 rpm. Flywheels have been used in applications such as

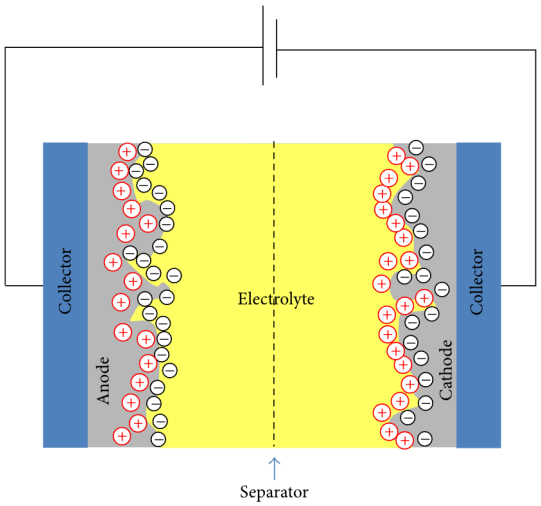


Figure 2.8: Schematic of a supercapacitor (source: (Saleem et al., 2016)).

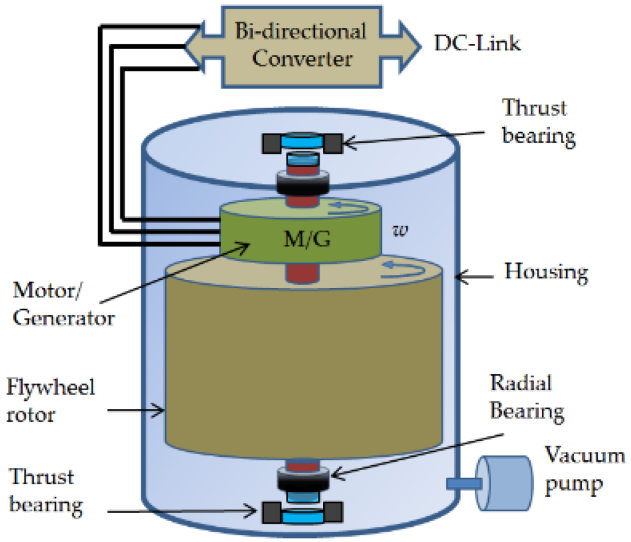


Figure 2.9: Structure and components of a flywheel (source: (Amiryar and Pullen, 2017)).

uninterruptable power supply systems, hybrid vehicles and power quality improvement by storing and releasing energy with specific power up to 1500 Wkg^{-1} (Arani et al., 2017). Flywheels offer superior lifetime and safety characteristics but have limited energy density and need to deliver and recover energy in short periods due to having a high self-discharge rate (Gür, 2018).

2.4.3 Batteries

The Lead-acid battery is a mature, dependable technology but offers low specific energy and specific power. They have an efficiency of around 70% and discharging below 80% will affect its lifetime significantly (May et al., 2018). Nickel-Cd batteries are superior to Lead-acid batteries in terms of capacity and lifetime. In comparison with Lithium-ion batteries, both Lead-acid and Nickel-Cd batteries offer limited energy densities and inferior system efficiencies. Consequently, the two types of batteries are considered less suitable for coastal shipping applications.

Figure 2.10 illustrates the schematic of the first Lithium-ion battery (Goodenough and Park, 2013). Lithium-ion batteries provide the highest energy density with reasonable lifetime and power density (Kim et al., 2015a). Among all the possible ESS solutions, Lithium-ion batteries are the most widely used in the automotive industry and have been developed extensively over the past decade with increasing energy density and reducing production cost (Lu et al., 2013; Jaguemont et al., 2016). They are still costly at present, and a shortage of Lithium and some of the transition metals currently used in Lithium-ion batteries may become an issue in the future. However, a significant shortage of Lithium is unlikely in the near future, and Lithium-ion batteries will probably continue to dominate electrochemical energy storage for years to come (Nitta et al., 2015; Curry, 2017).

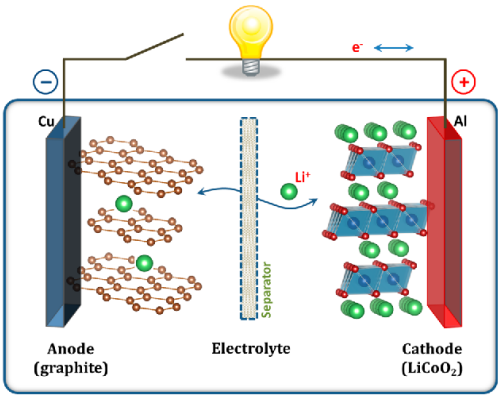


Figure 2.10: Schematic of the first Lithium-ion battery ($LiCoO_2$/Li + electrolyte/graphite) (source: (Goodenough and Park, 2013)).

2.4.4 Battery-powered ships

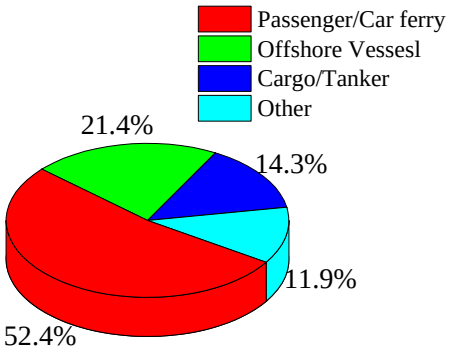
All-electric battery-powered vehicles are becoming an important option for the decarbonisation of road transport by utilising grid power (Anandarajah et al., 2013; Andwari et al., 2017). However, currently all battery vehicles have limited ranges due to limited battery energy density; hybrid plug-in configurations with internal combustion engines or H_2 fuelled PEMFCs have emerged with extended ranges (Sabri et al., 2016; Ehsani et al., 2018). Batteries are also being widely used in non-plug-in hybrid configurations with internal combustion engines or fuel cells to optimise propulsion system performance through load levelling (Cuma and Koroglu, 2015).

In recent years, progress has been made with the development of battery-powered ships. Figure 2.11 shows the battery-powered ship numbers (Figure 2.11b) trend and battery ship categories (Figure 2.11a) from 2011 to the first quarter of 2017 (Alnes et al., 2017). Passenger/car ferries are the main ship type adopting battery power. The main reason for this is passenger/car ferries typically operate on short routes with relatively low power demands, which helps alleviates the energy density issue of batteries. Moreover, operating on relatively fixed routes makes access to charging infrastructures less challenging.

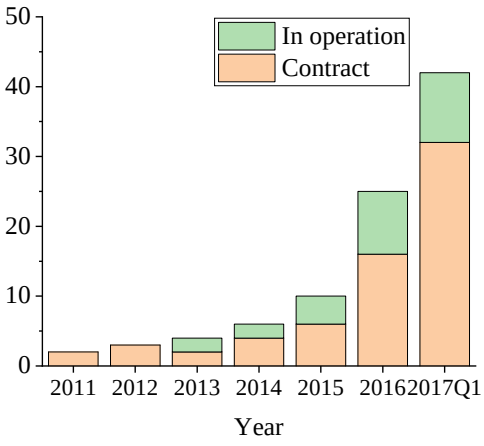
Batteries have also been used in hybrid propulsion systems to allow the diesel engines or fuel cells to operate at optimised load, i.e. load levelling to improve efficiency. Caledonian Ferries has three hybrid vessels, each having two banks of 800 kWh batteries (Figure 2.12a) (Geertsma et al., 2017). The battery banks are charged by utilising shore generated electricity and optimise the loading of the diesel generating sets. It has been reported that a 35% fuel saving was demonstrated during sea trials owing to the hybrid diesel/battery propulsion system (Breijs and Amam, 2016; Geertsma et al., 2017). The gird electricity directly accounted for a 24% fuel saving. The other 11% fuel saving was achieved by levelling the diesel engine loads by using the batteries and controlled using an energy management system specifically developed for this application (Breijs and Amam, 2016).

The Norwegian ferry *MF Ampere* (Figure 2.12b), which came into operation in 2015, is the world's first large-sized all-electric battery-powered car ferry (Alnes et al., 2017). This vessel undertakes 56 journeys per day, each which is of 5.6 km, and is powered by a 1.04 MWh Lithium-ion battery which is recharged by two 410 kWh shore charging stations located at each end of its journey (Skjong et al., 2015; Geertsma et al., 2017). It should be noted due to limited battery energy density, a purely battery-powered propulsion system is limited to ships like *MF Ampere* that require very limited range.

To date, the world's largest maritime battery application that has been an-



(a)



(b)

Figure 2.11: Battery ship statistics and trend (data source: (Alnes et al., 2017)).

nounced is a hybrid LNG-electric cruise ship with a battery capacity of 10 MWh (Figure 2.12d) (Corvus, 2019b). Another large hybrid LNG-battery ship with planned battery capacity of 6.1 MWh was announced in early 2019 (Figure 2.12c) (Corvus, 2019a).

Batteries have also been used with solar energy. Several car carriers with extensive upper deck area have been equipped with photovoltaic panels which generate enough energy to satisfy about 8% of the service load (Barnes, 2014). The experimental craft *Planet Solar* has 8.5 t of Lithium-ion batteries in its two hulls with solar cells to recharge them (Visa et al., 2016).

(a) **(b)** **(c)** **(d)**

Figure 2.12: Battery-powered ships. (a) hybrid diesel-battery *Caledonian MacBrayne* ferry *MV Hallaig* ((Geertsma et al., 2017)). (b) Norwegian all-electric battery-powered ferry *MF Ampere* (Alnes et al., 2017). (c) *Havila Kystruten*'s hybrid gas-battery ship (Corvus, 2019a). (d) World's largest battery maritime application: 10 MWh with LNG engines (Corvus, 2019b).

2.4.5 Battery challenges

2.4.5.1 Energy density

Thackeray et al. (2012) indicate that battery energy density should reach 300–750 $\mathrm{Wh\,kg^{-1}}$ for road vehicles (see Figure 2.13). This would allow electric-powered vehicles to attain a 300–400 mile range between recharges to better match the performance of internal combustion engine powered vehicles. Limited battery energy density is also an issue for maritime applications, especially for those smaller ships typically powered by 4-stroke medium or high-speed diesel engines with minimal volume and weight margin within their propulsion systems. Compared to a traditional diesel system, for a coastal bulk carrier with a voyage length of 100 nautical miles, at an operational speed of 10 knots, the battery system would require approximately 2.5 and 6 times more volume and mass than the installed diesel engines respectively if replaced by a state-of-the-art Lithium-ion battery pack with an energy density of 250 $\mathrm{Wh\,kg^{-1}}$ and 300 $\mathrm{Wh\,L^{-1}}$. Nevertheless, more advanced battery technologies, such as Lithium-O_2, could probably solve the volumetric and weight constraints by offering approximately ten times higher energy density than the Lithium ion battery option (Rahman et al., 2014).

2.4.5.2 Production cost

One of the main obstacles to battery power is the high production cost of the batteries (Schmidt et al., 2017). Batteries need to be replaced when they reach their calendrical life limits or cycling limits. The replacement costs will also contribute to the through-life cost of battery-powered ships. Nevertheless, such high prices are expected to decrease (Larcher and Tarascon, 2015; Alnes et al., 2017). From a long-term standpoint, e.g. 10 or 20 years, batteries can potentially deliver better economic performance over the traditional diesel system if and when the battery price falls below 150 $\mathrm{\$kWh^{-1}}$, which is the long-term goal for commercialisation of electric vehicles (see Figure 2.14) (Nykvist and Nilsson, 2015; Alnes et al., 2017).

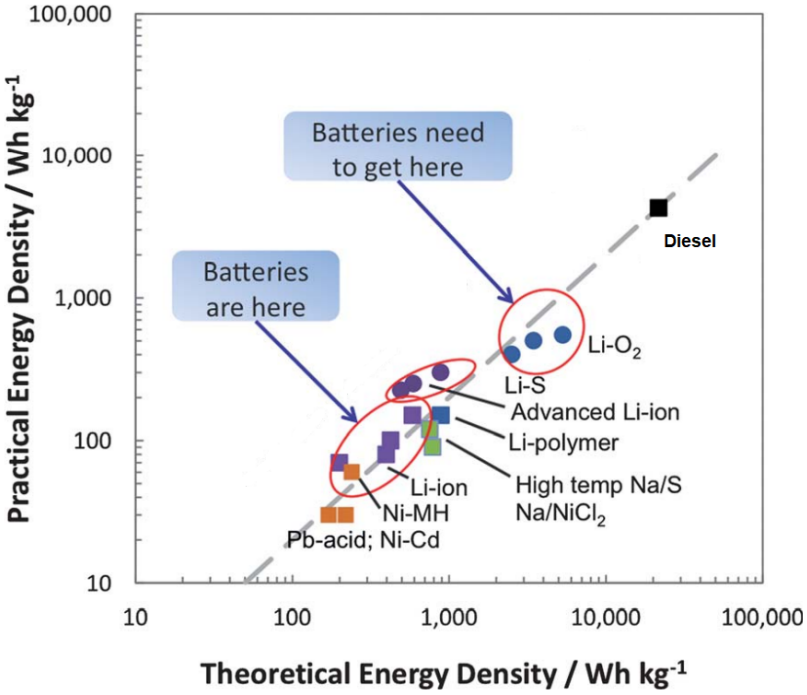


Figure 2.13: Battery theoretical and practical energy densities (source: (Thackeray et al., 2012)).

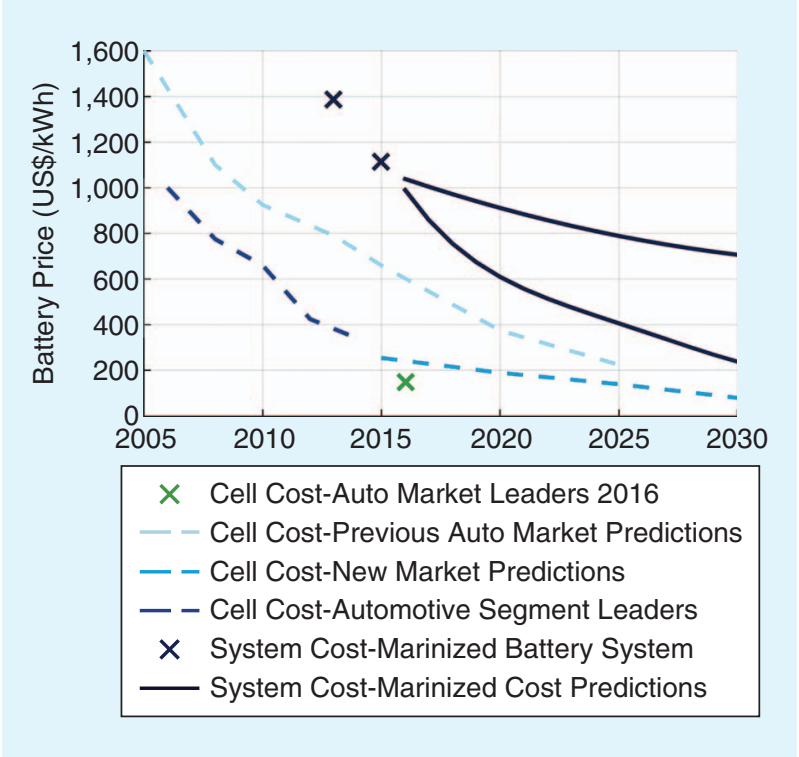


Figure 2.14: Predicted battery price trend (source: (Alnes et al., 2017)).

2.4.5.3 Safety

Lithium-ion batteries could fail due to reasons such as thermal stressing, physical damage, charge and discharge failures and short circuit (Doughty and Roth, 2012). These failures could lead to fires or explosions. There have been accidents caused by Lithium-ion batteries in transport applications including cars and aeroplanes (Wang et al., 2012). For shipboard applications, Rao et al. (2015) proposed safety measures for utilising Lithium-ion batteries including distributed battery space, flame-resistant shell materials, structural and thermal separation of adjacent battery spaces, a suitable fire extinguishing system (e.g. heptafluoropropane type).

The family of Lithium-ion batteries is vast, and each type may have its own failure characteristics so it is essential to design the safety and protection function of battery system in accordance with the particular characteristics when integrating them into ships (DNV GL, 2016). With increasing energy and power density, designing a safe battery could become more challenging as more highly reactive materials and higher operating voltage levels are used. Therefore, it is necessary to prioritise safety features throughout the design and manufacturing processes. Special attention would also need to be paid to safety when batteries are integrated into propulsion systems.

2.4.5.4 Recycling

Lithium-ion batteries contain valuable material, including Co, Li, Mn, and Ni (Zeng et al., 2014). Although Lithium-ion batteries are not generally classified as dangerous waste, they could be flammable and also have a negative impact on the environment. The study by Dewulf et al. (2010) shows that the recycling scenarios could lead to 51.3% natural resource savings and approximately 50% reduction in energy demand. According to Larcher and Tarascon (2015), the production process of Lithium-ion batteries is energy-intensive, i.e. more than 400 kWh is needed to manufacture a 1 kWh Lithium-ion battery whilst the production of 1 kWh of electricity from coal produces 1 kg of CO_2. Swain (2017) reviewed the technologies for

recovery and recycling of Lithium from used batteries and proposed that hydrometallurgical recycling of Lithium-ion batteries should be the focus. It is necessary to develop a more advanced recycling system for Lithium-ion batteries and other future battery types to achieve more sustainable battery power development with a lower overall carbon footprint, fewer pollutants and more resources recycled.

2.4.5.5 Charging infrastructures

For Electric Vehicles (EV), the shortage of charging infrastructure is no longer a major issue as more and more standardised charging ports are being built (Neubauer and Wood, 2014; Burnham et al., 2017). However, ships typically have much higher power and energy demands compared to EV and would need charging infrastructure with much higher capacities to be developed along with the battery-powered ships. The availability of charging infrastructures has a direct influence on a ship's battery system arrangement. The installed battery capacity has to be doubled if only one of the two ends of the route has charging infrastructure. Ships usually only remain for a limited time at ports. Therefore, the capability must exist for the batteries to be charged during that period. Hence, fast charging could be necessary, which could have a negative impact on the local electricity grid. Muratori (2018) considered the charging impact of plug-in hybrid-electric vehicles on the residential grids. However, similar research work can rarely be found for maritime applications. The *MF Ampere* has two additional battery banks installed on each end of its route to minimise the charging impact to the local grid.

Batteries, as a type of energy storage device, do not themselves emit GHG into the atmosphere. However, electricity from the power grid is generated with a certain size of carbon footprint, depending on the type of power generation method. The current global average electricity CO_2 intensity of approximately 500 g CO_2 kWh^{-1} is required to reduce to 100 g CO_2 kWh^{-1} by 2050 (Tran et al., 2012). Figure 2.15 shows the grid electricity Global Warming Potential (GWP) trends in the EU from 1990 to 2016. Actual GWP decreasing trends can be observed at both the EU and country levels. As mentioned in Section 2.2.2, the specific CO_2 emission of a typical

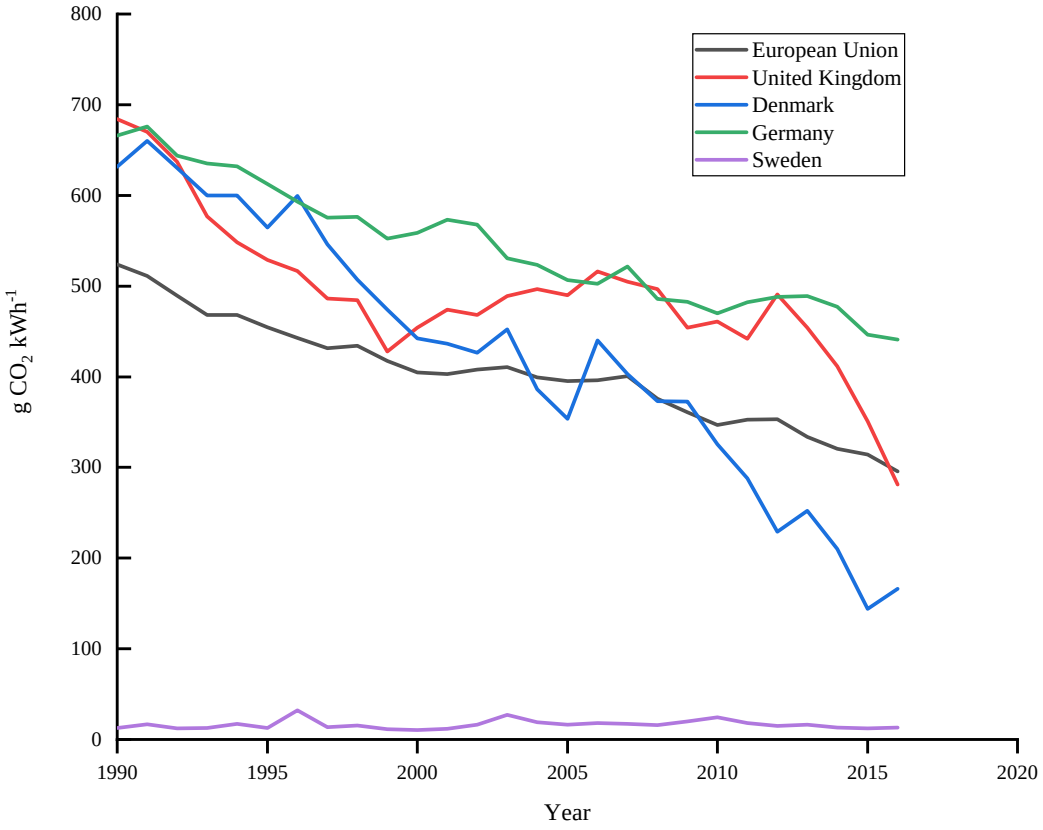


Figure 2.15: Grid electricity GWP in the European Union (Eurostat, 2019).

medium-speed marine diesel engine is approximately 600 g CO_2 kWh^{-1}. The use of shore-generated electricity would reduce the GWP emissions by 50% with the EU average grid electricity GWP of 296 g CO_2 kWh^{-1} in 2016. More GHG savings can be achieved by countries like Denmark with even lower grid electricity GWP. For countries with grid electricity GWP well above 600 g CO_2 kWh^{-1} (Li et al., 2017), using battery power would not realise GHG benefits.

2.5 Review of hybrid-electric propulsion systems

Having identified that diesel-based propulsion systems are not a viable powering option for future coastal shipping with increasingly stringent regulations on emissions in Section 2.2, the scope of this section is focused on electric propulsion system topologies which are more suitable for fuel cell (Section 2.3) and battery (Section 2.4) power.

Shipboard electric propulsion has been around for more than 100 years (Ådnanes, 2003). The availability of power electronics and variable speed converters makes modern all-electric propulsion systems feasible for large scale commercial marine applications (Woud and Stapersma, 2002; Hansen and Wendt, 2015). Figure 2.16 details the evolution of shipboard all-electric propulsion systems. Since the 1980s, the use of power electronics has become a common method for commercial ships such as cruise vessels to improve fuel efficiency (McCoy, 2002; Kim et al., 2015b). Integrated Full Electric Propulsion (IFEP) (Figure 2.16a) was adopted by *Queen Elizabeth 2* to replace the steam turbines for better fuel efficiency. Such a system topology has also been used by other ships such as platform supply vessels with highly variable load profiles.

More recently, the emerging energy storage technologies (mainly Lithium batteries) enable IFEP with even better fuel efficiency (i.e. by load levelling to allow more efficient operation of the diesel engines), and the capability of utilising greener shore-generated power when larger battery modules are installed onboard (see Figure 2.16b) (Geertsma et al., 2017). Such a hybrid diesel/battery system has been adopted by ships such as *MV Hallaig* (see Figure 2.12a). However, the overall efficiency improvement that can be offered by such systems are limited, and would depend on the application, load characteristics and control strategies (Breijs and Amam, 2016; Geertsma et al., 2017).

One of the major limitations of AC based diesel-electric propulsion systems (Figure 2.16a and 2.16b) is that the diesel prime-movers have to operate at fixed speeds, which limits the overall fuel efficiency due to high SFC when the engine operates under part-load conditions (Geertsma et al., 2017). To further improve fuel efficiency of these systems, a DC hybrid power system (Figure 2.16c) is a viable option (Zahedi et al., 2014; Herrera et al., 2015). In a DC power system, the diesel engines can operate at optimised speeds rather than fixed speeds to achieve better part-load efficiency. However, to utilise these types of DC systems, challenges such as fault protection and system stability would need to be addressed to allow reliable operations (Zadeh et al., 2013; Herrera et al., 2015).

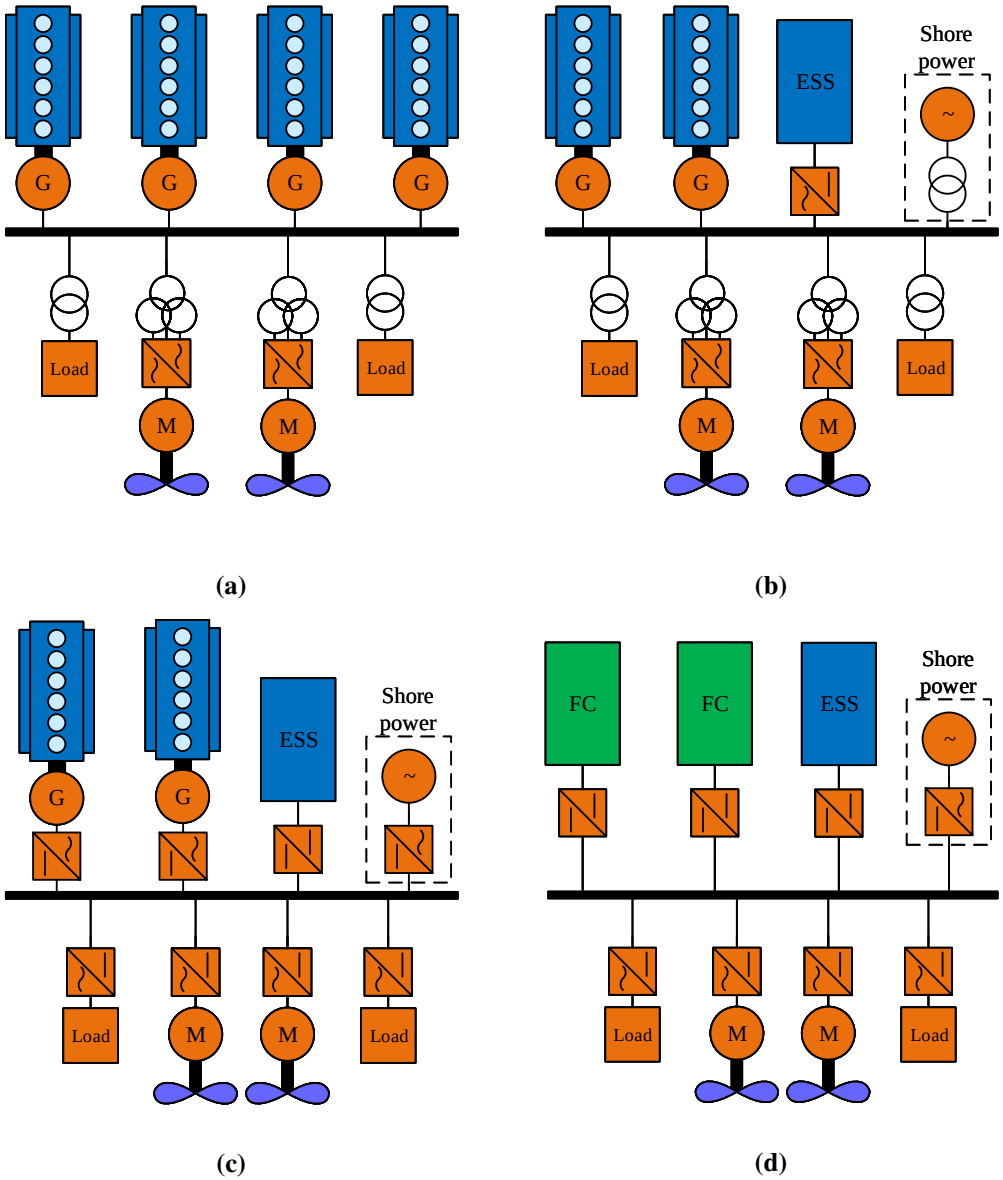


Figure 2.16: Evolution of ship electric propulsion system (Geertsma et al., 2017). (a) Integrated full electric propulsion AC power system applied on cruise ships and offshore vessels. (b) Hybrid diesel/battery electric with shore power and AC power system. (c) Hybrid diesel/battery with shore power and DC power system. (d) Hybrid fuel cell/battery with shore power and DC power system.

Replacing the diesel generators and their rectifiers in Figure 2.16c with fuel cells, a hybrid fuel cell/battery DC power system would be offered as in Figure 2.16d. Such a DC system would seem a natural fit since the fuel cells and batteries provide power outputs that are both DC (Sulligoi et al., 2016; Skjong et al., 2016). The fully battery-powered ferry *MF Ampere* (Figure 2.12b) has been constructed

with such a DC system (Skjong et al., 2015). Nevertheless, challenges similar to those encountered in the diesel engine based DC systems still apply to the fuel cell/battery powered system. It is possible to use a conventional AC power distribution system for fuel cell and battery powered ships but this would introduce further conversion losses when the DC power is converted to AC.

2.6 Summary

This literature review has been focused on the future direction of marine power and propulsion systems for coastal shipping. Diesel engines and their associated improvement technologies, alternative fuels, fuel cell technologies, energy storage systems and marine hybrid-electric propulsion system topologies have been reviewed.

As determined by Stephens et al. (2016), low-temperature fuel cells offer great potential as a means of replacing fossil fuels with renewable energy for transportation applications. Although there are still major challenges such as high production costs, the PEMFC is commercially available, and the production costs are expected to decrease rapidly (Chen et al., 2015; Fletcher et al., 2016; Ehsani et al., 2018). Such fuel cells can offer desirable life-cycle emission performance provided that the H_2 can be produced through use of clean and low-cost energy (Singh et al., 2015; Ahmed et al., 2016; Gurz et al., 2017; Bicer and Dincer, 2018). The limited system-level energy density of H_2 is one of the main obstacles hindering its application to wide ranging maritime operation (van Biert et al., 2016; Raucci, 2017). However, for coastal ships operating on short routes, e.g. *SF-BREEZE*, the volumetric impact of H_2 storage requirements could be alleviated, but at the cost of the need for more frequent replenishment of H_2 (Pratt and Klebanoff, 2016).

Existing batteries could be a viable powering solution for ships operating on short routes and have access to low-carbon grid power. However, challenges such as energy density, lifetime, safety, grid power GWP need to be considered before applications (Thackeray et al., 2012; Larcher and Tarascon, 2015; Wu and Bucknall,

2016; Doughty and Roth, 2012). Flywheels are not suitable for ship power due to very limited energy density and high self-discharging rates (Amiryar and Pullen, 2017). Supercapacitors could be useful for transient loads; however, with limited energy density, they are not ideal to be main power sources (Saleem et al., 2016).

To decarbonise the coastal shipping, the GHG emissions need to be considered holistically, as well as the constraints such as high costs and low energy densities compared to conventional plants. Both the H_2 powered PEMFC and Lithium battery would appear as zero or near zero-emission if the scope is only limited to the ship. However, the carbon intensities, measured by Global Warming Potential, of both H_2 and electricity, can vary greatly depending on the energy sources. The uncertainties from the energy supply side, power source manufacturing and practical ship design requirements need to be well addressed holistically to achieve balanced alternative propulsion system design.

The hybridisation of PEMFC, battery and grid power in a DC system architecture could potentially provide balanced propulsion system designs for coastal ship operating on short routes and have access to clean and renewable energy sources.

Chapter 3

Problem formulation

3.1 Overview

In Chapter 2, a range of technology options to replace diesel-based systems are considered. Based on this review, a Proton Exchange Membrane Fuel Cell (PEMFC), Lithium battery and shore power in hybrid plug-in configuration will be selected as the primary candidate technologies to decarbonise coastal shipping engaged on relatively short transits for their potential to eliminate carbon emissions. This chapter identifies research gaps (Section 3.2) that exist in integrating PEMFC and battery technology to marine propulsion and power systems. Development of energy management strategies for hybrid PEMFC and battery power systems will also be explored. Section 3.3 provides details of the candidate ship that will be the basis for this research. Section 3.4 summarises this chapter.

3.2 Gap analysis

As shown in Figure 3.1, designing and optimising hybrid propulsion systems with multiple power sources requires consideration of both the system optimisation and control design aspects (Hu et al., 2015b; Silvas et al., 2016; Valera-García and

Atutxa-Lekue, 2018). Conventional commercial ship propulsion systems typically utilise diesel engines as the main power source (Woud and Stapersma, 2002), which makes for a less complex propulsion system design than that for hybrid propulsion systems. Having identified a plug-in hybrid propulsion system with PEMFC and Lithium battery as the preferred configuration for this project, the research gaps in the power source sizing optimisation and optimal control of such hybrid propulsion systems need to be identified and clearly defined. Power source sizing is intended to optimise the propulsion system during the design phase prior to operations. However, control strategies for such systems should be considered in the sizing process. The aim of optimal control is the satisfying of specific operational objectives such as minimum fuel consumption.

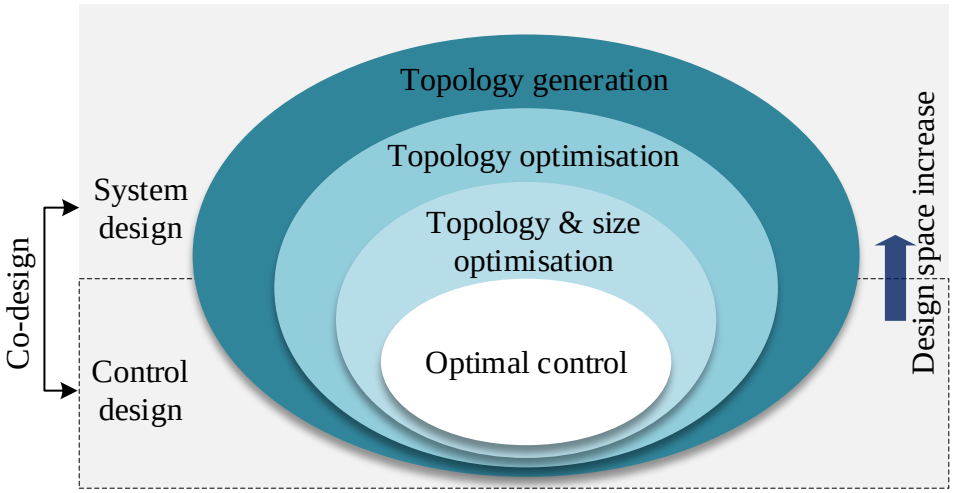


Figure 3.1: Propulsion system level design and its multi-layers (Silvas et al., 2016).

3.2.1 Hybrid propulsion system optimisation

Table 3.1 provides a list of research works relevant to hybrid propulsion design optimisation. The scope is extended to include road vehicles, aircraft, submarines as well as ships so as to include any concepts and methods that have been used in parallel industries and may prove to be of value of this research. The hybrid propulsion system design optimisation problem can be mathematically formulated as constrained optimisations. Optimisation schemes such as a Genetic Algorithm or a Particle Swarm Optimisation have been proposed and applied to solve such

problems (Erdinc and Uzunoglu, 2012). A minimising optimisation problem with a single objective function can be stated as to find a vector of decision variables:

$$x = [x_1, x_2, ..., x_n] \tag{3.1}$$

which minimises the objective function:

$$f = f(x) \tag{3.2}$$

subject to the constraints:

$$g_i(x) \leq 0 \tag{3.3}$$

for $i = 1, 2, ..., m$, and:

$$l_j(x) = 0 \tag{3.4}$$

for $j = 1, 2, ..., n$, where g_i is the i^{th} inequality constraint and l_j is the j^{th} equality constraint. For problems with multiple objectives, the objective functions can be expressed as:

$$f = [f_1(x), f_2(x), ..., f_z(x)] \tag{3.5}$$

where $f_k(x), k \in [1, z]$ is the k^{th} objective function.

For road vehicle applications, optimisation methods have been widely adopted to optimise hybrid propulsion system designs. Kim and Peng (2007) proposed a combined power source sizing and energy management for a hybrid fuel cell and battery road vehicle using a DIRECT optimisation algorithm and a pseudo-Stochastic Dynamic Programming Energy Management System (EMS) with a scalable system model. Their proposed method is a single-objective approach focusing on minimising H_2 consumption without considering any degradation of the power sources and life-cycle GHG emissions. (Hu et al., 2015a) applied convex optimisation to size the power sources of a hybrid PEMFC/battery road vehicle to minimise H_2 fuel and battery degradation costs. However, the degradation of the PEMFC, which could potentially influence the system design, were not considered.

Table 3.1: Selected hybrid-electric propulsion system optimisation literature.

Reference	Application	System topology	Optimisation objective	Decision variables	Constraint	Optimisation method
(Valera-García and Atutxa-Lekue, 2018)	Ship	Hybrid DG/battery	Fuel consumption Battery capacity Battery energy consumption	DG power	Power balance, DG, battery	NSGA-II
(Mashayekh et al., 2012)	Ship	Hybrid DG/battery	Total generation cost	DG power output Initial battery SOC	Power balance, DG and battery	Dynamic dispatch
(Zhu et al., 2018)	Ship	Plug-in hybrid DG/battery	Fuel consumption GHG emission Life-cycle cost	Engine parameters battery configuration	Safety, Minimum power Maximum battery number, SOC	NSGA-II MOPSO
(Bassam et al., 2016)	Ship	Hybrid FC/battery	Total cost	Battery and fuel cell numbers	Unconstrained	Exhaustive search
(Zhu et al., 2019)	Ship	Plug-in hybrid DG/battery	Fuel consumption GHG emissions	Engine parameters Battery configuration	Power balance, DG, battery, Ship speed	NSGA-II
(Ribau et al., 2014)	Road vehicle	Plug-in hybrid FC/battery	Cost Fuel consumption GHG emission	Components sizing	Vehicle performance	NSGA-II
(Xu et al., 2015)	Road vehicle	Hybrid FC/battery	Fuel economy Performance degradation	Components sizing	Vehicle performance	Not given
(Zhang et al., 2017)	Road vehicle	Plug-in hybrid battery/SC	Cost Weight	Components sizing	Vehicle performance	NSGA-II
(Song et al., 2014)	Road vehicle	Plug-in hybrid battery/SC	System cost Battery capacity loss	Components sizing	Battery capacity, power source voltage	NSGA-II
(Song et al., 2015)	Road vehicle	Plug-in hybrid battery/SC	Life cycle cost SC cost	Components sizing and EMS	Vehicle performance	NSGA-II
(Hu et al., 2015a)	Road vehicle	Hybrid FC/battery	H_2 and ESS costs	Components sizing and EMS	Power balance, Power source constraints	Convex optimisation
(Kim and Peng, 2007)	Road vehicle	Hybrid FC/battery	H_2 consumption	Components sizing and EMS	Fuel cell and battery limits, power balance	DIRECT algorithm
(Skinner et al., 2007)	Submarine	To be determined	Motor efficiency & size Propeller efficiency Energy consumption	System configuration	Load	GA
(Xie et al., 2018)	Aircraft	Hybrid engine/battery	Fuel consumption Flight duration	Fuel mass Battery mass	Mission requirements	NSGA-II

In practice, the design of a hybrid propulsion system often needs to consider the trade-off between multiple objectives due to potential conflicts amongst the objectives (Sulaiman et al., 2018). Ribau et al. (2014) developed a sizing methodology optimising fuel consumption, cost and life-cycle GHGs concurrently based on the well-established Non-dominated Sorting Genetic Algorithm (NSGA-II) (Deb et al., 2002) for a plug-in hybrid fuel cell/battery bus operating in Portugal. Their results suggest that 67% of the life-cycle GHG emissions can be reduced for the application scenario. However, different energy properties and driving cycles can influence the system design. The optimal design can vary significantly for different energy sources and driving cycles. Song et al. (2014, 2015); Zhang et al. (2017) adopted the same optimisation routine to optimise the designs of plug-in hybrid battery/supercapacitor road vehicles with different optimisation objectives.

Earlier work by Skinner et al. (2007) adopted a multi-objective genetic algorithm to optimise the integrated electric propulsion system for a naval submarine. The multi-objective genetic algorithm was proven to be effective in dealing with multi-objective and high dimensional design problems. There were 16 objective functions, and the design space had 9 discrete and continuous design variables. (Xie et al., 2018) proposed using NSGA to size the hybrid-electric propulsion system for retrofitting a mid-scale aircraft. Fuel consumption and flight duration were the two objectives on which the design was focussed.

For shipboard applications, when the decision space has limited dimensions and sizes, it is feasible to search sweeping through the design space. Bassam et al. (2016) proposed a power source sizing methodology for a hybrid PEMFC/battery powered ferry to minimise a lumped cost, including initial and operational costs through exhaustive search over the sizing variable spaces. However, such an approach did not consider any volumetric and weight constraints, which could invalidate the solutions. The optimal energy management problem for each sizing combination was solved using a classical Proportional-Integral controller (Motapon et al., 2013). Mashayekh et al. (2012) adopted a similar search approach for their diesel/battery hybrid propulsion system battery sizing determination. The opti-

mal energy management solution is obtained by using dynamic dispatch for a normalised load profile.

When the design variable space is continuous or is of large dimension and size, it would be impossible or computationally expensive to loop through the variable space. Valera-García and Atutxa-Lekue (2018) applied NSGA-II (Deb et al., 2002) to optimise three objectives (fuel consumption, battery capacity and battery energy consumption) concurrently for a platform supply vessel. Zhu et al. (2018) adopted NSGA-II (Deb et al., 2002) and Multiple Objective Particle Swarm Optimisation (MOPSO) (Coello and Lechuga, 2002) to optimise their plug-in hybrid diesel/battery propulsion systems with several constraints applied to determine the diesel system design parameters and battery module configuration. The objectives include fuel consumptions, GHG emissions and life-cycle costs. A rule-based EMS was used in their nested optimisation flow. A similar approach was used in the work of Zhu et al. (2019), but with the number of objective functions reduced to two, focusing on fuel consumption and GHG emissions.

Multi-objective optimisation schemes such as NSGA-II have been successfully applied to optimise plug-in and non-plug-in hybrid propulsion systems of road vehicles, submarines and aircraft. Such schemes have also been adopted to optimise the designs of plug-in hybrid diesel/battery ships. However, no work on optimising plug-in hybrid fuel cell/battery propulsion systems focusing on life-cycle GHG emissions and power source energy costs could be found in the literature. Moreover, coastal ships, such as ferries, typically have limited volume and weight margins for their propulsion systems. These constraints need to be accounted for in the topology optimisation phase.

It is worth mentioning that the design of a hybrid propulsion system is a nested problem, i.e. the system topology and control strategies need to be considered concurrently (Silvas et al., 2016). Also, note that alternative power sources, such as fuel cells and batteries, have limited lifetimes and could degrade at different rates under different control strategies. Although existing works have applied energy

management strategies, such as rule-based strategy when optimising the hybrid system designs, none of the existing studies have considered degradation of the fuel cells and batteries concurrently.

3.2.2 Energy management strategy

In a hybrid propulsion system, an effective EMS is crucial to managing power flows among multiple power sources (Martinez et al., 2016; Kalikatzarakis et al., 2018). Table 3.2 presents a compilation of energy management studies for hybrid-electric propulsion systems deemed relevant to this research. The EMS determines actions taken by the hybrid propulsion system under certain operating conditions. However, it is a challenge to develop an effective EMS for hybrid systems if future power profiles are unknown.

The research into EMS for hybrid propulsion systems is primarily driven by road vehicle applications. Sulaiman et al. (2015) provided a comprehensive review of the main EMS categories for hybrid fuel cell road vehicles. Their review indicates that rule-based, fuzzy logic, Equivalent Consumption Minimisation Strategy and wavelet-based load sharing are the main EMS streams for hybrid fuel cell road vehicles.

Caux et al. (2010) proposed an on-line fuzzy energy management system for the propulsion system of a hybrid fuel cell/supercapacitor road vehicle. Their fuzzy system parameters were optimised for two standard driving cycles using a genetic optimisation algorithm. The results suggest that the energy consumption when under the control of their fuzzy EMS is approximately 30% worse than that of the optimal EMS acquired via dynamic programming. An adaptive fuzzy-logic EMS was proposed for a plug-in hybrid/battery road vehicle later by Khayyam and Bab-Hadiashar (2014). However, the performance of that EMS was not compared against one derived by dynamic programming, but a 10% fuel consumption reduction has been observed when compared to the original EMS. Chen et al. (2017) proposed an on-line predictive fuzzy-logic EMS for a plug-in hybrid engine/battery

road vehicle. Their study tackles the uncertainties of driving cycles by combining dynamic-neighbourhood particle swarm optimisation and fuzzy logic. The energy consumption that could be achieved is up to 9.7% (when compared to their charge-depleting and charge-sustaining EMS).

Fletcher et al. (2016) adopted stochastic dynamic programming to generate an optimal EMS for a hybrid fuel cell/battery road vehicle, accounting for the fuel cell degradation characteristics generalised from experimental results. Their EMS was able to reduce the cost by 12.3% through prolonging fuel cell lifetime. The main contributions from this study are: (1) addressing the importance of considering power source degradation characteristics for EMS design and optimisation; (2) applying stochastic dynamic programming to generate an optimal EMS based on historical load profiles. However, the accuracy of stochastic dynamic programming is limited by its resolution due to 'the curse of dimensionality' (Sutton and Barto, 1998). Wang et al. (2019) proposed more sophisticated fuel cell and battery degradation models based on the degradation mechanisms. However, Wang et al. (2019) only used deterministic dynamic programming to generate an optimal EMS for a specific load profile based on these degradation models. It should be noted that an EMS generated by deterministic dynamic programming is a useful as an off-line EMS against which to benchmark other on-line EMSs, but it is not practical to apply it to actual unknown load profiles.

Recently, machine learning EMS have started to emerge for road vehicles. Muñoz et al. (2017) presented a neural network EMS for a hybrid fuel cell and battery road vehicle based upon supervised learning. With a target EMS for specific driving cycles generated through optimisation approaches, the neural network was subsequently trained to achieve levels of performance similar to that achieved by the target EMS. The actual performance of such an EMS for unknown future driving cycles is not clear. Murphey et al. (2012a,b) applied neural network for power demand prediction and used dynamic programming to generate an EMS. However, actual power demands could vary in practical driving conditions, which would influence the performance of the pre-generated EMS.

Table 3.2: Selected hybrid-electric propulsion system energy management literature.

Reference	Application	System topology	EMS objective	Methodology	Degradation considered?
(Khayyam and Bab-Hadiashar, 2014)	Road vehicle	Plug-in hybrid engine/battery	Fuel consumption	Neuro-fuzzy logic	No
(Song et al., 2015)	Road vehicle	Plug-in hybrid battery/supercapacitor	Battery loss	Rule-based based on DP	Yes, battery
(Fletcher et al., 2016)	Road vehicle	FC/battery hybrid	Total running cost	Stochastic DP	Yes, fuel cell
(Wang et al., 2019)	Road vehicle	FC/battery hybrid	Overall lifetime cost	Deterministic DP	Yes, fuel cell and battery
(Caux et al., 2010)	Road vehicle	FC/battery hybrid	Fuel consumption	Fuzzy logic	No
(Martel et al., 2015)	Road vehicle	Plug-in hybrid FC/battery	Overall operational cost	DP	Yes, battery
(Chen et al., 2017)	Road vehicle	Plug-in hybrid engine/battery	Total energy cost	Fuzzy logic	No
(Murphey et al., 2012a,b)	Road vehicle	Hybrid engine/battery	Total fuel cost	DP and NN	No
(Xiong et al., 2018)	Road vehicle	Plug-in hybrid engine/battery	Total system loss	RL	No
(Liu et al., 2017)	Road vehicle	Hybrid engine/battery	Fuel consumption and SOC	RL	No
(Wu et al., 2018)	Road vehicle	Hybrid engine/battery	Fuel cost	Deep RL	No
(Hu et al., 2018)	Road vehicle	Hybrid engine/battery	Fuel cost and SOC	Deep RL	No
(Zou et al., 2016)	Tracked vehicle	Hybrid engine/battery	Fuel consumption and SOC	RL	No
(Skjong et al., 2017)	Ship	Hybrid engine/battery	Operational efficiency	Mixed-integer linear programming	No
(Kanellos, 2013)	Ship	Hybrid engine/battery	Operational cost	DP	No
(Al-Falahi et al., 2018)	Ship	Hybrid engine/battery	Fuel consumption	GWO optimisation	No
(Kalikatzarakis et al., 2018)	Ship	Plug-in hybrid engine/battery	Fuel consumption	ECMS	No

For a hybrid system with an energy storage device, the problem, in essence, is a sequential-decision making problem, i.e. what actions or controls should be taken over the power cycles to deliver performance that approximates the optimal objectives (e.g. minimum operational costs or emissions). Reinforcement Learning (RL), as an approach for solving such sequential-decision making problems, has been proposed for hybrid diesel engine and battery road vehicles. Hu et al. (2018) and Wu et al. (2018) implemented Deep Q-Network (DQN) to generate an EMS for standard driving cycles. It is worth noting that, using a limited number of driving cycles to train an RL agent could lead to the generated EMS only being able to perform as per design under specific driving cycles. Xiong et al. (2018) proposed solving the optimal power split problem using Q-learning with the Kullback-Leibler divergence as an indicator as whether to update the EMS over time. Their results suggest that updating an EMS over time may further reduce fuel consumption. However, there is a lack of implementation of reinforcement learning algorithms for fuel cell/battery hybrid propulsion systems. Moreover, these road vehicle implementations only focus on single power source control (either the engine or the fuel cell stack in discrete action spaces). In contrast, marine propulsion systems would require control over multiple power sources for redundancy and reliability considerations.

For shipboard applications, it is rare to find an intelligent EMS based on RL. Kalikatzarakis et al. (2018) presented 'Equivalent Consumption Minimisation Strategies' for shipboard applications with a diesel engine in hybridisation with battery and shore power. Their results indicate that a 6% fuel saving can be achieved compared to the rule-based method. However, the actual fuel consumption achieved is not clear for actual ship operations since only a limited number of power profiles were analysed in this study. Bassam et al. (2017) proposed a multi-scheme EMS with a mix of several sub-EMSs in different states for a hybrid fuel cell passenger ship based on an operational profile of eight hours duration. Choi et al. (2016) implemented a load-following EMS for their hybrid fuel cell and battery powered boat, in which the fuel cells operate at a designated power ouput while the batteries provide any additional power demands. Han et al. (2014) proposed a rule-based

EMS tuned by a typical load cycle for a passenger boat.

The research efforts mentioned above have successfully developed hybrid propulsion system EMSs for shipboard applications. However, the existing EMSs were developed using a limited number of load profiles. Understanding of the actual EMS performance for as yet unknown voyages in the future has not been made clear in existing works that can be found from the literature. Novel approaches using reinforcement learning algorithms and historical power profiles over the long term do not exist in the literature for marine applications.

3.2.3 Identified research gaps

3.2.3.1 Design

There is no clear systematic integrating methodology that exists for designing and managing the alternative powering solutions for coastal ships. It is true that fuel cells operating on H_2 are emission-free or near emission-free. However, the production processes of H_2 can itself be carbon-intensive. The CO_2 emissions which have been transferred to other phases of the fuel production cycle could be even higher than that from conventional diesel engine based solutions. In addition, most of the current alternative power sources, such as fuel cells and batteries, do have certain disadvantages and limitations such as high production costs and low lifetime. These drawbacks and limitations could make alternative power systems extremely uneconomic to operate. In addition, the cost and CO_2 impact of fuel production vary significantly between different countries and regions of the world, which has to be considered from the beginning of the power system design phase. The identified research gap on integrating fuel cells and batteries for coastal shipping is:

- There is a lack of holistic sizing methodologies for the design of plug-in hybrid PEMFC/battery propulsion systems, considering life-cycle GHG emissions, costs and shipboard constraints.

3.2.3.2 Operation

For hybrid road vehicles it is typical to use standard driving cycles to develop the EMS. However, for hybrid ships, such standard cycles do not exist, and actual power demands over different voyages may vary significantly due to factors such as sea states, weather and ship cargo load conditions. In recent years, continuous monitoring of power demand over the long term provides a potential new approach to advance the EMS development for such vessels (Eriksen et al., 2018). The identified research gaps on intelligent EMS for plug-in hybrid PEMFC/battery propulsion system are:

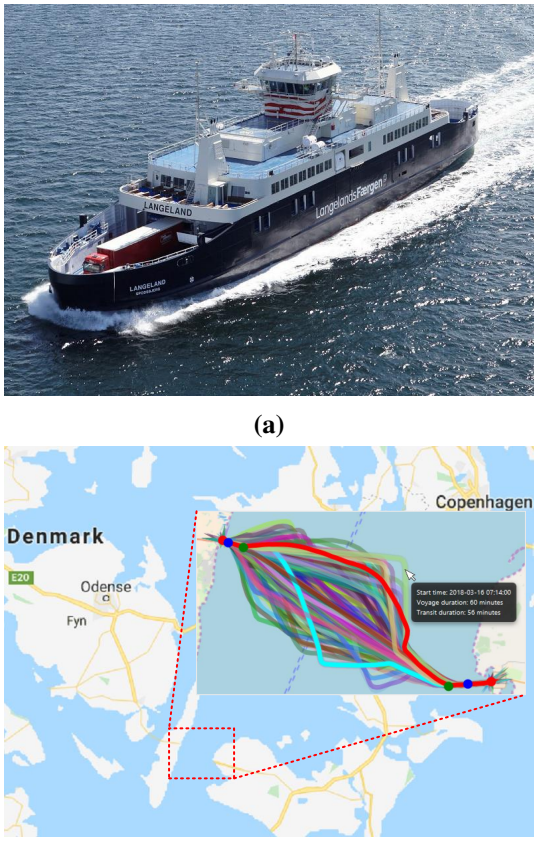
- Cost-effective reinforcement learning EMS trained by large-scale stochastic power profiles collected via continuous monitoring.
- EMS in continuous state spaces using deep neural networks as function approximators.
- Multiple fuel cell cluster control in continuous state and action spaces.

3.3 Case ship

Considering the quality and quantity of the data available, the proposed hybrid propulsion system optimisation methodology and reinforcement learning based energy management strategies will be applied to a coastal ferry with a route as shown in Figure 3.2 (Eriksen et al., 2018). Table 3.3 depicts the case ship's specifications. The original propulsion system featured an integrated full electric propulsion configuration with a total installed power capacity of 4370 kW (five diesel generator sets, with each prime mover rated at 874 kW). The annual operating duty is 300 days, and the ship operates between two fixed ports accomplishing 16 voyages (8 round trips) per day, with each voyage being of 60 min duration. It is assumed that battery charging can be carried out in both ports of the defined route, and hydrogen will be replenished overnight and never during the operational period.

Table 3.3: Case ship specifications.

Parameters	Value
Ship type	Ferry
Gross tonnage	4500
Power system configuration	Integrated full electric propulsion
Installed engine power	4370 kW
Fuel tank volume	140 m^3
Daily voyage number	16 (8 rounds)
Average voyage time	1 h



(a)

(b)

Figure 3.2: Case ship (a) and its route (b) (Eriksen et al., 2018).

Figure 3.3 presents sample power profiles of the case ship collected using monitoring and measuring equipment at the time step of 15 s. It can be observed that most of the time the actual engine power being delivered is less than the power capacity rating of two diesel generators (3.3c). High power profiles can also be observed, but the probability of them occurring is much lower than time spent operating at the moderate and low power profiles. Unfavourable low load operation,

which leads to low fuel economy, can also be observed as in 3.3d. The energy efficiency for this ship can be potentially improved by integrating energy storage systems for load levelling.

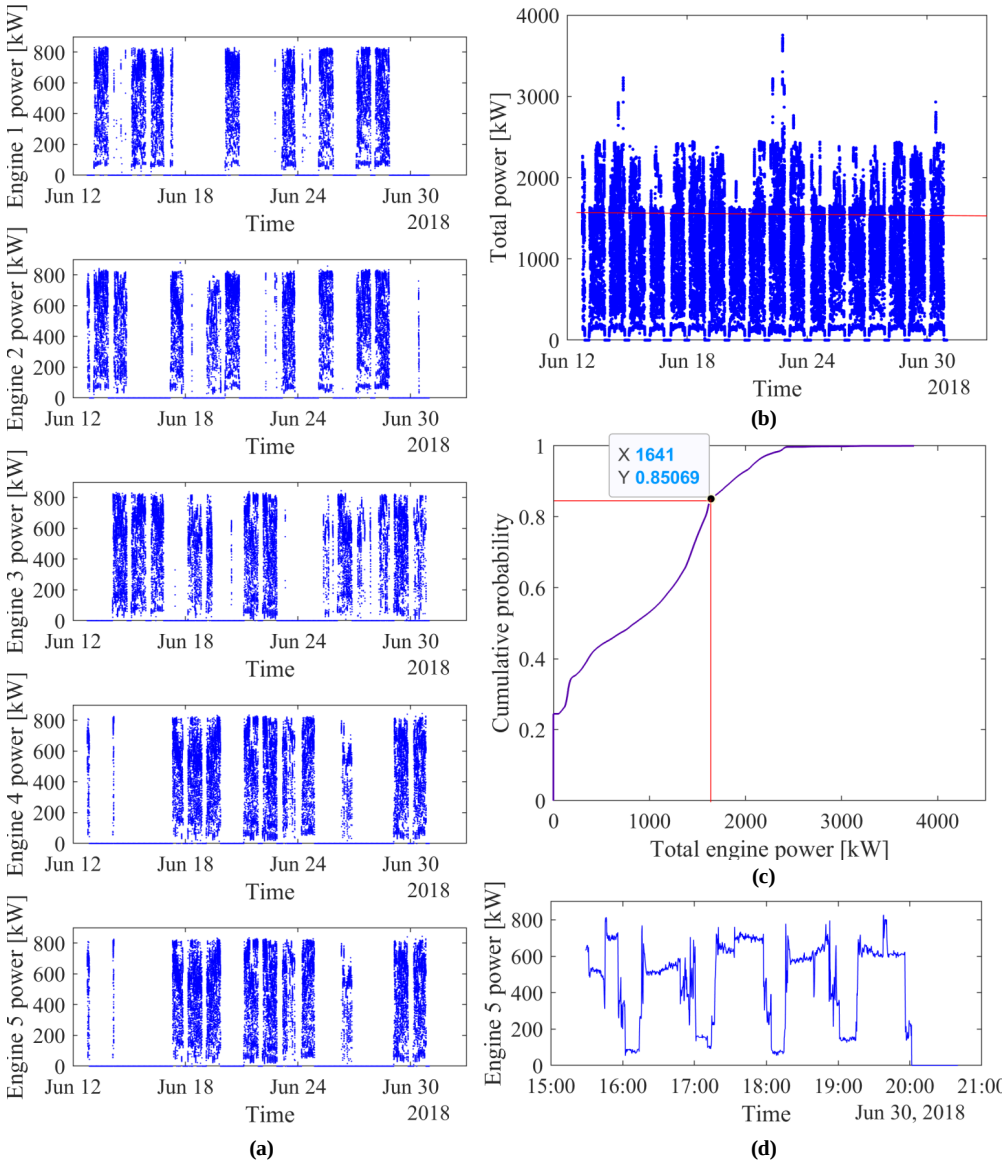


Figure 3.3: Continuous monitoring collected raw power data of the case ship. (a) individual engine power and (b) total delivered engine power; (c) cumulative probability of total delivered engine power below specific values, note that the probability of power below 1641 kW is 0.85, in other words, most of the time, the power demands are below the capacity of two generator sets; and (d) power delivered by Engine 5 in 5 continuous sample voyages, frequent unfavourable low load operations can be observed.

3.4 Summary

In this chapter, the research gaps in integrating fuel cells and batteries have been identified. There is a lack of holistic design methodology dealing with the trade-off between life-cycle GHG emissions and costs. The proposed two-layer multi-objective hybrid propulsion system optimisation method has been introduced to deal with the challenge of considering variable energy properties and optimal hybrid system control strategies concurrently. Such a design methodology is intended to achieve balanced hybrid propulsion system designs at the design phase.

This chapter also identified research gaps in the development of intelligent energy management strategies for hybrid fuel cell and battery propulsion systems. Reinforcement learning and deep reinforcement learning algorithms have been proposed to solve the optimal energy management problems using continuous monitoring data. Unlike the application scenarios for road vehicles, hybrid marine propulsion systems are subject to more complex operating conditions and could therefore require high dimensional control over multiple fuel cell stacks. The reinforcement learning and deep reinforcement learning based energy management strategies are aimed at improving the cost-effectiveness and energy efficiency in the operation phase.

Chapter 4

System modelling and optimisation

4.1 Overview

In this chapter, a holistic design methodology optimising the power source sizing for coastal ship plug-in hybrid PEMFC/battery propulsion systems based upon an energy efficiency model is proposed. Exiting studies on optimising plug-in hybrid PEMFC/battery are mainly focused on road vehicle applications. For coastal ship applications, life-cycle GHG emissions, costs and shipboard constraints need be considered to achieve environmentally and economically feasible sizing solutions. Such a methodology deals with the research gap associated with the integration of a PEMFC and battery in a hybrid propulsion system for coastal shipping identified in Chapter 3.

This sizing methodology optimises system performance while balancing the trade-off between life-cycle GHG emissions and average voyage costs while considering shipboard constraints, power and the properties of energy sources. The power source sizing problem is solved using constrained mixed-integer multi-objective optimisation in the external layer. The global optimum energy management strategies for an averaged operating profile are obtained from deterministic dynamic programming in the inner layer while considering power source degradations in the sizing

algorithm. Section 4.2 details the system model development. In this chapter Section 4.3 describes the sizing methodology. Section 4.4 implements the proposed methodology to the case coastal ferry to investigate the feasibility and benefit potential of adopting the plug-in hybrid PEMFC/battery propulsion system. A sensitivity analysis is presented in Section 4.5. Section 4.6 summarises the chapter.

4.2 System modelling

The models used in this chapter were coded in Matlab script to achieve fast simulation speed which is essential for system optimisation and EMS development (using reinforcement learning). The PEMFC, battery and converter models are developed to have the flexibility of being calibrated by experimental data. Such that the proposed methodology can be adopted with other types of power sources or converters, depending on the actual devices would be used for a project.

4.2.1 Modelling purpose and requirement

The purpose of modelling the plug-in hybrid PEMFC/battery system is to establish a scalable framework to investigate the influences of not easily predictable energy properties (i.e. life-cycle GHG emissions and prices) on the system design via multi-objective optimisation. Such a model will also be used as a basis for reinforcement learning energy management strategy development in the subsequent chapters. Figure 4.1 presents the single line diagram of the system model. Note that the focus of this research project is on system level optimisation and energy management. Therefore, a quasi-steady-state model has been developed. Such an approach has been validated and has been widely adopted in sizing optimisation and energy management studies (Hu et al., 2015a; Song et al., 2015; Wu et al., 2018; Xiong et al., 2018).

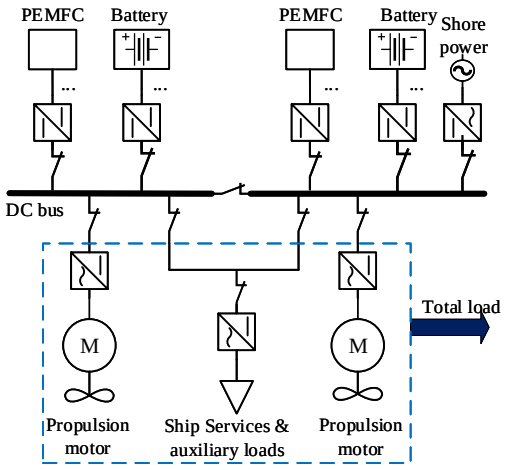


Figure 4.1: Single line diagram of the plug-in hybrid PEMFC/battery system.

4.2.2 Modelling assumptions and limitations

When using the proposed sizing methodology and reinforcement learning based energy management strategies based on the system model, the following assumptions and limitations apply and are deemed acceptable for the purposes of developing the model.

4.2.2.1 Modelling assumptions

The following assumptions will apply throughout this and the following chapters:

- Suitable power converters exist that can match the efficiency curves adopted in this study for the power source sizes under consideration.
- Shore power is immediately available when the ship arrives at the ports; in reality, extra time would be required to establish the electrical connections which could prolong the necessary docking time in port. However, it should be noted that automated charging solutions using industrial robots that can minimise connection time are commercially available (ABB, 2018). When the ship is in port, the shore electricity is used to power the auxiliary ship loads as well as charging the battery, i.e. cold ironing.
- As the case ship is continuously in service on operational days (Eriksen et al.,

2018), the PEMFC is not switched off but maintained at an idling setting when in port. In other words, the PEMFC is only switched off at the end of the operational day. This is such that excessive degradation due to unnecessary on/off cycling is avoided. Also, the power required for idling has been included in the load profiles.

- H_2 and electricity available on both ends of the case ship route have identical properties considering the case ship only operates domestically in Denmark with a voyage duration of 1 h.

4.2.2.2 Modelling limitations

- The models that will be employed do not allow the examination of detailed transient performance which is beyond the scope of this research.
- The models cannot be used to investigate the influences on performance of fuel cell and battery operating temperatures.
- The models are limited to ships plying routes that have access to H_2 replenishing and battery charging infrastructures.
- The models would need re-calibration if different types of power source are used.

4.2.3 System overview

To help determine the optimal design of a plug-in hybrid PEMFC and battery propulsion system for coastal ships, a quasi-steady-state model with PEMFC and lithium batteries that can be adjusted to scale has been developed. Figure 4.2 presents the system schematic of a plug-in hybrid PEMFC battery propulsion system. Such a layout is designed for coastal ships sailing on short routes with access to battery charging and H_2 replenishing facilities. The distribution system on a DC ship eliminates the need for dc to ac inverters within the main propulsion distribution system. Although there are still practical challenges to implementing such a DC distribution system, it directly matches the proposed power sources which are DC (Hansen and Wendt, 2015; Higier et al., 2017).

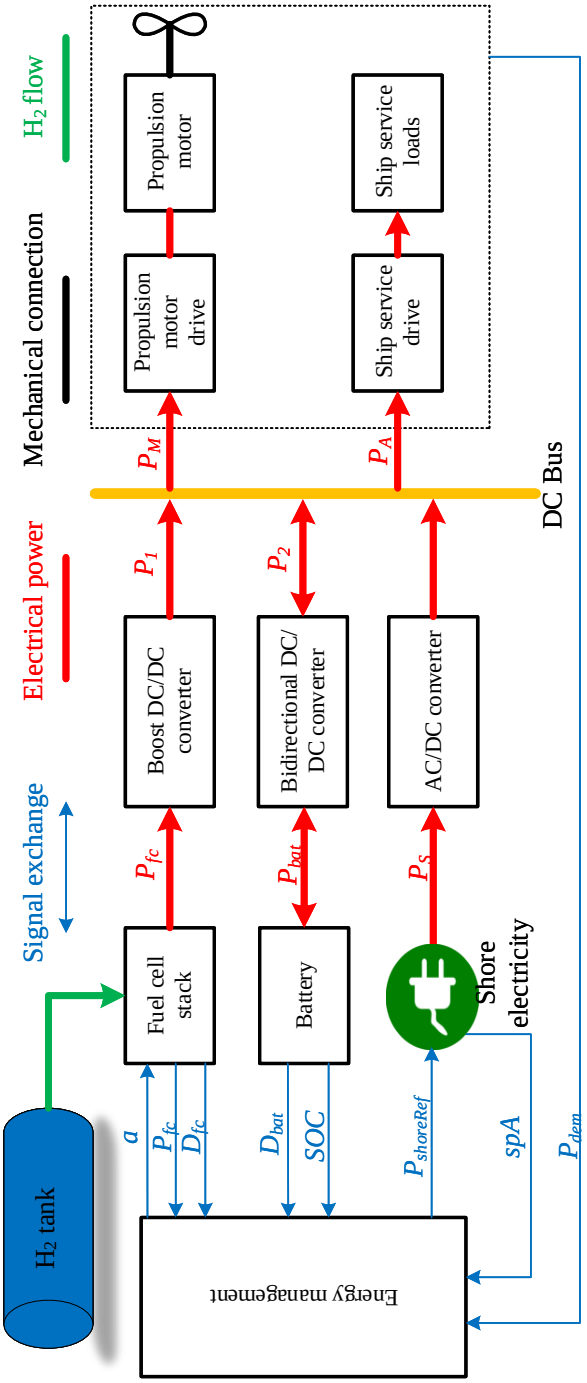


Figure 4.2: Hybrid propulsion system schematic.

The two main operational modes that are envisaged during the ship's routine are sailing and port modes. In sailing mode, either cruising or manoeuvring, the battery will work as an energy buffer to optimise the fuel cell loading and reduce the need for severe reaction to power transients at the PEMFC. Note that when the ship is in sailing mode the battery can undergo charging from the PEMFC or discharge power to the ship and propulsion. When the ship is berthed, i.e. in port mode, the shore connection charges the battery and powers the ship's electrical loads, while the PEMFC is idling during the operational turn around or shut down at the end of the operational day.

The energy management system manages the power split between the power sources by monitoring power demand, fuel cell power level, battery State of Charge (SOC) and shore power availability, then determines the fuel cell power change for the next time step. For each time step, the power supplied by the power sources has to satisfy:

$$P_1 + P_2 + P_s = P_{dem} \tag{4.1a}$$

$$P_1 = P_{fc}\eta_1 \tag{4.1b}$$

$$P_{fc} = \begin{cases} \geq 0, & \text{sailing mode} \\ = 0, & \text{port mode} \end{cases} \tag{4.1c}$$

$$P_s = \begin{cases} = 0, & \text{sailing mode} \\ \geq 0, & \text{port mode} \end{cases} \tag{4.1d}$$

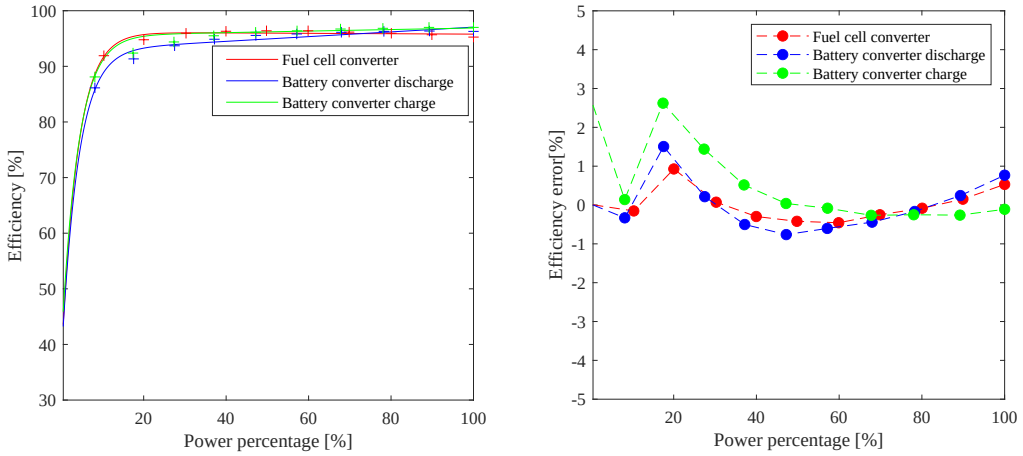
$$P_2 = \begin{cases} P_{bat}\eta_2, & P_{bat} \geq 0 \text{ for battery discharging or idling} \\ P_{bat}/\eta_3, & P_{bat} < 0 \text{ for battery charging} \end{cases} \tag{4.1e}$$

where P_1 and P_2 are the fuel cell and battery power after the power converters respectively; P_{fc} and P_{bat} are fuel cell stack power output and battery power output respectively; P_{dem} is the ship's lumped loads (including propulsive and auxiliary). Note that P_{fc} is determined by the EMS when in sailing mode and the fuel cell will not be providing any power to the system when the ship is on shore supply. Also

note that P_{bat} in port mode is calculated in line with the maximum charging rate neglecting the internal resistance. The actual charging rate would be slightly below the maximum charging rate due to the battery's resistance (will be identified in Section 4.2.7.3). η_1, η_2 and η_3 are the fuel cell and battery converter (charging and discharging) efficiencies respectively.

4.2.4 Power converter efficiency models

The power converter efficiency models output energy efficiency with the input of per unit power to the converters. Note that the converter efficiency models are representative of achievable characteristics and can be easily updated with real data when actual converter performance data is available for actual engineering applications. Such an approach is valid and has been widely adopted in similar studies (Hu et al., 2015a; Kalikatzarakis et al., 2018). As shown in Figure 4.3a, the power conversion efficiency is plotted against the percentage power output for each of the conversion modes. These efficiency models are based upon experimental data (Kanstad et al., 2019; Tseng et al., 2016). The maximum error of any of these best fit efficiency curves is 3%, which occurs at the 20% loading condition (Figure 4.3b). In high load regions, the percentage error is less than 1%.

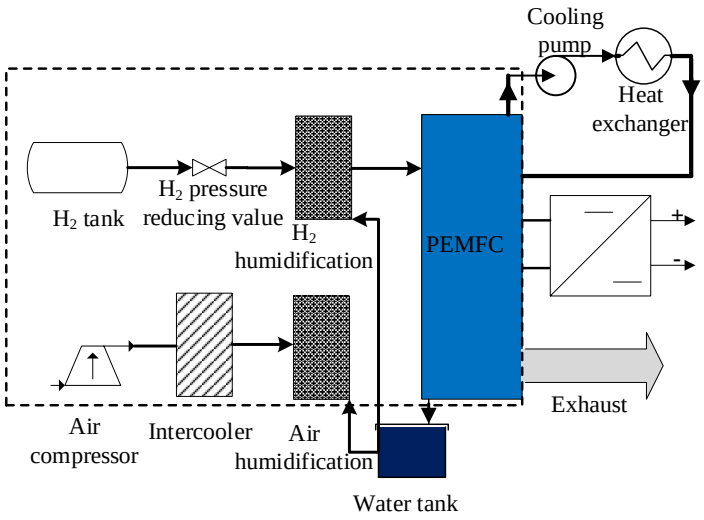


(a) **(b)**

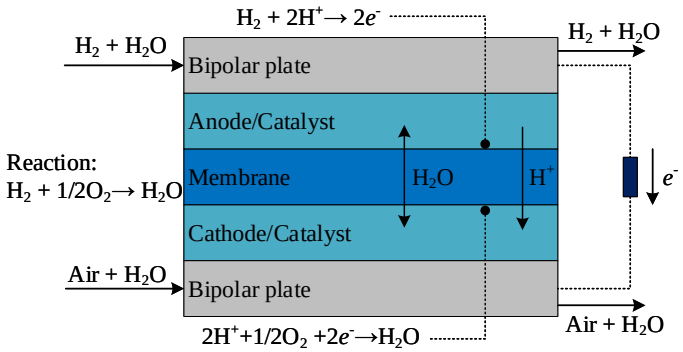
Figure 4.3: Power converter efficiency: (a) power converter efficiency curves fitted from experimental results and (b) errors of fitted efficiency curves compared to experimental data from (Kanstad et al., 2019; Tseng et al., 2016).

4.2.5 PEMFC model

To obtain fuel cell system specific hydrogen consumption and fuel cell degradation rates under different operating conditions, a system level PEMFC model has been developed and calibrated. Figure 4.4a represents the schematic of the PEMFC stack system, including ancillary components (Pukrushpan, 2003; Dicks and Rand, 2018). Figure 4.4b represents a single cell within the stack. The fuel cell stack model takes account of power consumed by the auxiliary equipment. Individual cell outputs are connected in series and parallel to form the stack output.



(a)



(b)

Figure 4.4: Fuel cell schematic: (a) PEMFC system and (b) single cell.

4.2.5.1 PEMFC modelling assumptions

- All the PEMFC's individual cells perform in exactly the same manner within a stack, and any internal connection resistance can be considered negligible.
- The individual PEMFC cells are one dimensional with cell current uniformly distributed. The humidified reactant gases are in equilibrium with liquid water. Water presents as vapour at the membrane-electrode interface and the water activity is uniform across the membrane and is in equilibrium state. Gas convection is neglected by assuming there is no pressure gradient between anode and cathode. There are no internal currents nor fuel crossover losses (Abdin et al., 2016).
- The PEMFC is properly cooled and operates at constant temperature, the temperature being uniform across individual cells. However, the operating pressure will vary with stack loading to match the cathode inlet air compressor operating line and improve efficiency at high load.

4.2.5.2 PEMFC single cell model

The single cell model considers three overpotentials, i.e. ohmic overpotential V_{ohm}, activation overpotential V_{act} and convection overpotential V_{con}. Due to the three overpotentials, the single cell output voltage V_{cell} varies at different current densities (corresponding to different loadings). The single cell governing equations are described in Table 4.1. The PEMFC model is calibrated using the results acquired from (Yan et al., 2006). The molar flow rates of O_2, H_2 and H_2O, i.e. n_{O_2}, n_{H_2} and $n^k_{H_2O}$ were calculated using the method provided in (Abdin et al., 2016). Model adjustable parameters such as β,i_L,γ_m, i_0^{Ref} and α are estimated by minimising the difference between model voltage output and experimental results with Particle Swarm Optimisation function provided in Matlab. Once these parameters are calibrated, the model is verified by running the model in other operating states.

Table 4.1: PEMFC governing equations.

Description	Equation	Reference
Oxygen partial pressure Hydrogen partial pressure	$p_{O_2} = \frac{n_{O_2}}{n_{O_2}/0.21+n_{H_2O}^{cat}} P_{cat}$ $p_{H_2} = \frac{n_{H_2}}{n_{H_2}+n_{H_2O}^{an}} P_{an}$	(Pukrushpan et al., 2002)
Open circuit voltage	$E_{oc} = 1.229 - 0.85 \times 10^{-3}(T - 298.15) + \frac{RT}{2F} \times \left(\ln p_{H_2} + \frac{1}{2}\ln p_{O_2}\right)$	(Dicks and Rand, 2018)
Concentration overpotential	$E_{act} = \frac{RT}{\alpha_{an}F}\ln\left(\frac{i}{i_0^{an}}\right) + \frac{RT}{\alpha_{cat}F}\ln\left(\frac{i}{i_0^{cat}}\right)$ $i_0^k = \gamma_M \exp\left[\frac{-\Delta Gc}{R}\left(\frac{1}{T} - \frac{1}{T_{ref}}\right)\right] i_0^{k,ref}, k = an, cat$	(Abdin et al., 2016)
Ohmic overpotential	$E_{ohm} = \frac{t_m i}{\sigma_m}$ $\sigma_m = (0.005139\lambda_m - 0.003260)\exp\left[1260\left(\frac{1}{303} - \frac{1}{T}\right)\right]$ $\lambda_m = \begin{cases} 0.043 + 17.81a_m - 39.85a_m^2 + 36.0a_m^3, & \text{for } 0 < a_m \le 1 \\ 14 + 1.4(a_m - 1), & \text{for } 1 < a_m \le 3 \\ 16.8, & \text{for } a_m > 3 \end{cases}$ $a_m = (a_{cat} + a_{an})/2$ $a_k = \frac{\varphi_k P_k}{P_{sat}(T)}$, k = an, cat $\log_{10}(P_{sat}(T)) = -7.90298 \times (T_s/T - 1) + 5.02808 \times \log_{10}(T_s/T)$ $- 1.3816 \times 10^{-7}\left[10^{11.344\times(1-T/T_s)-1}\right]$ $+ 8.1328 \times 10^{-3}\left[10^{3.49149\times(1-T_s/T)}\right] + \log_{10} P_s$ in which $T_s = 373.15\,\text{K}$ and $P_s = 101\,324.4\,\text{Pa}$.	(Pukrushpan et al., 2002) (Musio et al., 2011)
Concentration overpotential	$E_{con} = \beta i\left(c_1\frac{i}{i_L}\right)^{c_2}$ $c_1 = \begin{cases} (7.16 \times 10^{-4}T - 0.622)\left(\frac{p_{O_2}}{0.1173} + p_{sat}\right) + (-1.45 \times 10^{-3}T + 1.68), \\ \text{for } \left(\frac{p_{O_2}}{0.1173} + p_{sat}\right) < 2\text{atm} \\ (8.66 \times 10^{-5}T - 0.068)\left(\frac{p_{O_2}}{0.1173} + p_{sat}\right) + (-1.6 \times 10^{-4}T + 0.54), \\ \text{for } \left(\frac{p_{O_2}}{0.1173} + p_{sat}\right) \ge 2\text{atm} \end{cases}$	(Ang et al., 2010)

4.2.5.3 PEMFC stack model

By connecting the individual fuel cell outputs in series and parallel, the PEMFC stack model is formulated (Figure 4.5). Note that the equivalent circuit resistances could vary with cell stack density (see Table 4.1). It is assumed that all the cells within the stack perform uniformly. The total electrical power output of the stack is calculated by:

$$P_{fc} = V_{fc} I_{fc} - P_{com} - P_{aux} \tag{4.2}$$

where V_{fc} and I_{fc} are fuel cell stack voltage and current output respectively, and $V_{fc} = n_1 V_{cell}$, $I_{fc} = m_1 I_{cell} = m_1 i A_m$ (m_1 is number of strings in parallel, n_1 is number of cells in series, A_m is the active area of membrane electrode assembly). P_{com} is the power consumed the by air compressor:

$$P_{com} = \frac{1}{\eta_c \eta_m} \dot{m}_{air} c_p T \left[CR^{\left(\frac{\gamma_a - 1}{\gamma_a}\right)} - 1 \right] \tag{4.3}$$

where η_c is the air compressor efficiency, obtained from (Dicks and Rand, 2018) and normalised with respect to fuel cell loading. η_m is the compressor motor efficiency, and it is assigned a constant value of 0.96 in this study. $c_p = 1004\,\mathrm{J\,kg^{-1}\,K^{-1}}$ is the inlet air specific heat capacity. CR is the cathode air compression ratio which increases along the compressor operating line when fuel cell load level increases. γ_a is the air specific heat ratio. P_{afc} is balancing of plant power, and is a linear function of fuel cell power (Kim and Peng, 2007):

$$P_{afc} = k_1 P_{fc} + c_2 \tag{4.4}$$

where k_1 and c_2 are constant coefficients. The fuel cell stack net efficiency is calculated by:

$$\eta_{fc} = \frac{P_{fc}}{\dot{m}_{H_2} LHV} \times 100\% \tag{4.5}$$

where LHV is the Lower Heating Value of H_2, $\dot{m}_{H_2}$ is hydrogen mass flow rate which can be calculated from cell H_2 molar flow rate n_{H_2}, molar mass and total cell number.

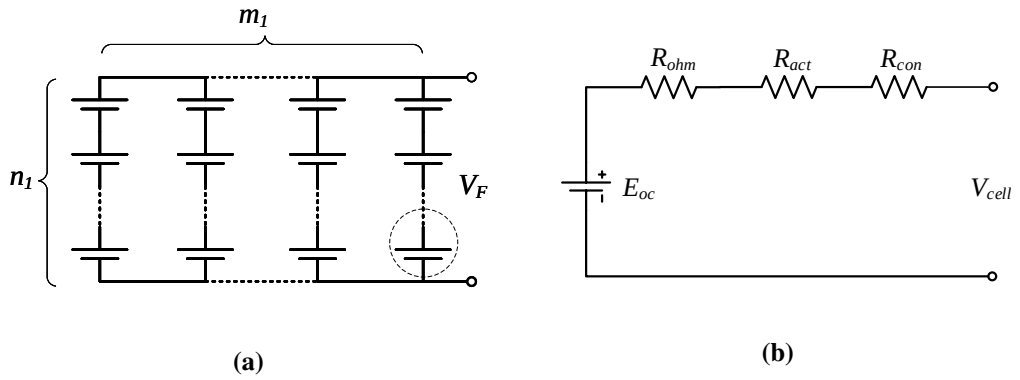


Figure 4.5: PEMFC equivalent circuit: (a) PEMFC module schematic and (b) PEMFC single cell equivalent circuit.

Figure 4.6a shows the single cell power at different current densities and corresponding power consumed by the air compressor and other auxiliary loads. The fuel cell stack net efficiency and specific H_2 consumption across 0-100% fuel cell load are shown in Figure 4.6b. The maximum fuel cell stack net efficiency is about 52%; the net efficiency at full load is 36%. As in Figure 4.6c, the fuel cells operating pressure increases along the compressor operating line with the increase of fuel cell load (Dicks and Rand, 2018). The fuel cell model outputs specific H_2 consumption with the input of fuel cell power percentage.

4.2.5.4 PEMFC degradation model

Compared to marine diesel engines, PEMFCs generally have a much shorter operational life. Factors such as power transients, cycling frequency and loading conditions can influence the rate of degradation. Considering the high production costs of PEMFCs, it is necessary to include the degradation characteristics in both the design development and operating phases of the ship to obtain the overall optimal cost performance. Table 4.2 details the PEMFC's single cell degradation rates used in this study (Fletcher et al., 2016; Chen et al., 2015). The cell degradation resulting from one voyage is:

$$D_{fc} = D_{low} + D_{high} + D_{transient} + D_{cycle} \tag{4.6}$$

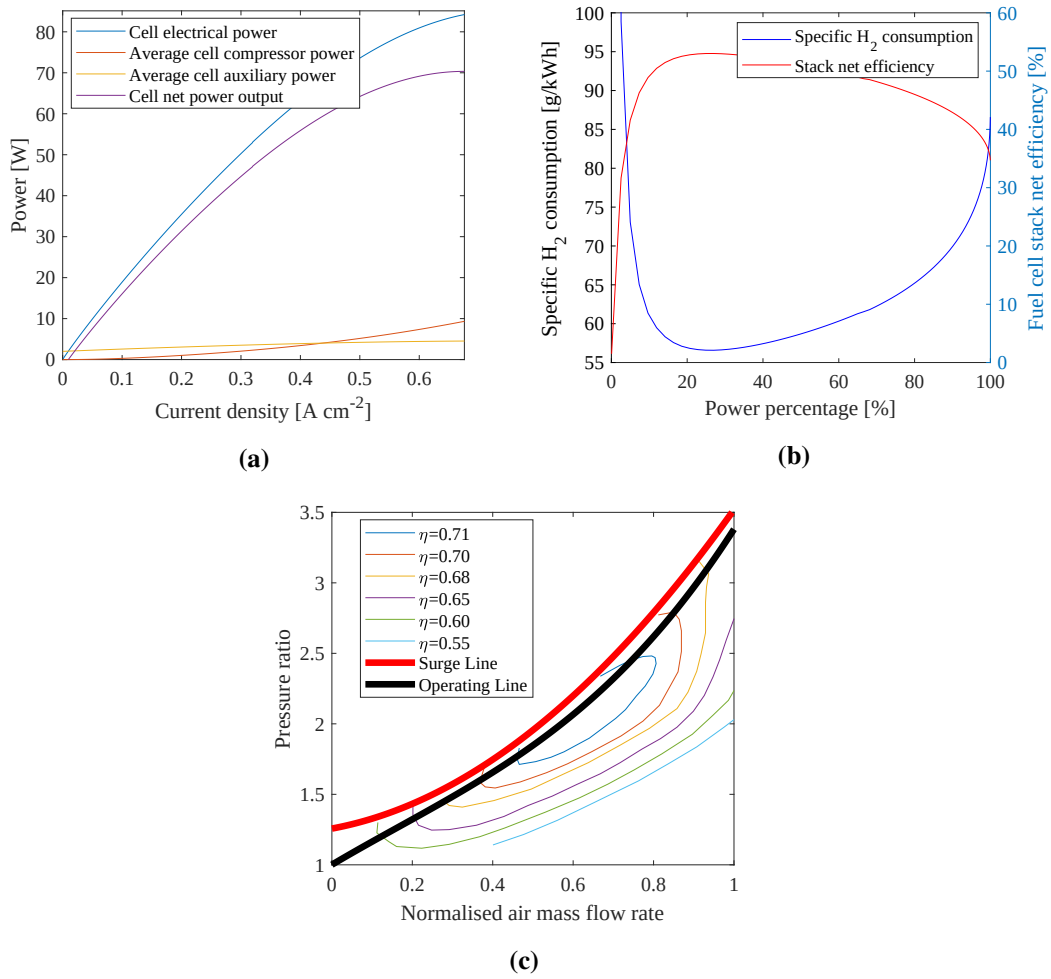


Figure 4.6: (a) Single cell power and auxiliary power consumption at different operating current densities, note that the compressor operates following the operating line in (c), the cell net power is cell electrical power subtracted by compressor and auxiliary loads and the net power output drops down rapidly after 0.68 $\mathrm{A\,cm^{-2}}$ hence regions beyond this point are not considered as useable region, (b) specific H_2 consumption and net fuel cell stack efficiency in different fuel cell power fractions, highest efficiency appears around 20% power region and (c) compressor map from (Dicks and Rand, 2018).

where D_{low}, D_{high}, $D_{transient}$ and D_{cycle} are cell voltage degradation caused by: low power operation, high power operation, power transient and start/stop cycling, respectively. These parameters will differ across different fuel cell types, depending on the design and actual operating parameters, and the model will require recalibration using experimental data before use.

Table 4.2: PEMFC cell degradation rates (Fletcher et al., 2016; Chen et al., 2015).

Operating condition	Degradation rate
Low power (0-80% rated power)	$10.17\,\mu\mathrm{V/h}$
High power ($>$ 80% rated power)	$11.74\,\mu\mathrm{V/h}$
Transient loading	$0.0441\,\mu\mathrm{V}/\Delta\mathrm{kW}$
Start/stop	$23.91\,\mu$V/cycle

4.2.6 PEMFC model validation

Figure 4.7 presents the fuel cell validation results against actual experimental data. Note that the validation data is acquired from Yan et al. (2006), and the experimental data in Figure 4.7b, i.e. at the operating pressure of 2 atm for both anode and cathode was used to calibrate the adjustable model parameters. The operating temperature is fixed at 80 °C. A good agreement between the model output and experimental results can be observed in Figure 4.7a, 4.7c and 4.7d. The root-mean-squared error between the model output and experimental data is 0.031 (the normalised root-mean-squared error is 3.5%). It is worth mentioning that increasing operating pressure brings higher cell voltage output, especially in high current density load regions. The degradation model is directly fitted from experimental data and can be deemed valid for this study.

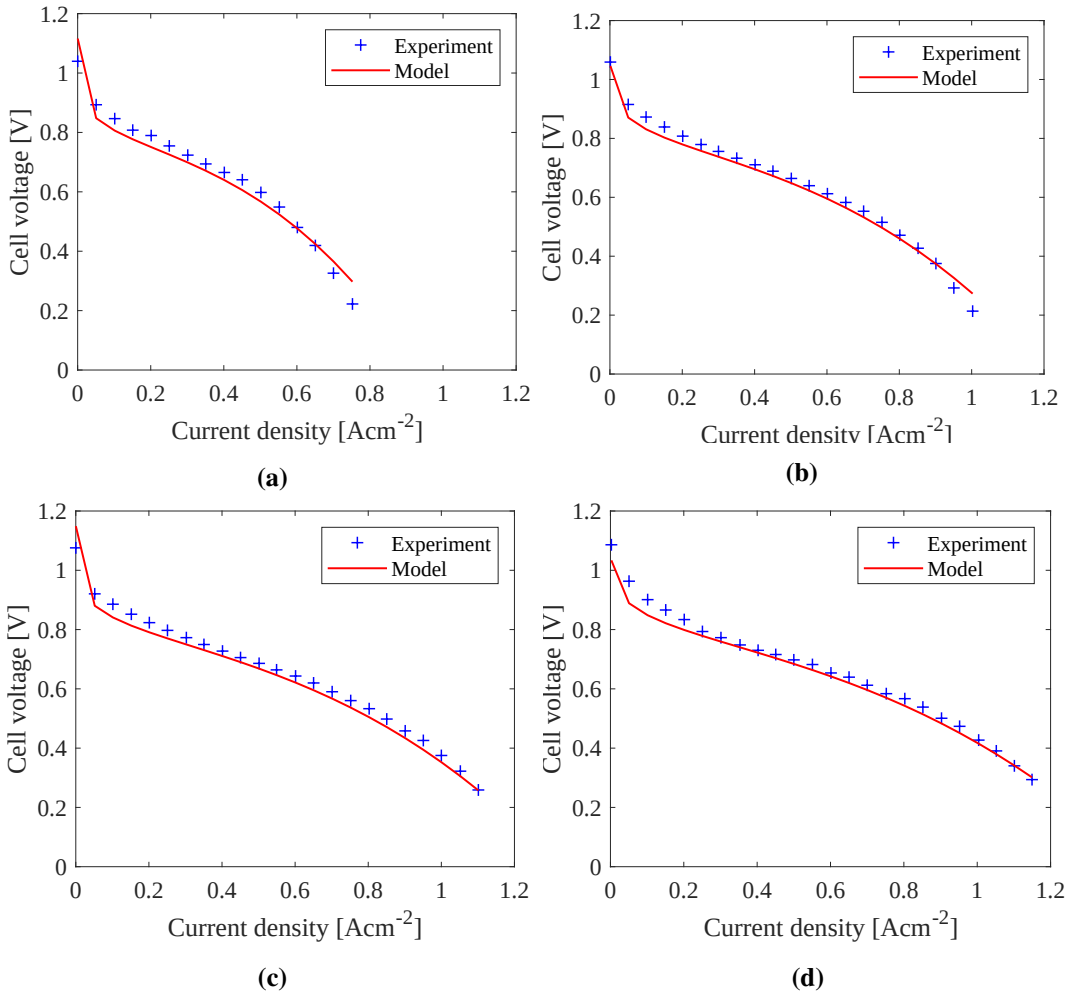


Figure 4.7: PEMFC model calibration and validation using experimental data from (Yan et al., 2006). The operating temperature is set at 80 °C for all the operating pressures. From (a) to (d), the operating pressures are 1 atm, 2 atm, 3 atm and 4 atm, respectively. Note that the experiment data of 2 atm is used to calibrate the model adjustable parameters.

4.2.7 Lithium-ion battery model

To accurately represent battery performance across its capacity range, an equivalent circuit calibrated using experimental data from (Zheng et al., 2016) has been developed. Figure 4.8 shows the battery model equivalent circuit adopted in this study. The battery cells are connected in series and parallel to form the battery module (Hu et al., 2015a):

$$P_{bat} = I_B V_B \tag{4.7}$$

where I_B, V_B are battery module current and voltage respectively. Note that $V_B = n_2 V_t$, and V_t is the battery cell terminal voltage, which is a function of battery cell open circuit voltage $V_{oc}(SOC)$ (V_{oc} is a function of battery SOC), R_0, R_1 and C_1:

$$V_t = V_{oc} - i_0 R_0 - V_c \tag{4.8}$$

$$\dot{V}_c = -\frac{V_c}{R_1 C_1} + \frac{i_0}{C_1} \tag{4.9}$$

Over a period, the battery SOC changes as:

$$SOC_2 = SOC_1 - \eta_b \int_{t_1}^{t_2} C(t)dt \tag{4.10}$$

where V_c is the voltage over the capacitor C_1; $C(t)$ is battery charge rate (unit is C, 1 C corresponds to the current which discharges the cell from full capacity to zero capacity in 1 h) at time t; note $C(t)$ is positive for battery discharging and $C(t)$ is negative for battery charging; η_b is the battery coulombic efficiency.

4.2.7.1 Lithium-ion battery modelling assumptions

- All battery cells perform exactly the same within a battery module, and resistance of the internal connections can be deemed as negligible (Hu et al., 2015a).

- All the battery cells are properly cooled and work at the temperature of the experiment carried out to identify model parameters.

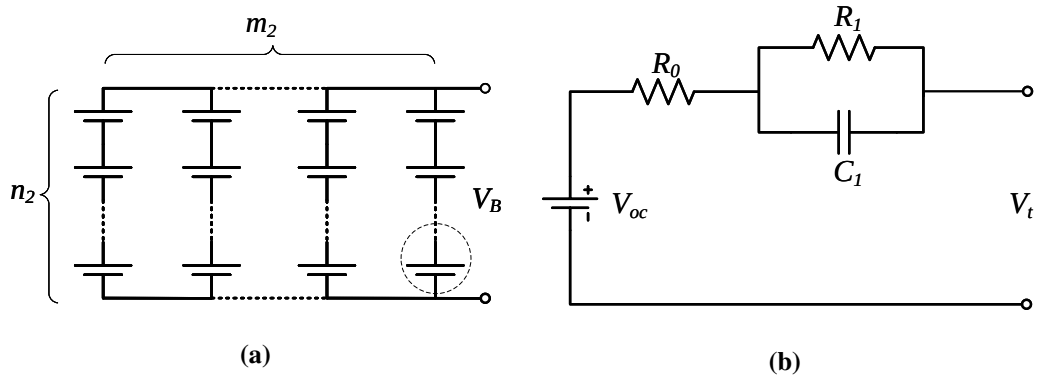


Figure 4.8: Battery equivalent circuit: (a) battery module schematic and (b) battery single cell equivalent circuit.

- Batteries can be fast-charged at a C-rate of up to 3 C in port. This assumption is based on the requirement that the case ship has only very limited port time. Lower charging rate would inevitably need more time in port and would not fit the operational profiles. When the ship is in normal operation, in sailing mode, the maximum allowable C-rate is 6 C but only in extremis. It is recommended not to exceed the C-rate beyond 3 C under discharge conditions to prolong battery life Corvus (2017).

4.2.7.2 Open circuit voltage

The low-current OCV-SOC experimental data from (Zheng et al., 2016) was applied to identify the OCV-SOC map. Such low-current tests can be used to approximate the OCV-SOC mapping by charging and discharging the cell with very low current (e.g. 0.05 C) (Dai et al., 2012; Xing et al., 2014). In the OCV-SOC test of Zheng et al. (2016), the battery was initially charged at a constant voltage until the charging current reduces to 0.01 C, indicating the cell had been effectively charged to 100% SOC. Subsequently, the cell was fully discharged until the terminal voltage is reduced to 2.5 V. The cell was then charged with a constant current of 0.05 C until the battery terminal voltage rises 4.2 V. In this way, the experimental results were acquired to identify the OCV-SOC mapping, i.e. $V_{oc}(SOC)$. Note that due to the presence of hysteresis (Roscher et al., 2011; Eichi and Chow, 2012), the charging OCV ($V_{oc,ch}$) tends to be higher than the discharging OCV ($V_{oc,dis}$) (see Figure 4.9).

Therefore, an average V_{oc} is obtained via:

$$V_{oc}(SOC) = \frac{V_{oc,dis}(SOC) + V_{oc,ch}(SOC)}{2} \tag{4.11}$$

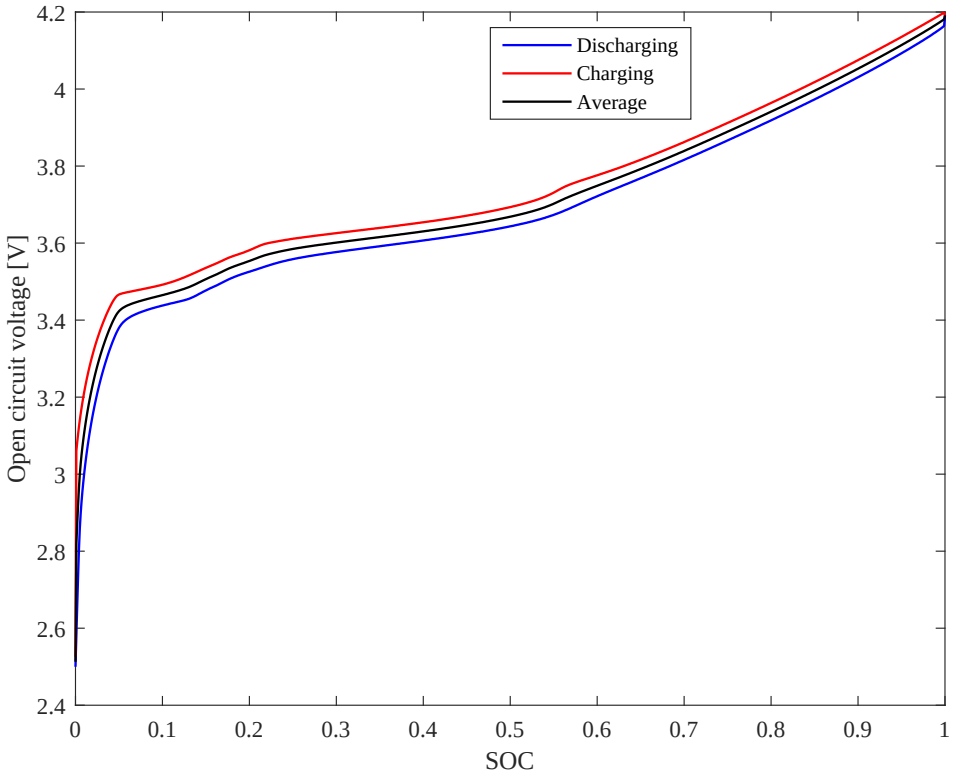


Figure 4.9: Battery cell open circuit voltage.

4.2.7.3 Equivalent circuit parameters

Having identified the open circuit voltage map $V_{oc}(SOC)$ in Section 4.2.7.2, R_0, R_1 and C_1 were identified using a least square algorithm with the dynamic stress test (DST) experiment data (at 25 °C) from (Zheng et al., 2016). As the influence of battery temperature is not the focus of this study, all these parameters were identified using experimental data at 25 °C. The identified parameters are presented in Table 4.3. There was good agreement between the model output and experimental data (Figure 4.10). Due to the large time step and the main focus of this study being energy efficiency and emissions, as opposed to dynamic performance, the direct current internal resistance ($R_0 + R_1$) of the battery model is used in the subsequent simulations (Kim et al., 2012).

Table 4.3: Battery parameters.

R_0 [Ω]	R_1 [Ω]	C_1 [F]
0.074	0.022	1624.628

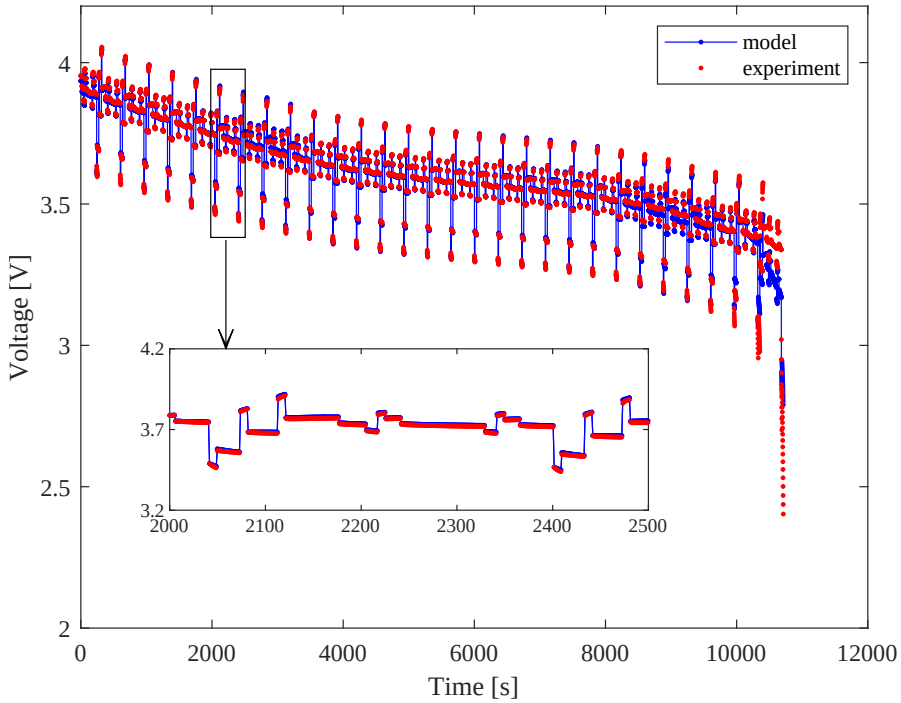


Figure 4.10: Battery model calibration with the dynamic stress test.

4.2.8 Lithium-ion battery model validation

To validate the battery cell model, the urban driving schedule (UDS), i.e. a time series of battery currents normalised from an UDS load cycle, from (Zheng et al., 2016) was applied. As depicted by Eq. 4.8, the battery cell model outputs terminal voltage V_t with given current i_0. The battery model outputs are compared against the experimental measurements of battery terminal voltage as in Figure 4.11. A good agreement between the model outputs and experimental results can be observed.

4.3 Multi-objective propulsion system optimisation

Based upon the hybrid fuel cell and battery propulsion system model, the proposed multi-objective propulsion system sizing methodology optimises the power source

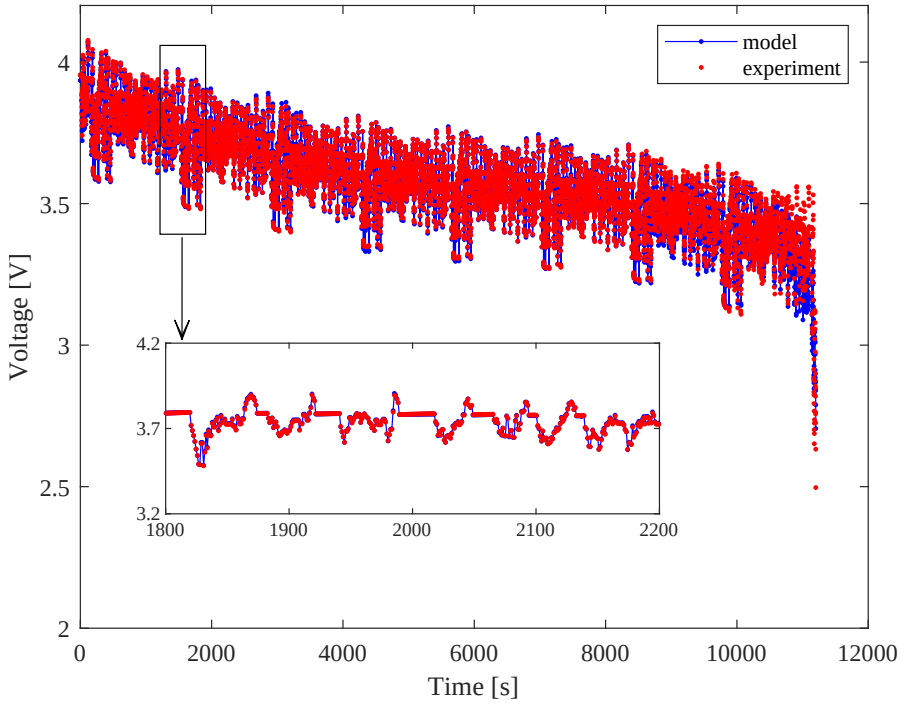


Figure 4.11: Battery model validation using urban driving schedule.

sizes as shown in Figure 4.12. The two-layer multi-objective design methodology looks to solve the power source sizing problem by considering both emission and economic performances. With an average operating profile normalised from route performance data obtained from an actual ship of similar capacity or other routines such as scaled experiments, the inner layer solves the optimal power split problem for each power source sizing combination passed from the external layer to minimise average voyage cost.

The optimisation method adopted for the external layer is Non-dominated Sorting Genetic Algorithm II, based upon the work of Deb et al. (2002). Such an approach has been successfully applied in the field of hybrid propulsion system sizing optimisation (see Table 3.1). The external layer optimises the average voyage cost and GWP emissions concurrently. The decision variables of the external layer are a vector of the fuel cell and battery module sizing parameters. The trade-off between the two objectives needs to be determined manually based on the Pareto front.

4.3.1 External layer

Both fuel cell and battery modules are emission-free if the scope is limited only to the ship. However, looking beyond the ship itself, the production of H_2 and generation of shore power will have specific carbon footprints. The high prices, shorter life and lower energy/power densities are significant drawbacks of these technologies when compared to conventional diesel engine based solutions. To balance the trade-off between emission and economic performance, a multi-objective power source sizing methodology considering both emission performance and average voyage cost is formulated as follows:

$$\text{minimise: } F(Z) = [f_1, f_2]^T \tag{4.12}$$

$$\text{subject to: } \begin{cases} K_1 P_{dg} \leq P_{fc}^R + P_{bat,1C} \\ K_2 P_{dg} \leq P_{fc}^R + P_{bat,3C} \\ V_x \leq V_D \\ W_x \leq W_D \end{cases} \tag{4.13}$$

where $Z = [m_1, n_1, m_2, n_2]^T$ is a vector of power source sizing integer variables, which is constrained by its lower and upper limits (Z_1 and Z_2); $F(Z)$ is a vector of objective functions: f_1 is the average voyage cost including fuel cell and battery degradation costs, hydrogen cost and shore electricity costs, f_2 is voyage GWP emission comprising through life GWP emissions from hydrogen production and shore electricity generation:

$$f_1 = \sum_{t=1}^{T} \psi_{fc}(x_t + \frac{u_t}{2}) P_{fc} \Delta t \, \sigma_{H_2} + \sum_{t=1}^{T} \delta_{fc}(x_t + \frac{u_t}{2}) P_{fc} \sigma_{fc} + \sum_{t=J}^{T} P_{sh} \Delta t \, \sigma_e + \sum_{t=1}^{T} \delta_{bat} B \sigma_{bat} \tag{4.14}$$

$$f_2 = \sum_{t=1}^{T} \psi_{fc}(x_t + \frac{u_t}{2}) P_{fc} \Delta t \, \upsilon_{H_2} + \sum_{t=J}^{T} P_{sh} \Delta t \, \upsilon_e \tag{4.15}$$

where ψ_{fc} is the fuel cell specific hydrogen consumption function (see Figure 4.6b), x_t is the fuel cell power fraction in time step t and u_t is the fuel cell power change in time step t (note that H_2 consumption is calculated by an averaged fuel cell power

level in one time step), T is the final time step; δ_{fc} is the fuel cell degradation function and is determined by the degradation rates in Table 4.2; P_{sh} is shore power when the ship is in port; δ_{bat} is an averaged battery degradation rate in each time step, assuming the batteries are guaranteed to last for a specific period; σ_{H_2}, σ_e, σ_{fc} and σ_{bat} are H_2, shore electricity, fuel cell and battery prices respectively; υ_{H_2} and υ_e are H_2 and shore electricity GWP respectively; B is battery capacity.

V_x and W_x are hybrid system volume and weight (including both power sources and fuels) respectively; V_D and W_D are original diesel system volume and weight respectively; P_{fc}^R, $P_{bat,1C}$, $P_{bat,3C}$ and P_{dg} are rated fuel cell power output, battery output power at 1 C discharge rate, battery output power at 3 C discharge rate and diesel generator rated power output in the original system. The first two constraints guarantee that the hybrid system can provide sufficient power. K_1 is the highest frequency statistically weighted number of diesel engines required online. The hybrid system should be able to provide sufficient power without need for fast charging/discharging the batteries in the normal operation scenarios. Therefore, the battery power delivered at 1 C together with the rated fuel cell power should not be less than the capacity provided by the K_1 number of diesel generators. K_2 is the number of installed diesel generators in the original diesel-electric system. Note that K_2 is usually larger than K_1 in an IFEP system for reasons of redundancy. Since the batteries can provide higher C-rates than 1 C, such capabilities are ideal for providing extra redundancy without installing additional power sources (e.g. extra generators in the original IFEP system). The battery power delivered at recommended maximum C-rate (3 C, see Section 4.2.7.1 for battery modelling assumptions) together with the fuel cell rated power output should not be less than the total installed diesel generator power. The last two constraints in Eq. 4.13 ensure that the hybrid's system volume and weight do not exceed those of the original diesel-electric system.

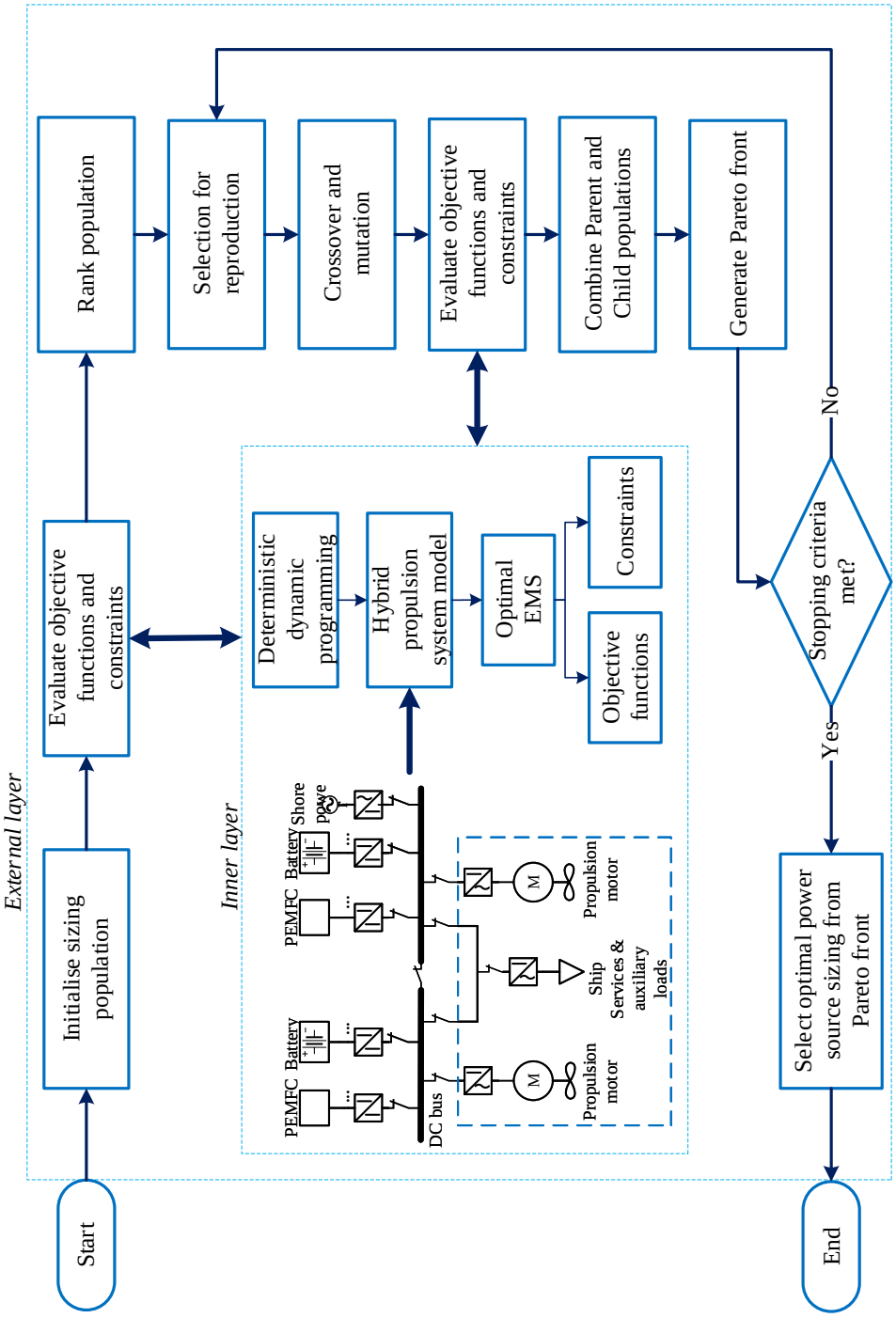


Figure 4.12: Proposed holistic design methodology.

4.3.2 Inner layer—optimal energy management for the average operating profile

The inner layer of the sizing process is solved using a process of reverse iteration to find the optimal energy management strategy π for the average power profile for each feasible power source sizing combination Z using Deterministic Dynamic Programming (DDP) (Sundström et al., 2010). As a special case of stochastic dynamic programming, the transition from one state to the next is deterministic in DDP; hence the computation time is significantly reduced when compared to that of stochastic dynamic programming. DDP requires complete knowledge of voyage power demands during all time steps, which is not realistic for an on-line strategy but can be used to obtain an optimal off-line strategy to benchmark the quality of the on-line strategy. The optimal strategy $\pi^*(u|x)$ minimises the objective function f_1 subject to the constraints listed in Eq. 4.16:

$$u \in U \tag{4.16a}$$

$$SOC_{min} \leq SOC \leq SOC_{max} \tag{4.16b}$$

$$P_{bat,min} \leq P_{bat} \leq P_{bat,max} \tag{4.16c}$$

$$P_{bat,port} - P_{dem,port} + P_{sh,max} \geq 0 \tag{4.16d}$$

$$\begin{cases} x = 0, & \text{port mode} \\ 0 \leq x \leq 1, & \text{sailing mode} \end{cases} \tag{4.16e}$$

where U is a set of fuel cell power change ratios. x is fuel cell power expressed in per unit. $P_{sh,max}$ is the maximum available shore power; when $P_{bat,port}$ is negative this indicates the battery charging power in port mode. $P_{dem,port}$ is the ship's power demand in port mode. Note that $P_{bat,min}$ is a negative value, and corresponds the maximum battery charging power. With the optimal energy management strategy solved, the second objective function, that of the GWP emissions, can then be calculated.

4.4 Case studies and discussion

4.4.1 Simulation settings

In this section, the proposed hybrid fuel cell and battery propulsion system model and power sources sizing methodology are applied to the case ship with the specification as presented in Table 3.3. The original system featured an integrated full electric propulsion configuration with a total installed diesel powered capacity of 4370 kW. It is assumed that the annual operating duty is 300 days, and the ship operates between two fixed ports with 16 voyages (8 round trips) per day—each voyage taking approximately one hour. It is also assumed that the batteries can be charged at both ends of the defined voyage, and hydrogen needs to be replenished at the end of the operational day. The simulations were implemented in Matlab 2019a on a workstation with two Intel Xeon E5-2683 V3 processors and 64 GB memory.

The case ship (Figure 3.2a) and its route (Figure 3.2b) have been detailed in Chapter 3 (Eriksen et al., 2018). The continuous monitoring power data, compiled from measurements obtained from the original diesel powered ship upon which the case ship is based, acquired from (Eriksen et al., 2018) was segregated into voyage power profiles by judging the ship's speed and location. An average operating profile based upon these profiles was generated for sizing and optimisation, as shown in Table 4.4. Table 4.5 shows the grids of state and action spaces for the inner layer DDP. Note the fuel cell power change fraction is defined by considering fuel cell power ramp up/down limits (Rouholamini and Mohammadian, 2016).

The original time step of the raw data used to generate the power profile was 15 s. The same time step was adopted in this study. Since the problem is solvable with the available computational power, the time step was not increased. The DDP energy management solver in the inner layer can generate an off-line strategy in less than 1 s with the above-mentioned settings and time step. Fletcher et al. (2016) reported that their Stochastic Dynamic Programming algorithm took approximately 6 h to generate an on-line strategy using a quad-core processor. Nevertheless, it is

worth noting that the application of DDP is limited to on-line strategy due to the requirement of complete profile information being available before a solution can de derived.

Table 4.4: Average operating profile for sizing.

Parameter	Value	Unit
Port time	810	s
Departing time	375	s
Sailing time	2025	s
Approaching time	375	s
Sailing power	1454	kW
Power at shore	283	kW

Table 4.5: State and action space grid setting for dynamic programming.

Parameter	Grid length	Range
Power demand	50	0–4400 kW
SOC	0.05	0–1
Fuel cell power level	0.02	0–1
Shore power availability	–	0 or 1
Fuel cell power change fraction	0.02	$[-0.04, -0.02, 0, 0.02, 0.04]$

Table 4.6 details the parameters adopted for the case studies. The power sources' parameters were acquired from manufacturer's specifications. It should be noted that the optimal design would most likely be different for different geographical locations as they may have different energy supply tariffs and fuel costs. Since H_2 GWP and price can depend on the method of and energy source for H_2 production, H_2 produced by electrolysis using two different energy sources (by nuclear power generation and wind power generation) are considered in two case studies. The first case study optimises they system for the ship's original operating area (Denmark), whereas the second case study investigates the optimising of the hybrid system should the ship operate in UK waters with UK energy tariffs.

Table 4.6: Case study parameters.

Category	Parameter	Value	Unit	Reference
Diesel system	Original diesel generator system volume	294	m^3	(Caterpillar, 2018)
	Diesel system price	500	$\$kW^{-1}$	
	Diesel system lifetime	20	year	
Marine gas oil	Original diesel generator system weight	156	t	(BunkerIndex, 2017) (Kristensen, 2012)
	Marine gas oil price	0.64	$\$kg^{-1}$	
	Marine gas oil GWP	3.2	kg CO_2 kg^{-1}	
PEMFC system	PEMFC price	830	$\$kW^{-1}$	(Raucci, 2017)
	PEMFC volumetric specific power	128	$kW\,m^{-3}$	(Ballard, 2019)
	PEMFC gravimetric specific power	200	$kW\,t^{-1}$	
	H_2 tank volume	0.17	$m^3\,kg^{-1}$ H_2	(Choi et al., 2016)
	H_2 tank weight	28.5	$kg\,kg^{-1}$ H_2	
Battery system	Battery price	800	$\$kWh^{-1}$	(Ovrum and Dimopoulos, 2012)
	Battery lifetime	1.5	year	(Stroe et al., 2015)
	Battery volumetric specific energy	91.8	$kWh\,m^{-3}$	(Corvus, 2017)
	Battery gravimetric specific energy	80.6	$kWh\,t^{-1}$	
	Battery SOC higher limit	0.25		
	Battery SOC lower limit	0.90		
	Battery maximum C-rate	6	C	
UK electricity	Shore electricity price	0.16	$\$kWh^{-1}$	(Eurostat, 2019)
	Shore electricity GWP	0.28	kg CO_2 kWh^{-1}	
Denmark electricity	Shore electricity price	0.09	$\$kWh^{-1}$	(Eurostat, 2019)
	Shore electricity GWP	0.17	kg CO_2 kWh^{-1}	
Nuclear-powered electrolysis H_2	H_2 price	3.50	$\$kg^{-1}H_2$	(Acar and Dincer, 2014)
	H_2 GWP	2.40	kg CO_2 kg^{-1}	
Wind-powered electrolysis H_2	H_2 price	8.24	$\$kg^{-1}$ H_2	(Acar and Dincer, 2014)
	H_2 GWP	1.50	kg CO_2 kg^{-1}	

4.4.2 Case study 1

For this case study (Denmark), detailed fuel cell and battery sizing combinations with regards to the cost and emission objectives are presented in Figure 4.13 for the two considered two H_2 scenarios (Table 4.7).

Table 4.7: Case study 1 scenarios.

Scenario	Electricity	H_2
1-a	Denmark	nuclear power generation
1-b		wind power generation

In the scenario of H_2 produced using nuclear-powered high temperature electrolysis (Figure 4.13a), the minimum voyage GWP that can be achieved is about 155 kg CO_2, at a voyage cost of \$610. This minimum GWP corresponds to a PEMFC power of 4320 kW and a battery capacity of 37 kW h. Note that the fuel cell power at this point is very close to the original total diesel engine power of 4370 kW, and the battery capacity is comparatively negligible when compared to the installed fuel cell power. Such an observation suggests that, under the parameters governing this case study, to achieve extremely low GWP emissions, a large fuel cell plant combined with a comparatively small capacity battery would be the optimum solution. However, it would be more reasonable to select the design point around fuel cell power of 2915 kW and battery capacity of 600 kW h, which corresponds the knee point (voyage cost of \$480 and GWP of 170 kg) of the Pareto front as in Figure 4.13a–3. Such a design point avoids unfavourably large cost sacrifice to achieve minimal improvement in emission performance.

For the scenario of H_2 produced using wind-powered electrolysis (Figure 4.13b), an even lower GWP can be achieved (105 kg CO_2) due to the GWP of this H_2 approach (1.5 kg CO_2 kg^{-1}) is much lower than that from the nuclear-powered approach. However, such a low GWP emission would lead to a voyage cost of \$860. It is also true that a large fuel cell plant leads to lower emissions. A large battery plant can achieve lower average voyage cost (minimum \$615 per voyage). In practice, when choosing the design point, other factors need to be considered with reference to these Pareto fronts.

Note the Pareto front in Figure 4.13b–3 appears more linear compared to that in Figure 4.13a–3. This is mainly due to the nuclear power generated H_2 is much cheaper than that generated by wind power; increasing the fuel cell power from 2400 kW to 2915 kW would rapidly bring down the voyage GWP to 170 kg CO_2 with increased fuel cell efficiency (see Figure 4.6b). However, beyond this knee point, increasing the fuel cell power would lead to fuel cell degradation costs more dominating and the GWP saving could be achieved by increasing installed fuel cell becomes more challenging.

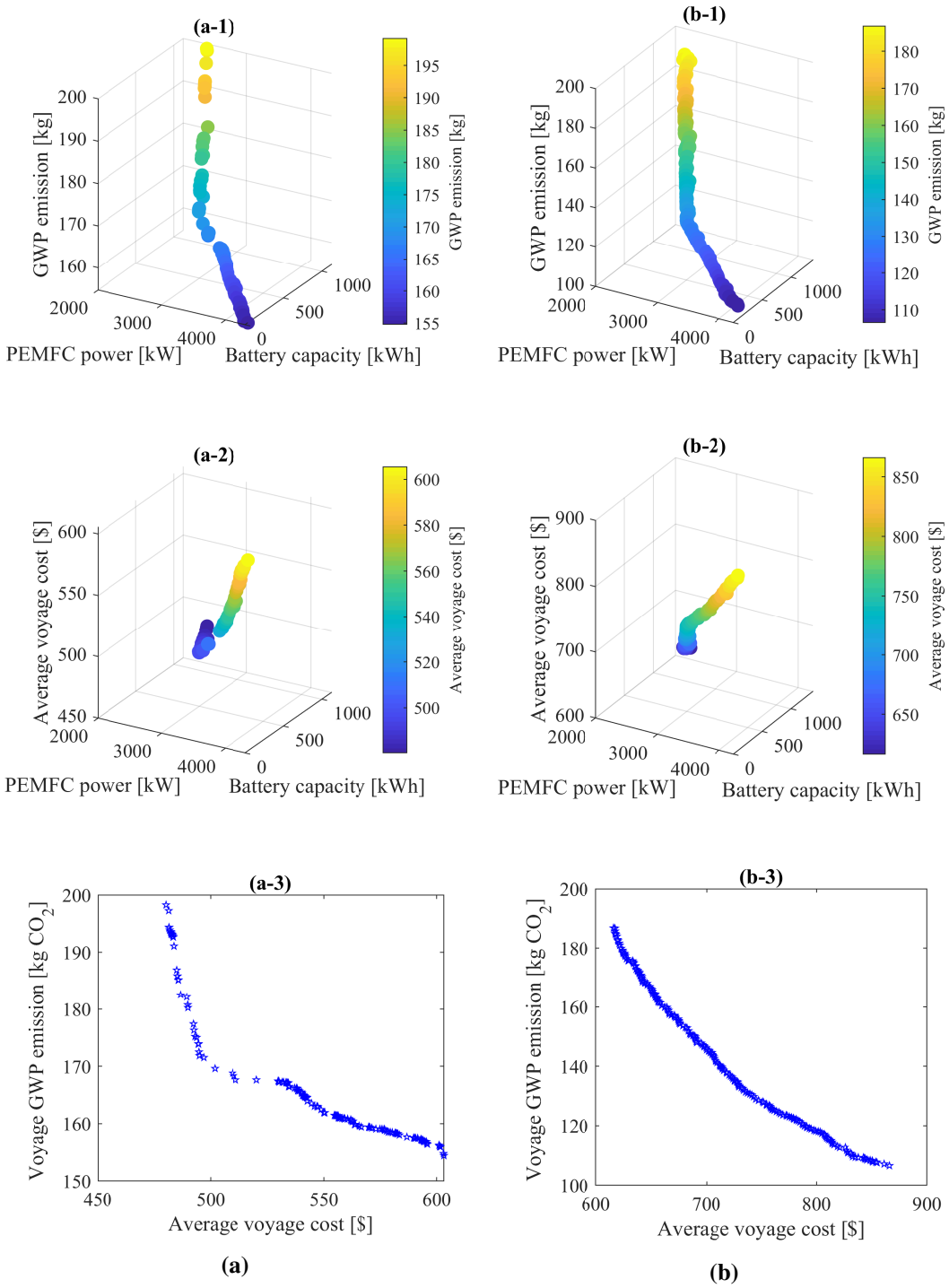


Figure 4.13: Case study 1 (Denmark electricity) fuel cell and battery sizing combinations with regards to the cost and emission objectives: (a) scenario of H_2 via nuclear-powered high temperature electrolysis and (b) scenario of H_2 via wind-powered electrolysis.

4.4.2.1 Optimal energy management

To further illustrate the optimal energy management strategy (Section 4.3.2), based on the Pareto front shown in Figure 4.13b, a design point of fuel cell power of 2940 kW and battery capacity of 581 kWh is selected for demonstration. Figure 4.14 shows the optimal strategy for the selected power source design under the averaged operating profile (Table 4.4). The battery provides power solely during departure and deliver a small portion of power while cruising. As this operating profile is averaged for sizing and does not include power transients, the optimal strategy does not charge the battery during sailing (i.e. the battery power flow is never negative during sailing). When the ship is in port, the fuel cell is switched to idle while the batteries are charged by shore power. Note that such a strategy is solved by DDP, and the entire profile is known in advance.

To better understand the compositions of voyage cost and emission of the selected design point discussed in Section 4.4.2.1, Table 4.8 shows the cost and emission breakdown of the proposed hybrid and original systems. H_2 cost contributes more than 50% of the average cost. The fuel cell degradation cost is a significant source as well. The total GWP emission is reduced by 85% with 226% cost increase compared to the original diesel-electric system. The hybrid system is more expensive than the conventional diesel engine based system. However, the Pareto front could vary significantly if parameters such as fuel cell price or H_2 price decrease.

Table 4.8: Voyage cost and GWP emission breakdown

	Cost [$]	GWP emission [kg CO_2]
Fuel cell	238	-
Hydrogen	368	67
Battery	65	-
Shore electricity	39	123
Total	*710*	*190*
Diesel engine	46	-
Diesel fuel	168	838
Shore electricity	4	4
Total	*218*	*842*

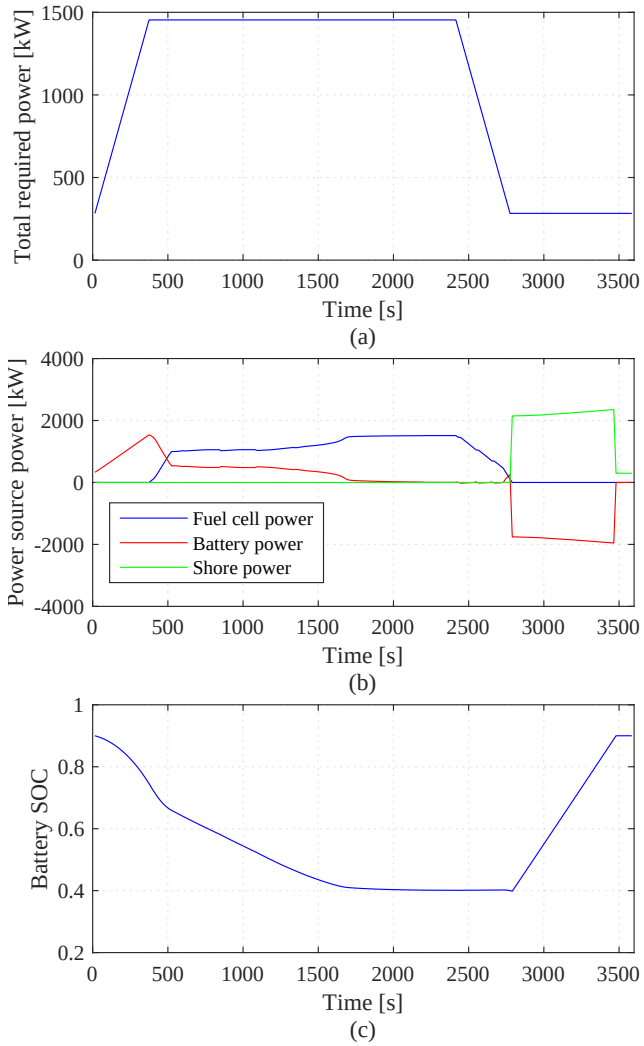


Figure 4.14: DDP solved energy management strategy for averaged power profile of selected power source sizing: (a) averaged power profile, (b) power distribution between power sources and (c) battery SOC trajectory.

4.4.2.2 Optimisation methodology validation

To validate the proposed multi-objective design methodology, sizing results generated by random search over the decision variable space (i.e. the sizing parameters of the PEMFC and batteries) are compared against the Pareto front generated by the optimisation solver as in Figure 4.15. The parameters used for this validation are identical to the ones applied for the Denmark–wind power generated H_2 case. Solutions violating the constraints (see Eq. 4.16) are filtered out from the random search solution space, leaving only the feasible ones as presented in Figure 4.15

(red circles). The lower boundary of the random search solutions matches well with the Pareto front. Such that, the optimisation methodology is validated. It is worth noting that the random search is also a feasible approach, but it would require more time compared to the optimisation solver.

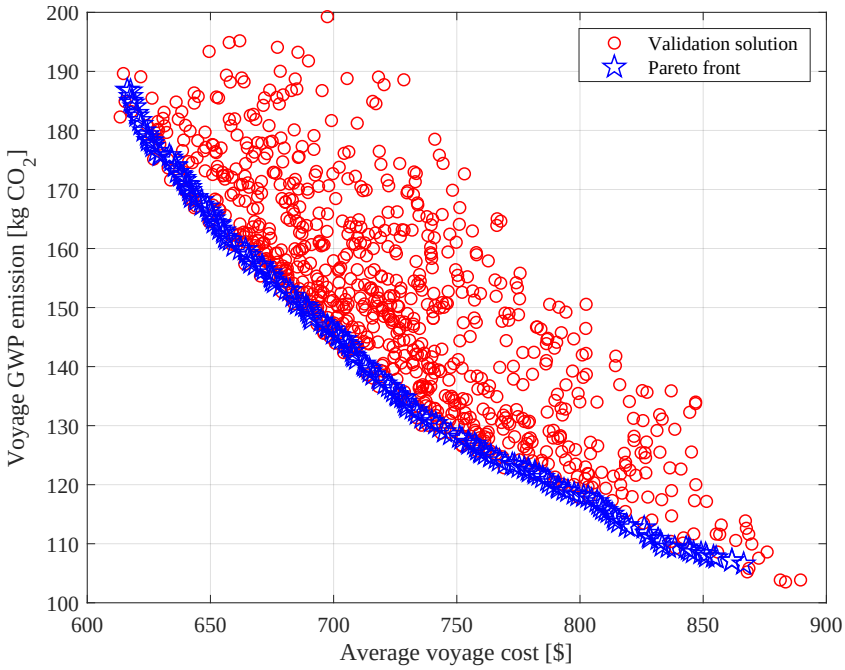


Figure 4.15: Multi-objective optimisation methodology validation against random search over the sizing parameters.

4.4.3 Case study 2

To understand how the proposed hybrid system should be optimised if it operates in UK waters and uses UK grid power, further investigations of the two H_2 scenarios (Table 4.9) are shown in Figure 4.16a and 4.16b. Similar to the Denmark case, larger fuel cells lead to lower GWP in both H_2 scenarios. Due to UK non-household electricity costs being higher than that of Denmark, the minimum voyage costs are \$510 (Figure 4.16a) and \$670 (Figure 4.16b) respectively. As in Figure 4.16a–1&2, increasing the fuel cell power from 2900 to 3090 kW (corresponding battery capacity is 425 kW h) would bring down the voyage GWP down from 215 to 200 kg CO_2. However, it is more challenging to further reduce voyage GWP by increasing installed fuel cell power beyond this knee point. As shown in Figure 4.16b, the knee

Table 4.9: Case study 2 scenarios.

Scenario	Electricity	H_2
2-a	UK	nuclear power generation
2-b		wind power generation

point is around the fuel cell power of 3060 kW and the battery capacity of 450 kW h. This knee point provides an average voyage cost of $750 and GWP of 170 kg CO_2. Comparing Figure 4.16a and 4.16b, wind power generated H_2 would provide much lower voyage GWP (minimum 110 kg CO_2), though the cheapest sizing solution is still more expensive than $600 (maximum voyage cost appeared in Figure 4.16a).

4.4.4 Summary of case studies

Figure 4.17 provides an overview above mentioned case studies. The two considered H_2 scenarios would both reduce GWP emission significantly. The maximum voyage GWP for the wind-powered electrolysis generated H_2 scenario is 290 kg CO_2 for the UK case (65% GWP emission reduction compared to the original diesel-electric configuration), whereas it is 185 kg CO_2 for the Denmark case. However, for the nuclear power based H_2 scenario, the minimum cost increases would be 120% and 135% for Denmark and the UK respectively. Adopting wind power generated H_2 can bring the voyage GWP close to 100 kg CO_2 for both countries. However, the minimum cost increase would be 181% and 207% for Denmark and the UK respectively.

4.5 Sensitivity studies

In this section, sensitivity studies are implemented to further investigate the impacts from parameters which are subject to change due to uncertainties. The parameters included in these sensitivity studies are PEMFC price, battery price, H_2 price. In the first three studies (Section 4.5.1, 4.5.2 and 4.5.3), one of these parameters is varied within a range while the others are kept constant as defined in Section 4.4.2, i.e. the Denmark case with wind power generated H_2. In Section 4.5.4, the battery and

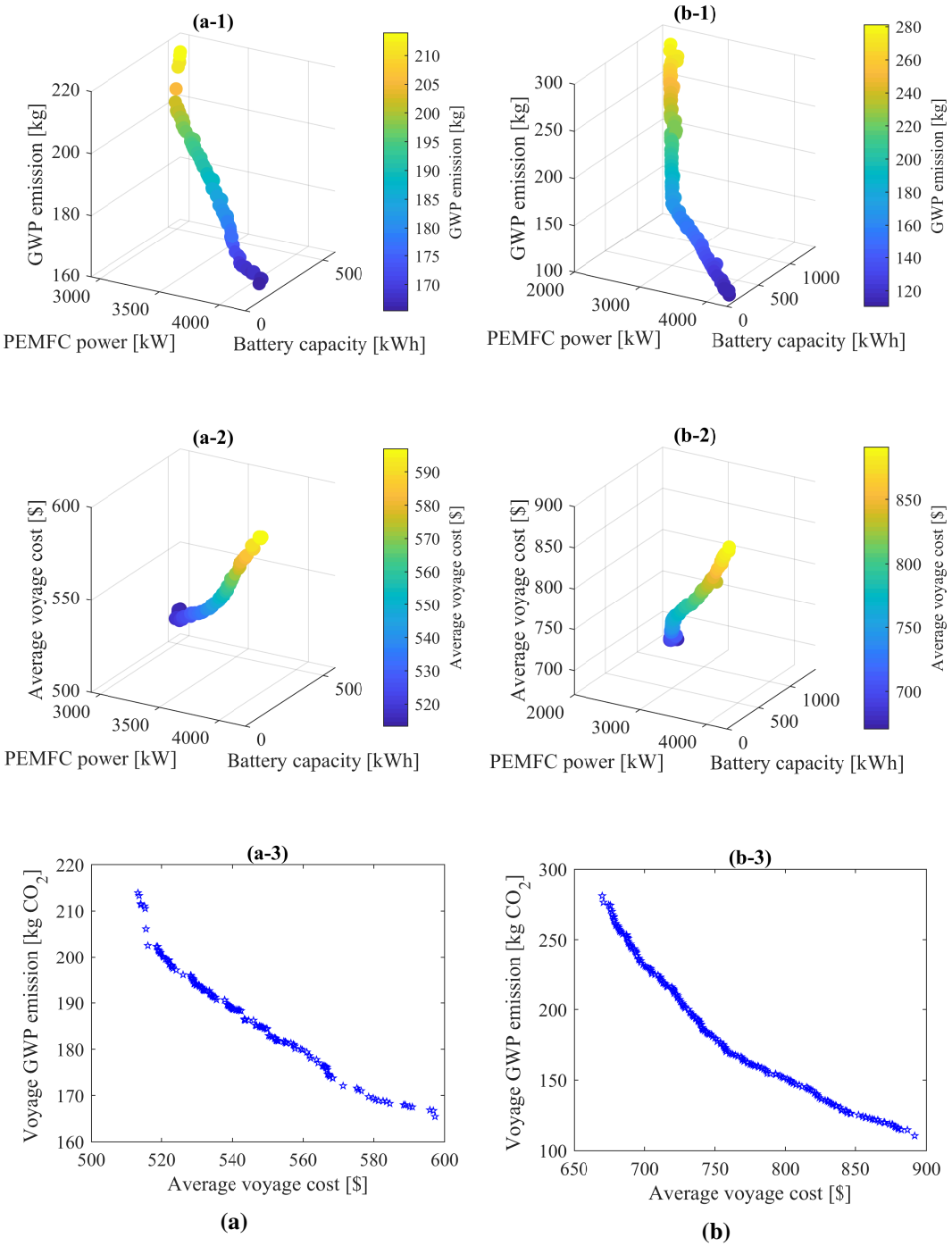


Figure 4.16: Case study 2 (UK electricity) fuel cell and battery sizing combinations with regards to the cost and emission objectives: (a) scenario of H_2 via nuclear-powered high temperature electrolysis and (b) scenario of H_2 via wind-powered electrolysis.

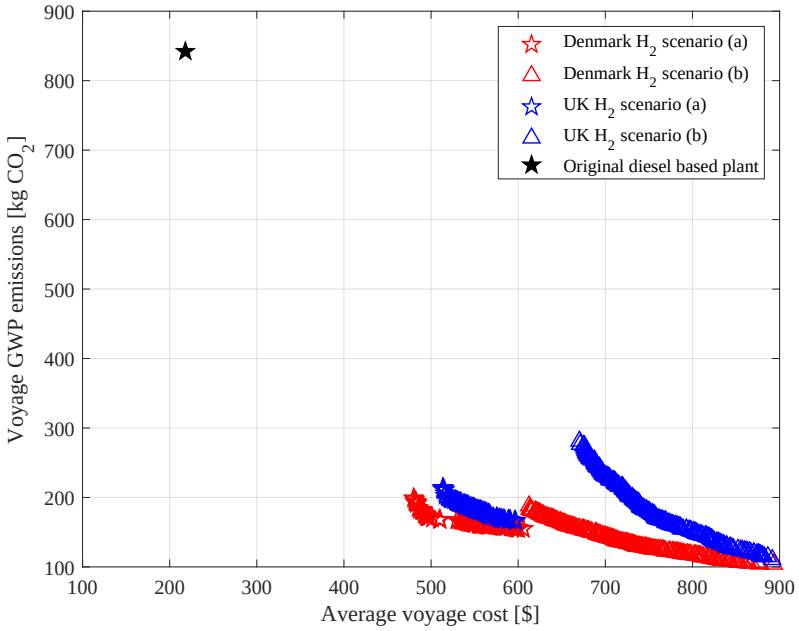


Figure 4.17: Combined Pareto fronts of the case studies.

PEMFC prices are varied concurrently.

4.5.1 PEMFC price

Although the PEMFC price adopted the previous sections is defined as 830 \$kW^{-1}, it is expected to decrease over time and might reach the level around 50 \$kW^{-1} with increasing annual production rate (US Department of Energy, 2018). Therefore, the PEMFC price is varied from 50 to 950 \$kW^{-1} to investigate the PEMFC price impact on the vessel's economic and GWP emission performance. Figure 4.18 presents the Pareto fronts of these PEMFC price scenarios. As grid electricity and H_2 GWP are fixed at 0.17 kg CO_2 kWh^{-1} and 1.50 kg CO_2 kg^{-1} H_2 respectively, the voyage GWP could be achieved for the seven considered PEMFC price scenarios are all between 100 and 200 kg CO_2. If the PEMFC price of 50 \$kW^{-1} is achievable, the minimum average voyage cost would be around \$430, which is approximately 2 times of the original diesel based plant. Note that PEMFC price of 50 \$kW^{-1} is lower than of the diesel based plant (500 \$kW^{-1}). However, its lifetime is much shorter compared to that of diesel engines (assumed as 20 years in this study). The H_2 cost (price of 8.24 \$kg^{-1}) is another main contributor to this

high voyage cost. If available PEMFC price is 950 \$kW^{-1}, the minimum voyage cost would be \$640 per voyage with battery as the dominating power source for this design point.

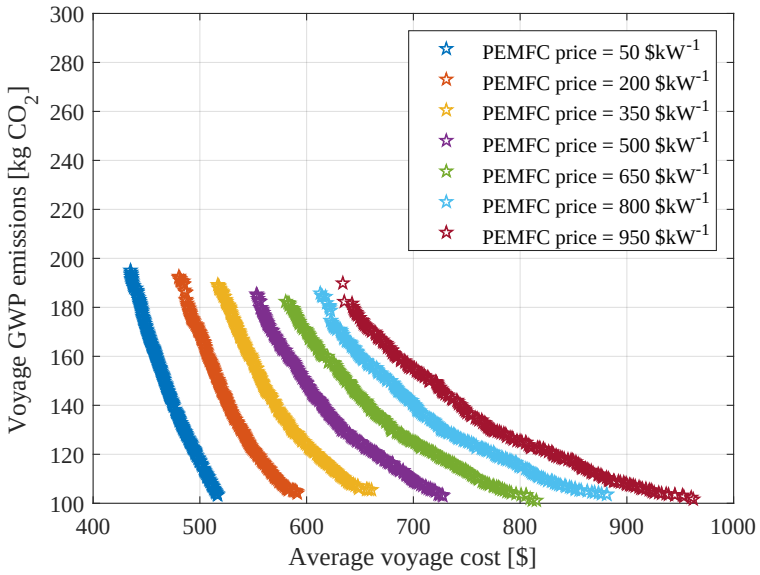


Figure 4.18: PEMFC price sensitivity study.

4.5.2 Battery price

As predicted by Larcher and Tarascon (2015), the system level production cost of Lithium batteries is expected to decrease to a level below 200 \$kWh^{-1} in the coming decade (see Figure 2.14). However, these price predictions may change due to many factors, such as market demand. The battery price sensitivity is implemented to investigate the impacts of battery price on the design of hybrid PEMFC/battery propulsion system. As in Figure 4.19, Pareto fronts of seven battery price scenarios are presented. Note that the PEMFC price is fixed as 830 \$kW^{-1}. The battery price of 50 \$kWh^{-1} would yield a minimum average voyage cost of \$480. However, such a voyage cost would still be more than two times of that of the original diesel-based system, which is mainly due to the high prices of the PEMFC and H_2.

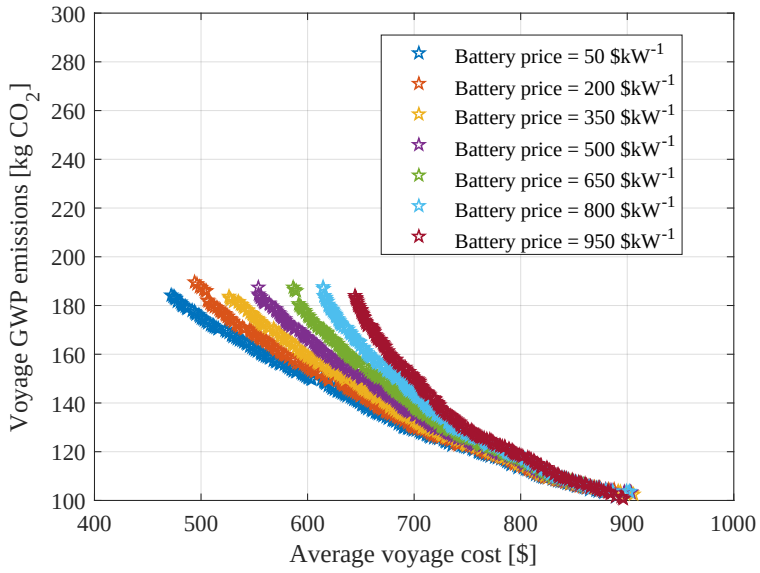


Figure 4.19: Battery price sensitivity study.

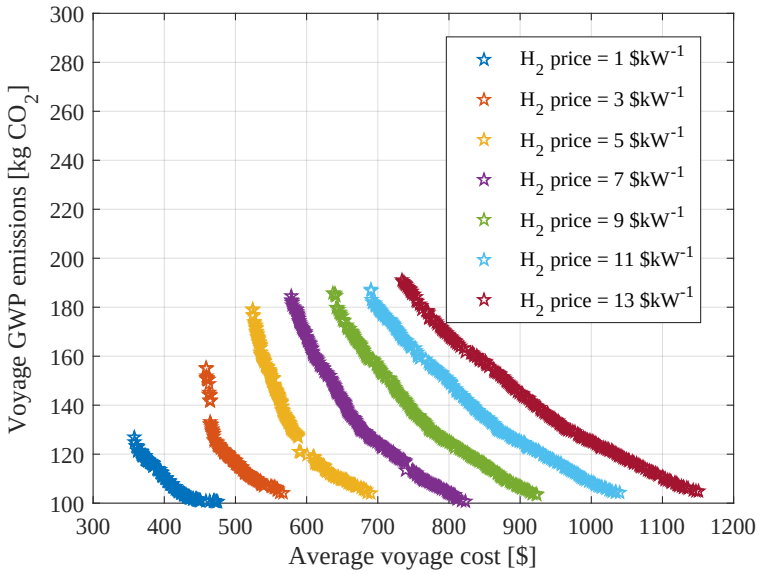


Figure 4.20: H_2 price sensitivity study.

4.5.3 H_2 price

Figure 4.20 shows the sensitivity study on H_2 price. As the PEMFC and the battery are still costly (830 \$kW^{-1} and 800 \$kWh^{-1} respectively) with limited lifetime, an extreme low H_2 price of 1 \$kW^{-1} would result in a minimum voyage cost of \$350 (corresponds 130 kg CO_2 GWP per voyage). A high H_2 price of 13 \$kW^{-1} would

lead to a minimum voyage cost of $740. Note that the maximum voyage GWPs for the low H_2 price scenarios are lower than those of the high H_2 price scenarios, which is due to PEMFC is the more dominating with better fuel efficiency on the upper left part of the cheaper Pareto fronts; while batteries are more dominating in the scenarios of high H_2 prices.

4.5.4 PEMFC and battery prices

In Sections 4.5.1 and 4.5.2, the influences of PEMFC and battery prices were investigated independently. Driven by road vehicle applications, both the PEMFC and battery prices are expected to decrease over time (US Department of Energy, 2018; Larcher and Tarascon, 2015). In this section, the fuel cell and battery prices are varied simultaneously. Figure 4.21 presents the Pareto fronts of the considered fuel cell and battery price scenarios. Note that a low fuel cell price corresponds to a low battery price. Such a setting is designed to simulate both the fuel cell and battery prices decrease over time. The PEMFC price of 50 $\$kW^{-1}$ and battery price of 50 $\$kWh^{-1}$ would yield a minimum voyage cost of $289, which is only 33% higher than that of the diesel-based plant but would reduce the voyage GWP emission by

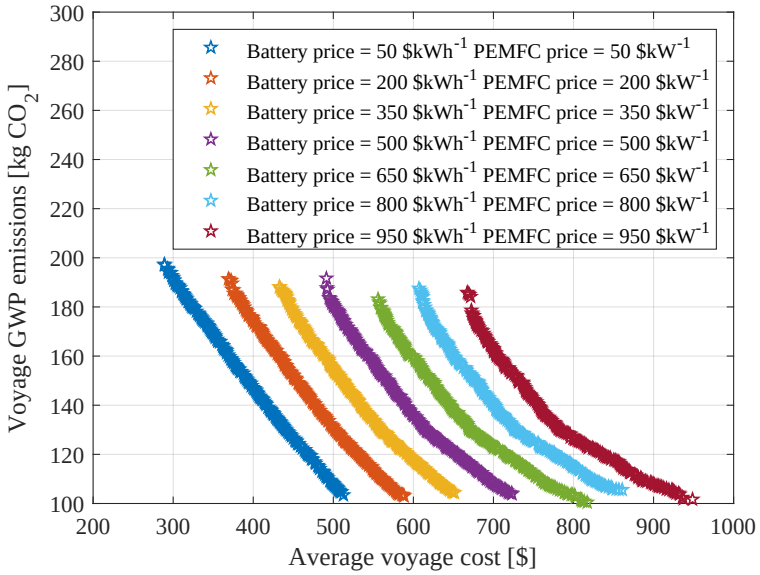


Figure 4.21: PEMFC and battery prices sensitivity study.

77%. Note the H_2 price of 8.24 \$kW^{-1} is relatively high, a reduced H_2 price would make the plug-in hybrid PEMFC/battery system even more competitive.

4.6 Summary

In this chapter, a quasi-steady-state plug-in hybrid PEMFC/battery system model has been developed for power source sizing and energy management strategy development purposes. A multi-objective design methodology has been applied to optimise the PEMFC/battery hybrid system design with the objectives of minimising average voyage costs and voyage Global Warming Potential emissions in two case studies.

Power source degradations and optimal energy management between multiple power sources are considered in the optimisation. It is evident that such a PEMFC and battery hybrid system can significantly reduce the GWP emissions, but is still constrained by high costs, mainly from H_2 cost of production and fuel cell degradation. The multi-objective power source sizing methodology can be used to guide alternative propulsion system design considering the influences of energy life-cycle properties and power source degradation characteristics. For the two cases considered with two H_2 scenarios, a minimum 65% GWP emission reduction can be achieved by utilising the plug-in hybrid fuel cell and battery propulsion system configuration. The developed model, design methodology and case studies have been published in (Wu and Bucknall, 2020).

The sensitivity studies further explore the impacts of power source and energy prices on the power source sizing and performances. Both the PEMFC, battery and H_2 prices are expected to decrease to be competitive with conventional diesel-based power solutions. Nevertheless, the GWP emission reduction potentials offered by the plug-in hybrid PEMFC/battery system are significant.

Chapter 5

Reinforcement learning based energy management strategies

5.1 Overview

In Chapter 4, the plug-in hybrid PEMFC/battery system was been optimised by the proposed methodology considering the trade-off between average voyage cost and voyage GWP emissions for the case ship. However, this sizing and operational configuration of the hybrid system is constrained by high costs, due to power source degradation and energy costs. It would therefore be beneficial to improve the operational cost-effectiveness of such hybrid systems.

This chapter aims to formulate the practical optimal energy management problem for shipboard plug-in hybrid fuel cell and battery propulsion systems as a Markov Decision Process (MDP) which can be solved using reinforcement learning (RL). Section 5.2 provides an introduction to RL and its mathematical framework– MDP. Section 5.3 mathematically formulates the energy management problem with MDP and introduces the two RL agents that will be applied. The agent training processes are detailed in Section 5.4. The RL EMS performance is assessed in Section 5.5. Section 5.6 summaries this chapter.

5.2 Reinforcement learning

5.2.1 An introduction to reinforcement learning

Reinforcement learning (RL) is an approach to solving goal-directed problems by trial-and-error interactions with the environment, which is similar to how an untutored and unguided human learns to perform a new task (Sutton and Barto, 2018). In contrast to many other approaches such as supervised learning, RL explicitly considers the whole problem of a goal-directed agent interacting with an uncertain environment. By interaction with the environment, the sensation of environment states and observation of reward signals returned from the environment, RL aims to achieve defined goals by finding an optimal action map which maximises cumulative rewards (Sutton and Barto, 1998).

To explain how RL works, the following terminologies need to be introduced (Sutton and Barto, 2018):

- *Agent:* An agent in RL is the decision-making component which determines actions with given environment states. For example, the *AlphaGo* computer programme is an RL agent specialised in playing Go games (Silver et al., 2016). *AlphaGo* acts like a master Go player, making decisions by observing current board states.

- *Environment:* The other component the agent interacts with is called the environment, which comprises everything outside the agent. In the *AlphaGo* example, the environment includes the other player and the game states represented by the board with stones of the two players.

- *State:* States represent the environment status. For example, the State of Charge (SOC) represents the state of battery capacity. Another example of the state could be current ship speed or total required power of a sailing ship.

- *Action:* The agent interacts with its environment by taking actions. For a con-

trol problem of fuel cells, the action could be increasing or decreasing power output. For an unmanned surface vehicle navigation problem, the action could be turning to starboard or to port by a certain degree angle.

- *Reward:* The reward defines the goal in an RL problem and is a measurement of how good or bad an event could be to the agent. Such a signal is returned from the environment to the agent. In biological systems, rewards could be analogous to the experiences of pleasure or pain. In an RL problem, the value of the immediate reward is an indication of how the agent ought to behave. The immediate reward is the primary signal of altering a policy.

- *Policy:* A policy defines the agent's behaviour at given environment states. In other words, a policy is a mapping from observed states of the environment to actions to be taken by the agent. The training of an RL agent aims to find an optimal policy which maximises the cumulative rewards. For example, in the optimal control problem of a hybrid fuel cell/battery propulsion system, the policy is effectively the energy management strategy.

- *Episode:* For a problem with a finite time frame, a training episode starts from time step zero and ends when a terminal state or the horizon is reached. The agent interacts with its environment in each time step of an episode. The training episodes are repeated until a rational policy has been found by the agent.

5.2.2 Markov Decision Process

A Markov Decision Process (MDP) is a stochastic control process in discrete time space, which provides a mathematical framework to model sequential-decision making problems (Puterman, 2014). MDPs are idealised mathematical frameworks of RL problems (Sutton and Barto, 2018). Such a process can be represented by a tuple (S,A,P,R), where S is a finite set of states s, A is a finite set of actions a, P is a set of state transition probabilities, i.e. $p_{ss',a} = \mathbb{P}[s_{t+1} = s' | s_t = s, a_t = a]$, and R

is a reward function $r_{ss',a} = \mathbb{E}[r_{t+1}|s_t = s, a_t = a]$. The subscript t denotes state s, action a or reward r at time step t.

The action-value function, which is also called the Q function, for an episodic task with finite horizon of T, is the expected return of taking action a in state s following a policy $\pi(s|a)$:

$$Q(s,a) = \mathbb{E}\left[\sum_{k=0}^{T} \gamma^k r_{t+k} | s_t = s, a_t = a, \pi\right] \tag{5.1}$$

Solving an MDP is to find an optimal policy π^*:

$$\pi^*(s) = \arg\max_a \mathbb{E}\left[\sum_{k=0}^{T} \gamma^k r_{t+k} | s_t = s, a_t = a\right] \tag{5.2}$$

which leads to the optimal action-value function (Sutton and Barto, 2018):

$$Q^*(s,a) = \max_\pi \mathbb{E}\left[\sum_{k=0}^{T} \gamma^k r_{t+k} | s_t = s, a_t = a\right] \tag{5.3}$$

where $\gamma \in [0,1]$ is the discount rate. As in Figure 5.1, at time step t, in current state s_t, the agent takes action a_t under the policy $\pi(s|a)$ and observes the resulting next state s_{t+1} and immediate reward r_{t+1} returned from the environment.

5.3 Reinforcement learning based energy management strategy

This section formulates the optimal energy management problem of the plug-in hybrid PEMFC/battery system with MDP and introduces two RL agents which will be applied to solve the formulated MDP. Figure 5.1 shows the detailed MDP agent-environment interaction framework for the energy management problem. The environment of the MDP framework includes the hybrid PEMFC/battery system model (developed in Chapter 4) and historical voyage data (Eriksen et al., 2018).

The objective of developing the EMS is to find an optimal policy π^* to achieve

minimum voyage costs. Note that high operational costs are a major limitation of the hybrid system (see Chapter 4). The optimal policy π^* is a mapping of system states to actions, i.e. actions to be taken in certain situations. The voyage cost comprises two parts, i.e. the degradation costs from power sources (fuel cells and batteries) and the costs of consumed fuel and energy (H_2 and electricity). Table 5.1 summarises the RL terminologies, which will be used for the optimal energy management problem in the subsequent sections.

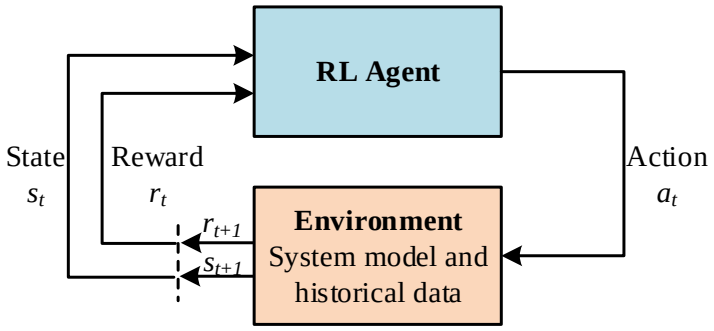


Figure 5.1: MDP agent-environment interaction framework.

Table 5.1: Summary of RL terminologies in the optimal energy management problem.

Terminology	Description
Agent	Reinforcement learning algorithm
Environment	Hybrid system model and historical power profiles
States	System states, including current PEMFC power level, battery SOC, power demand and shore power availability
Action	Fuel cell power change
Reward	A function of constraints and costs incurred in one time step
Policy	Energy management strategy of EMS

5.3.1 States

In the optimal energy management problem, the states represent the current system status. In the proposed system, such states are characterised by shore power availability, spA, system power demand, P_{dem}, fuel cell power level $x \in [0,1]$, and battery $SOC \in [0,1]$. spA is binary, i.e. $spA = 0$ when the ship is sailing and $spA = 1$ when the ship is in port. It is assumed that the transition from transit to port, as regards

battery charging, is instantaneous, i.e. the shore power is applied immediately when the ship is in port. Although x, SOC and P_{dem} are continuous physical parameters they are however divided into discrete grids, such that the gridded state space can be formulated by looping through all possible state combinations. As each of the state parameters has a finite dimension, the total number of states is the product of the four state dimensions. Each possible state is assigned a unique state index sequentially (i.e. from 1 to the total number of states). At time step t, the exact state of the system:

$$s_{actual}(t) = [spA(t), P_{dem}(t), x(t), SOC(t)]^T \tag{5.4}$$

is converted into state index $s(t)$, which is an integer. Note that the environment knows the actual states $s_{actual}(t)$ and $s_{actual}(t+1)$ which results from taking action $a(t)$, but only communicates with the agent using state indices. Such a communication format is designed intentionally so that the agent can record the learning process into tables.

5.3.2 Action space

In reinforcement learning, the agent interacts with the environment by taking actions in relation to the system states. The action taken by the agent is the control of fuel cell power change within each time step in this study. The action space is defined as a tuple of possible fuel cell power level changes:

$$A = [a_1, a_2, ..., a_m, ..., a_{n-1}, a_n]^T \tag{5.5}$$

where $a_1 < 0$ is the maximum decrease and $a_n > 0$ is the maximum increase of fuel cell power output in a time step. $a_m = 0$ indicates there is no change and the fuel cell output power remains constant; all other values of a represent changes of power within the range of (a_1, a_n). The environment overrides an action when the resulting fuel cell power output would be negative or greater than the rated power. When action $a_t \in A$ is chosen from the action space at time step t, the fuel cell power

level at $t+1$ will be:

$$x_{t+1} = \begin{cases} 0, & x_t + a_t < 0 \\ 1, & x_t + a_t > 1 \\ x_t + a_t, & \text{else} \end{cases} \tag{5.6}$$

5.3.3 Reward

The environment returns reward signal r_{t+1} to the agent when action a_t is taken by the agent. The value of r_{t+1} represents how cost-effective a_t is at state s_t:

$$r_{t+1} = \begin{cases} -1, & \text{if } s_{t+1} \text{ is infeasible} \\ -1, & \text{if } p_{fc} + a_t \notin [0,1] \\ \tanh\left(\frac{1}{cost_{t+1}}\right), & \text{else} \end{cases} \tag{5.7}$$

where the negative reward of -1 means the agent is penalised if the next state is not feasible or fuel cell power override will occur; the $\tanh\left(\frac{1}{cost_{t+1}}\right)$ function normalises the cost $cost_{t+1}$ to a reward signal in the range of $[0,1]$ elsewhere. Note that the next state is not feasible if the battery is over charged/discharged or C-rate exceeds the system limit or fuel cell power is not reduced to zero when the ship is in port (fuel cells are not switched off to avoid unnecessary start/stop cycling degradations). $cost_{t+1}$ is the cost incurred in one time step Δt due to action a_t if the next state is feasible:

$$cost_{t+1} = \psi_{fc}(x_t + \frac{a_t}{2})P_{fc}\Delta t\,\sigma_{H_2} + \delta_{fc}(x_t + \frac{a_t}{2})P_{fc}\sigma_{fc} + P_{sh}\Delta t\,\sigma_e + \delta_{bat}B\sigma_{bat} \tag{5.8}$$

i.e. the sum of H_2 cost, fuel cell degradation cost, battery average degradation cost and shore power cost (only when the ship is in port), where σ denotes price. The sub-scripts H_2, fc, e and bat denote H_2, fuel cell, electricity and battery prices respectively. Note the cost $cost_{t+1}$ is unpenalised since the negative reward -1 includes a penalty. To better understand the impact of infeasible actions, a penalised cost is also introduced in the following case study. The penalised cost is $cost_{t+1} +$

1 whenever the next state is not feasible, or agent action is overridden or early termination occurs.

5.3.4 Environment

The environment of the reinforcement learning comprises two parts, i.e. the hybrid propulsion system model (see Chapter 4) and historical power profiles collected using continuous monitoring of required power demand. Algorithm B.1 in Appendix B depicts how the environment of the optimal energy management problem is formulated. Using the historical voyage power profiles, in each learning episode, the environment randomly samples one power profile from the historical data with which the agent interacts. Note that the environment would carry out an early termination of an episode if the agent fully discharges the battery ($SOC < 0$) or over-charges the battery ($SOC > 1$). Normal termination occurs when the final targeted time step has been reached. An episode is successful if the agent manages to achieve all the required time steps and recharge the battery to a SOC of SOC_H to be fully prepared for next voyage; otherwise, the episode terminates and is recorded as having failed.

5.3.4.1 System model

It is assumed that the original diesel-electric system is replaced by a plug-in hybrid PEMFC and battery system, as described in Figure 5.2 (also see Chapter 4). The ship specifications are presented in Table 3.3. The original system featured an integrated full electric propulsion configuration with a total installed diesel engine power of 4370 kW. The ship operates between two fixed ports with 8 round trips (16 voyages) per day—each voyage between the two ports takes approximately 1 h (Eriksen et al., 2018). It is assumed the ship's batteries can be recharged at both ports, and the shipboard H_2 storage needs to be replenished once per day outside of operational hours.

The intended fuel cell power and battery capacity for the alternative plug-in hybrid PEMFC and battery propulsion system are 2940 kW and 581 kWh respec-

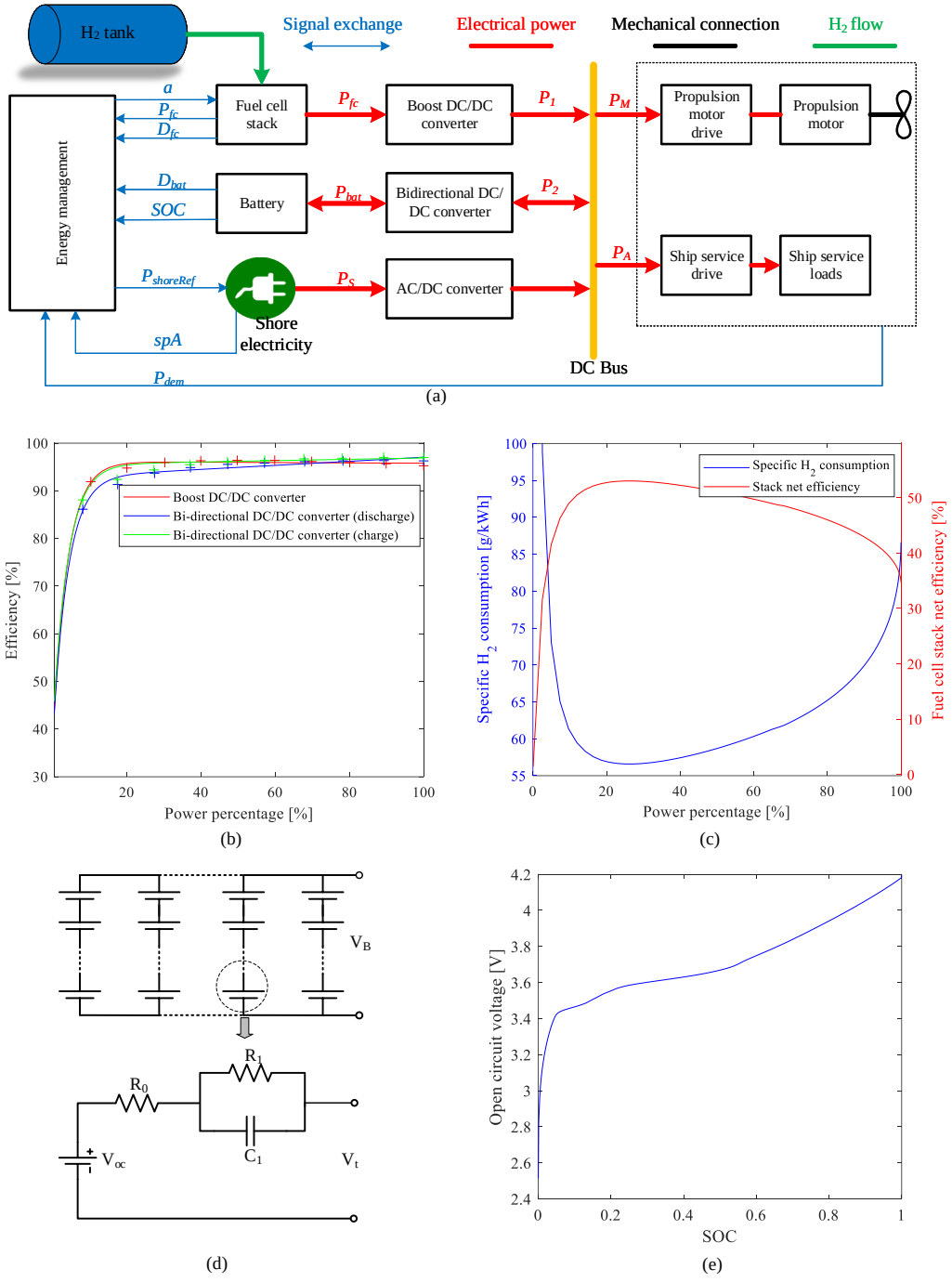


Figure 5.2: Plug-in hybrid PEMFC and battery propulsion system model. (a) top-level system model, (b) power converter efficiency, (c) PEMFC specific H_2 consumption and system efficiency, (d) battery stack model and individual cell and (e) battery individual cell voltage in a function of battery SOC. Note that this model has been detailed in Chapter 4.

tively. The system is capable of delivering a regular service power of 4683 kW and peak power of 6720 kW, corresponding to battery C-rates of 3 and 6 respectively. Note that the system sizing has been optimised in Chapter 4 for the scenario when the ship was operational under the energy tariffs applicable to Denmark using wind power to supply hydrogen production. The adopted H_2 Global Warming Potential (GWP), electricity GWP, H_2 price and electricity price are set at 1.5 kg $CO_2 kg^{-1}$, 0.166 kg $CO_2 kWh^{-1}$, 8.240 $\$kg^{-1}$, and 0.089 $\$kWh^{-1}$ respectively. The battery's limits of SOC are set to upper and lower limits are limited to upper and lower values of 0.90 and 0.25 respectively, and the maximum C-rate is 6 for discharge. Note that the SOC limits are soft constraints, meaning they can be exceeded if deemed necessary. The battery needs to be charged to a SOC of 0.9 prior to departure. A starting SOC of 0.90 affords the system the flexibility to excessive power from the fuel cells if and when required. SOC below 0.25 should be avoided to provide minimum charge conservation, as well as extend battery life (Omar et al., 2014).

5.3.4.2 Historical data

The case ship and its route have been shown in Figure 3.2. The historical power profiles applied to the agent training were acquired from (Eriksen et al., 2018) (1081 voyages in total, from 1 July 2018 to 31 August 2018). Another dataset (392 voyages in total) collected over a different period (from 1 September 2018 to 30 September 2018) will be used for EMS validation. The datasets were first segregated into voyages determined by the ship's speed and location. The original time step of the power profiles is 15 s and remains unchanged. Figure 5.3 shows 8 randomly selected sample power profiles in the training dataset. The original power values were smoothed with a Gaussian-weighted moving average filter to reduce measurement noise. The moving average window of the Gaussian filter is 4, and the standard deviation is calculated from $1/5$ of the total window width. Although the power profiles follow a specific pattern in general, each of them varies from the others. Such an observation suggests that the MDP environment is stochastic and uncertain (see Figure A.1 in Appendix A).

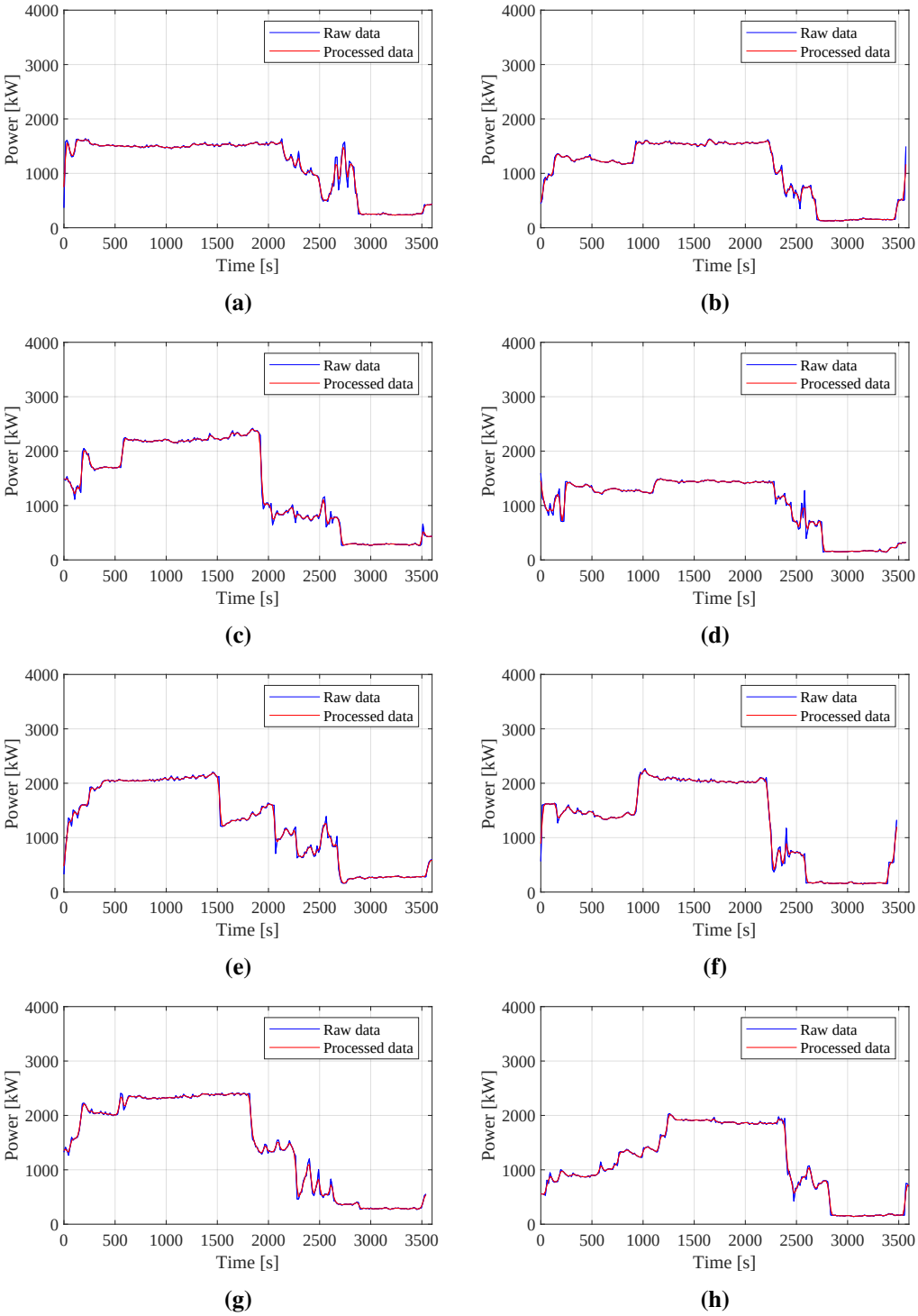


Figure 5.3: Sample power profiles from the training dataset.

5.3.5 Q-learning agent

The Q-learning (Algorithm B.2 in Appendix B), proposed by Watkins (1989), is a model-free approach for solving MDPs, i.e. transition probabilities P are not considered directly during agent training. It is also an off-policy RL method, i.e. the action-values are updated using the next state and the greedy action. When updating the action-value function, the agent acts greedily by choosing the action that maximises the next action-value function:

$$Q(s,a) \leftarrow Q(s,a) + \alpha \left[r + \gamma Q(s', \arg\max_{a} Q(s',a)) - Q(s,a) \right] \tag{5.9}$$

where s' is next state. However, the maximisation operations involved in the construction of policy and the ε-greedy action selection processes can lead to poor learning performance with maximisation bias in stochastic environments (van Hasselt, 2010; van Hasselt et al., 2015).

5.3.6 Double Q-learning agent

This study takes advantage of Double Q-learning (a variant of Q-learning) to learn optimal energy management strategies for the sequential power split problem between multiple power sources (Sutton and Barto, 2018). Algorithm B.3 in Appendix B shows the Double Q-learning agent (van Hasselt, 2010). The Double Q agent reduces the maximisation bias by using two action-value estimates, Q_1 and Q_2. For each update, with 0.5 probability, Q_2 is used to determine maximising action while Q_1 updates its value:

$$Q_1(s,a) \leftarrow Q_1(s,a) + \alpha \left[r + \gamma Q_2(s', \arg\max_{a} Q_1(s',a)) - Q_1(s,a) \right] \tag{5.10}$$

Otherwise Q_2 is updated with Q_1 and Q_2 being switched. Both the learning rate α and ε of the ε-greedy policy decrease linearly with the increase of learning episodes and stabilise at fixed values after rate decaying episode number N_d.

5.4 Agent training

The objective of the on-line EMS is to minimise the overall voyage cost in an environment that is not pre-known. The intent of such an on-line EMS is to manage the power flows within the hybrid power system effectively when applied for future, and as yet unknown, voyages. The learning process is an episodic task. In each episode, the environment randomly samples one of the historical voyage power profiles for the agent to interact with to learn a policy minimising the voyage cost. This process repeats until the average episode reward converges. Related historical power profiles need to be collected before the beginning of the agent training procedure. These profiles will be an inherent part of the RL environment. Note that each profile is unique although there will be similarities.

The RL agent training and policy application follow the procedure presented in Figure 5.4. Note that the RL training parameters, such as the learning rate α and the probability of exploration ε at a time step, require careful tuning to achieve a strategy with adequate performance:

- The agent should be able to complete the training voyages without early terminations.
- Achieve minimum voyage cost with the minimum of constraint violations.

Once the training has converged, the learned policy, i.e. the strategy of the EMS, needs to be validated using a different set of power profiles. In the application phase, a battery over-discharge protection function ensures the battery modules are not over-discharged. This protection mechanism is beyond the MDP agent-environment interaction framework (Figure 5.1) and is not enabled during agent training (see Figure 5.4), such that the agent can learn from penalties during training without external interventions. Actions leading to penalties would be avoided due to their lower Q values in corresponding states.

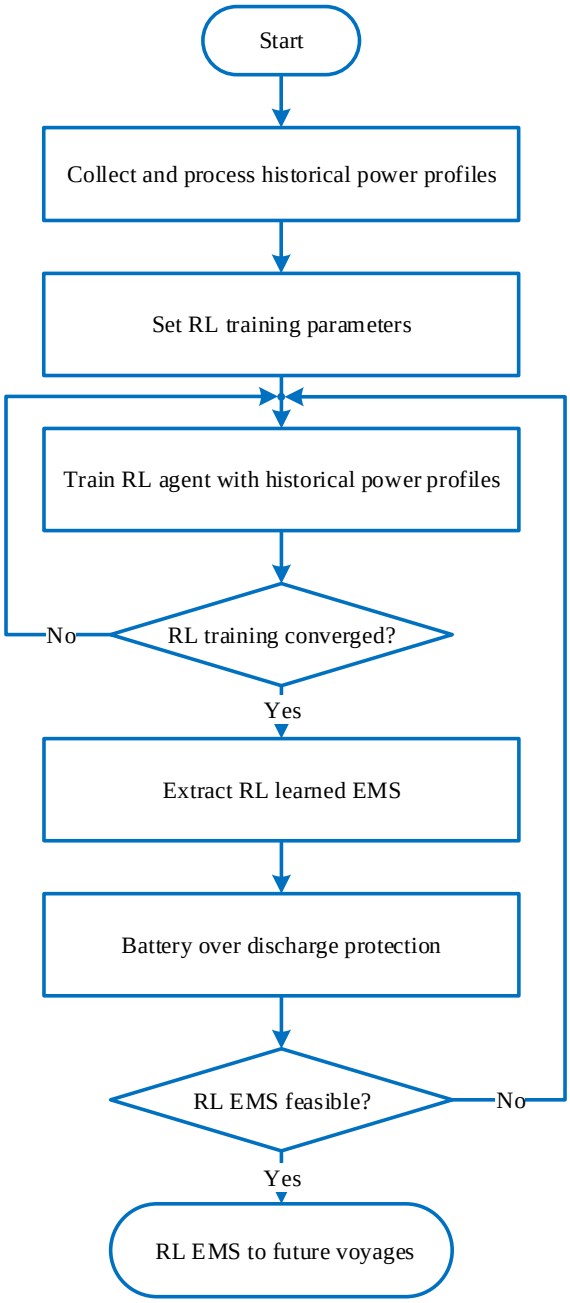


Figure 5.4: Reinforcement learning agent training and policy application procedure.

5.4.1 Training settings

Table 5.2 provides the detail of the state and action space grids for both the Deterministic Dynamic Programming (DDP) strategy (resolution 1) and Double Q strategy. The DDP implementation is based upon the work of Sundström et al. (2010). The results obtained using DDP are used to evaluate the quality of the strategy generated by the Double Q agent. Therefore, the grids are defined identically in the two algorithms to initially allow a fair comparison between on-line and off-line strategies. Note that developing a strategy by DPP requires complete knowledge of the profile, which is not possible for actual applications. Therefore, a DDP strategy is only valid as an off-line benchmark to assess the performance of other on-line strategies.

To further investigate the potential for cost reduction, the DDP strategy SOC grid length was further refined to 0.0125. However, such a refined SOC resolution was not implemented in the Double Q strategy due to 'the curse of dimensionality' (Sutton and Barto, 2018), which would make the problem impossible to solve with the available computational resources.

Table 5.2: State and action space grids.

Parameter	Grid resolution	Range	Unit
Power demand	50	0–4400	kW
SOC	5	0–100	%
Fuel cell power level	0.02	0–1	pu
Shore power availability	–	0 or 1	
Fuel cell power change fraction	0.02	$[-0.04, -0.02, 0, 0.02, 0.04]$	pu

5.4.2 Training

Table 5.3 shows the parameters used to train the Double Q agent. The parameter ε represents the probability of exploration at a time step. The learning rate α determines to what degree the temporal difference is acquired: $\alpha = 1$ suggest that only the most recent information is learned, $\alpha = 0$ nothing new has been learned. Both α and ε decrease linearly from their initial values whilst the training episode number

is less than N_s. Such settings reflect the need for the agent to explore less frequently and learn more cautiously when enough experience has been gained, while a more aggressive and bold learning style is preferred at the outset to quickly gain experience. As the energy managemnet problem is formulated with an average episode length of 240, and the costs incurred in all steps are of equal importance, the discount rate γ is set at 1 (i.e. un-discounted). It is worth mentioning that careful tuning of these parameters is necessary to balance the conflict between *exploration and exploitation* (Sutton and Barto, 2018).

Table 5.3: Reinforcement learning hyper parameters.

Parameter	Description	Value
α_{init}	Initial learning rate	1.0
$\Delta\alpha$	Learning rate decaying rate	3.3×10^{-6}
ε_{init}	Initial ε	1.0
$\Delta\varepsilon$	ε decaying rate	3.3×10^{-6}
γ	Discount rate	1.0
N_s	Episode α and ε stabilises	3.0×10^5

Figure 5.5 shows the learning process of the RL agent. It is interesting that the mean episode reward decreases to -12 after 0.6×10^5 episodes. This decrease suggests that initially the policy being learned was divergent before the agent was able to learn towards a convergent policy. The training was terminated after 5×10^5 episodes (4.8 h on an Intel i7-4790 processor using single thread in Matlab 2019a).

The mean episode reward stabilised at a value of 88 after about 3×10^5 episodes of training (Figure 5.5a), while the maximum episode reward stabilised at around 120. Such stabilisation suggests that the algorithm has converged. The average success rates (see Algorithm B.1) were close to 100% after convergence. Note that this rate is not exactly 100% (Figure 5.5b) which is mainly due to a small exploration probability (1.0×10^{-3}) that still exists and a minor fraction of training voyages with high power demand that vary significantly from other voyages. In Figure 5.5c, both the actual episode cost and penalised episode cost increases rapidly in the first 1×10^5 episodes. The reason for that is early termination frequently occurs and at the initial stage of the training. In other words, the agent could not complete

the majority of the training voyages in the initial stages of training (also see the mean episode steps in Figure 5.5d) due to the policy's tendency to drain the battery aggressively from the beginning.

As the training progressed, the agent managed to complete most of the training voyages from 2×10^5 episodes onwards. Also, the average voyage cost starts to decrease after 2×10^5 episodes. The actual cost and penalised cost (including the penalties caused by exceeding the constraints) overlap with each other, suggesting non-feasible actions have been reduced to a minimum. In summary, the agent appears to first complete voyages, then learn to minimise voyages costs (maximum reward) due to the reward setup. In contrast, as shown in Figure 5.6, with the same hyperparameter settings, the Q agent failed to converge to a policy with reasonable performance, owing to the presence of maximisation biases throughout the learning process (see Eq. 5.9). These biases cause over-estimation of the action-value function, which leads to unstable training in Q-learning. The double Q-learning reduces such biases by using two Q-functions.

Note that the environment is highly stochastic, with a small fraction of training voyages with high power demand that vary significantly from other voyages. The learned policy fails to fulfil the final battery SOC constraint of $SOC = SOC_H$ in less than 0.5% of the 1081 total training voyages. This failure suggests that an override function would be necessary to make the learned policy fully compliant with the final battery state constraint. A battery over-discharge protection, as in Figure 5.4, was proved to be effective. This protection was realised by forcing the fuel cell to increase power by 5% of rated power in one time step when the battery SOC drops below the lower limit (0.25) (Rouholamini and Mohammadian, 2016).

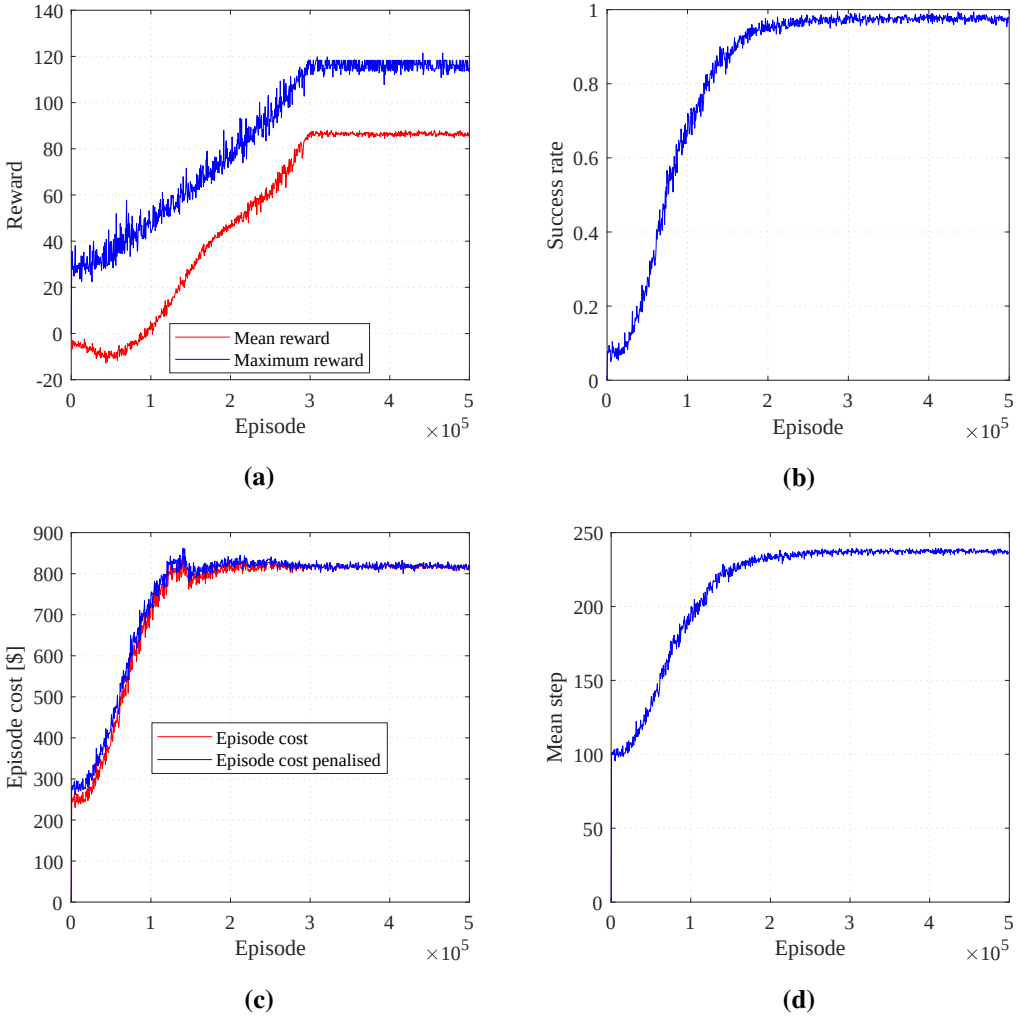


Figure 5.5: Double Q agent training process. The values are calculated every 500 episodes. (a) average reward, (b) maximum reward, (c) average penalised and unspecialised costs and (d) average episode steps.

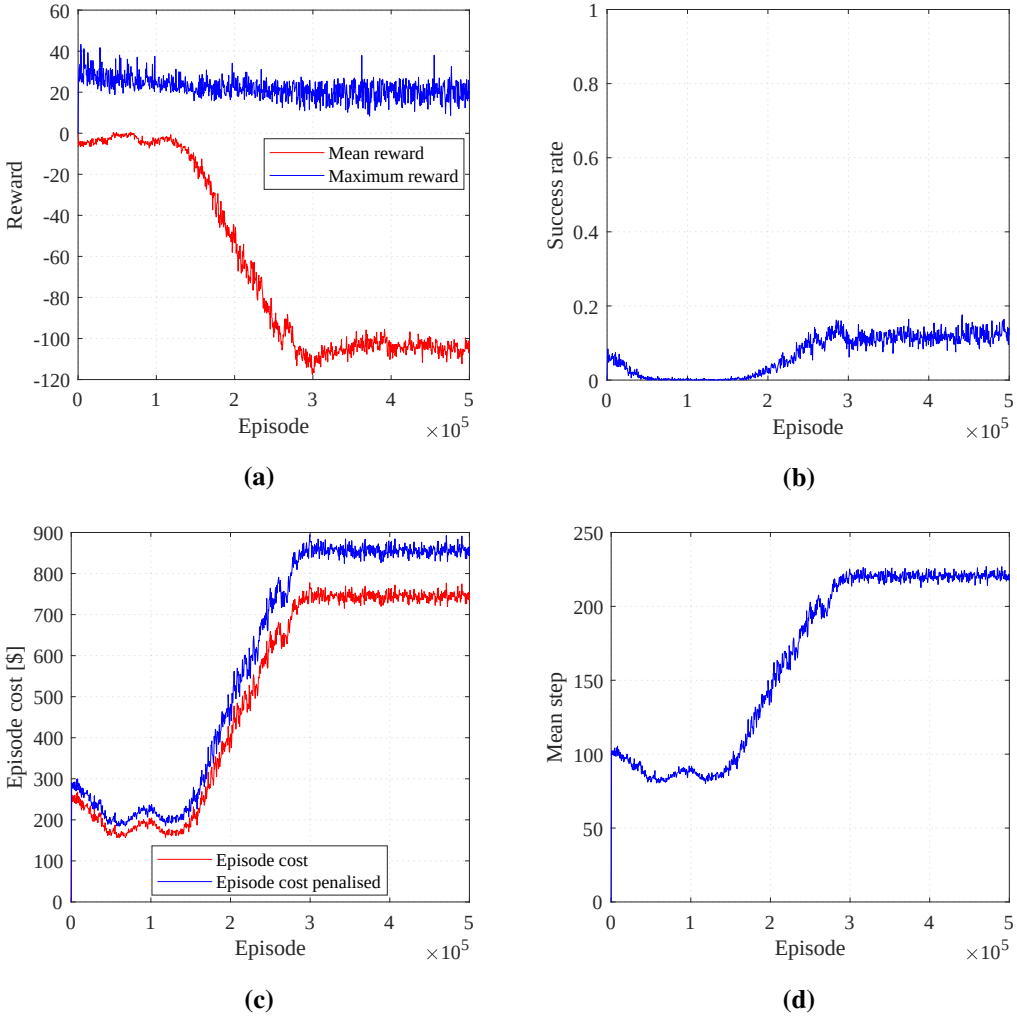


Figure 5.6: Q agent training process. The values are calculated every 500 episodes. (a) average reward, (b) maximum reward, (c) average penalised and unspecialised costs and (d) average episode steps. The Q agent failed to converge to a policy with reasonable cost performance and the constraints were violated frequently in late stage of training.

5.5 Results

5.5.1 Overview of results

Table 5.4 details the two different datasets which will be used in this section. Dataset A is used to train the agent to generate the strategy of the EMS. Once the training of the agent has converged, the strategy is verified by removing the random exploration ε adopted in the training phase. Subsequently, the EMS performance is validated using the dataset B, which have not been applied to the agent in the training phase. The strategy is a 4-dimensional action map over the four state parameters. With the system state observed, the optimal action of fuel cell power control can then be found from the action map.

Table 5.4: Datasets of load profiles and their purposes. Dataset A is used to train the agent to generate the strategy of the EMS. The EMS is then applied to load profiles in dataset B to validate the EMS performance in unseen voyages.

Dataset	Start date	End date	Voyage number	Purpose
A	01/07/2018	31/08/2018	1081	Training/verification
B	01/09/2018	30/09/2018	381	Validation

The learned policy was then applied to the training voyages and a set of validation voyages. As depicted in Table 5.5, for the training voyages, the Double Q strategy achieved 96.6% cost minimisation performance of the off-line strategy solved by DDP (knowing complete profiles before solving), both with the SOC grid resolution of 0.05. Note that state space resolution also limits the accuracy of DDP (Wang et al., 2015). A refined SOC grid resolution of 0.0125 yields an average voyage cost of $740.0 for the training dataset. The Double Q strategy achieves 89.0% cost minimisation performance of the refined DDP solution. For the validation voyages, similar performance was achieved. The DDP strategy results presented in the following strategy analysis sections are all solved with SOC resolution of 0.0125.

Figure 5.7 presents the voyage cost achieved by the Double Q strategy in comparison with that solved via DDP, for the training (Figure 5.7a) and validation (Fig-

Table 5.5: Double Q and DDP strategy average voyage costs comparison.

	DDP_1 [$]	DDP_2 [$]	RL [$]	DDP_1/RL [%]	DDP_2/RL [%]
SOC resolution	0.0125	0.0500	0.0500	-	-
Training voyages	740.0	803.1	831.8	89.0	96.6
Validation voyages	724.9	789.4	815.0	88.9	96.9

ure 5.7b) voyages. The Double Q strategy has achieved satisfactory cost performance (only 3.2% higher than DDP strategy) in validation voyages without prior knowledge of future power demand. Note that some voyages in the training dataset have much higher power demand, yielding a maximum Double Q strategy voyage cost close to $1600.0.

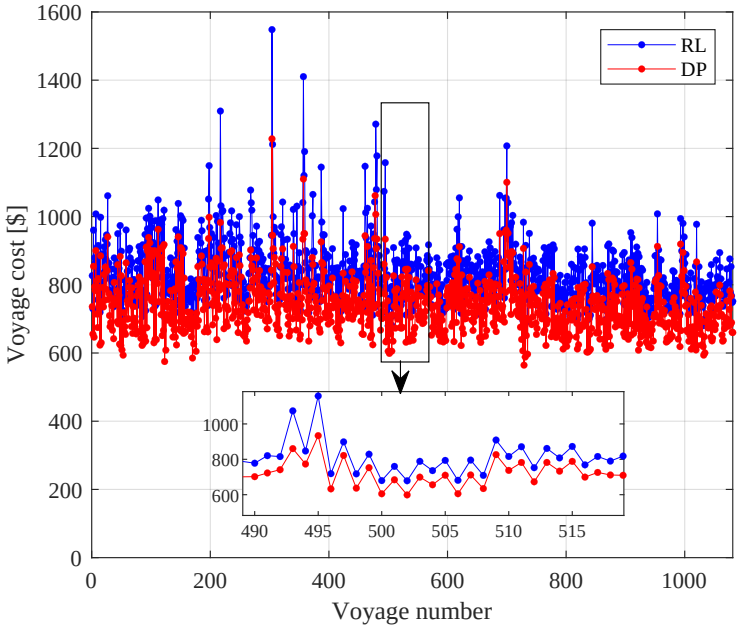
To verify the strategy performance learned by the Double Q learning agent, the Double Q strategy was applied directly (without any exploration) to the training voyages with over-discharge protection enabled. Such a process will be referred to as verification in the following content. Applying the strategy to a set of validation voyages will be referred to as EMS validation. Table 5.6 provides a summary of the sample voyages with low, moderate and high power demand, which will be discussed in the following analysis.

Table 5.6: Summary of sample voyages.

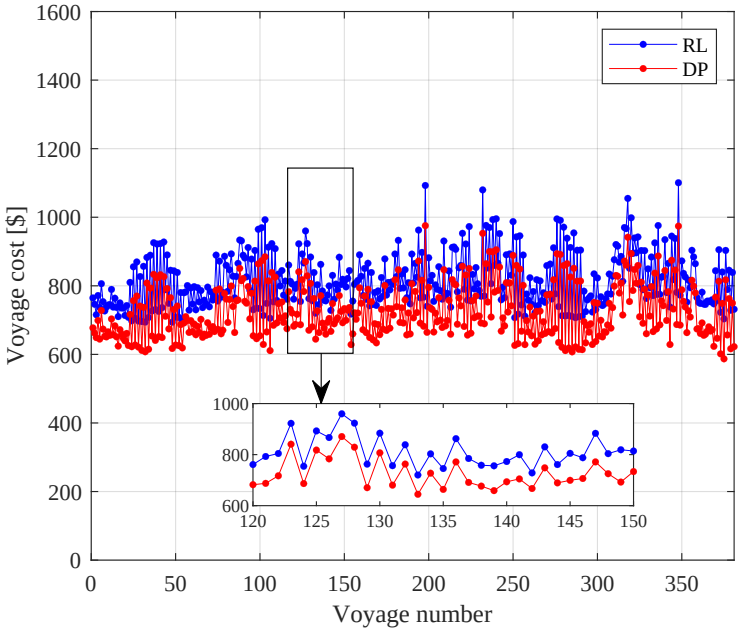
Category	Profile	Average power [kW]	Peak power [kW]	Voyage time [s]
Training	Training sample 1	904.2	1615.3	3585
	Training sample 2	1086.3	1836.8	3735
	Training sample 3	2040.3	3320.4	3165
Validation	Validation sample 1	1036.8	1487.0	3555
	Validation sample 2	1167.0	2060.0	3555
	Validation sample 3	1597.8	2752.7	3555

5.5.2 EMS verification

In this section, the Double Q-learning agent generated EMS is applied to three sample voyages in the training dataset to evaluate the EMS performance in different operation scenarios. The three sample voyages are with low, moderate and heavy



(a)



(b)

Figure 5.7: Voyage costs: (a) training voyages and (b) validation voyages. The DDP costs are obtained with a SOC resolution of 0.0125, while it is 0.05 for the Double Q strategy.

power demand, respectively. Details of the voyage cost and emission compositions are presented. Note that the objective of the EMS is to minimise voyage costs. The voyage emissions are calculated based on electricity usage and H_2 consumption figures obtained using the models presented in Chapter 4.

5.5.2.1 Training sample 1 with low power demand

Figure 5.8 shows the DDP and Double Q strategies for sample verification voyage 1. This voyage has comparatively low overall power demand in the training dataset. It starts with relatively high power demand (1600 kW). During cruising, the power demand stays around 1000 kW. Note that to solve for the DDP strategy requires complete knowledge of the power profiles in advance. The Double Q strategy only takes actions in each time step by observing current system states. The PEMFC power trajectory in the DDP strategy (Figure 5.8a) is relatively smoother than that of the Double Q strategy (Figure 5.8b). The Double Q strategy tends to adjust the PEMFC power more frequently within a narrow power band, which could be due to limited knowledge of future power demand. Such behaviour leads to higher PEMFC degradation (see Table 4.2) and H_2 costs (see Table 5.7). Also, the Double Q strategy rapidly discharges the battery to an SOC of 0.4 (at 950 s) after departure and then gradually recharges the battery. In contrast, the minimum battery SOC in the DDP strategy is 0.3 and occurs just before shore charging commences (2800 s).

Table 5.7 details the voyage cost and emission breakdowns of the verification sample voyage 1. The DDP strategy yields a voyage cost of \$585.2, which is 85.3% of the Double Q strategy voyage cost. The Double Q strategy leads to higher costs from PEMFC degradation and H_2 consumption. It is worth noting that the voyage GWP emission of the DDP strategy is 11.9% higher than that of the Double Q strategy which is due to the trade-off between voyage cost and GWP emission.

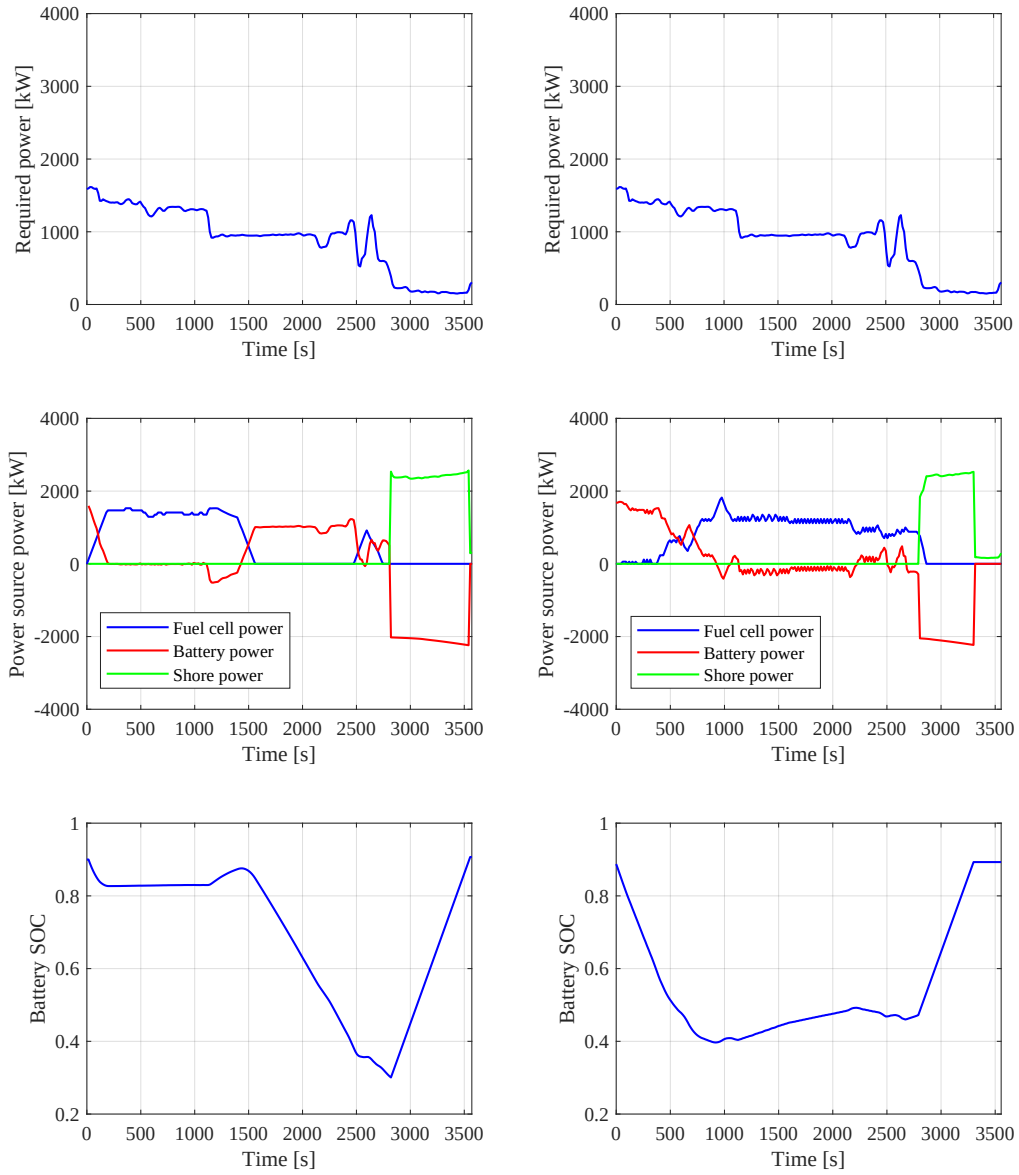


Figure 5.8: DDP and Double Q energy management strategies for training sample voyage 1 with low power demand: (a) optimal off-line strategy solved by DDP, (b) on-line strategy solved by the Double Q agent.

Table 5.7: Comparison of DDP and Double Q strategy voyage costs and GWP emissions for training sample voyage 1.

Training voyage 1	Voyage cost			Voyage GWP Emission		
	DDP [$]	Double Q [$]	$\frac{\text{DDP}}{\text{Double Q}}$ [%]	DDP [kg]	Double Q [kg]	$\frac{\text{DDP}}{\text{Double Q}}$ [%]
PEMFC	196.8	252.6	77.9	-	-	-
Battery	64.3	64.3	100.0	-	-	-
Electricity	44.6	31.3	142.4	83.4	58.6	142.4
H_2	279.5	337.7	82.8	50.9	61.5	82.8
Total	585.2	686.0	85.3	134.3	120.1	111.9

5.5.2.2 Training sample 2 with moderate power demand

Sample voyage 2 is a typical voyage with moderate power demand in the training dataset. Figure 5.9 compares the off-line DDP strategy (Figure 5.9a) and on-line Double Q strategy (Figure 5.9b) for this voyage. For both Double Q and DDP strategies, in the departure phase (0-800 s), the batteries provide most of the power from the beginning, while the fuel cells come online after a delay. The minimum SOC of the DDP strategy for this voyage is approximately 0.25 (at 2850 s). As the Double Q agent does not exactly know the future power demand and the strategy is generic, the Double Q strategy tends to adjust fuel cell power more frequently. Also, the fuel cells delay being switched to idle until shore power is available, which is because the agent does not know in advance if shore power is available, and the environment was designed to force the fuel cell power to decrease to zero only after shore power was being delivered. Note that, because the ship only stays in port for a short period between voyages, the batteries need to be charged at high C-rates, which could pose additional requirements on the charging infrastructure.

Table 5.8 depicts the cost and GWP emission breakdowns for sample voyage 2 in the training dataset. The Double Q strategy achieves 89.8% cost performance of that of the DDP strategy. Nevertheless, the Double Q strategy yields better GWP emission performance, which has also been observed in sample voyage 1 (Section 5.5.2.1). The H_2 costs account for 55.4% and 56.3% of the total voyage costs for the DDP and Double Q strategies, respectively. PEMFC degradation costs are the

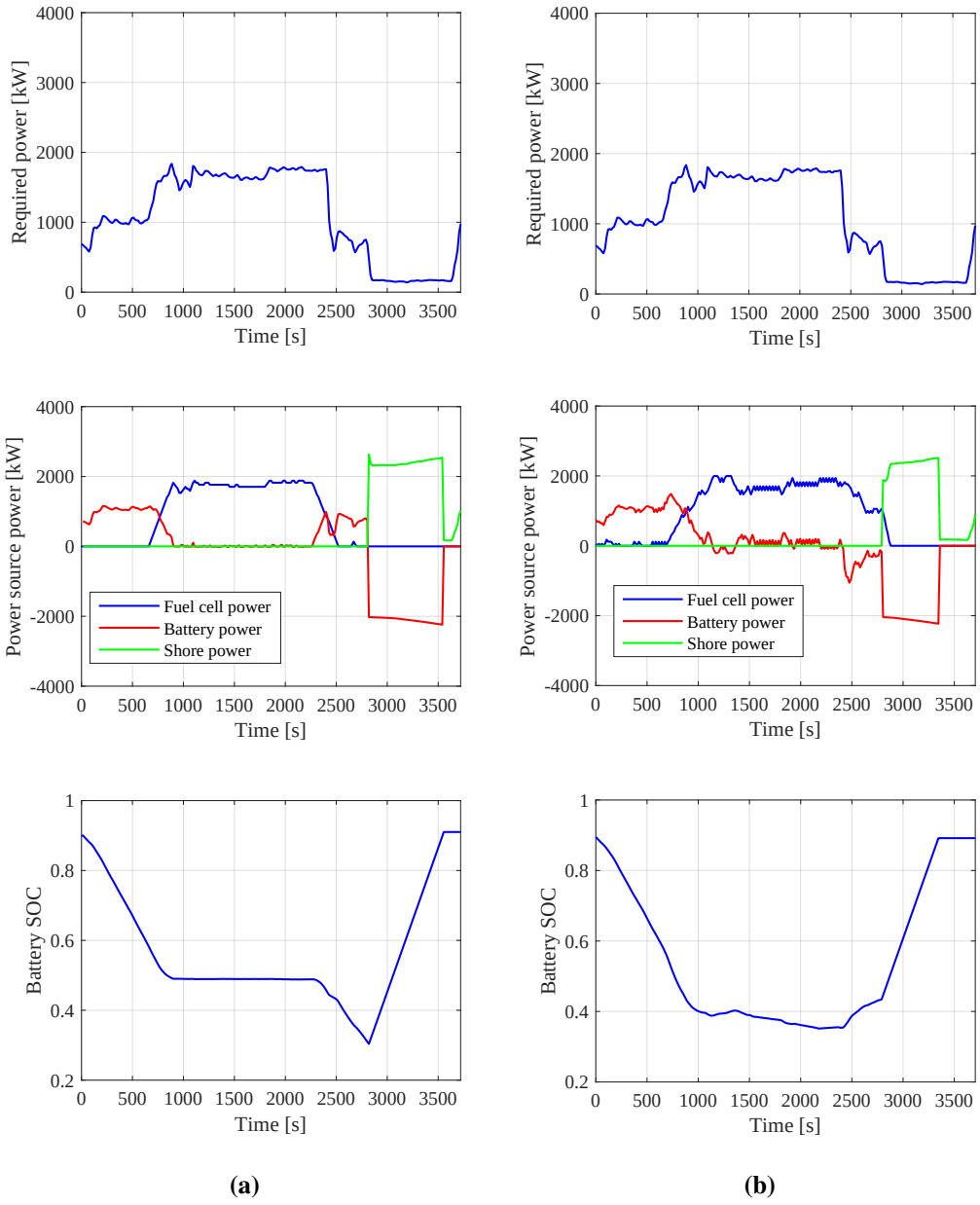


Figure 5.9: DDP and Double Q energy management strategies for training sample voyage 2 with moderate power demand: (a) optimal off-line strategy solved by DDP, (b) on-line strategy solved by the Double Q agent.

second highest cost source in both strategy results.

Table 5.8: Comparison of DDP and Double Q strategy voyage costs and GWP emissions for training sample voyage 2.

Training voyage 2	Voyage cost			Voyage GWP Emission		
	DDP [$]	Double Q [$]	$\frac{\text{DDP}}{\text{Double Q}}$ [%]	DDP [kg]	Double Q [kg]	$\frac{\text{DDP}}{\text{Double Q}}$ [%]
PEMFC	206.1	246.4	83.6	-	-	-
Battery	67.0	67.0	100.0	-	-	-
Electricity	45.5	34.5	132.1	85.1	64.4	132.1
H_2	395.8	447.5	88.5	72.1	81.5	88.5
Total	714.4	795.3	89.8	157.1	145.9	107.7

5.5.2.3 Training sample 3 with high power demand

As mentioned in Section 5.4.2, the Double Q agent failed to provide a strategy to complete the voyage in less than 0.5% of the training voyages as a consequence of final battery SOC constraint being exceeded. When these failed voyages were examined after the training process it was noted that they had much higher power demand compared to the typical voyages in the training dataset. Figure 5.10 presents a sample profile when it is known that the ship was heavily laden (corresponds the voyage with maximum cost in Figure 5.7a), and its optimal EMS solved via DDP (Figure 5.10a). Unlike the profile discussed in Section 5.5.2.2, the fuel cell power ramps up immediately after departure for this profile, in contrast to the more normal situation where significant increases in fuel cell power output are delayed as shown in a typical profile similar to Figure 5.9b. Without the battery over-discharge protection, the Double Q strategy tends to discharge the battery rapidly to a *SOC* below 0.25 after departure from the port. Figure 5.10b illustrates how the battery over-discharge protection function actuates to minimise the impact and shows how such an override function is effective when tackling voyages with very high power demand.

Table 5.9 presents a detailed comparison between the DDP and Double Q strategies in terms of voyage cost and GWP emissions. Such a high power profile is unusual in the training dataset. The DDP strategy would generate a voyage cost of $1228.0, which is 71.8% higher than that of sample voyage 2 (discussed in

Section 5.5.2.2). As a result of the battery over-discharge protection being triggered at 450 s, the PEMFC degradation cost of the Double Q strategy is less than that of the DDP strategy as frequent fuel cell power adjustments have been avoided by action overrides. However, the Double Q strategy outputs a much higher H_2 cost (36% higher), which is due to the PEMFC being forced to run at very high load regions where the fuel efficiency is reduced.

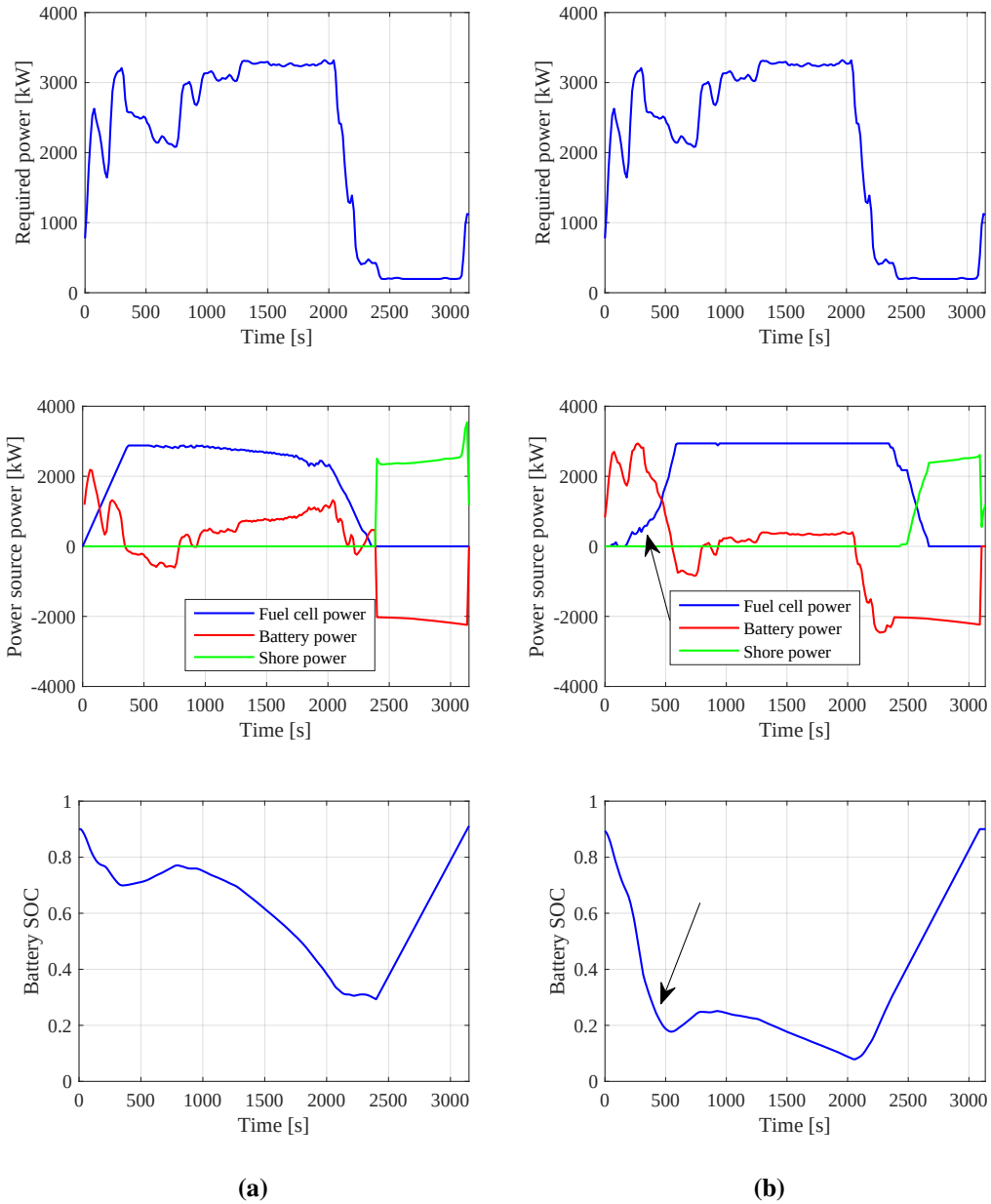


Figure 5.10: DDP and Double Q energy management strategies for training sample voyage 3 with high power demand: (a) optimal off-line strategy solved by DDP, (b) on-line strategy solved by the Double Q agent.

Table 5.9: Comparison of DDP and Double Q strategy voyage costs and GWP emissions for training sample voyage 3.

Training voyage 3	Voyage cost			Voyage GWP Emission		
	DDP [$]	Double Q [$]	$\frac{\text{DDP}}{\text{Double Q}}$ [%]	DDP [kg]	Double Q [kg]	$\frac{\text{DDP}}{\text{Double Q}}$ [%]
PEMFC	242.8	259.0	93.8	-	-	-
Battery	56.7	56.7	100.0	-	-	-
Electricity	46.5	32.9	141.4	87.0	61.5	141.4
H_2	881.9	1199.9	73.5	160.5	218.4	73.5
Total	1228.0	1548.5	79.3	247.5	279.9	88.4

5.5.3 EMS validation

5.5.3.1 Validation sample 1 with low power demand

Figure 5.11 shows the comparison between the DDP and Double Q strategies of a sample validation voyage with comparatively lower power demand. The Double Q strategy (Figure 5.11b) discharges the battery modules quickly down to a *SOC* of 0.4 in the first 1000 s, and maintains the fuel cell power output to a narrow region during sailing. The batteries satisfy significant transients in the departing and approaching phases. In contrast, the DDP strategy only discharges the battery rapidly at the beginning of the voyage (0-550 s). Similar trends have been observed in the sample training voyage (Figure 5.9).

Table 5.10 describes the detailed cost and GWP emission breakdowns of the validation sample voyage 1. The voyage cost of the Double Q strategy is 12.8% higher than that of the DDP strategy. Nevertheless, the Double Q EMS performs 10.1% better in terms of GWP emission. Such an observation reflects the trade-off between voyage costs and GWP emissions. Note that similar observations have been found in the training sample voyages (see Sections 5.5.2.1 and 5.5.2.2).

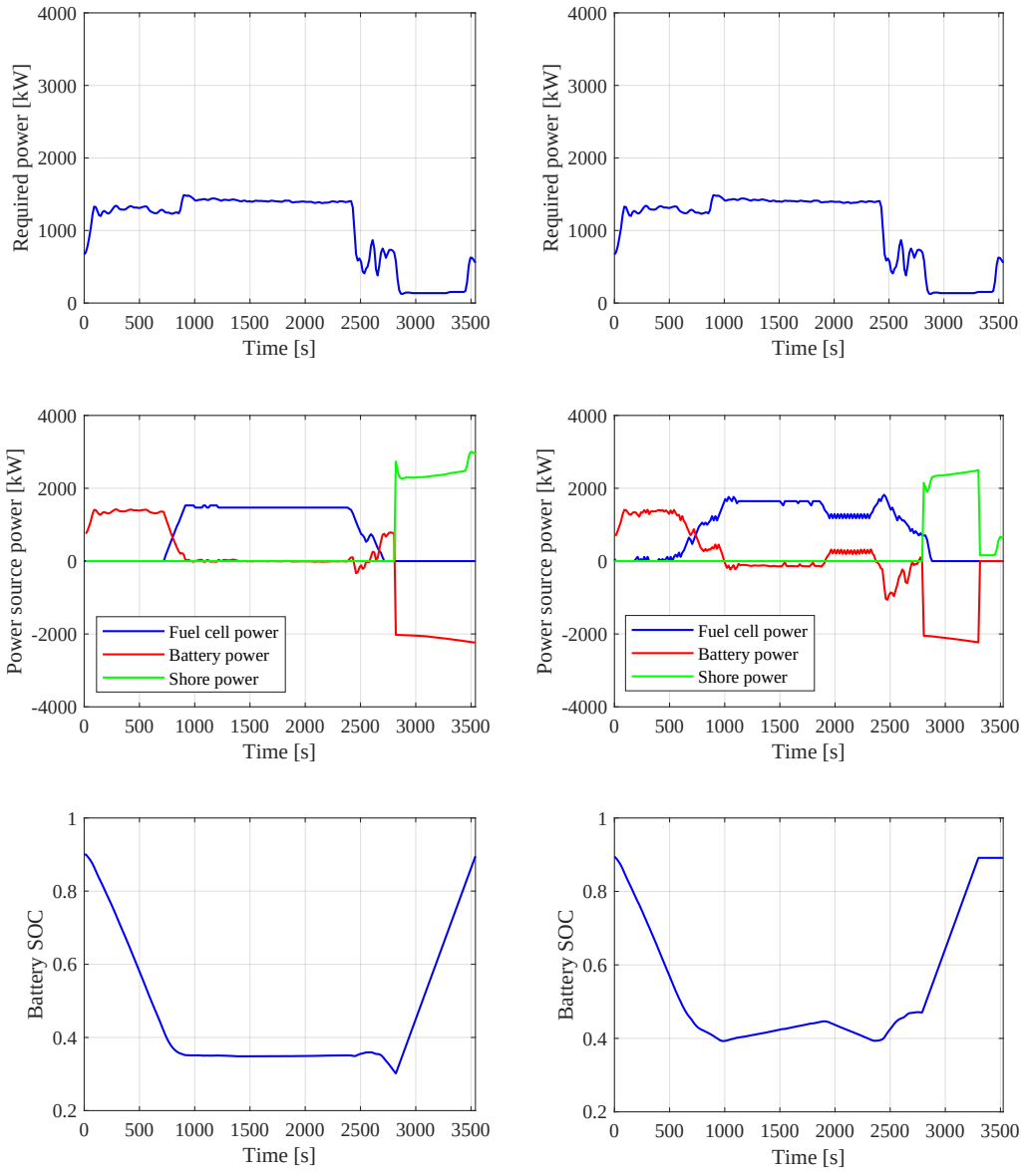


Figure 5.11: DDP and Double Q energy management strategies for validation sample voyage 1 with low power demand: (a) optimal off-line strategy solved by DDP, (b) on-line strategy solved by the Double Q agent.

Table 5.10: Comparison of DDP and Double Q strategy voyage costs and GWP emissions for validation sample voyage 1.

Validation voyage 1	Voyage cost			Voyage GWP Emission		
	DDP [$]	Double Q [$]	$\frac{\text{DDP}}{\text{Double Q}}$ [%]	DDP [kg]	Double Q [kg]	$\frac{\text{DDP}}{\text{Double Q}}$ [%]
PEMFC	208.6	244.7	85.2	-	-	-
Battery	63.7	63.7	100.0	-	-	-
Electricity	44.4	31.2	142.2	82.9	58.3	142.2
H_2	345.2	406.8	84.9	62.8	74.1	84.9
Total	661.9	746.5	88.7	145.8	132.4	110.1

5.5.3.2 Validation sample 2 with moderate power demand

Figure 5.12 presents the DDP and Double Q strategies of a sample profile with moderate power demand from the validation dataset. In Figure 5.12a, as the complete profile is known before solving the DDP strategy, the DDP strategy only adjusts PEMFC power output when necessary. As in Figure 5.12b, the Double Q strategy adjusts PEMFC power more frequently due to uncertainty regarding the power demand in the next time steps. Such a pattern has also been observed in the first two training sample profiles.

The Double Q strategy voyage cost is 11.2% higher than that of the DPP strategy, which is due to frequent PEMFC power adjustments and higher H_2 consumption. Note that the Double Q strategy still performs better than the DDP strategy in terms of GWP emissions (Table 5.11).

Table 5.11: Comparison of DDP and Double Q strategy voyage costs and GWP emissions for validation sample voyage 2.

Validation voyage 2	Voyage cost			Voyage GWP Emission		
	DDP [$]	Double Q [$]	$\frac{\text{DDP}}{\text{Double Q}}$ [%]	DDP [kg]	Double Q [kg]	$\frac{\text{DDP}}{\text{Double Q}}$ [%]
PEMFC	211.7	239.6	88.4	-	-	-
Battery	63.7	63.7	100.0	-	-	-
Electricity	43.9	32.4	135.6	82.0	60.5	135.6
H_2	411.9	477.6	86.3	75.0	86.9	86.3
Total	731.2	813.2	89.9	157.0	147.4	106.5

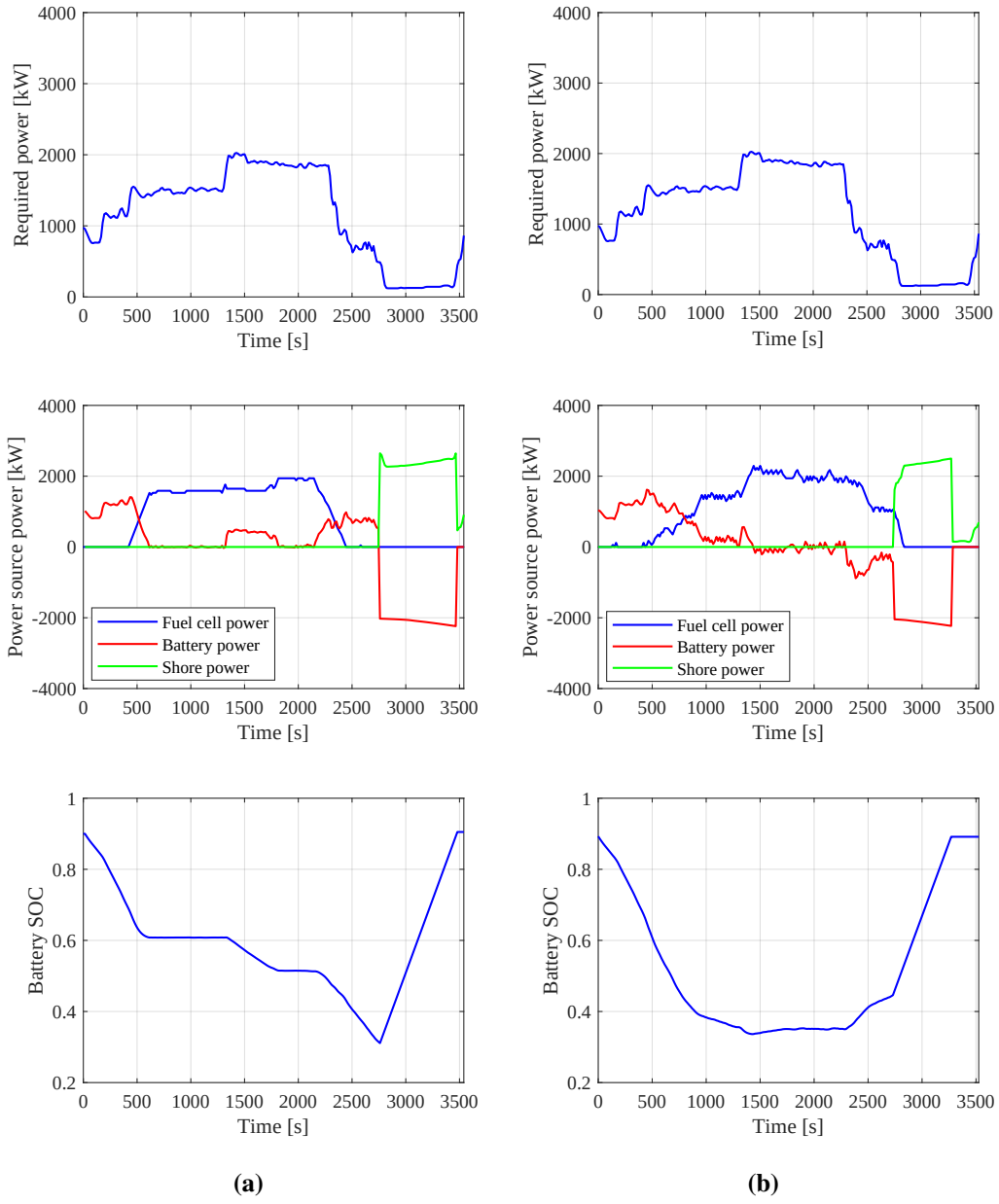


Figure 5.12: DDP and Double Q energy management strategies for validation sample voyage 2 with moderate power demand: (a) optimal off-line strategy solved by DDP, (b) on-line strategy solved by the Double Q agent.

5.5.3.3 Validation sample 3 with high power demand

As discussed in Section 5.5.2.3, the Double Q agent failed in training voyages with extremely high power demand. Nevertheless, the Double Q strategy managed to complete all the validation voyages without triggering the battery over-discharge protection function. Figure 5.13 compares the DDP and Double Q strategies. As in Figure 5.13b, the Double Q strategy discharges the battery rapidly to a *SOC* of 0.4 after departure with a delay before the PEMFC provides any power output. In contrast, the DDP strategy (Figure 5.13a) ramps the PEMFC output immediately at departure in response to such a high load profile.

The voyage cost of the DDP strategy is 89.9% of its RL counterpart (Table 5.12). It is worth noting that the GWP emissions produced by the two strategies are very close to each other (0.7% difference). Although the Double Q strategy consumes more H_2 than the DDP strategy, it requires much less shore generated electricity compared to the DDP strategy.

Table 5.12: Comparison of DDP and Double Q strategy voyage costs and GWP emissions for validation sample voyage 3.

Validation voyage 3	Voyage cost			Voyage GWP Emission		
	DDP [$]	Double Q [$]	$\frac{\text{DDP}}{\text{Double Q}}$ [%]	DDP [kg]	Double Q [kg]	$\frac{\text{DDP}}{\text{Double Q}}$ [%]
PEMFC	256.2	257.4	99.5	-	-	-
Battery	63.7	63.7	100.0	-	-	-
Electricity	50.4	36.9	136.3	94.1	69.1	136.3
H_2	605.2	734.9	82.3	110.2	133.8	82.3
Total	975.5	1093.0	89.3	204.3	202.8	100.7

Table 5.13 summaries the Double Q-learning strategy performance in comparison with DDP strategy. Wu et al. (2018) reported that, in a non-stochastic environment with a single power profile, their Q-learning agent achieved 89.0% fuel economy compared to dynamic programming policy in their road vehicle-related study. The Double Q-learning strategy presented in this study has achieved 89.0%, and 88.9% cost performance of refined DDP strategy results in training and validation datasets, respectively. The Double Q-learning agent presented in this chapter

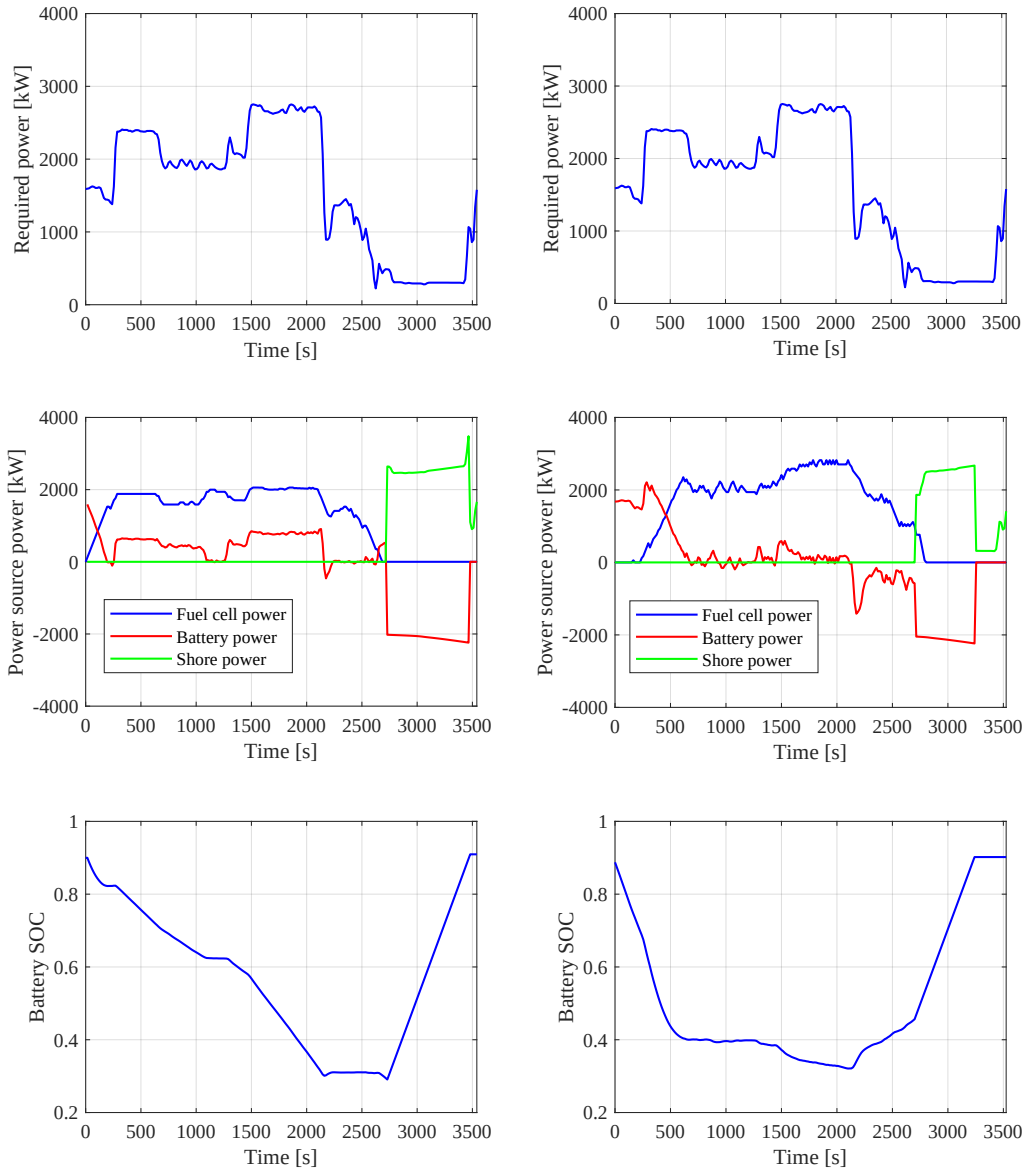


Figure 5.13: DDP and Double Q energy management strategies for validation sample voyage 3 with high power demand: (a) optimal off-line strategy solved by DDP, (b) on-line strategy solved by the Double Q agent.

can achieve near-optimal cost performance for the case ship in a stochastic environment. Although the objective of the Double Q-learning strategy was designed to minimise voyage costs, due to the trade-off between costs and GWP emissions, the Double Q-learning strategy perform even better than the DDP strategy in terms of GWP emissions (approximately 6% less GWP emissions for both training and validation datasets). More H_2 usage would result in higher voyage costs but lower GWP emissions.

Table 5.13: Summary of Double Q and DDP strategy voyage costs and GWP emissions.

Category	Profile	Voyage cost			Voyage GWP emission		
		DDP [$]	Double Q [$]	$\frac{\text{DDP}}{\text{Double Q}}$ [%]	DDP [kg]	Double Q [kg]	$\frac{\text{DDP}}{\text{Double Q}}$ [%]
Training	Sample 1	585.2	686.0	85.3	134.3	120.1	111.9
	Sample 2	714.4	795.3	89.8	157.1	145.9	107.7
	Sample 3	1228.0	1548.2	79.3	247.5	279.9	88.4
Average	all training profiles	740.0	831.8	89.0	161.4	152.0	106.1
Validation	Sample 1	661.9	746.2	88.7	145.8	132.4	110.1
	Sample 2	731.2	813.0	89.9	157.0	147.4	106.5
	Sample 3	975.5	1092.7	89.3	204.3	202.8	100.7
Average	all validation profiles	723.5	813.8	88.9	158.4	149.2	106.2

5.6 Summary

This chapter has formulated the optimal energy management problem of the plug-in hybrid PEMFC/battery system using the novel approach of Markov Decision Process. The formulated Markov Decision Process has been solved using reinforcement learning agents in discrete state and action spaces. With continuous monitoring data collected from the case ship, a Double Q reinforcement learning based energy management strategy has been proposed. The Double Q agent has been trained adequately with one dataset of 1081 training voyages and subsequently validated using another dataset of 381 voyages over different periods.

Without prior knowledge of future power demand, the Double Q agent can achieve a cost-performance similar to that solved by dynamic programming with the identical settings in state and action spaces. Such a similarity indicate that the Double Q agent is effective in dealing with stochastic environments by reducing

maximisation biases. Also, such performance suggests that reinforcement learning is a viable approach to solve the optimal power split problem in a hybrid propulsion system, provided that enough historical data has been collected. In contrast, the Q agent which introduces maximisation biases fails to achieve satisfactory performance.

It can also be observed that Double Q strategy tends to adjust fuel cell power output more frequently, which could be due to two reasons: (1) the agent does not have certainty as to what will be the next power demand (i.e. the environment is stochastic) and (2) the state space was defined with limited resolution. Refining the action and state spaces could possibly reduce the amplitude of these fuel cell power adjustments. Both approaches would require other RL approaches with function approximators to deal with the increases in action and state spaces.

Chapter 6

Deep reinforcement learning based energy management strategies

6.1 Overview

In the previous chapter, the optimal energy management problem of the plug-in hybrid PEMFC/battery system has been formulated and solved using reinforcement learning algorithms in discrete state and action spaces. However, the accuracy of such an approach is limited by the resolutions of the two spaces without applying function approximators.

This chapter aims to further improve the cost-effectiveness of reinforcement learning based energy management strategies by refining the state space with deep neural networks as function approximators. The optimal energy management problem of the plug-in hybrid PEMFC/battery propulsion system will be solved using Deep Q-Network (DQN) (Mnih et al., 2015) and Double DQN agents (van Hasselt et al., 2015). Section 6.2 details the DQN and Double DQN agents. Section 6.3 reshapes the reward function proposed in Chapter 5 by removing unnecessary training steps in port. Section 6.4 depicts the training processes of the agents. Section 6.5 assesses and discusses the EMS performance.

6.2 Deep reinforcement learning agents

6.2.1 Deep Q-Network

For reinforcement learning problems with large or continuous state spaces, function approximators are typically needed to generalise from previously encountered states which are similar in some sense to current ones (Sutton and Barto, 2018). A function approximator can be linear or non-linear (Boyan and Moore, 1995; Sutton et al., 2000). However, the training process of reinforcement learning agents can be unstable or even diverge when a non-linear function approximator such as a neural network is used (Tsitsiklis and Van Roy, 1997). Lin (1993) developed the concept of 'experience replay' to store the agent experience into a memory pool to train a reinforcement learning agent with a neural network. Later work of Mnih et al. (2013) proposed deep Q-learning using a deep neural network with convolution layers to approximate high dimensional raw pixel state inputs. Mnih et al. (2015) further improved the deep Q-learning agents by adding target networks to improve training stability. Mnih et al. (2015)'s Deep Q-Network (DQN) achieved performance levels comparable to professional human game testers in 49 Atari 2600 games.

Figure 6.1 presents the detailed DQN agent-environment interaction framework. The DQN is a model-free, off-policy reinforcement learning algorithm (Mnih et al., 2015). The agent maintains an experience memory pool with capacity M, storing the most recent M transition sequences. A transition sequence, collected via agent-environment interaction, is denoted by:

$$\phi = (s_t, a_t, s_{t+1}, r_{t+1}) \tag{6.1}$$

i.e. at time step t, in state s_t, the agent performs action a_t (following ε-greedy policy) and observes next environment state s_{t+1} and a reward signal r_{t+1} is returned from the environment. In each agent training step, a mini-batch with capacity D is randomly sampled from the experience memory pool such that previous experiences

can be used effectively. In addition, the random sampling breaks the correlations of consecutive samples which can lead to unstable neural network training. The DQN agent includes two deep neural networks with identical structure, i.e. the Q-network $Q(s,a;\theta)$ parametrised by θ, and the Q-target network $\hat{Q}(s,a;\theta^{-})$ parametrised by θ^{-}. These neural networks approximate the action-value function with state (s) inputs for all actions (a) in the action space A.

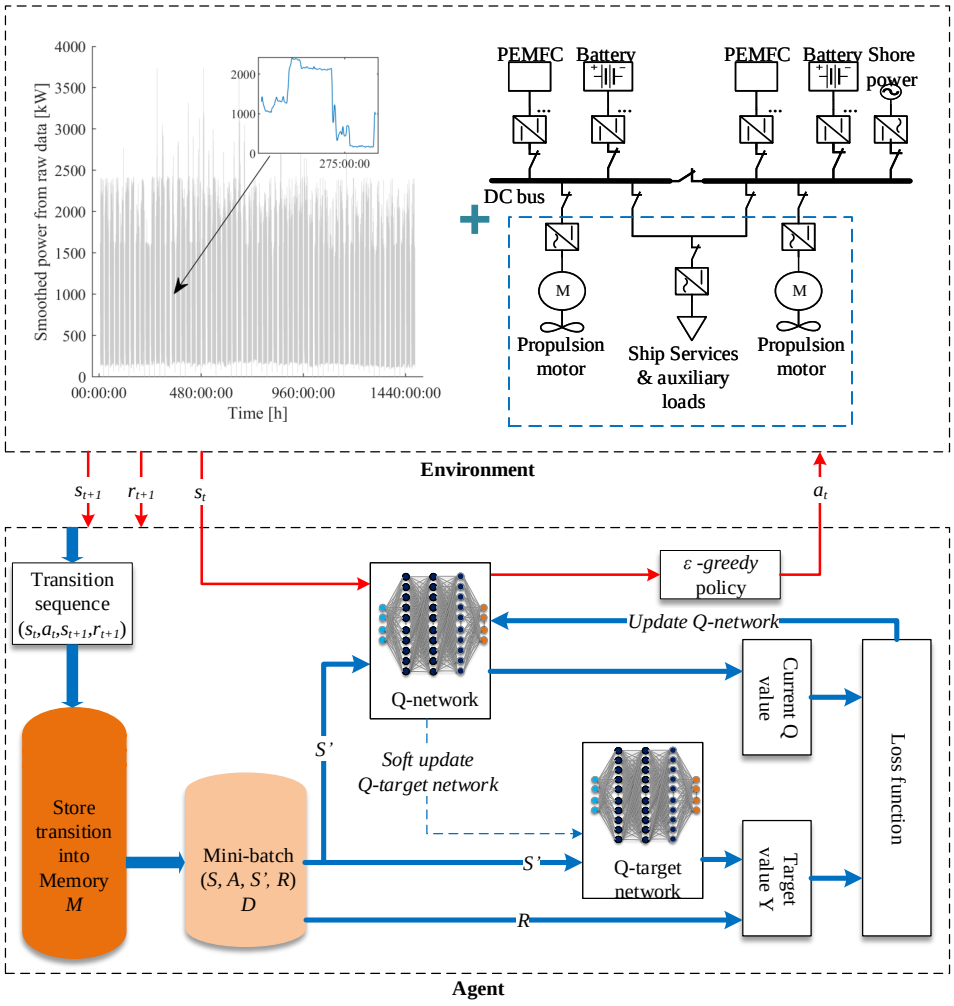


Figure 6.1: Deep Q-Network agent and environment schematic.

As an improvement to the Deep Q-network of Mnih et al. (2013), the additional Q-target network enhances the agent training stability by providing fixed target ac-

tion value y_j for non-terminal states:

$$y_j = r_{j+1} + \gamma \max_{a'} \hat{Q}\left(s_{j+1}, a'; \theta^-\right) \tag{6.2}$$

where j denotes $j-th$ sample in the mini-batch. Note that the terminal state is defined as states with battery over charged ($SOC > 1$) or discharged beyond capacity ($SOC < 0$). In the original DQN algorithm of Mnih et al. (2015), the Q-target network is updated periodically, while in this work it is soft-updated at each training step to further improve training stability:

$$\theta^- \leftarrow \tau\theta + (1-\tau)\,\theta^- \tag{6.3}$$

where $\tau \ll 1$ (Lillicrap et al., 2015).

Two types of loss functions (i.e. Mean Squared Error (MSE) and Huber loss (Huber, 1992)) are employed independently in this study to investigate the influences of the loss function over training stability and EMS quality in a stochastic environment. The MSE loss is defined as the mean squared error of the temporal difference (denoted by δ) between the action values given by the Q-network and the targets y_j ($j \in [1, D]$) over a mini-batch:

$$L(\theta) = \frac{1}{D}\sum_{j=1}^{D} \delta_j^2 \tag{6.4}$$

where the temporal difference δ_j of $j-th$ sample in the mini-batch is:

$$\delta_j = y_j - Q\left(s_j, a_j; \theta\right) \tag{6.5}$$

In the work of Mnih et al. (2015), δ_j was clipped to between -1 and +1 to improve the DQN algorithm stability. Such a technique corresponds to using an absolute value loss function for temporal differences outside $(-1, 1)$. Note that the clipping reduces the chances of overestimations for the action-value function when values given by the networks are noisy over large ranges. In Chapter 5, the Q-

learning agent failed due to overestimations caused by the maximisation operation which approximated the expected action value. The concept of error clipping may provide a new approach to dealing with overestimations in the stochastic environment.

Instead of clipping the error term, the Huber loss, which performs similar function has been employed in this study. The Huber loss is calculated by (Huber, 1992):

$$L(\theta) = \frac{1}{D}\sum_{j=1}^{D}\sigma_j \tag{6.6}$$

where:

$$\sigma_j = \begin{cases} \frac{1}{2}\delta_j^2, & \text{if } \left|\delta_j\right| < 1 \\ \left|\delta_j\right| - \frac{1}{2}, & \text{otherwise} \end{cases} \tag{6.7}$$

The Huber loss is the mean squared error when the temporal difference δ_j is small ($\left|\delta_j\right| < 1$) but acts like the mean absolute error ($|\delta| - \frac{1}{2}$) when the difference is large, which makes it more robust when overestimations of action-value function may degrade the agent training.

Figure 6.2 shows the neural network structure for the Q-network and the Q-target network. The neural networks are configured with two fully-connected hidden layers. The Q-network is trained by minimising the loss function $L(\theta)$ with respect to its parameters θ. The optimiser adopted in this study is the Adam optimiser (Kingma and Ba, 2014). The neural networks output action-value function values for each possible action with given state inputs. Note that the continuous signals (battery state of charge, power demand and fuel cell per unit power) are not discretised as in Chapter 5. Instead, these states are used as direct inputs to the Q-network such that the environment states can be accurately represented by continuous actual values. The state inputs are forward propagated sequentially from the input layer via hidden layers to output Q-values for all actions. Note that each neuron of the output layer corresponds to an action in the action space.

As depicted in Algorithm B.4 in Appendix B, the DQN agent training starts

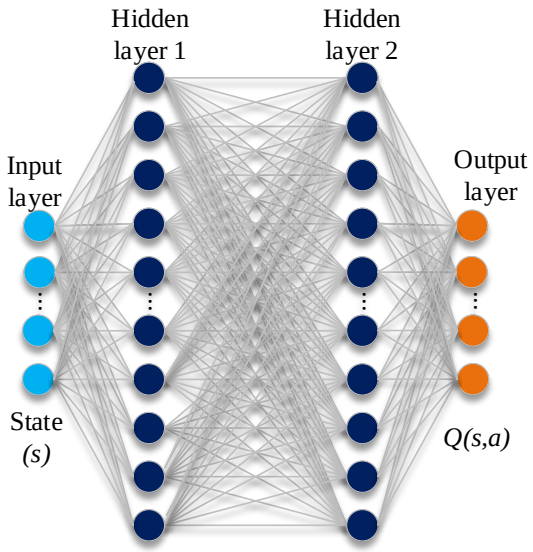


Figure 6.2: Deep neural network with 2 fully-connected hidden layers.

with randomly initialised network parameters θ and θ^{-}. The exploration probability ε of the ε-greedy policy decreases linearly from a large initial value with the increase of training episode number, and is fixed at a small final value in the later stage of the training (i.e. training episode $n > N_d$). Note that completely random explorations initially fill the experience memory pool before the neural network training starts. The Q-network is trained every Z steps to gain sufficient experience.

6.2.2 Double Deep Q-Network

The results in Chapter 5 suggest that maximisation biases introduced during the construction of the action-value function can lead to poor learning performance if such biases are not addressed properly. The Double Q-learning agent achieved satisfactory performance using two Q-functions, while, with the same hyperparameter settings, the Q-learning agent diverged. It is not clear whether the DQN agent (as a deep variant of Q-learning) can succeed in the highly stochastic environment based on recorded historical power profiles. Therefore, the author has also explored solving the energy management problem with Double DQN.

The Double DQN (Algorithm B.5 in Appendix B) is proposed by van Hasselt

et al. (2015) based on the concept of Double Q-learning (van Hasselt, 2010) and DQN (Mnih et al., 2015). In Double Q-learning, two Q-functions are used to reduce the overestimations by decomposing the maximisation in the target into action selection and action evaluation (van Hasselt, 2010; van Hasselt et al., 2015). Without introducing additional neural networks to DQN, the Double DQN utilises the Q-target network to evaluate the maximising action (i.e. $\arg\max_a \left(Q\left(s_{j+1},a;\theta\right)\right)$) given by the Q-network (see Figure 6.3) such that, the target value is calculated by:

$$y_j = r_j + \gamma Q\left(s_{j+1}, \arg\max_a \left(Q\left(s_{j+1},a;\theta\right)\right);\theta^-\right) \tag{6.8}$$

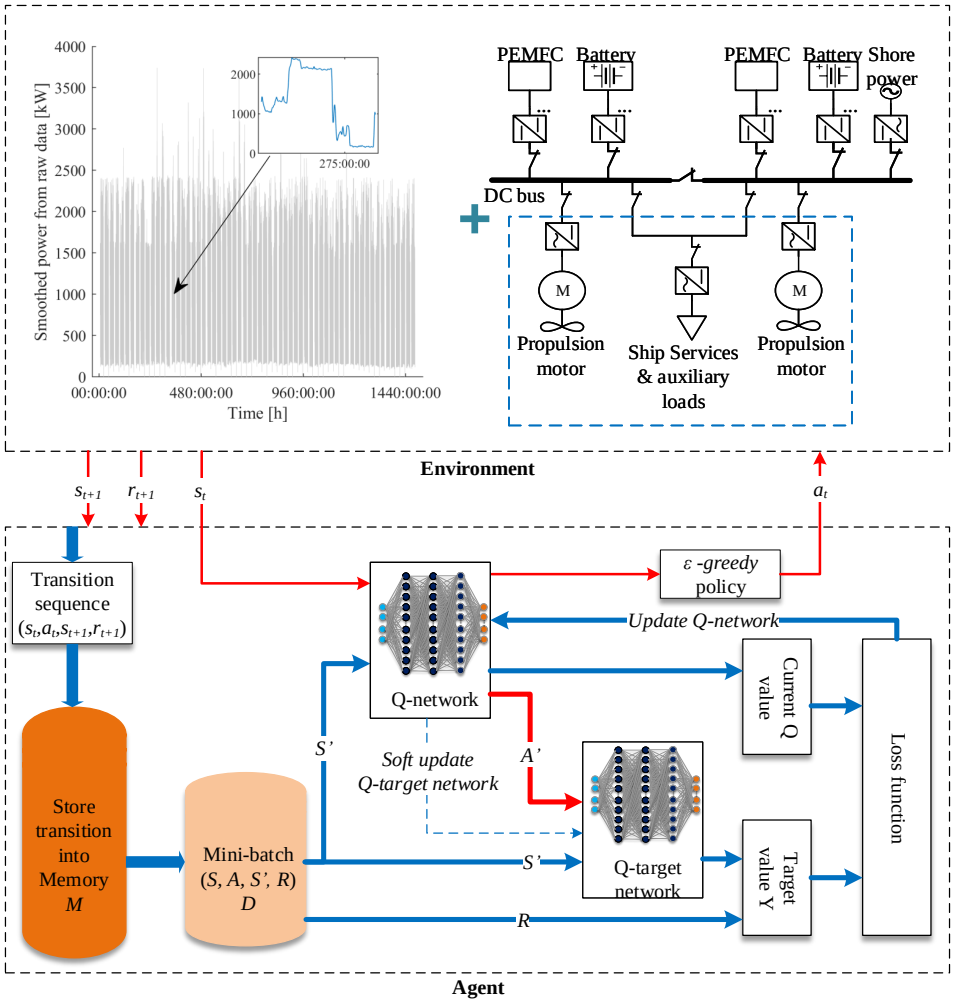


Figure 6.3: Double Deep Q-Network agent and environment schematic.

6.3 Environment update

6.3.1 Reward function

In Chapter 5, the agent was trained throughout all time steps of the voyages, i.e. from the first moments of departure to last moments in ports. However, the ship was designed to operate purely on shore based electricity when in port mode (i.e. cold ironing, see Eq. 4.1). Although the Double Q agent has demonstrated its ability to maintain zero fuel cell power state in port mode, removing the training steps in port mode could potentially simplify the training processes due to the cold ironing logic (see Section 4.2.2.1) would only utilise shore provided power. Control of the fuel cell when in port mode appears unnecessary. Also, it has been observed that, the Double Q agent struggled to maintain final battery SOC constraint in some high power profiles. In practice, it would be feasible to increase the port time slightly to get the battery charged to SOC_H. Therefore, the reward function of the environment is reshaped as:

$$r_{t+1} = \begin{cases} -1, & \text{spA=0, if } s_{t+1} \text{ is infeasible} \\ -1, & \text{spA=0, if } p_{fc} + a_t \notin [0,1] \\ \tanh\left(\dfrac{1}{cost_{t+1}}\right), & \text{spA=0, else} \\ \displaystyle\sum_{k=t+1}^{K} \tanh\left(\dfrac{1}{cost_k}\right), & \text{spA=1} \end{cases} \tag{6.9}$$

where when shore power is available ($spA = 1$), the environment returns a summed reward of all the costs incurred in port mode of the current episode. In sailing mode, i.e. $spA = 0$, the reward function is defined identically as in Chapter 5 (see Section 5.3.3). Note that K is the time step when the entire profile is completed; and $cost_k = \infty$ if $k > T$ (i.e. extra time required to charge the battery), otherwise $cost_k$ is calculated as described in Chapter 5 (Eq. 5.8).

6.3.2 State space

Previously, in Chapter 5, the four-dimensional state space was discretised to state indices (Section 5.3.1) to store the action-value function into tables indexed by discrete state indices. Such a discretisation process is necessary for tabular RL approaches. However, the discretisation process and its resolution limit the quality of the generated policy (Sutton and Barto, 2018). In this chapter, discretisation of state space has been removed. The actual state space:

$$s(t) = [spA(t), p_{dem}(t), x(t), SOC(t)]^T \tag{6.10}$$

is directly applied to represent the environment states, where spA denotes the shore power availability ($spA = 0$ for sailing mode, $spA = 1$ for port mode), P_{dem} is normalised system power demand by dividing the actual power demand in kW by 1500 (i.e. $p_{dem} = \frac{P_{dem}}{1500}$, such that the power demand input to the Q-network is around 1), $x(t)$ fuel cell per unit power level at time step t ($x \in [0,1]$), and $SOC \in [0,1]$ denotes battery state of charge (SOC).

6.3.3 Action

The action space is defined as a tuple of fuel cell power level changes:

$$A = [a_1, a_2, ..., a_m, ..., a_{n-1}, a_n]^T \tag{6.11}$$

where $a_1 < 0$ is the maximum decrease and $a_n > 0$ is the maximum increase of fuel cell output in a time step, $a_m = 0$ means maintaining current power level; all other values of a represent changes of power within the range of (a_1, a_n).

In sailing mode, the environment overrides an action that would result the fuel cell power output becoming negative or higher than the rated power. When action $a_t \in A$ is chosen from the action space at time step t, the fuel cell power level at

$t+1$ will be:

$$x_{t+1} = \begin{cases} 0, & x_t + a_t < 0 \\ 1, & x_t + a_t > 1 \\ x_t + a_t, & \text{else} \end{cases} \tag{6.12}$$

In port mode, the agent is not required to control the fuel cell. The environment would force the fuel cell power to decrease to zero if fuel cell power is not zero. Note that the environment would extend episode length whenever necessary to charge the battery SOC to SOC_H (the power demand would be extrapolated from the last power demand that appears in original power profile). Such settings vary from the ones defined in Chapter 5, in which the agents were required to explore actions in port mode to maintain cold ironing.

6.4 Agent training

The agents were trained on a workstation with two Intel Xeon E5-2683 V3 processors running on Windows 10. The environment and the agent were coded in Python. The agent's neural networks were built and trained with PyTorch v1.20. Each agent was trained with 10 different random seeds for reproducibility. During training, the agent policy performance was assessed by calculating the average values and standard deviations across the 10 instances running with different random seeds. Note that as the neural networks are relatively small, only one CPU thread is assigned to each running instance to avoid training speed degradation due to unnecessary parallelisation.

Also, the actual policy performance was periodically tested (every 100 training episodes) with 10 random training voyages during training. Note that in test mode, the $\varepsilon-$greedy exploration probability was set at 0 with battery over-discharge protection enabled (disabled in training mode). Once the training of all the 10 instances was completed, the agent with the lowest episode cost was chosen to generate detailed EMS results in the following sections.

6.4.1 Neural network settings

Figure 6.4 illustrates the neural network configuration for the Q and Q-target networks. The environment state inputs are processed by the input layer with four neurons with Rectified Linear Unit (ReLU) activation function. Two fully-connected hidden layers are configured with 256 neurons each. Note that both hidden layers are applied with an ReLU activation function, while no activation function is applied to the output layer to allow negative action-value outputs. The neural network outputs five Q-values, corresponding to the 5 actions in the action space, respectively.

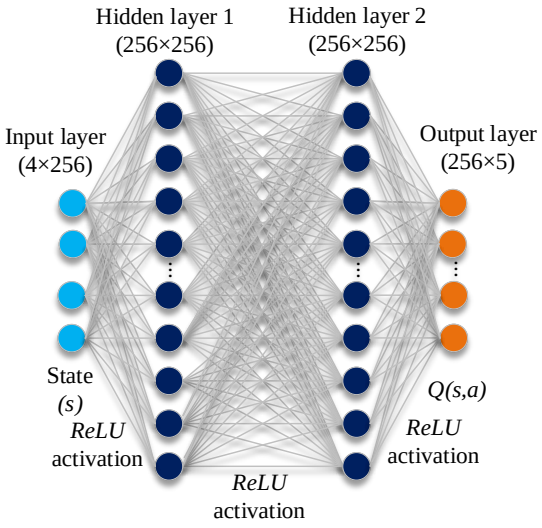


Figure 6.4: Q-network and Q-target network settings.

6.4.2 Hyperparameter settings

Table 6.1 details the hyperparameter settings used for the four agent-loss function combinations. The policy is updated every 32 transition sequences (ϕ). In each training step, a mini-batch with 32 transition sequences is randomly sampled from the experience memory with a capacity of 1×10^6. Such a mini-batch is applied to train the Q-network using an Adam optimiser. The learning rate of the Adam optimiser is fixed at 0.0001 throughout the training. The exponential decay rates

of first and second moment estimates (β_1 and β_2) are set at 0.9 and 0.999 respectively (Kingma and Ba, 2014). Note that the Q-target network is soft-updated with a soft-update weight of $\tau = 0.001$ in each training step. The exploration probability ε of $\varepsilon-$greedy policy starts with 1 and fixes at 0.05 after 5×10^3 episodes of training. Note that these parameters require careful tuning to achieve satisfactory performance.

Table 6.1: Hyperparameter settings.

Parameter	Description	Value
B	Mini-batch size	32
M	Experience memory size	1×10^6
τ	Target network update weight	0.001
γ	Discount factor	1
Z	Policy update frequency	32
α	Learning rate of Adam optimiser	0.001
β_1	Exponential decay rate for the first moment estimates of Adam optimiser	0.9
β_2	Exponential decay rate for the second moment estimates of Adam optimiser	0.999
ε_0	Initial exploration probability	1
ε_f	Final exploration probability	0.05

6.4.3 Training

Two agents, i.e. DQN and Double DQN, have been tested in this study. Also, two types of loss functions, i.e. MSE and Huber losses are tested with the two deep reinforcement learning agents.

6.4.3.1 MSE loss

Figures 6.5 and 6.6 illustrates the training processes with MSE loss function for the DQN and Double DQN agents respectively. Both agents were trained over 1.2×10^4 episodes. Every 100 training episodes, the energy management strategy is tested by sampling 10 random training profiles. The DQN agent training diverged without finding an effective strategy. Although the Double DQN performed slightly better than the DQN, it was not particularly successful.

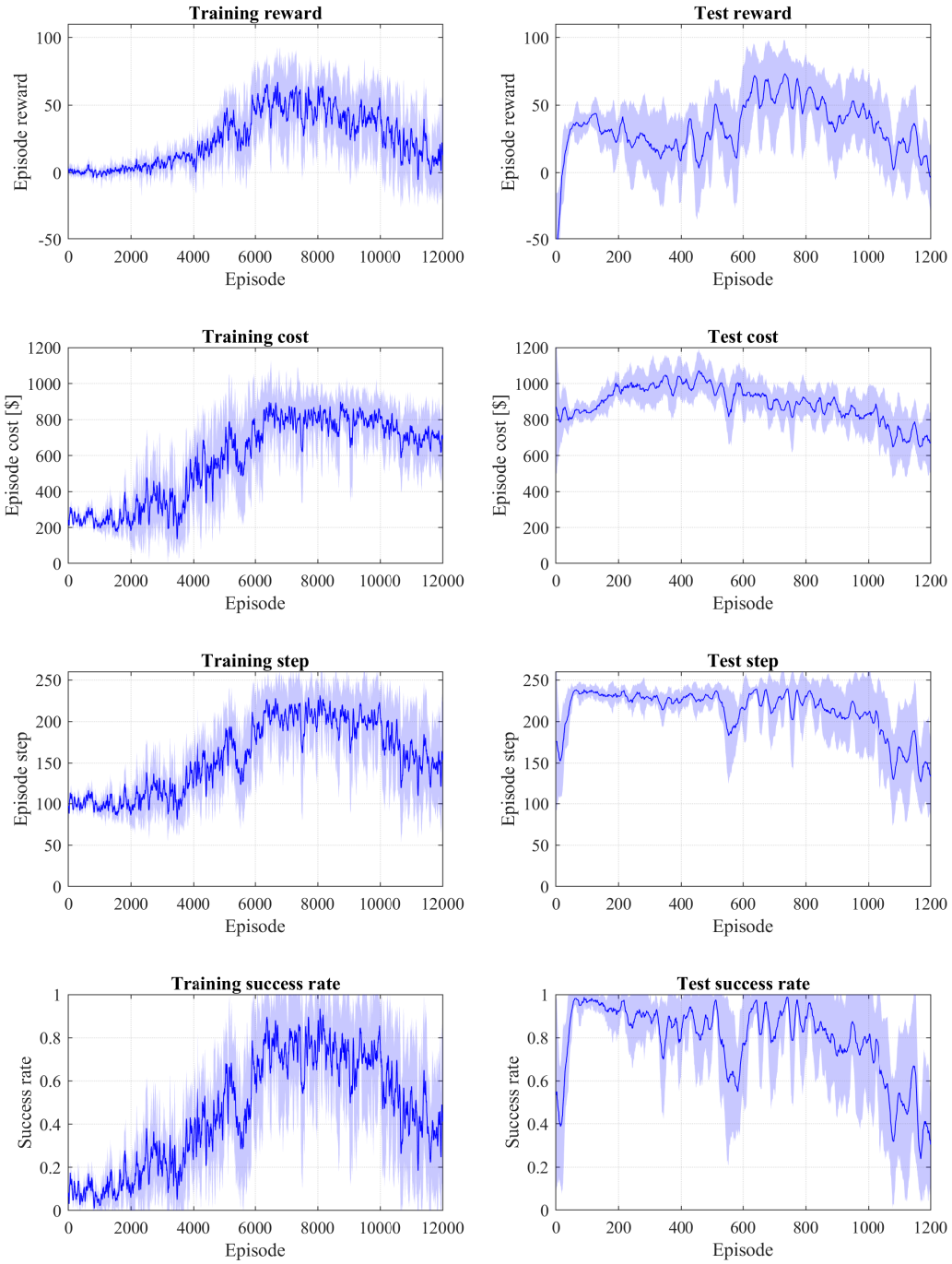


Figure 6.5: DQN agent training and testing with MSE loss function. The deep blue lines are moving average values across 10 instances running with different random seeds. The light blue shadows are the confidence bounds calculated by mean values $\pm$ standard deviations across the 10 instances.

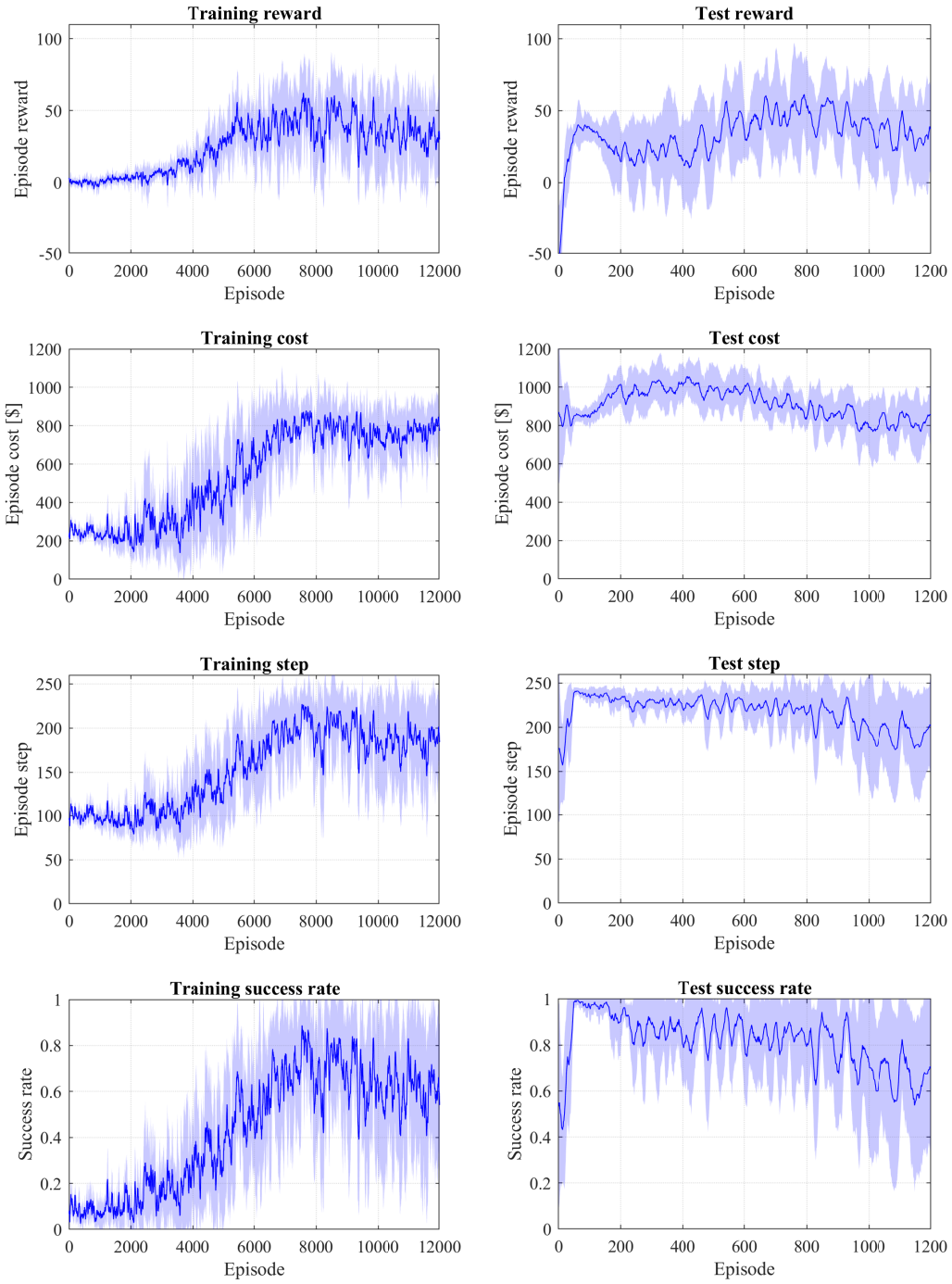


Figure 6.6: Double DQN agent training and testing with MSE loss function. The deep blue lines are moving average values across 10 instances running with different random seeds. The light blue shadows are the confidence bounds calculated by mean values $\pm$ standard deviations across the 10 instances.

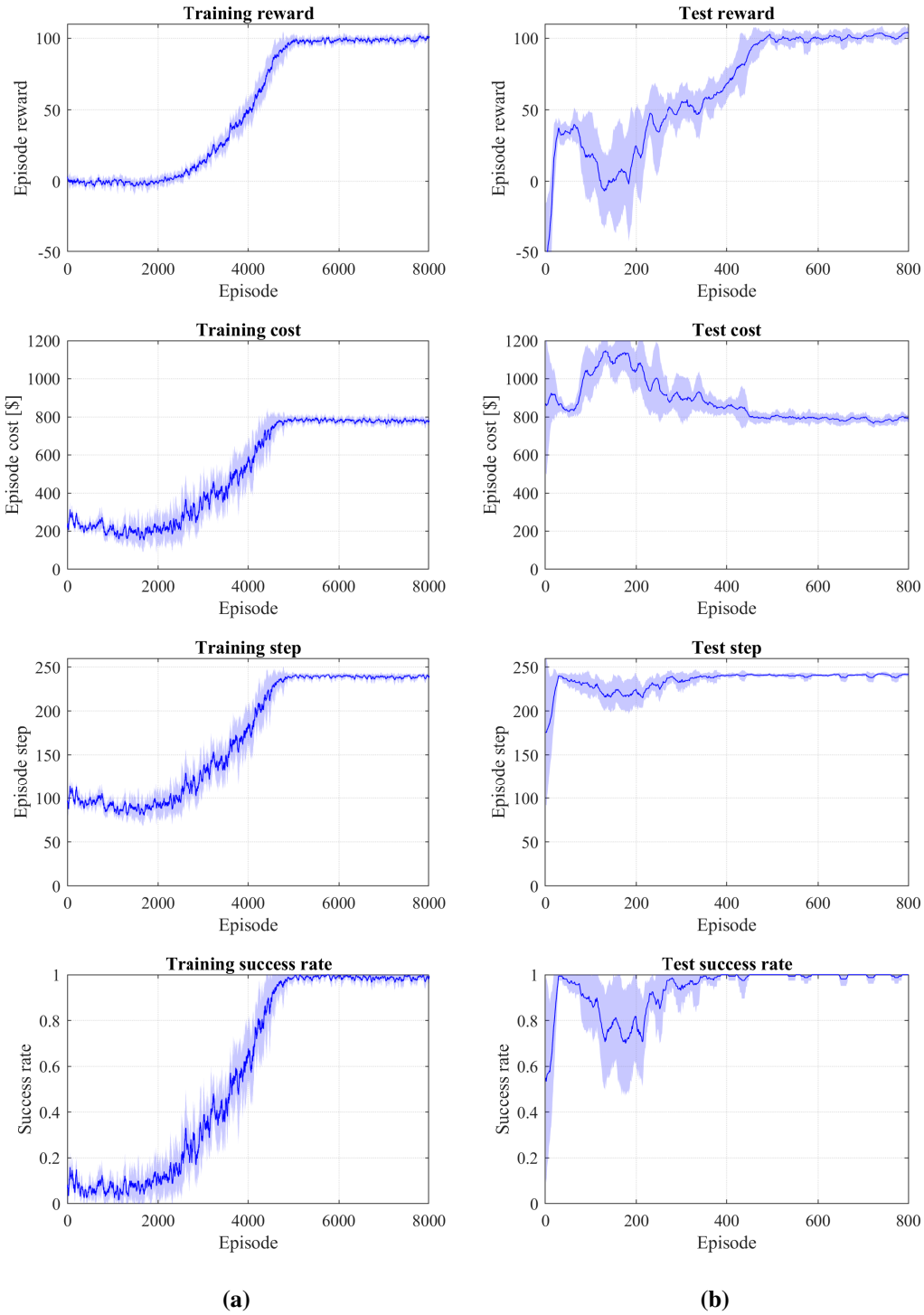


Figure 6.7: DQN agent training and testing with Huber loss function. The deep blue lines are moving average values across 10 instances running with different random seeds. The light blue shadows are the confidence bounds calculated by mean values $\pm$ standard deviations across the 10 instances.

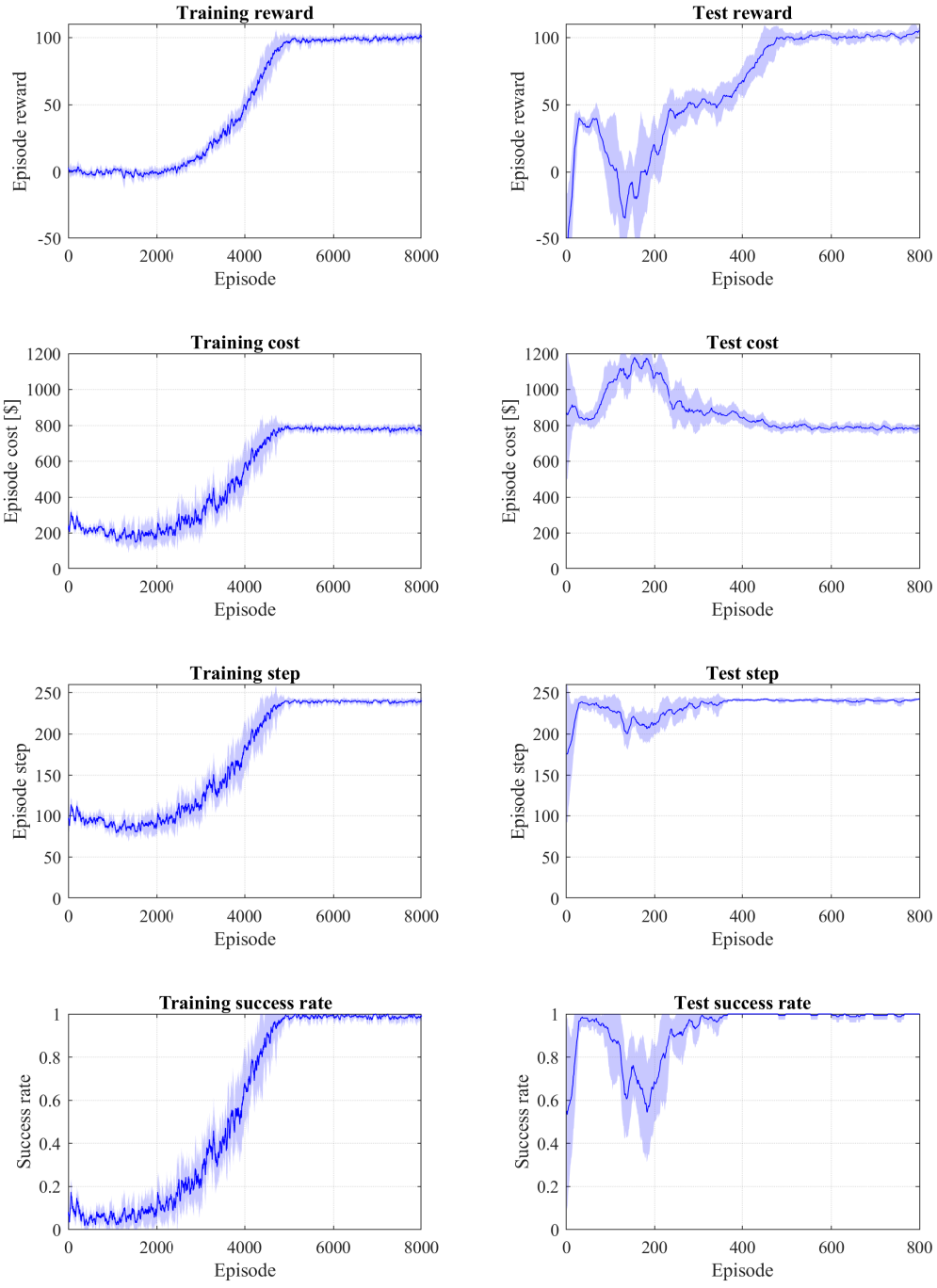


Figure 6.8: Double DQN agent training and testing with Huber loss function. The deep blue lines are moving average values across 10 instances running with different random seeds. The light blue shadows are the confidence bounds calculated by mean values $\pm$ standard deviations across the 10 instances.

6.4.3.2 Huber loss

In contrast, with the Huber loss function, the training processes of the two agents are much more stable (Figures 6.7 and 6.8). Note that both agents are trained across 8000 training episodes. The Double DQN training is more consistent across the 10 running instances with different random seeds. In test mode, the moving average line of the Double DQN strategy converged to a value of around \$780 (6.6b). Similar performance has been observed for the DQN agent (6.5b).

6.5 Results

The results presented in this section were acquired with the Double DQN agent trained by the Adam optimiser with Huber loss function. The training and validation dataset and sample voyages are identical to those used in Chapter 5. Table 6.2 compares the average voyage cost of the strategies generated by Double DQN, discrete Double Q-learning and off-line DDP (with a SOC resolution of 0.0125). The strategy generated by the Double DQN achieves average costs of \$782.5 and \$768.9 for the training and validation voyages respectively (as mentioned in Chapter 5). The off-line DDP strategy average voyage cost is 94.6% and 94.3% of those of the Double DQN strategy, for the training and validation datasets respectively. It is worth mentioning that the DDP strategy is acquired for each voyage independently by proving complete power profiles before solving, representing the best that could theoretically be achieved but requires pre-existing knowledge of power profiles. Therefore, the DDP strategy can only be used as a benchmark to assess other on-line EMS performance. Compared to the Double Q strategy in discrete state space, the Double DQN strategy further reduces the average voyage costs by approximately 6.0% with continuous state space. Note that the Double Q strategy is obtained with a SOC resolution 0.05, while it is continuous for the Double DQN strategy.

The computation time required by the Double DQN agent to generate a strategy is approximately 27 min using a single thread of an Intel Xeon E5-2683 V3

Table 6.2: Double DQN, Double Q and DDP strategy average voyage costs comparison.

	DDP_1 [$]	Double Q [$]	Double DQN [$]	$\frac{DDP_1}{\text{Double Q}}$ [%]	$\frac{DDP_1}{\text{Double DQN}}$ [%]
SOC resolution	0.0125	0.05	Continuous	-	-
Training voyages	740.0	831.8	782.5	89.0	94.6
Validation voyages	724.9	815.0	768.9	88.9	94.3

processor (18 min on an Intel i7-4790 processor). The Double Q agent requires 288 min to generate a strategy using a single thread of an Intel i7-4790 processor. Compared to the Double Q agent, the Double DQN agent managed to achieve a 6.0% improvement in cost performance with 93.8% less computational resource required.

To examine the strategy performance generated by the Double DQN and Double Q agents in detail, the power distributions between the power sources for 6 sample voyages (3 training and 3 validation voyages, see Table 5.6) are discussed in this section.

6.5.1 Training voyages

6.5.1.1 Training sample 1 with low power demand

Figure 6.9 presents the Double Q (Figure 6.9a) and Double DQN (Figure 6.9b) strategies for training sample voyage 1. This voyage has relatively low power demand in the training dataset. The voyage starts with a power demand around 1600 kW. The Double DQN strategy utilises the battery only and rapidly discharges the battery to a SOC of 0.35 (0–750 s). The Double Q strategy utilised both the fuel cell and battery during this period and tends to adjust the fuel cell power output frequently. After the initial 750 s and during cruising, both strategies tend to adjust fuel cell output frequently, while the power demand is relatively constant. However, the Double DQN strategy fuel cell power adjustments are less frequent compared to that of the Double Q strategy. Moreover, the Double Q strategy tends to use the battery more aggressively, i.e. the battery SOC is mostly maintained in the vicinity of 0.35, while the Double Q strategy is more conservative. At 950 s, the Double

DQN strategy discharges the battery to a SOC of 0.4, and thereafter starts to charge the battery gradually.

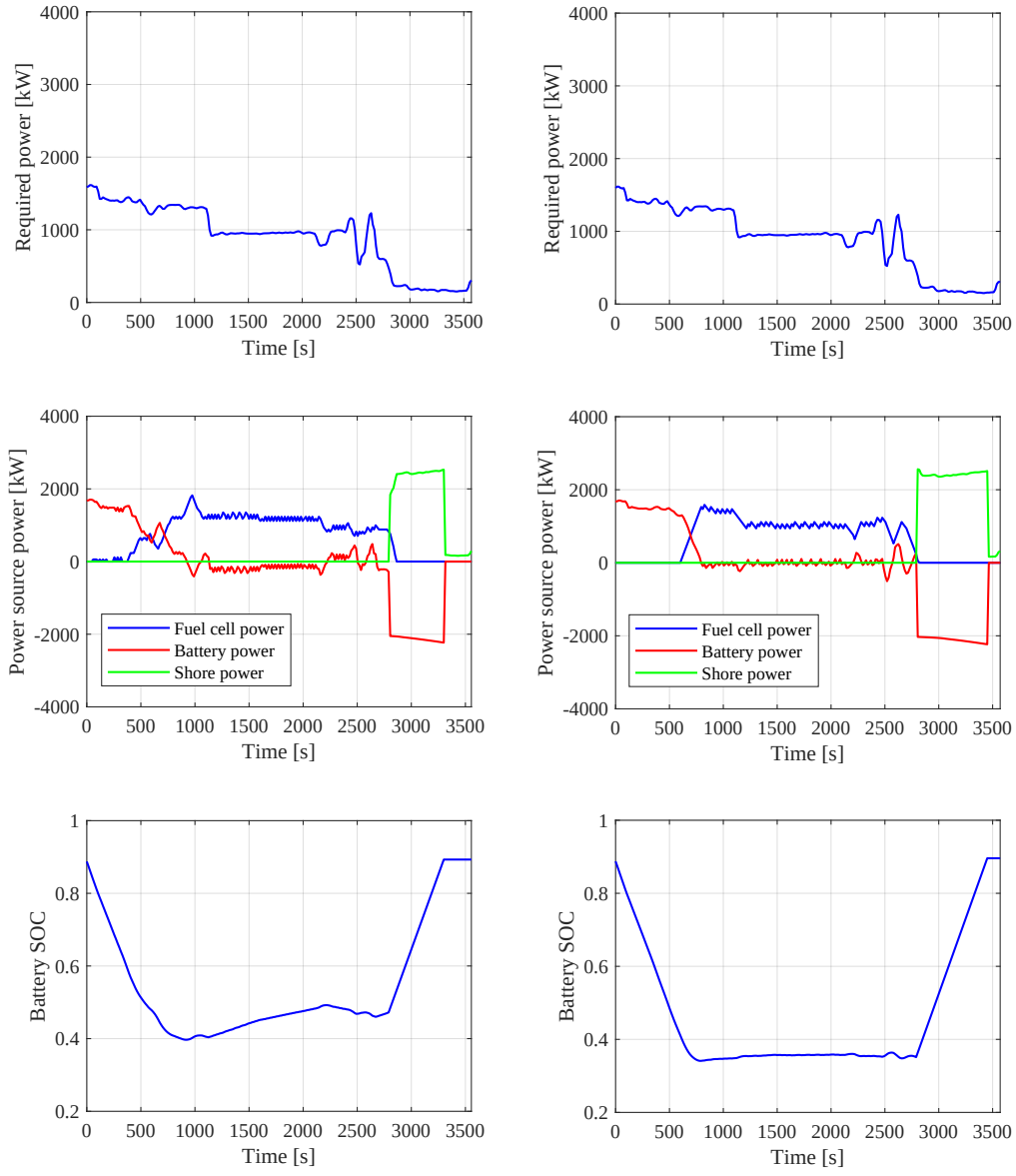


Figure 6.9: Double Q and Double DQN energy management strategies for training sample voyage 1 with low power demand: (a) Double Q strategy and (b) Double DQN strategy.

Table 6.3 details the voyage cost and GWP emission breakdowns of the sample training voyage 1. The voyage cost achieved by the Double DQN strategy is \$627.3 (91.5% of the Double Q strategy). As the Double DQN strategy tends to use the batteries more aggressively, shore-generated electricity consumed by the Double DQN strategy is 28.1% more than that consumed by the Double Q agent. Higher shore electricity usage results in lower H_2 consumption. In addition, the Double DQN strategy adjusts fuel cell power less frequently. Hence its PEMFC degradation cost is 11% lower than that of the Double Q strategy. The battery degradation costs are the same for both strategies, as a simple averaged battery degradation cost model has been adopted.

Table 6.3: Double DQN and Double Q strategy voyage cost and GWP emission breakdowns of training sample voyage 1.

	Voyage cost			Voyage GWP Emission		
	Double DQN	Double Q	$\frac{\text{Double DQN}}{\text{Double Q}}$	Double DQN	Double Q	$\frac{\text{Double DQN}}{\text{Double Q}}$
	[\$]	[\$]	[%]	[kg]	[kg]	[%]
PEMFC	224.9	252.6	89.0	-	-	-
Battery	64.3	64.3	100.0	-	-	-
Electricity	40.1	31.3	128.1	75.0	58.6	128.1
H_2	298.0	337.7	88.2	54.2	61.5	88.2
Sum	627.3	686.0	91.5	129.3	120.1	107.7

6.5.1.2 Training sample 2 with moderate power demand

Figure 6.10 shows the Double Q and Double DQN strategies for a sample voyage in the training dataset with moderate power demand. As displayed in Figure 6.10b, the battery levels off the power transients throughout the voyage, while the fuel cell power adjustments are much less frequent compared to that of the Double Q strategy (Figure 6.10a). When approaching the port (2400-2800 s), the Double DQN strategy reduces the fuel cell power in advance, while the Double Q strategy only reduces fuel cell power when shore power availability is detected.

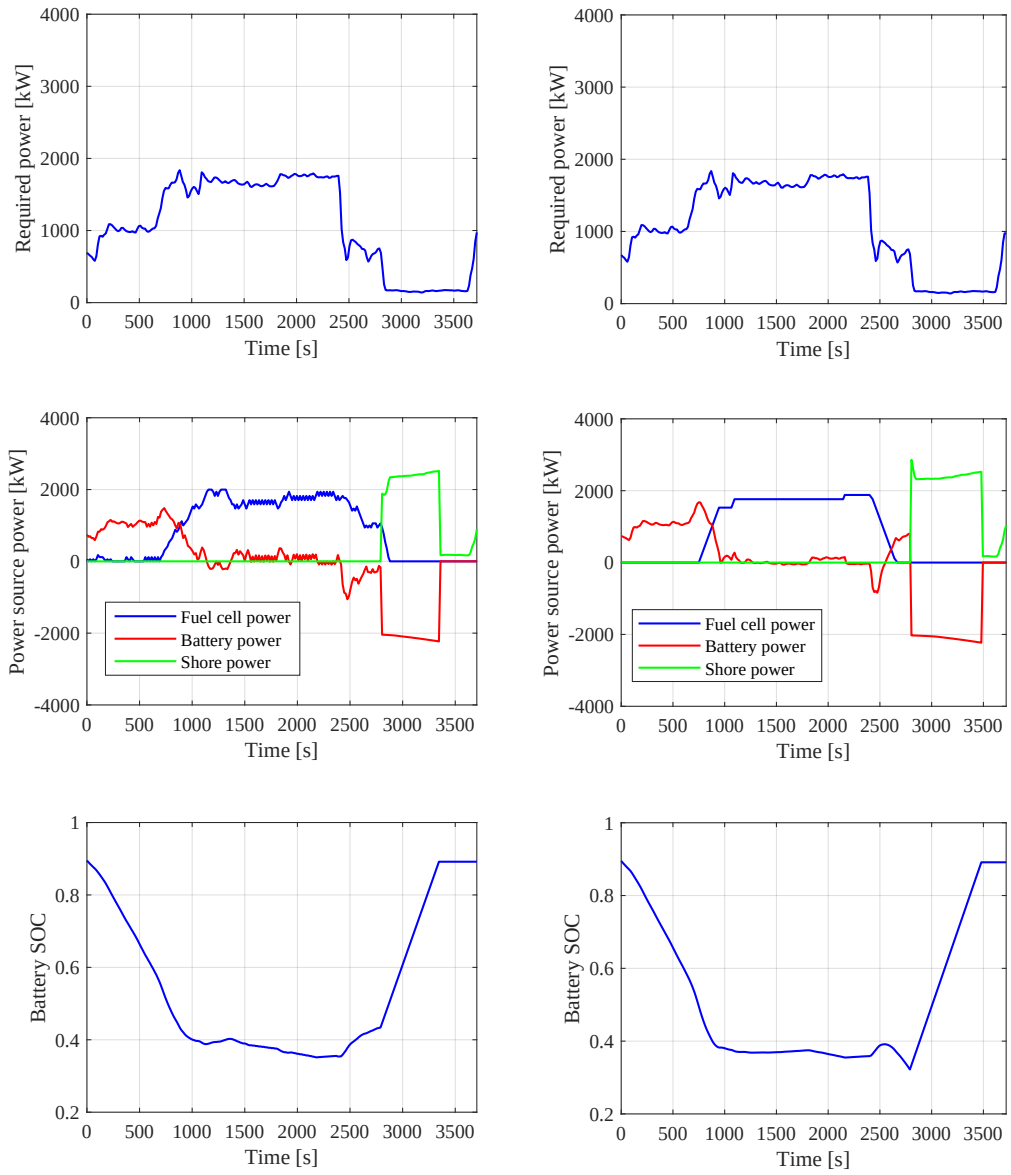


Figure 6.10: Double Q and Double DQN energy management strategies for training sample voyage 2 with moderate power demand: (a) Double Q strategy and (b) Double DQN strategy.

Table 6.4 depicts the cost and GWP emission breakdowns for training sample 2. The voyage cost achieved by the Double DQN is 91.2% of that of the Double Q strategy. Although the Double DQN strategy consumes more shore generated electricity, the degradation cost from PEMFC degradation is reduced by 16%, and the H_2 cost has also been reduced since the electricity consumption has been increased. Due to the trade-off between the voyage cost and GWP emission, the Double DQN emits 6.1% more GWP emissions. Increased shore-generated electricity consumption can lower the cost of the voyage but would increase the GWP emissions.

Table 6.4: Double DQN and Double Q strategy voyage cost and GWP emission breakdowns of training sample voyage 2.

	Voyage cost			Voyage GWP Emission		
	Double DQN [$]	Double Q [$]	$\frac{\text{Double DQN}}{\text{Double Q}}$ [%]	Double DQN [kg]	Double Q [kg]	$\frac{\text{Double DQN}}{\text{Double Q}}$ [%]
PEMFC	207.1	246.4	84.0	-	-	-
Battery	67.0	67.0	100.0	-	-	-
Electricity	43.0	34.5	124.8	80.4	64.4	124.8
H_2	408.2	447.5	91.2	74.3	81.5	91.2
Sum	725.3	795.3	91.2	154.7	145.9	106.1

6.5.1.3 Training sample 3 with high power demand

Figure 6.11 shows results for the power profile with the highest power demand in the training dataset. Such a profile is unusual. Both the Double Q and Double DQN strategies would trigger the battery over-discharge protection, i.e. when the battery SOC drops below the lower limit (0.25), the over-discharge protection function to increase fuel cell by 5% of rated fuel cell power until the SOC is restored to above 0.25. As the battery over-discharge protection is enabled in both strategies, the battery SOC trajectories are similar. The DDP strategy for this voyage presented in Chapter 5 requires the PEMFC output to be increased immediately after departure. With the battery over-discharge protection function enabled, both the Double DQN and Double Q agent can complete all voyages in the training and validation datasets.

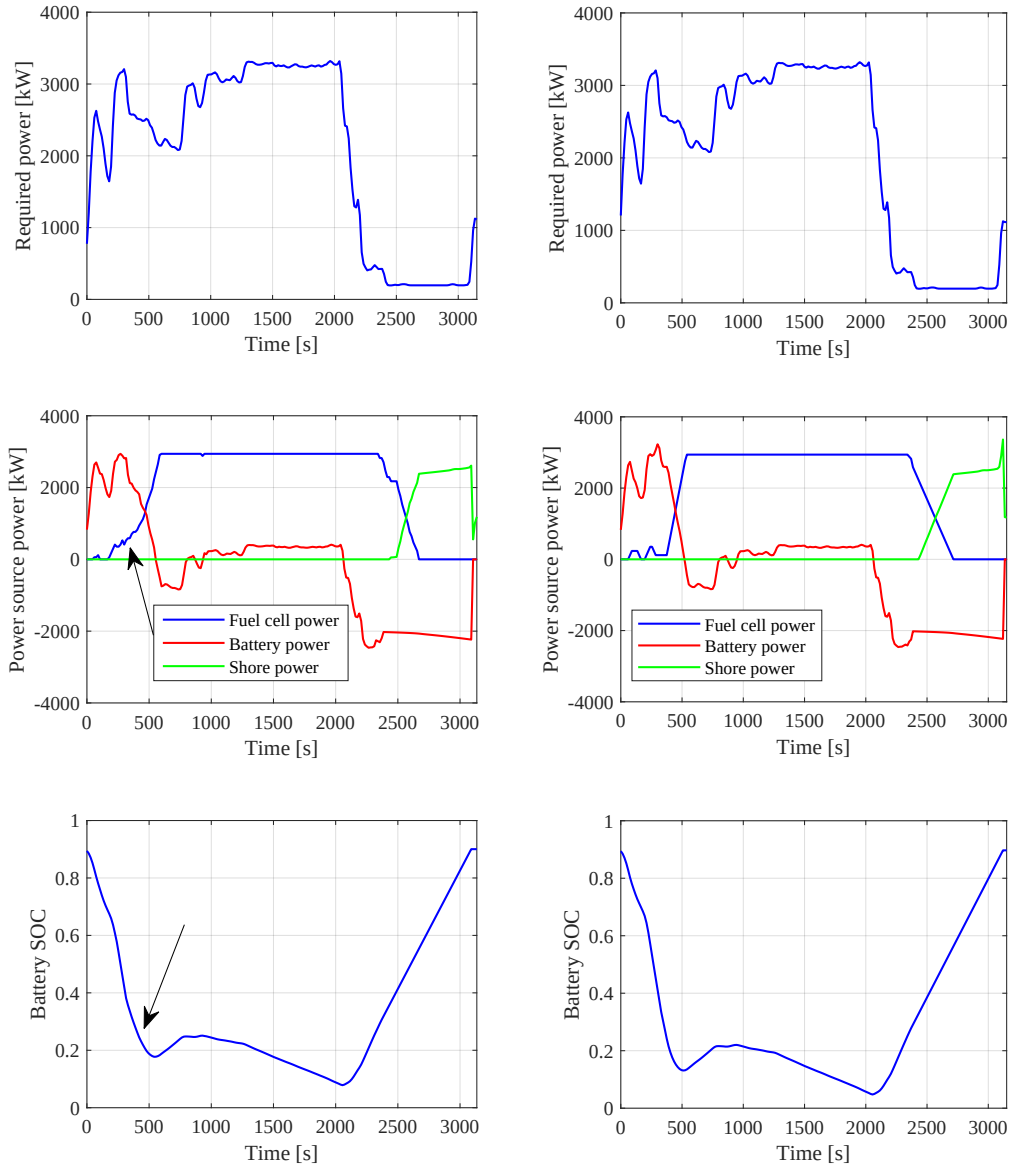


Figure 6.11: Double Q and Double DQN energy management strategies for training sample voyage 3 with extreme power demand: (a) Double Q strategy and (b) Double DQN strategy.

However, neither agent managed to develop a strategy for such an extreme power profile. Note that the Double DQN strategy is trained by 8000 episodes. Increasing the training episode number may find a strategy capable of tacking such extreme voyage profiles without triggering over-discharge protection. However, further experiments were not implemented since the overall Double DQN strategy performance is satisfactory.

Table 6.5 compares the cost and emission breakdowns of the Double DQN strategy with those of the Double Q strategy. As the battery over-discharge protection functions for most of the duration of this voyage, the cost and emission performance of the Double DQN strategy are similar to that of the Double Q strategy.

Table 6.5: Double DQN and Double Q strategy voyage cost and GWP emission breakdowns of training sample voyage 3.

	Voyage cost			Voyage GWP Emission		
	Double DQN [$]	Double Q [$]	$\frac{\text{Double DQN}}{\text{Double Q}}$ [%]	Double DQN [kg]	Double Q [kg]	$\frac{\text{Double DQN}}{\text{Double Q}}$ [%]
PEMFC	263.9	259.0	101.9	-	-	-
Battery	56.7	56.7	100.0	-	-	-
Electricity	34.8	32.9	105.8	65.1	61.5	105.8
H_2	1214.8	1199.9	101.2	221.1	218.4	101.2
Sum	1570.2	1548.5	101.4	286.2	279.9	102.3

6.5.2 Validation voyages

As the EMS is intended for use on future voyages for which, of course, there would be no predetermined data, the Double DQN strategy is applied to a set of validation voyages to examine its performance against load profiles that have not been experienced by the agent. Note that the validation voyages are not included in the training dataset.

6.5.2.1 Validation sample 1 with low power demand

Figure 6.12 compares the Double DQN strategy (Figure 6.12a) with the Double Q strategy (Figure 6.12b) for a validation sample voyage with low power demand. In

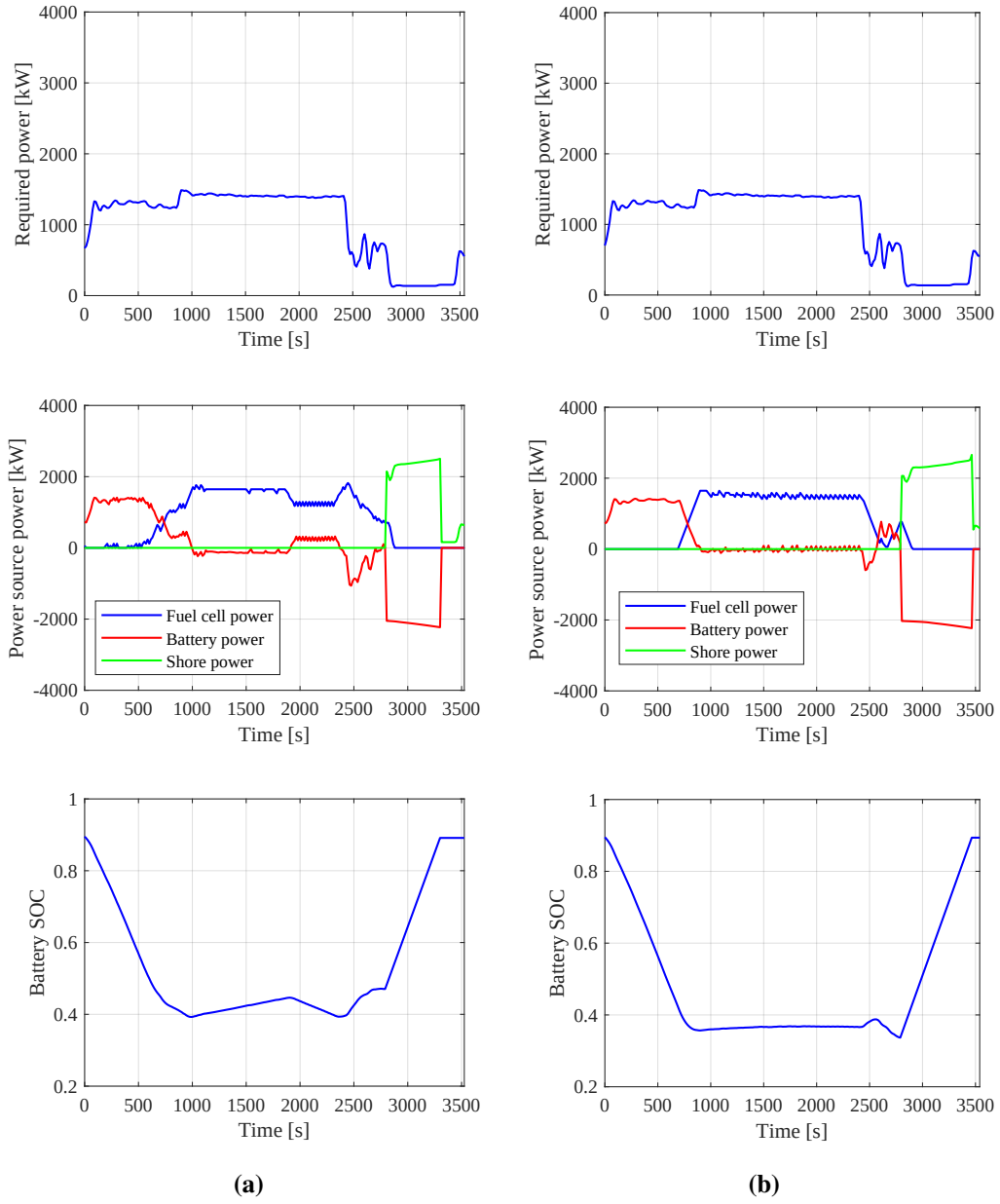


Figure 6.12: Double Q and Double DQN energy management strategies for validation sample voyage 1 with low power demand: (a) Double Q strategy and (b) Double DQN strategy.

Figure 6.12b, the Double DQN strategy delays the increase of the fuel cell power until the battery SOC has dropped to 0.36 (750 s). During cruising, the fuel cell power is maintained in a narrow band. However, the Double DQN strategy tends to adjust fuel cell power output frequently. Nevertheless, unnecessary large adjustments as in Figure 6.12a (1950–2300 s) have been avoided. Moreover, the minimum

battery SOC of the Double DQN strategy is 0.35, while it is 0.4 for the Double Q strategy.

Table 6.6 details the cost and GWP emission breakdowns for validation sample voyage 1. Similar to the voyages discussed in Sections 6.5.1.1 and 6.5.1.2, the Double DQN strategy reduces the voyage cost by 7.0%, while increasing the GWP emission by 6.9%. This is due to the conflict between voyage cost and GWP emission. PEMFC degradation cost is reduced by 8.3% by avoiding unnecessary PEMFC power adjustments and by making more use of the battery.

Table 6.6: Double DQN and Double Q strategy voyage cost and GWP emission breakdowns of validation sample voyage 1.

	Voyage cost			Voyage GWP Emission		
	Double DQN	Double Q	$\frac{\text{Double DQN}}{\text{Double Q}}$	Double DQN	Double Q	$\frac{\text{Double DQN}}{\text{Double Q}}$
	[$]	[$]	[%]	[kg]	[kg]	[%]
PEMFC	224.3	244.7	91.7	-	-	-
Battery	63.7	63.7	100.0	-	-	-
Electricity	40.0	31.2	128.2	74.8	58.3	128.2
H_2	366.5	406.8	90.1	66.7	74.1	90.1
Sum	694.5	746.5	93.0	141.5	132.4	106.9

6.5.2.2 Validation sample 2 with moderate power demand

Figure 6.13 illustrates the Double DQN strategy (Figure 6.13b) in comparison with the Double Q strategy (Figure 6.13a) for a sample voyage with moderate power demand from the validation voyage dataset. The Double DQN strategy starts ramping up the PEMFC output at 700 s. As in Figure 6.13b, the power trajectory of the PEMFC is much smoother compared to that in Figure 6.13a. The batteries absorb the small power transients by frequent charging and discharging. In addition, when approaching port (2300–2750 s), the Double DQN starts to decrease fuel cell power in advance. Such behaviour has not been observed with the Double Q strategy (e.g. Figure 6.13a).

Table 6.7 compares the voyage cost and GWP emission breakdowns of the two strategies for validation sample 2. The Double DQN strategy reduces the voyage cost by 8.4% for this voyage. As the PEMFC power adjustments are less frequent,

the PEMFC degradation cost of the Double DQN strategy is reduced by 14.7%. The Double DQN strategy increases the electricity cost by $9.6 but reduces the H_2 cost by $43.2. The Double DQN strategy tends to use more shore-generated electricity to achieve lower overall voyage cost. Such a tendency would increase the electricity cost slightly but would bring greater cost reduction from H_2 consumption. However, the Double DQN strategy increases voyage GWP emission by 7.0%.

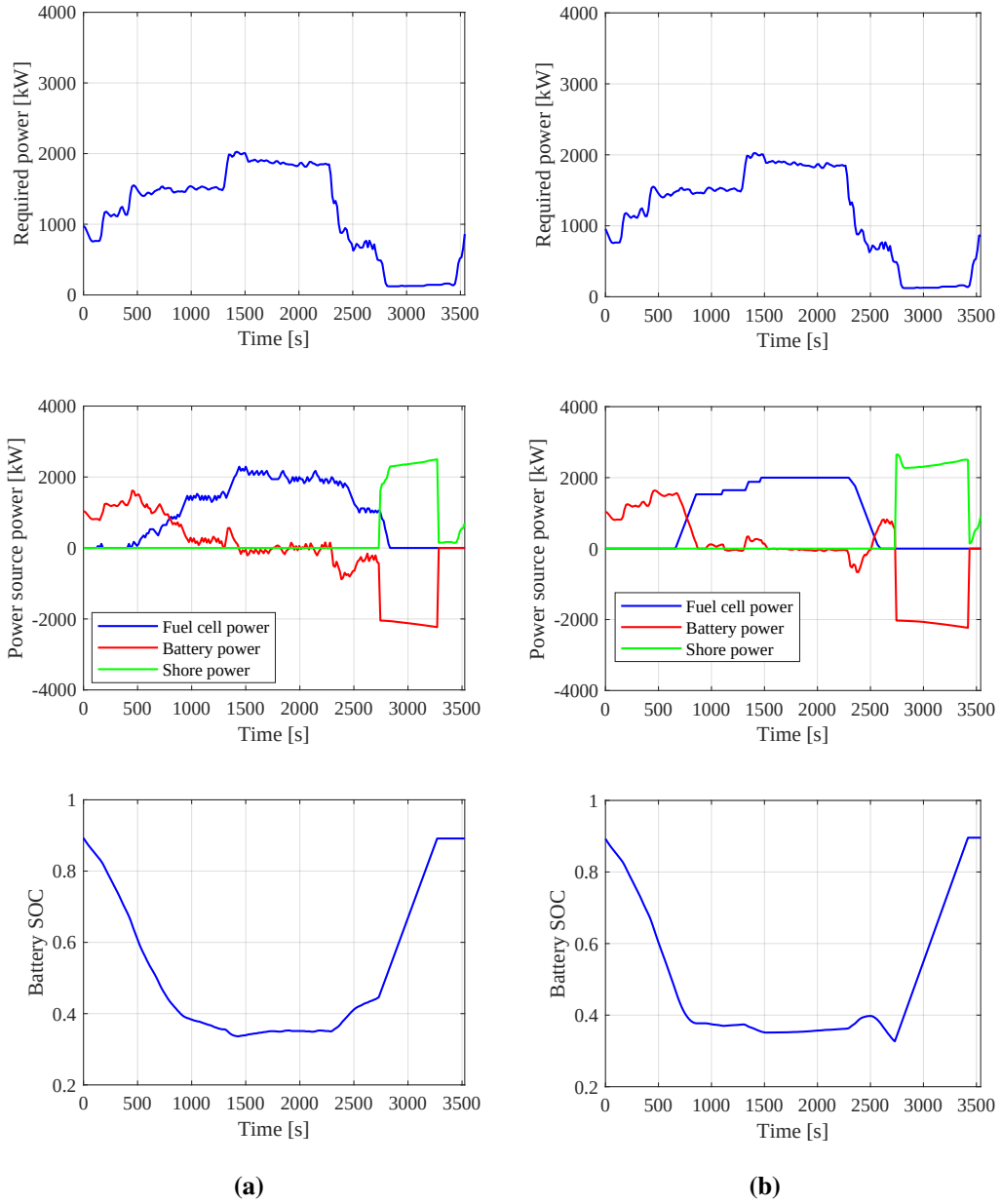


Figure 6.13: Double Q and Double DQN energy management strategies for validation sample voyage 2 with moderate power demand: (a) Double Q strategy and (b) Double DQN strategy.

Table 6.7: Double DQN and Double Q strategy voyage cost and GWP emission breakdowns of validation sample voyage 2.

	Voyage cost			Voyage GWP Emission		
	Double DQN [$]	Double Q [$]	$\frac{\text{Double DQN}}{\text{Double Q}}$ [%]	Double DQN [kg]	Double Q [kg]	$\frac{\text{Double DQN}}{\text{Double Q}}$ [%]
PEMFC	204.5	239.6	85.3	-	-	-
Battery	63.7	63.7	100.0	-	-	-
Electricity	42.0	32.4	129.9	78.6	60.5	129.9
H_2	434.4	477.6	91.0	79.1	86.9	91.0
Sum	744.7	813.2	91.6	157.7	147.4	107.0

6.5.2.3 Validation sample 3 with high power demand

Figure 6.14 details the Double DQN and Double Q strategies for validation sample 3. This sample has relatively high power demand (with an average power requirement of 1597.8 kW). Although the PEMFC power trajectories of the two strategies follow similar trends in general, the Double DQN strategy maintains the PEMFC power more consistently and only makes adjustments when significant power transients have been observed (e.g. at 1450 s). Also, the Double DQN discharges the battery to a SOC of around 0.26 (close to the lower SOC limit). In addition, the Double DQN decreases the PEMFC output in advance of reaching the port and reduces fuel cell power output to zero when shore power is available.

Detailed voyage cost and GWP emission breakdowns of the two strategies for this voyage are detailed in Table 6.8. The voyage costs of the Double DQN and Double Q strategies are $1056.7 and $1093.0, respectively, corresponding to a 3.3% voyage cost difference. The cost saving of 3.3% is lower compared to those voyages discussed in Section 6.5.2.1 and 6.5.2.2. The reason for the reduced cost saving is the PEMFC power trajectories of the two strategies follow very similar trends, while the Double DQN strategy only adjusts fuel cell power when necessary.

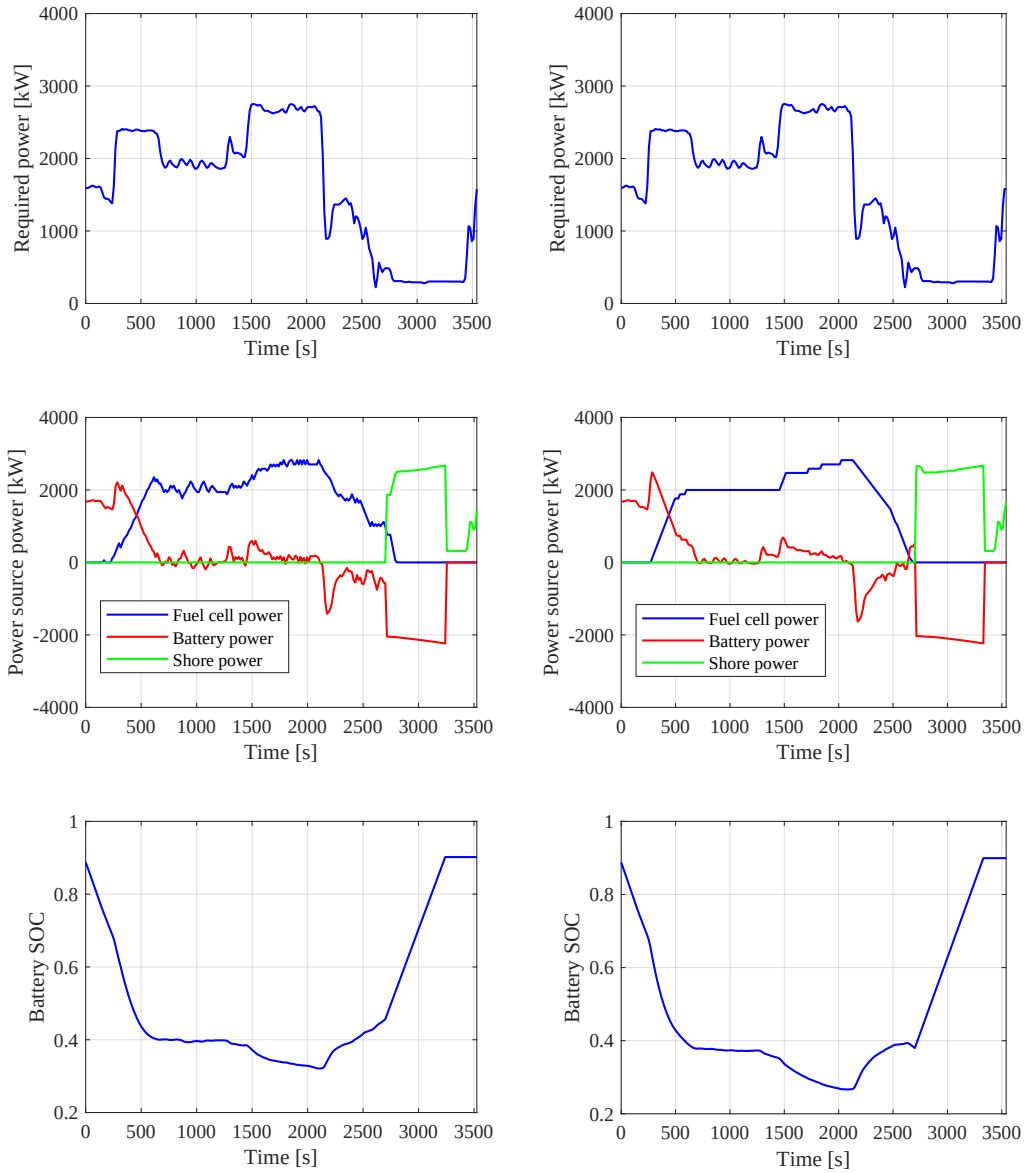


Figure 6.14: Double Q and Double DQN energy management strategies for validation sample voyage 3 with high power demand: (a) Double Q strategy and (b) Double DQN strategy.

Table 6.8: Double DQN and Double Q strategy voyage cost and GWP emission breakdowns of validation sample voyage 3.

	Voyage cost			Voyage GWP Emission		
	Double DQN [$]	Double Q [$]	$\frac{\text{Double DQN}}{\text{Double Q}}$ [%]	Double DQN [kg]	Double Q [kg]	$\frac{\text{Double DQN}}{\text{Double Q}}$ [%]
PEMFC	245.7	257.4	95.5	-	-	-
Battery	63.7	63.7	100.0	-	-	-
Electricity	43.5	36.9	117.9	81.4	69.1	117.9
H_2	703.7	734.9	95.8	128.1	133.8	95.8
Sum	1056.7	1093.0	96.7	209.5	202.8	103.3

Table 6.9 summaries the Double DQN strategy performance in comparison with that of the Double Q strategy. The Double DQN strategy further reduces the average costs for training and validation datasets by 5.9% and 5.5% respectively. The Double DQN strategy achieves average voyage costs of 105.7% and 106.0% of that of the DDP strategy in training and validation datasets respectively (see Table 6.2).

Table 6.9: Comparison of Double DQN and Double Q strategy average voyage costs and GWP emissions.

Category	Profile	Voyage cost			Voyage GWP emission		
		Double DQN [$]	Double Q [$]	$\frac{\text{Double DQN}}{\text{Double Q}}$ [%]	Double DQN [kg]	Double Q [kg]	$\frac{\text{Double DQN}}{\text{Double Q}}$ [%]
Training	Sample 1	627.3	686.0	91.5	129.3	120.1	107.7
	Sample 2	725.3	795.3	91.2	154.7	145.9	106.1
	Sample 3	1570.2	1548.2	101.4	286.2	279.9	102.3
Average	all profiles	782.5	831.8	94.1	159.7	152.0	105.0
Validation	Sample 1	694.5	746.2	93.1	141.5	132.4	106.9
	Sample 2	744.7	813.0	91.6	157.7	147.4	107.0
	Sample 3	1056.7	1092.7	96.7	209.5	202.8	103.3
Average	all profiles	768.9	813.8	94.5	157.5	149.2	105.5

6.6 Summary

The aim of this chapter was to further improve the cost-effectiveness of reinforcement learning based energy management strategies by extending discrete state space to be continuous. The environment of the agent-environment interaction framework has been improved by removing action explorations in port mode since cold ironing

has been deemed as one of the design requirements. Such an adjustment reduces training episode length.

Novel approaches using deep reinforcement learning have been proposed to solve the optimal energy management problem of the plug-in hybrid PEMFC/battery propulsion system in continuous state space. Two deep reinforcement learning algorithms, i.e. Deep Q-Network and Double Deep Q-network have been applied in light of the results in Chapter 5 suggesting overestimations of action-values could lead to divergence in agent training. In addition, two loss functions, Mean Squared Error and Huber loss functions, have been explored to deal with value overestimations in the stochastic environment.

The training processes of the two agents suggest that the Double Deep Q-Network performed slightly better than the Deep Q-Network with Mean Squared Error loss function applied. However, with the Mean Squared Error loss function, both agents have not achieved what could be considered satisfactory performance. When the Huber loss function, the Double DQN and DQN agents delivered similar performance.

The energy management strategy generated by the Double Deep Q-Network with Huber loss function was examined in detail in comparison to that generated by the Double Q agent. When compared to the Double Q energy management strategy developed in Chapter 5, a further 6% cost-performance improvement has been achieved by the Double Deep Q-Network with more than 90% computation time reduction. The cost reduction is achieved by more accurate PEMFC control and reduced H_2 consumption. However, the Double DQN based energy management strategy leads to a 5% increase in voyage GWP emission due to higher shore-generated electricity consumption.

Chapter 7

Deep reinforcement learning based continuous energy management strategies

7.1 Overview

In Chapter 6, the optimal energy management problem was solved in continuous state but discrete action spaces using DQN and Double DQN agents. Although both algorithms achieved voyage cost performance close to that of the off-line strategy solved by DDP, such algorithms are limited to small discrete action space (Mnih et al., 2015). Considering fuel cell power level is a continuous parameter, this chapter aims to extend the discrete action space to be continuous.

In addition, in marine propulsion (especially IFEP) systems, for redundancy considerations, it is usual to install multiple power sources that can be controlled independently. Therefore, this chapter will also explore the feasibility of controlling multiple fuel cell clusters using deep reinforcement learning algorithms. Instead of controlling each fuel cell stack independently, PEMFC stacks are grouped into clusters, and stacks within one cluster are controlled uniformly. Such a setting

simplifies the problem by avoiding very high dimensional action and state spaces.

Section 7.2 updates the optimal energy management problem to independent and continuous control of multiple fuel cell clusters. Section 7.3 introduces the Twin Delayed Deep Deterministic Policy Gradient (TD3) deep reinforcement learning algorithms (Fujimoto et al., 2018). Section 7.4 details the training process of the agent. Section 7.5 assesses the energy management strategy performance using voyage samples as discussed in previous chapters.

7.2 Optimal energy management problem reformulation

7.2.1 Action space

In the preceding chapters, the fuel cell stacks are controlled uniformly in one-dimensional discrete action space. In this chapter, the action space is extended to be multi-dimensional and continuous to control multiple PEMFC clusters concurrently. For the energy management problem with m PEMFC clusters, the action space is defined as:

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{m-1} \\ a_m \end{bmatrix} \tag{7.1}$$

where $a_k \in [a_{M-}, a_{M+}]$ is the $k-th$ ($k = 1, 2, \ldots, m-1, m$) PEMFC cluster per unit power adjustment. Note that $a_{M-} = -0.04$, $a_{M+} = +0.04$ are maximum allowed per unit power decreasing and increasing limits, respectively. Note that all the fuel cell stacks are controlled uniformly when $m = 1$.

7.2.2 State space

It is assumed that all PEMFC clusters are assigned with an equal PEMFC power, and the rated cluster power is $P_c = \frac{P_{fc,rated}}{m}$, where $P_{fc,rated}$ is the total installed fuel cell power (see Chapter 4). As each of the fuel cell clusters has its own power state, therefore the fuel cell state is extended to:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{m-1} \\ x_m \end{bmatrix} \tag{7.2}$$

where x_k is the $k-th$ fuel cell per unit power. The definitions of shore power availability spA, battery state of charge SOC, and power demand p_{dem} remain unchanged. Consequently, the new state space of the multi-stack energy management problem is:

$$s = \begin{bmatrix} x \\ SOC \\ spA \\ p_{dem} \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{m-1} \\ x_m \\ SOC \\ spA \\ p_{dem} \end{bmatrix} \tag{7.3}$$

With given action a, from current state s to next state s', the state transition is calculated by:

$$s' = \begin{bmatrix} x_1 + a_1 \\ x_2 + a_2 \\ \vdots \\ x_{m-1} + a_{m-1} \\ x_m + a_m \\ SOC' \\ spA' \\ p'_{dem} \end{bmatrix} \tag{7.4}$$

where the next battery SOC SOC' is calculated by the system model (see Chapter 4 and Figure 5.2), next shore power availability spA' and p'_{dem} are determined by the power profile. Note the fuel cell power override function still applies (see Eq. 6.12). It is worth mentioning that when calculating battery power, the total fuel cell power after the power converters, i.e. P_1 (see Eq. 4.1), is updated to:

$$P_1 = \sum_{k=1}^{m} P_c x_k \eta_{1,k} \tag{7.5}$$

where $\eta_{1,k}$ is uni-directional power converter efficiency (see Figure 4.3) of $k-th$ PEMFC cluster.

7.2.3 Reward function

Based on the reward function described in Chapter 6, as multiple PEMFC clusters are configured, the reward function is updated to:

$$r_{t+1} = \begin{cases} -1, & \text{spA=0, if } s_{t+1} \text{ is infeasible} \\ -1, & \text{spA=0, if any } x_i + a_{i,t} \notin [0,1] \\ \tanh\left(\dfrac{1}{cost_{t+1}}\right), & \text{spA=0, else} \\ \displaystyle\sum_{n=t+1}^{K} \tanh\left(\dfrac{1}{cost_n}\right), & \text{spA=1} \end{cases} \tag{7.6}$$

where, when shore power is available ($spA = 1$), the environment returns a summed reward of all the costs incurred in port mode of current episode. In sailing mode, i.e. $spA = 0$, the environment returns -1 if s_{t+1} is infeasible or one or more PEMFC cluster control actions are overridden. Note that the degradation and H_2 fuel costs for each PEMFC cluster are calculated independently using the scalable PEMFC model, then summed to further calculate the cost $cost_{t+1}$ incurred in one time step.

7.3 Twin Delayed Deep Deterministic Policy Gradient

Although Q-network based DQN and Double DQN agents achieved satisfactory performance in Chapter 6, these algorithms cannot be applied to problems with continuous or large action space due to maximisation operating in selecting optimal actions:

$$a_t \leftarrow \arg\max_a \left(Q\left(s_{t+1}, a; \theta\right)\right) \tag{7.7}$$

i.e. finding the greedy policy at every time step. Silver et al. (2014) proposed an actor-critic (see Figure 7.1) based Deterministic Policy Gradient (DPG). In DPG, the actor π_ϕ (i.e. the policy) parametrised by ϕ is updated by taking the gradient of the expected return $J(\phi)$ (Silver et al., 2014; Fujimoto et al., 2018):

$$\nabla_\phi J(\phi) = \mathbb{E}_{s\sim\rho^\pi}\left[\nabla_a Q^\pi\left(s,a\right)|_{a=\pi(s)} \nabla_\phi \pi_\phi(s)\right] \tag{7.8}$$

where $Q^\pi\left(s,a\right) = \mathbb{E}_{s\sim\rho_\pi, a\sim\pi}\left[R_t|s,a\right]$ is the expected return of performing action a in state s following policy π. Note that ρ^π is the state distribution which also depends

on the policy parameters. $R_t = \sum_{i=t}^{T} \gamma^{i-t} r(s_i, a_i)$ with a discount factor γ is the discounted sum of rewards which the agent aims to maximise. Such an architecture is applicable to continuous control problems (Silver et al., 2014).

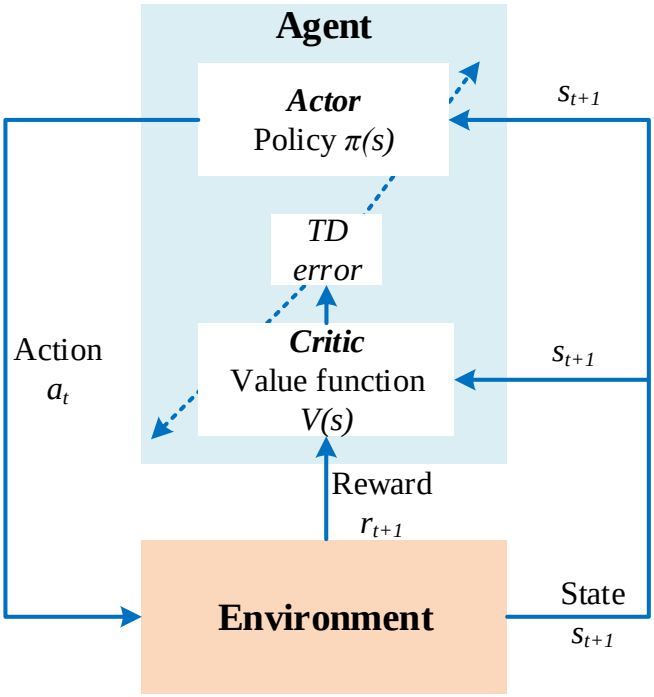


Figure 7.1: Actor-critic architecture (Sutton and Barto, 2018).

Later work of (Lillicrap et al., 2015) developed a Deep Deterministic Policy Gradient (DDPG) based on DPG. In DDPG, exploration noise $\mathcal{N}_t$ is added to the actor policy $\pi_\phi(s_t)$:

$$a_t = \pi_\phi(s_t) + \mathcal{N}_t \tag{7.9}$$

The concept of experience replays and target networks have also been included in DDPG to break sample correlations and to improve training stability. The critic, i.e. the Q-function $Q_\theta(s,a)$ parametrised by θ, is updated by minimising the loss:

$$L = \frac{1}{D} \sum_{j=1}^{D} \left(y_j - Q_\theta(s_j, a_j) \right)^2 \tag{7.10}$$

The policy π_ϕ is updated using the sampled policy gradient (see Eq. 7.8). In every

training step, both the actor and critic target networks are soft-updated with a soft-update rate τ.

Although DDPG can achieve satisfactory performance in some continuous control tasks, the overestimation bias can be problematic (Fujimoto et al., 2018). Fujimoto et al. (2018) proposed the Twin Delayed Deep Deterministic Policy Gradient (TD3) (Algorithm 7.1) addressing the function approximation errors and overestimation bias in actor-critic methods. The results in Chapters 5 and 6 have highlighted the problem of overestimations in the stochastic environment. Consequently, a novel approach using the TD3 algorithm is proposed in this chapter to solve the optimal energy management problem of the plug-in hybrid PEMFC/battery system.

Algorithm 7.1 Twin delayed policy deterministic policy gradients (TD3) agent (Fujimoto et al., 2018).

Initialise replay memory D to capacity M
Initialise critic networks Q_{θ_1}, Q_{θ_2}, and actor network π_ϕ with random parameters θ_1, θ_2, ϕ
Initialise target networks $\theta_1' \leftarrow \theta_1$, $\theta_2' \leftarrow \theta_2$, $\phi' \leftarrow \phi$
while $n < N_{max}$ **do**
    Initialise initial state s_1
    for $t = 1 : T$ **do**
        Select action a_t with exploration noise $a_t \sim \pi_\phi(s_t) + \varepsilon$, $\varepsilon \sim \mathcal{N}(0, \sigma)$
        Take action a_t, observe r_{t+1}, s_{t+1} and $terminationflag$
        Store transition $(s_t, a_t, r_{t+1}, s_{t+1})$ in replay memory
        Every Z steps sample random mini-batch of transitions (s, a, r, s') from mini-batch
        $\tilde{a} \leftarrow \pi_{\phi'}(s') + \tilde{\varepsilon}$, $\tilde{\varepsilon} \sim \text{clip}(\mathcal{N}(0, \tilde{\sigma}), -c, c)$
        Set $y = \begin{cases} r, & \text{if episode terminates} \\ r + \gamma \min_{i=1,2} Q'_{\theta_i}(s', \tilde{a}), & \text{otherwise} \end{cases}$
        Update critics $\theta_i \leftarrow \arg\min_{\theta_i} L_{\theta_i}$
        if $t \bmod d = 0$ **then**
            Update ϕ by the deterministic policy gradient:
            $\nabla_\phi J(\phi) = D^{-1} \sum \nabla_a Q_{\theta_1}(s, a)|_{a=\pi_\phi(s)} \nabla_\phi \pi_\phi(s)$
            Soft-update target network:
            $\theta_i' \leftarrow \tau\theta_i + (1-\tau)\theta_i'$
            $\phi' \leftarrow \tau\phi_i + (1-\tau)\phi'$
        end if
        Terminate if $terminationflag$
    end for
end while

In TD3, there are two critic networks, i.e. Q_{θ_1} and Q_{θ_2} and one actor network π_ϕ. The subscripts of θ_1, θ_2 and ϕ denote the neural network parameters. Correspondingly, there are two critic target networks Q'_{θ_1} and Q'_{θ_2} and one actor target network π'_ϕ. The superscript $'$ denotes target network. Three key improvements have been made to DDPG (Fujimoto et al., 2018).

The first improvement is clipped Double Q-learning for the actor-critic. When calculating the target value, the minimum value between the two critic estimates is selected:

$$y \leftarrow r + \gamma \min_{i=1,2} Q'_{\theta_i}\left(s', \tilde{a}\right) \tag{7.11}$$

so that the less biased Q-value estimate is used, which is similar to the concept of Double Q-learning (van Hasselt, 2010). Note that $\tilde{a}$ is given by the target actor network with a small amount of random noise added to the target action to smooth the value estimate by bootstrapping off of similar state-action value estimates (second improvement) (Fujimoto et al., 2018):

$$\tilde{a} \leftarrow \pi_{\phi'}\left(s'\right) + \tilde{\varepsilon}, \tilde{\varepsilon} \sim \text{clip}\left(\mathcal{N}\left(0, \tilde{\sigma}\right), -c, c\right) \tag{7.12}$$

where $\tilde{\varepsilon} \in (-c, c)$ is the added noise clipped from a Gaussian distribution $\mathcal{N}\left(0, \tilde{\sigma}\right)$ with a 0 mean value and a standard deviation of $\tilde{\sigma}$ such that the target is maintained close to the original action. Moreover, the actor π_ϕ parametrised by ϕ is updated less frequently than the critics, i.e. ϕ is updated every d critic updates such that accumulated errors can be reduced.

It is worth noting that the Huber loss function is adopted to update the critic networks, which is different from the Mean Squared Error used in (Fujimoto et al., 2018). This adjustment is made to achieve more stable agent training as the Mean Squared Error loss function in Chapter 6 led to diverged training processes. Consequently, the loss function for the critic is:

$$L_i\left(\theta\right) = \frac{1}{D}\sum_{j=1}^{D} \sigma_{i,j} \tag{7.13}$$

where:

$$\sigma_{i,j} = \begin{cases} \frac{1}{2}\delta_{i,j}^2, & \text{if } \left|\delta_{i,j}\right| < 1 \\ \left|\delta_{i,j}\right| - \frac{1}{2}, & \text{otherwise} \end{cases} \tag{7.14}$$

where $\delta_{i,j} = y_{i,j} - Q_{\theta_i}(s,a)$, $i = 1,2$, $j = 1,2,\ldots,D$. δ denotes temporal difference (see Chapter 6). i denotes the $i-th$ critic; j denotes the $j-th$ sample in the mini-batch with capacity D.

7.4 Agent training

The aims of this chapter are twofold: (1) uniform fuel cell control in a continuous action space and (2) multi-cluster fuel cell control in a continuous action space. Consequently, the agent is trained separately for each of the two scenarios:

- Uniform fuel cell control, $m = 1$, i.e. the fuel cells are controlled uniformly as in Chapters 5 and 6 but in continuous action space.
- Multi-cluster fuel cell control, $m = 4$, i.e. the fuel cells are distributed to multiple clusters and are controlled separately. The cluster number m is set to 4. This results in the number of total power sources is being five (four PEMFC clusters and one battery). This setting is to bring the system close to the original IFEP configuration with 5 diesel generators.

The agents were trained on a workstation with two Intel Xeon E5-2683 V3 processors (28 cores in total) running on Windows 10. The environment and the agent were coded in Python. The agent's neural networks were built and trained with PyTorch v1.20.

As the learning curve of the agent can be influenced by the random seeds (determining the appearance order of training power profiles to the agent) of the environment (Henderson et al., 2018) each agent was trained with 28 different random seeds for reproducibility. The agent policy performance was assessed by calculating the average values and standard deviations across all converged instances. Note

that as the neural networks are relatively small, only one CPU thread is assigned to each running instance to avoid training speed degradation due to unnecessary parallelisation.

The actual strategy performance was periodically tested (every 100 training episodes) with 10 random training voyages during training. Note that in test mode, no exploration noise was added; and battery over-discharge protection was enabled (disabled in training mode). Once the training of all the 28 instances was completed, the agent with the lowest episode cost was chosen to generate detailed energy management strategy results.

7.4.1 Neural network settings

Figure 7.2 illustrates the settings for the actor (Figure 7.2a) and critic (Figure 7.2b). The actor observes state inputs s and chooses action a. As in Figure 7.2a, the inputs to the actor are state vectors. Two fully-connected hidden layers forward propagate the state inputs followed by a fully-connected output layer. The input layer and the two hidden layers are activated by ReLU. The output layer is activated by a hyperbolic tangent function (tanh) then multiplied by a_M to match the action limits of $[-0.04, 0.04]$. As in Figure 7.2b, the critic receives the state and action inputs, and outputs the Q-value (see Section 7.3). Note that no activation function is applied to the output layer to allow free value estimates. Also note that, for uniform fuel cell control, the state and action space dimensions are 4 and 1, respectively; for the 4-cluster control, the state and action space dimensions are 7 and 4, respectively.

7.4.2 Hyperparameters

As the state and action space dimensions differ in the two scenarios to be investigated, the hyperparameters were tuned separately for the scenarios. Table 7.1 details the hyperparameter settings for the uniform PEMFC control energy management strategy. The agent is trained every Z time steps ($Z = 32$) instead of every time step to increase explorations. Note that the TD3 parameters σ, $\tilde{\sigma}$, c and d are adopted from (Fujimoto et al., 2018). The mini-batch size D is 32. The Adam

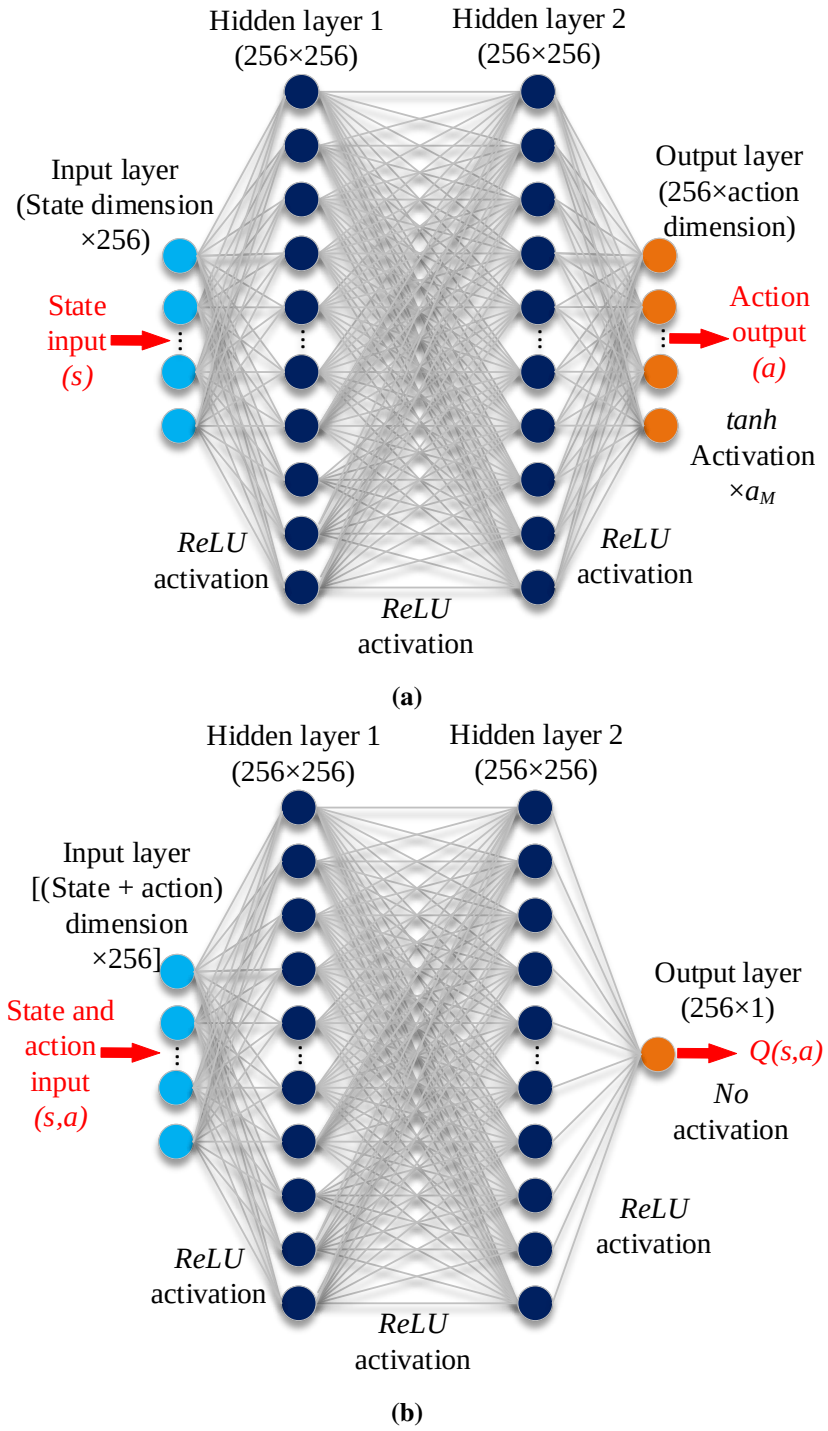


Figure 7.2: Neural network settings of TD3. (a) Actor network setting and (b) Critic network setting.

optimiser exponential decay rates for the first and second moment estimates adopts the values as suggested in (Kingma and Ba, 2014). Other parameters are tuned by trial and error to gain satisfactory performance.

Table 7.1: Hyperparameters for PEMFC uniform control.

Parameter	Hyperparameter	Value
B	Mini-batch size	32
M	Experience memory size	1000000
τ	Target network update weight	0.001
γ	Discount factor	1
Z	Update period	32
α	Learning rate of Adam optimiser	0.0001
β_1	Exponential decay rate for the first moment estimates of Adam optimiser	0.9
β_2	Exponential decay rate for the second moment estimates of Adam optimiser	0.999
σ	Standard deviation of exploration noise	0.1
$\tilde{\sigma}$	Standard deviation of policy noise	0.2
c	Policy noise clip factor	0.5
d	Delayed policy update step	2

The hyperparameters adopted for the 4-cluster energy management problem are listed in Table 7.2. The Adam optimiser setting follows the values suggested in (Kingma and Ba, 2014). The TD3 parameters σ, $\tilde{\sigma}$, c and d are adopted from (Fujimoto et al., 2018). The discount rate is $\gamma = 1$ as the energy management problem is an episodic task with a limited number of time steps. This setting is identical to that used in Chapters 5 and 6. Other parameters are tuned by trial and error to achieve satisfactory performance.

Table 7.2: Hyperparameters for PEMFC 4-cluster control.

Parameter	Hyperparameter	Value
B	Mini-batch size	128
M	Experience memory size	1000000
τ	Target network update weight	0.005
γ	Discount factor	1
Z	Update period	8
α	Learning rate of Adam optimiser	0.001
β_1	Exponential decay rate for the first moment estimates of Adam optimiser	0.9
β_2	Exponential decay rate for the second moment estimates of Adam optimiser	0.999
σ	Standard deviation of exploration noise	0.1
$\tilde{\sigma}$	Standard deviation of policy noise	0.2
c	Policy noise clip factor	0.5
d	Delayed policy update step	2

7.4.3 Training

7.4.3.1 Uniform fuel cell control

Figure 7.3 details the training process of the uniform fuel cell control EMS. The training was implemented on the workstation in 28 instances with different random seeds (each processor core was assigned a running instance). The policy was evaluated every 100 training episodes in test mode by randomly sampling 10 training voyages. The total number of training episodes is 8000. 7 out of the 28 instances diverged. It took 100 min for a converged instance to complete the training. The final test episode costs of the converged instances are around \$800.0. The policy can complete all training voyages with battery over-discharge protection enabled.

7.4.3.2 Multi-cluster fuel cell control

Figure 7.4 illustrates the training process for the 4-cluster fuel cell control EMS. The training was terminated at 8000 episodes. 2 out of the 28 instances diverged. It required 204 min for a converged instance to complete the training. Note the converged reward (around 74) of the 4-cluster strategy is lower than that of uniform EMS (around 100), which was mainly as result of more frequent fuel cell control action overrides occurring as the action space dimension increased. The voyage cost converged to a value slightly higher than \$800.0.

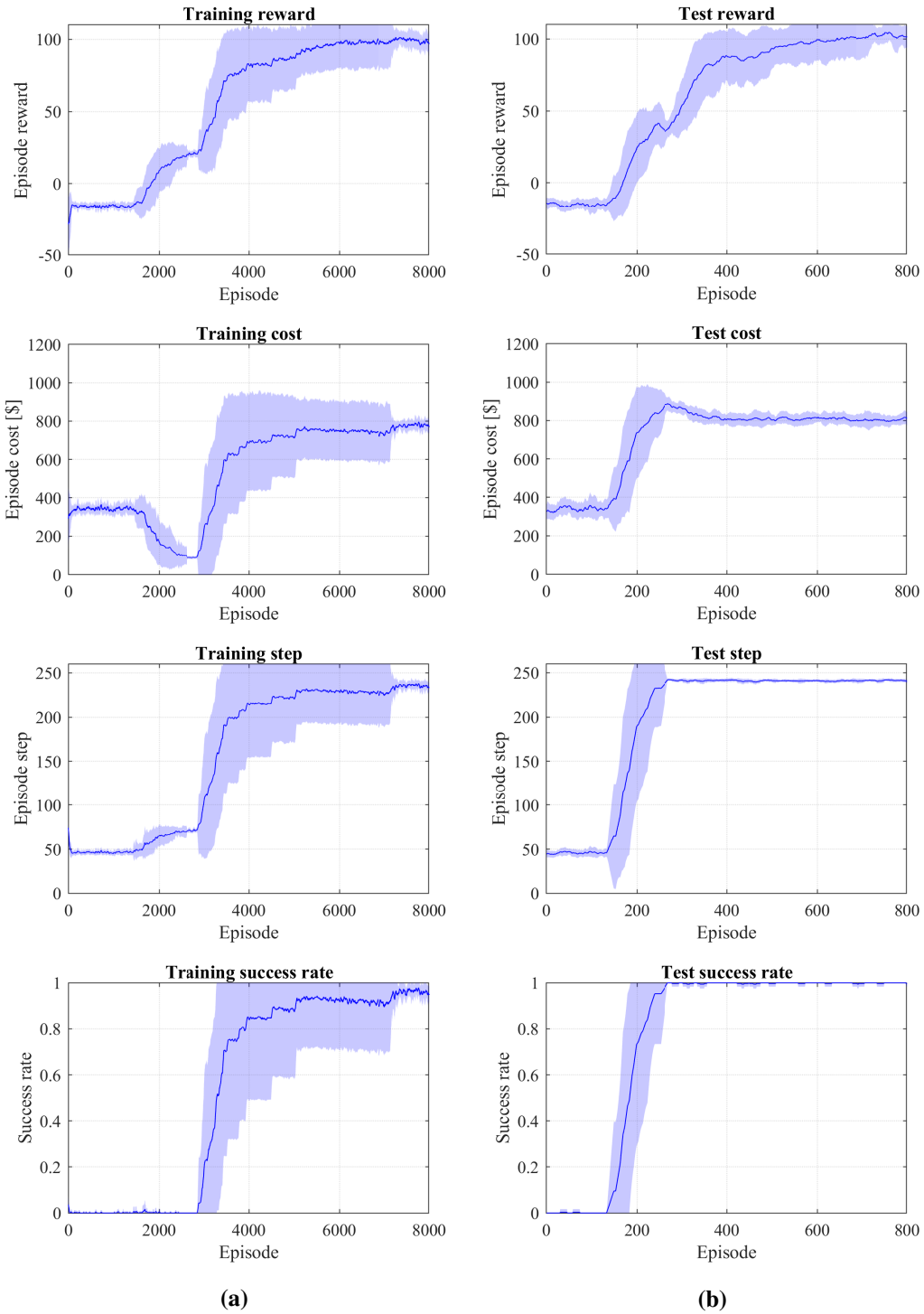


Figure 7.3: Uniform fuel cell control training and testing with Huber loss function. The deep blue lines are moving average values across 21 converged instances running with different random seeds. 7 diverged instances are not included. The light blue shadows are the confidence bounds calculated by mean values $\pm$ standard deviations across the 21 instances.

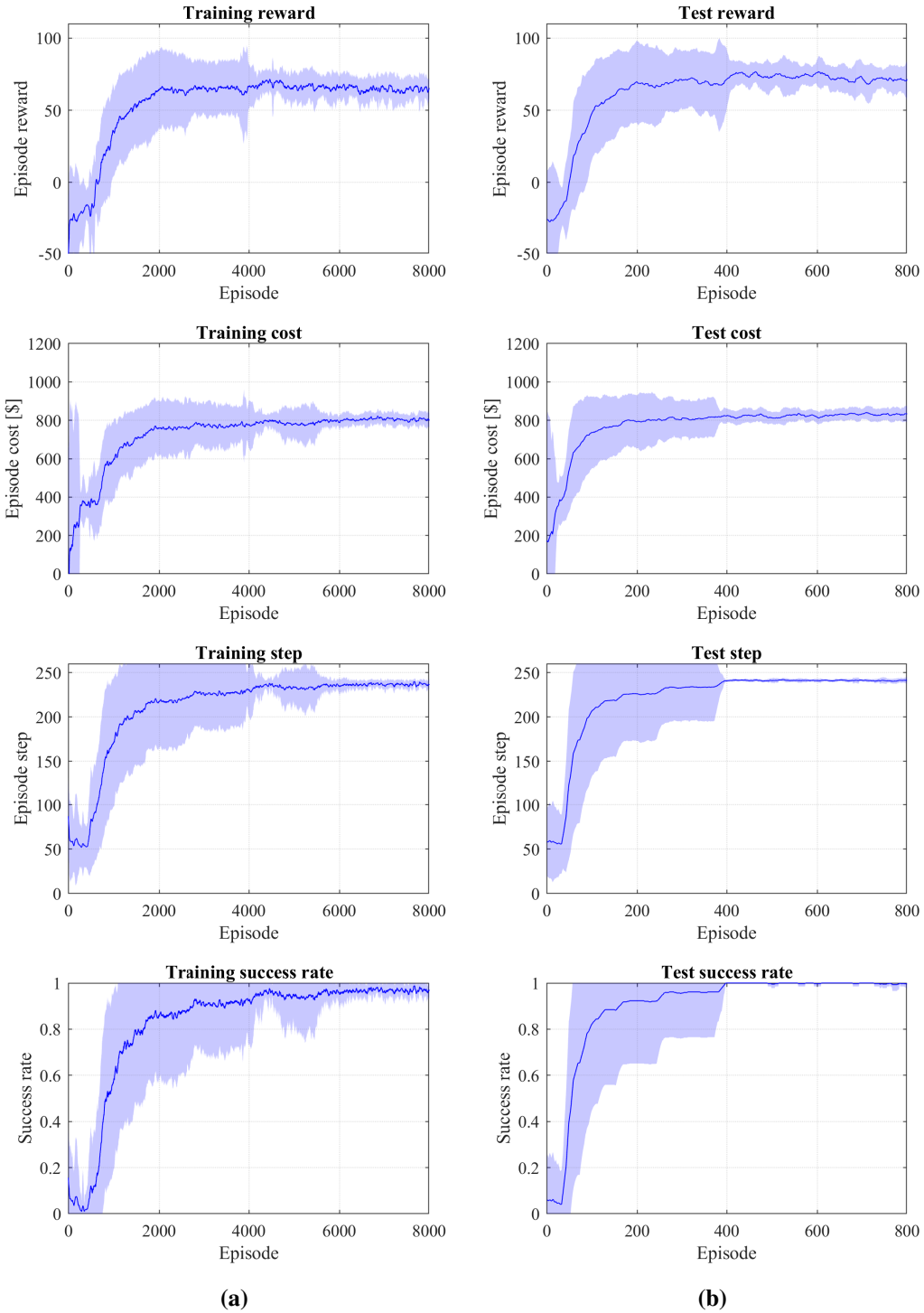


Figure 7.4: 4-cluster fuel cell control training and testing with Huber loss function. The deep blue lines are moving average values across 26 converged instances running with different random seeds. Two diverged instances are not included. The light blue shadows are the confidence bounds calculated by mean values $\pm$ standard deviations across the 26 instances.

7.5 Results

The learned strategy was first applied to the training voyages to verify the EMS cost performance without any policy noise. As the learned strategy is intended to achieve minimum voyage cost for un-predicted future voyages, the EMS was validated by applying to a set of un-predicted future voyages.

As each of the strategies generated with different random seeds vary slightly from others, for both scenarios, the strategies with minimum average voyage costs were selected to generate the detailed voyage power distributions. The training and validation voyage sets are identical, as used in Chapters 5 and 6. The power profile samples (see Table 5.6) are also identical to the ones discussed previously.

7.5.1 Training voyages

7.5.1.1 Training sample 1 with low power demand

Training voyage sample 1 is a voyage with relatively low power demand. Figure 7.5 compares the uniform TD3 (Figure 7.5b) and Double DQN (Figure 7.5a) strategies for this voyage. As in Figure 7.5b, the PEMFC power adjustments during cruising obtained from the TD3 strategy are smaller than those of the Double DQN. Although the two strategies follow similar trajectories in general, the TD3 strategy tends to discharge the battery to a lower SOC value (around 0.28). Consequently, as depicted in Table 7.3, the TD3 strategy leads to a % voyage cost reduction with lower PEMFC degradation and H_2 fuel costs, but a higher electricity cost. The GWP emission is increased by 4.1% due to an increase in electricity usage.

Figure 7.6 presents the 4-cluster energy management strategy for training sample voyage 1. After departure (0–800 s), the strategy tends to use all four clusters at low load; which differs from the Double DQN and TD3 uniform strategy. During cruising, cluster 3 is maintained with relatively lower power output, while the outputs from clusters 1 and 4 are almost identical. However, it should be noted that all cluster power outputs fluctuate in a small region during cruising. These fluctuations

Table 7.3: Comparison of TD3 and Double DQN strategy voyage costs and GWP emissions for training sample 1.

	Voyage cost			Voyage GWP Emission		
	TD3 [$]	Double DQN [$]	$\frac{\text{TD3}}{\text{Double DQN}}$ [%]	TD3 [kg]	Double DQN [kg]	$\frac{\text{TD3}}{\text{Double DQN}}$ [%]
PEMFC	222.9	224.9	99.1	-	-	-
Battery	64.3	64.3	100.0	-	-	-
Electricity	43.8	40.1	109.0	81.8	75.0	109.0
H_2	290.3	298.0	97.4	52.9	54.2	97.4
Sum	621.3	627.3	99.0	134.7	129.3	104.1

together with early PEMFC starts lead to 10.5% increase in PEMFC degradation cost (see Table 7.4). Consequently, the total voyage cost of the 4-cluster strategy is increased by 3.7%. Nevertheless, the voyage GWP emission is similar between the uniform and 4-cluster strategies.

Table 7.4: Comparison of TD3 uniform and multi-cluster strategy voyage costs and GWP emissions for training sample 1.

	Voyage cost			Voyage GWP Emission		
	TD3–multi [$]	TD3–uniform [$]	$\frac{\text{Multi}}{\text{Uniform}}$ [%]	TD3–multi [kg]	TD3–uniform [kg]	$\frac{\text{Multi}}{\text{Uniform}}$ [%]
PEMFC	246.2	222.9	110.5	-	-	-
Battery	64.3	64.3	100.0	-	-	-
Electricity	43.9	43.8	100.4	82.2	81.8	100.4
H_2	290.3	290.3	100.0	52.8	52.9	100.0
Sum	644.7	621.3	103.8	135.0	134.7	100.3

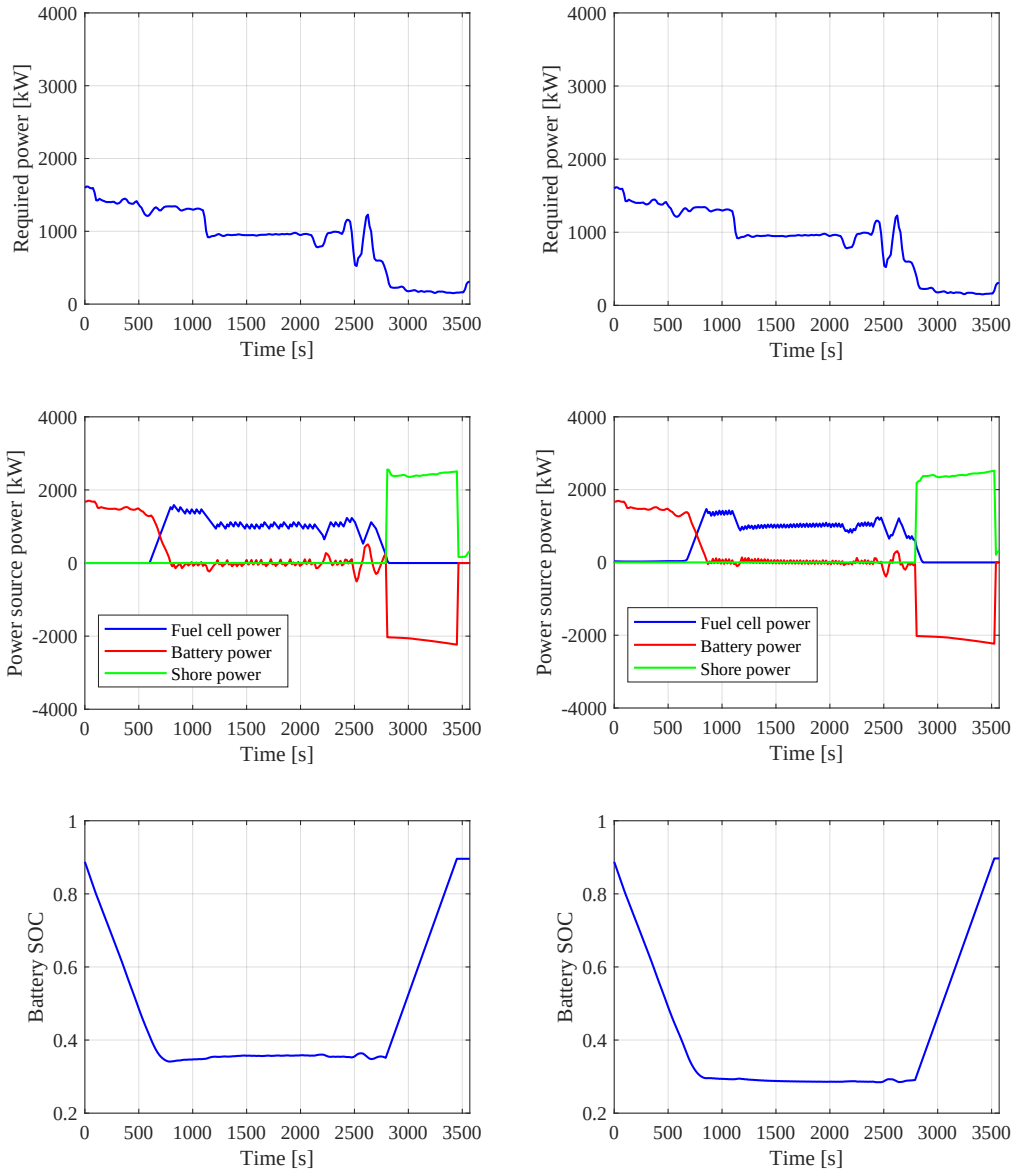


Figure 7.5: Double DQN and TD3 uniform energy management strategies for training sample voyage 1 with low power demand: (a) Double DQN strategy and (b) TD3 strategy.

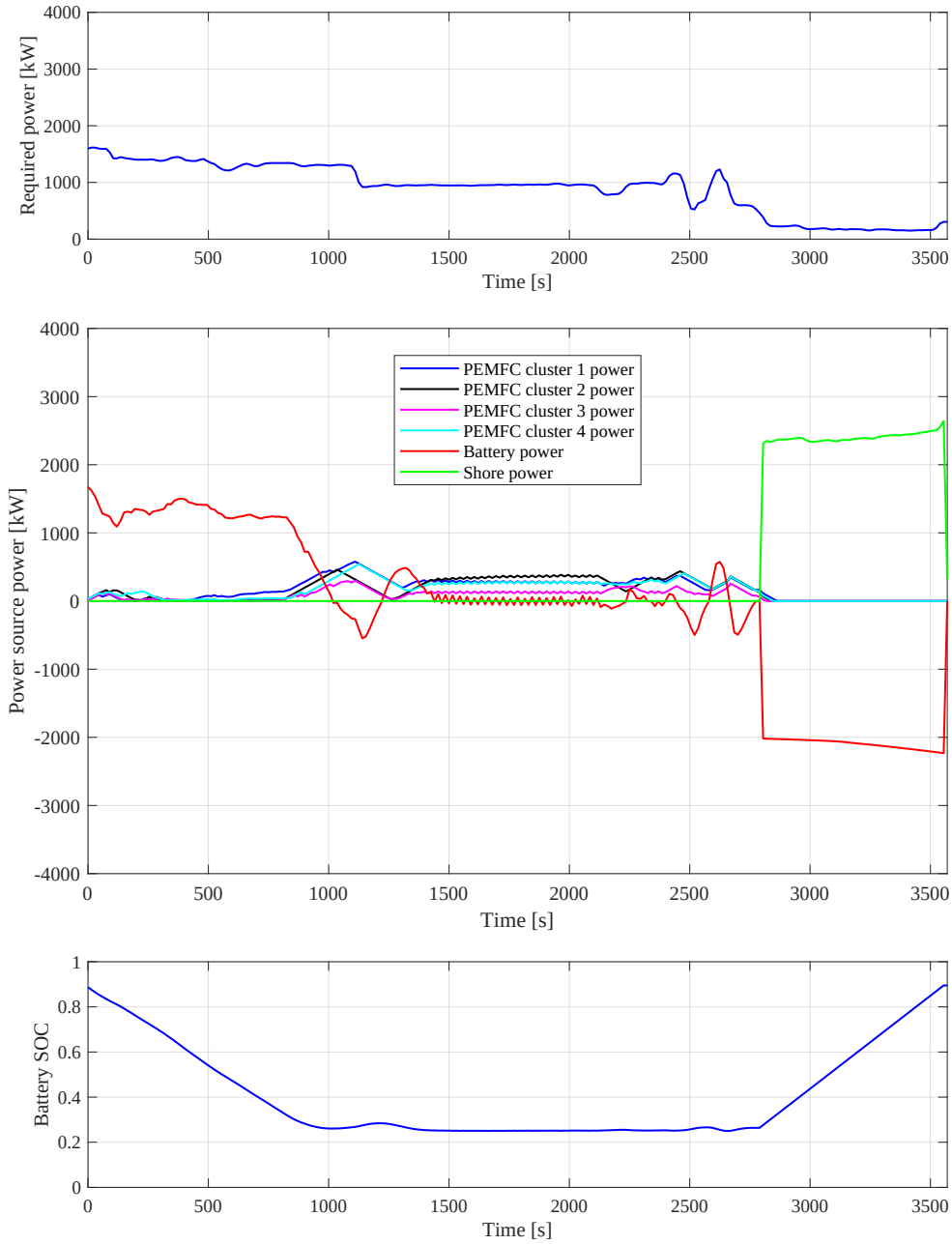


Figure 7.6: TD3 4-cluster energy management strategy for training sample voyage 1 with low power demand.

7.5.1.2 Training sample 2 with moderate power demand

Figure 7.7 shows the Double DQN strategy (Figure 7.7a) and TD3 uniform strategy (Figure 7.7b) for a sample voyage in the training dataset with moderate power demand. Both the TD3 and Double DQN strategies delay the time at which the fuel cell provides power to the system after departure with the batteries delivering the manoeuvring power. However, the TD3 strategy tends to use the batteries more aggressively, i.e. the battery SOC is maintained in lower values during cruising. Note that the TD3 strategy adjusts the PEMFC power output frequently by relatively small amounts.

As detailed in Table 7.5, the TD3 strategy results in a 1% higher voyage cost due to increased electricity and H_2 consumption. Such a pattern is different from the ones observed in Chapter 6. The Double DQN strategy with higher electricity consumption typically has lower H_2 consumption. The reason for this is that from 0–800 s the fuel cell power is maintained at a power level very close to zero, which corresponds to very low efficiency operation.

Table 7.5: Comparison of TD3 and Double DQN strategy voyage costs and GWP emissions for training sample 2.

	Voyage cost			Voyage GWP Emission		
	TD3 [$]	Double DQN [$]	$\frac{\text{TD3}}{\text{Double DQN}}$ [%]	TD3 [kg]	Double DQN [kg]	$\frac{\text{TD3}}{\text{Double DQN}}$ [%]
PEMFC	207.3	207.1	100.1	-	-	-
Battery	67.0	67.0	100.0	-	-	-
Electricity	46.1	43.0	107.2	86.2	80.4	107.2
H_2	411.9	408.2	100.9	75.0	74.3	100.9
Sum	732.2	725.3	101.0	161.2	154.7	104.2

Figure 7.8 shows the TD3 4-cluster strategy. As PEMFC cluster outputs are adjusted frequently, the PEMFC degradation cost is 6.4% higher than that of the uniform strategy (see Table 7.6). Although the 4-cluster strategy reduces both electricity and H_2 costs, the overall voyage cost of the 4-cluster strategy is 6.3% higher than that of the uniform strategy. Nevertheless, due to reduced electricity and H_2 consumptions, the 4-cluster EMS achieves 1.5% less GWP emissions.

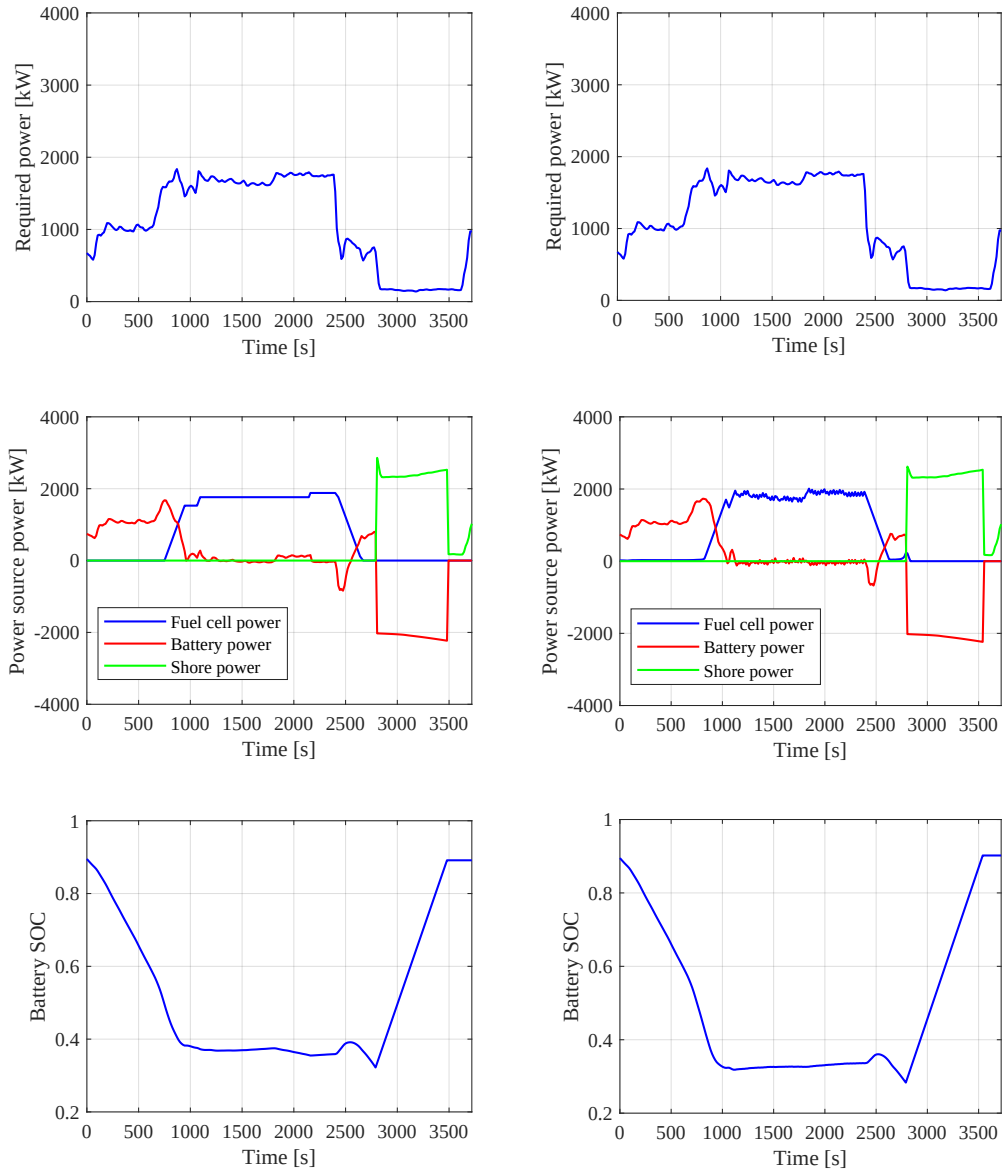


Figure 7.7: Double DQN and TD3 uniform energy management strategies for training sample voyage 2 with moderate power demand: (a) Double DQN strategy and (b) TD3 strategy.

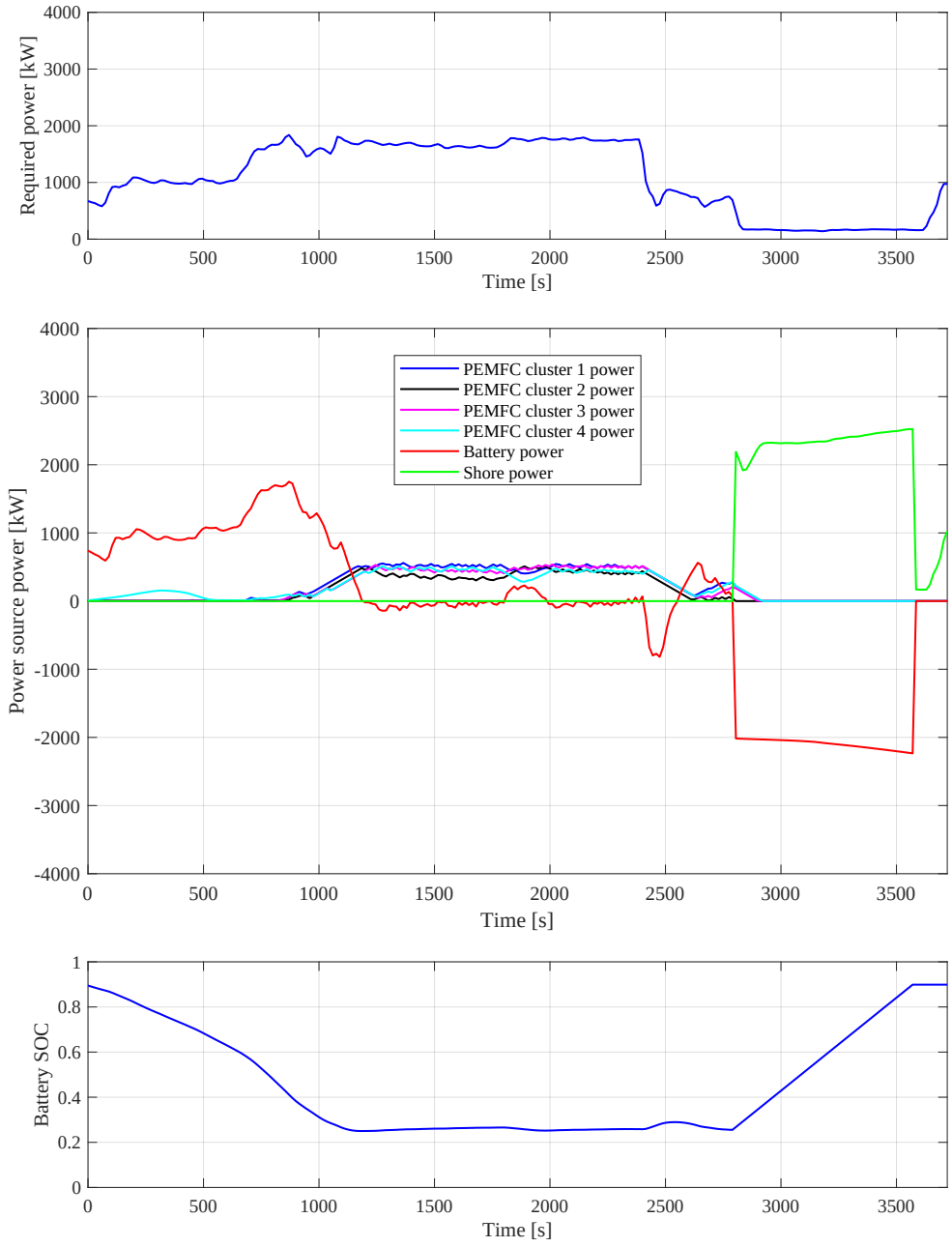


Figure 7.8: TD3 4-cluster energy management strategy for training sample voyage 2 with moderate power demand.

Table 7.6: Comparison of TD3 uniform and multi-cluster strategy voyage costs and GWP emissions for training sample 2.

	Voyage cost			Voyage GWP Emission		
	TD3–multi [$]	TD3–uniform [$]	$\frac{\text{Multi}}{\text{Uniform}}$ [%]	TD3–multi [kg]	TD3–uniform [kg]	$\frac{\text{Multi}}{\text{Uniform}}$ [%]
PEMFC	260.6	207.3	125.7	-	-	-
Battery	67.0	67.0	100.0	-	-	-
Electricity	45.5	46.1	98.7	85.1	86.2	98.7
H_2	405.2	411.9	98.4	73.8	75.0	98.4
Sum	778.2	732.2	106.3	158.8	161.2	98.5

7.5.1.3 Training sample 3 with high power demand

Figure 7.9 compares the uniform TD3 strategy (Figure 7.9b) with Double DQN strategy (Figure 7.9a) for an extremely high power profile in the training dataset. Both strategies triggered the battery over-discharge protection. The TD3 strategy results in 6.4% lower overall voyage cost (Table 7.7). The H_2 cost is reduced by increasing fuel cell power at an earlier stage such that the PEMFC can operate with higher efficiencies from 400 s to 800 s. With less H_2 consumption, the TD3 strategy voyage GWP is reduced by 5.3%.

Table 7.7: Comparison of TD3 and Double DQN strategy voyage costs and GWP emissions for training sample 3.

	Voyage cost			Voyage GWP Emission		
	TD3 [$]	Double DQN [$]	$\frac{\text{TD3}}{\text{Double DQN}}$ [%]	TD3 [kg]	Double DQN [kg]	$\frac{\text{TD3}}{\text{Double DQN}}$ [%]
PEMFC	248.4	263.9	94.1	-	-	-
Battery	56.7	56.7	100.0	-	-	-
Electricity	34.9	34.8	100.2	65.3	65.1	100.2
H_2	1130.8	1214.8	93.1	205.8	221.1	93.1
Sum	1470.8	1570.2	93.7	271.1	286.2	94.7

Figure 7.10 shows the TD3 4-cluster strategy for training sample voyage 3. All the clusters start to increase power output immediately after departure. However, the PEMFC power outputs are reduced during 150–350 s. Consequently, battery over-discharge protection is also triggered.

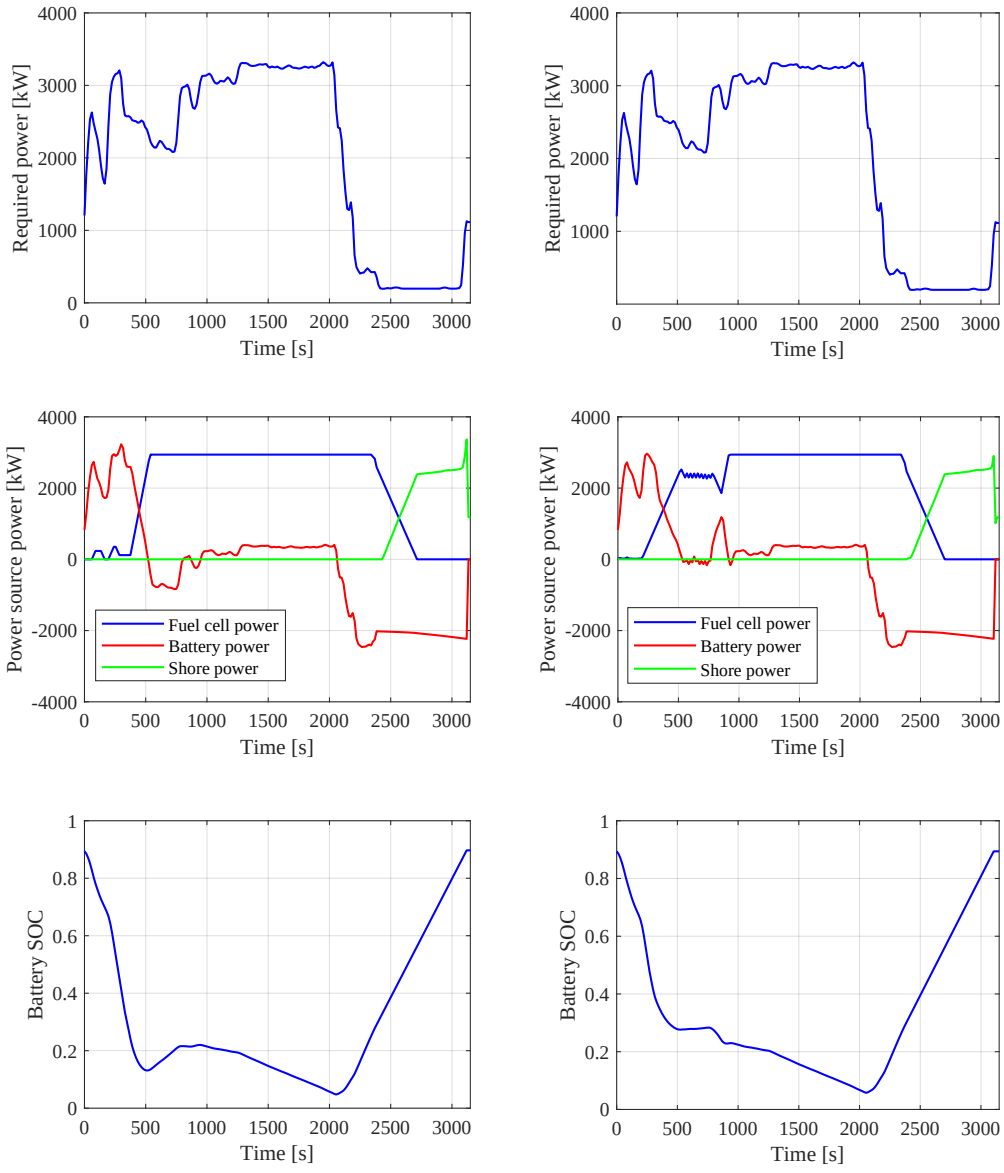


Figure 7.9: Double DQN and TD3 uniform energy management strategies for training sample voyage 3 with extreme power demand: (a) Double DQN strategy and (b) TD3 strategy.

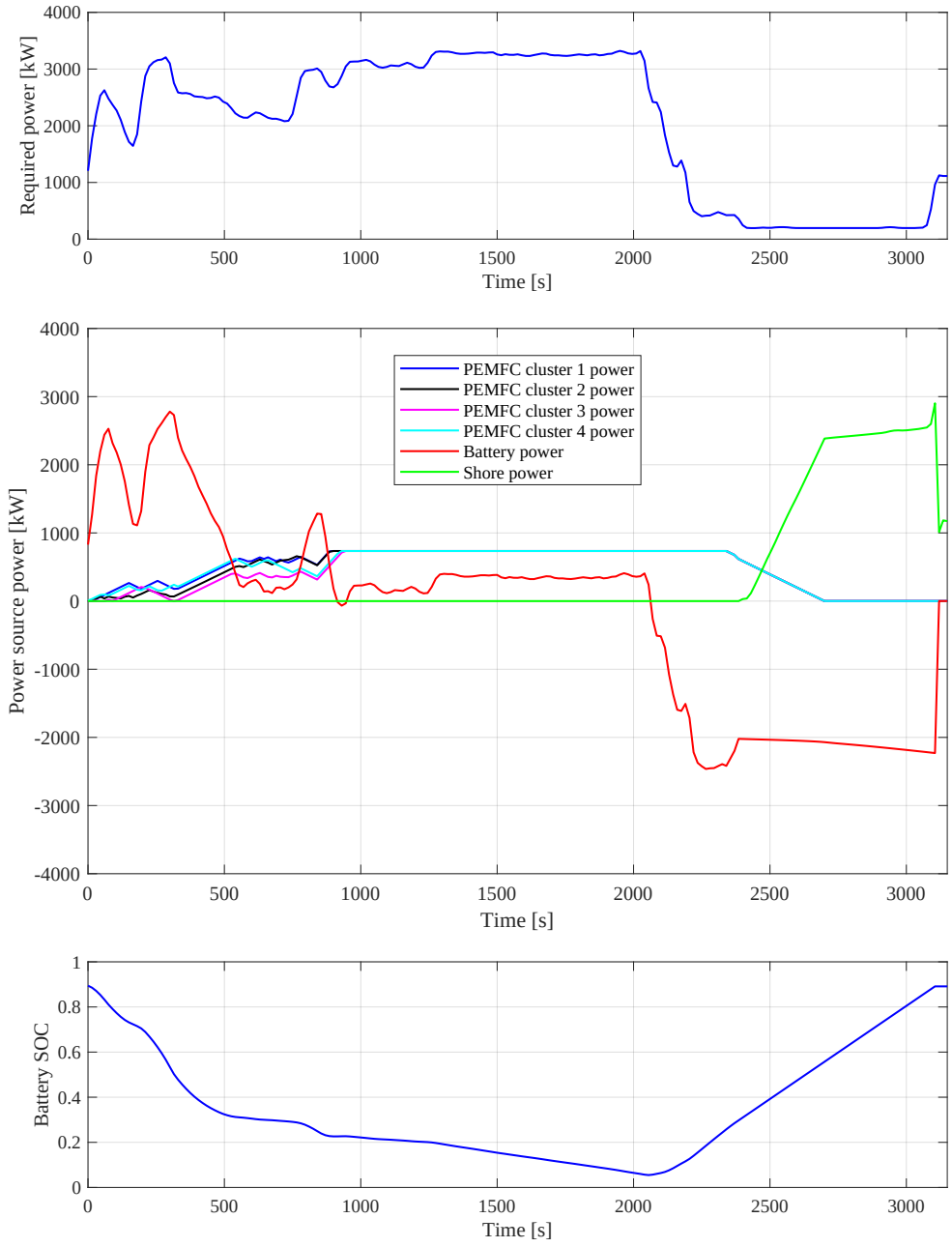


Figure 7.10: TD3 4-cluster strategy for training sample voyage 3 with extreme power demand.

Table 7.8: Comparison of TD3 uniform and multi-cluster strategy voyage costs and GWP emissions for training sample 3.

	Voyage cost			Voyage GWP Emission		
	TD3–multi [$]	TD3–uniform [$]	$\frac{\text{Multi}}{\text{Uniform}}$ [%]	TD3–multi [kg]	TD3–uniform [kg]	$\frac{\text{Multi}}{\text{Uniform}}$ [%]
PEMFC	264.3	248.4	106.4	-	-	-
Battery	56.7	56.7	100.0	-	-	-
Electricity	35.7	34.9	102.3	66.8	65.3	102.3
H_2	1105.3	1130.8	97.7	201.2	205.8	97.7
Sum	1462.1	1470.8	99.4	268.0	271.1	98.8

Table 7.8 details the voyage cost and GWP emission breakdowns of the two TD3 strategies for this training voyage with very high loads. The multi-stack strategy achieves slightly less voyage cost and emissions due to reduced H_2 consumption.

7.5.2 Validation voyages

In this section, the energy management strategies are applied to validation voyages which were not included in the training dataset such that the EMS performance for future unknown voyages can be assessed.

7.5.2.1 Validation sample 1 with low power demand

Figure 7.11 compares the Double DQN strategy (Figure 7.11a) with the uniform TD3 strategy (Figure 7.11b) for a validation sample voyage with low power demand. In Figure 7.11a, the Double DQN strategy delays to increase the fuel cell power output until the battery SOC has dropped to 0.36 (750 s). In contrast, the TD3 uniform strategy starts to ramp up fuel cell power later (at 800 s). Consequently, the minimum battery SOC of TD3 strategy is 0.26. During cruising, both strategies frequently adjust fuel cell power output.

Table 7.9 details the cost and GWP emission breakdowns for validation sample voyage 1. The uniform TD3 strategy voyage cost is 1.4% lower. This lower voyage cost is achieved as a result of lower PEMFC degradation H_2 costs. However, due to increased electricity consumption, the voyage GWP emission of the TD3 strategy

is 4.7% higher than that of Double DQN strategy.

Table 7.9: Comparison of TD3 and Double DQN strategy voyage costs and GWP emissions for validation sample 1.

	Voyage cost			Voyage GWP Emission		
	TD3 [$]	Double DQN [$]	$\frac{\text{TD3}}{\text{Double DQN}}$ [%]	TD3 [kg]	Double DQN [kg]	$\frac{\text{TD3}}{\text{Double DQN}}$ [%]
PEMFC	216.1	224.3	96.3	-	-	-
Battery	63.7	63.7	100.0	-	-	-
Electricity	44.0	40.0	110.1	82.3	74.8	110.1
H_2	361.2	366.5	98.6	65.8	66.7	98.6
Sum	685.1	694.5	98.6	148.1	141.5	104.7

Figure 7.12 presents the TD3 4-cluster strategy for validation sample voyage 1. Clusters 1 and 4 show similar trajectories with higher loads. The power output of clusters 2 and 3 are also similar but are lower compared to clusters 1 and 4. Due to early fuel cell starts and unnecessary power adjustments, the TD3 4-cluster strategy leads to a 19.6% higher PEMFC degradation cost as depicted in Table 7.10. The voyage GWP emission of the TD3 4-cluster strategy is lower by 2.1% due to reduced electricity and H_2 consumption.

Table 7.10: Comparison of TD3 uniform and multi-cluster strategy voyage costs and GWP emissions for validation sample 1.

	Voyage cost			Voyage GWP Emission		
	TD3–multi [$]	TD3–uniform [$]	$\frac{\text{Multi}}{\text{Uniform}}$ [%]	TD3–multi [kg]	TD3–uniform [kg]	$\frac{\text{Multi}}{\text{Uniform}}$ [%]
PEMFC	259.4	216.1	120.1	-	-	-
Battery	63.7	63.7	100.0	-	-	-
Electricity	43.7	44.0	99.2	81.7	82.3	99.2
H_2	347.4	361.2	96.2	63.2	65.8	96.2
Sum	714.3	685.1	104.3	144.9	148.1	97.9

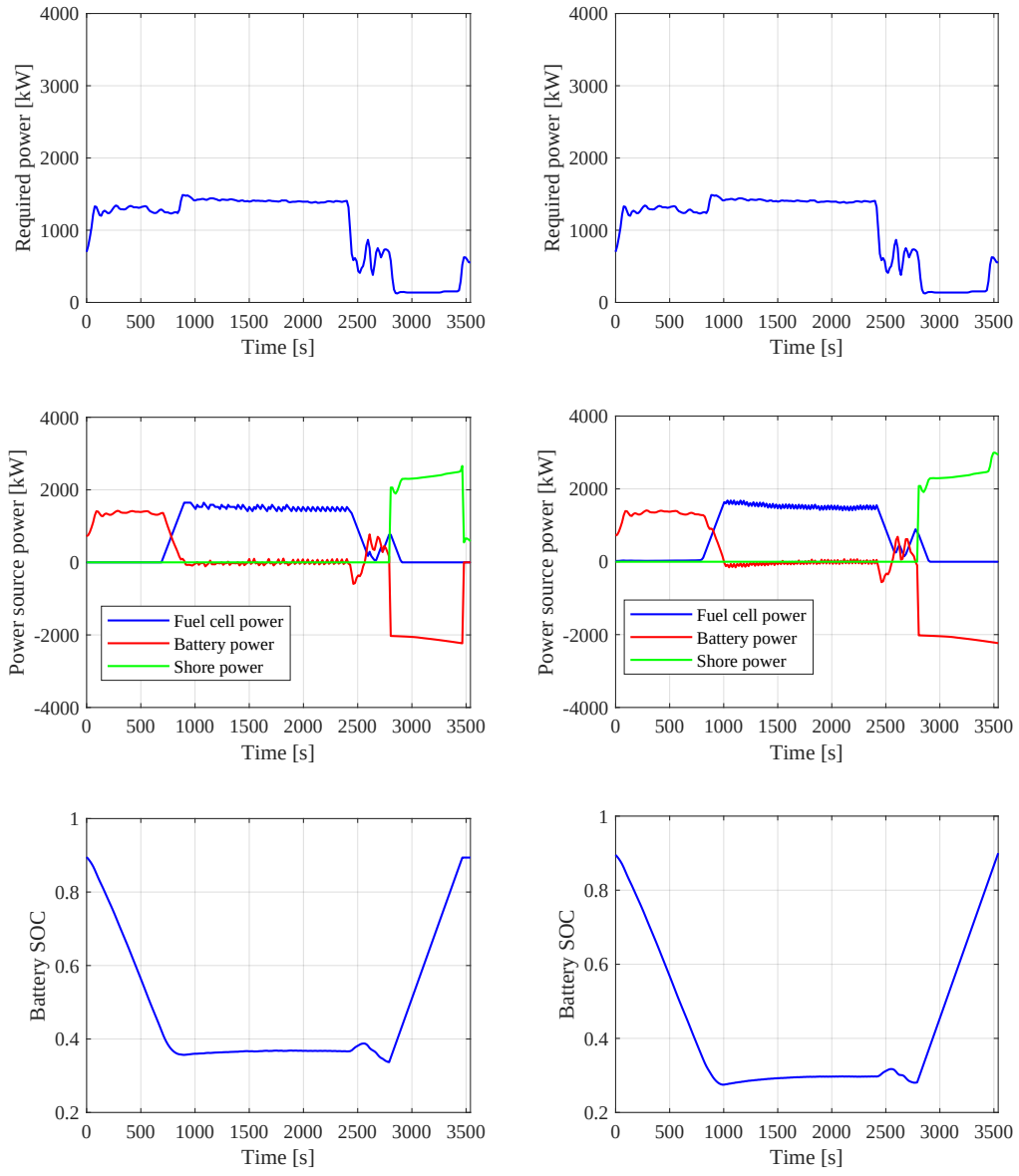


Figure 7.11: Double DQN and TD3 uniform energy management strategies for validation sample voyage 1 with low power demand: (a) Double DQN strategy and (b) TD3 strategy.

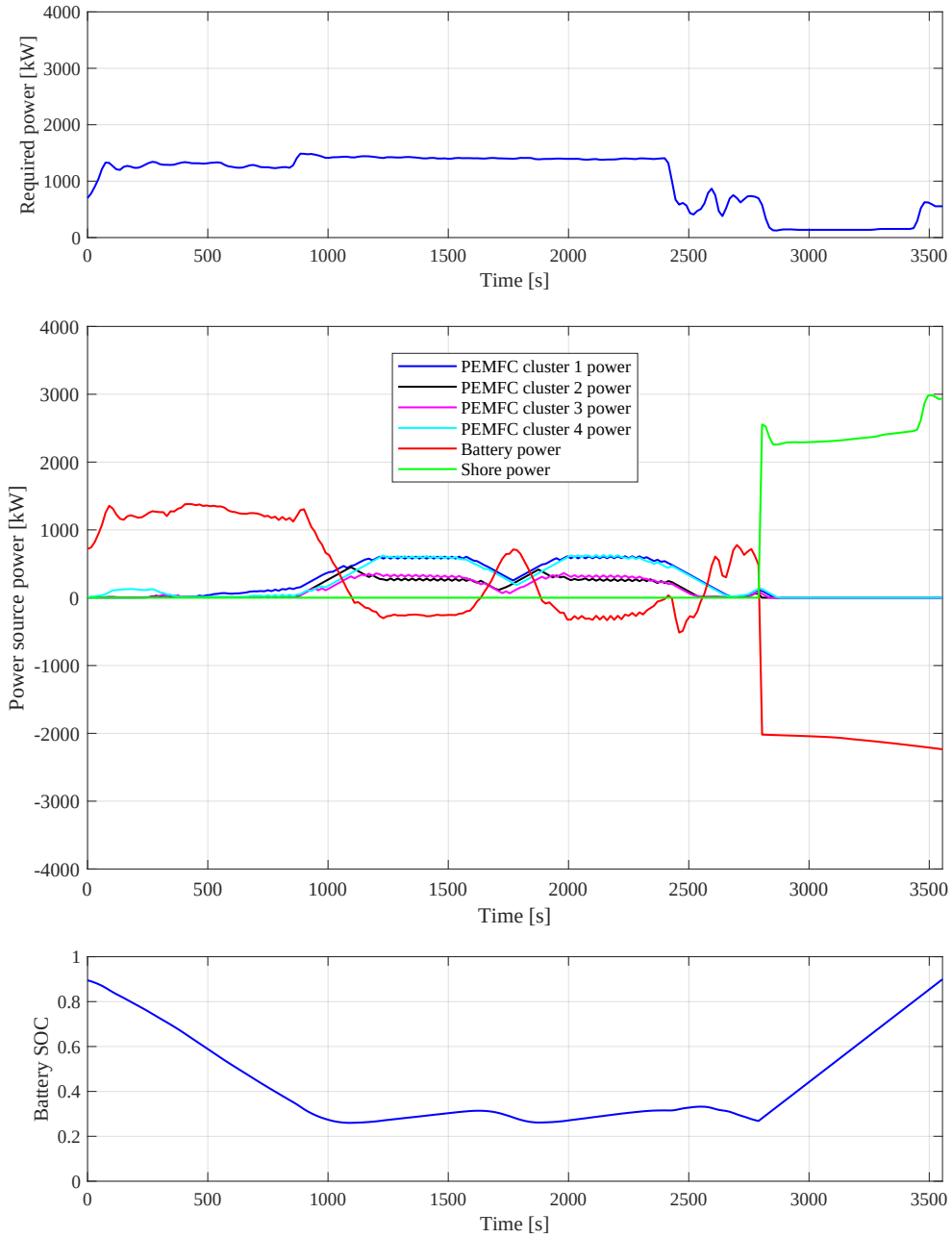


Figure 7.12: TD3 4-cluster strategy for validation sample voyage 1 with low power demand.

7.5.2.2 Validation sample 2 with moderate power demand

Figure 7.13 illustrates the Double DQN strategy (Figure 7.13a) when compared with the TD3 uniform strategy (Figure 7.13b) for a sample voyage with moderate power demand from the validation voyage dataset. The Double DQN strategy starts

ramping up the PEMFC output at 700 s. Similar behaviour has been observed for the TD3 uniform strategy. While the Double DQN maintains the PEMFC power output relatively constant, the TD3 strategy tends to adjust the power output with small oscillatory changes. The TD3 strategy leads to higher electricity and H_2 consumptions, hence the voyage GWP is also higher than that the of Double DQN strategy (see Table 7.11).

Table 7.11: Comparison of TD3 and Double DQN strategy voyage costs and GWP emissions for validation sample 2.

	Voyage cost			Voyage GWP Emission		
	TD3 [$]	Double DQN [$]	$\frac{\text{TD3}}{\text{Double DQN}}$ [%]	TD3 [kg]	Double DQN [kg]	$\frac{\text{TD3}}{\text{Double DQN}}$ [%]
PEMFC	244.8	204.7	119.6	-	-	-
Battery	63.7	63.7	100.0	-	-	-
Electricity	44.7	44.9	99.5	83.6	84.0	99.5
H_2	427.6	437.9	97.7	77.8	79.7	97.7
Sum	780.8	751.3	103.9	161.4	163.7	98.6

Figure 7.14 shows the TD3 4-cluster strategy for validation sample 2. As the PEMFC cluster outputs are adjusted frequently, the PEMFC degradation cost of the 4-cluster strategy is 19.6% higher than that of the uniform strategy (Table 7.12). Although the 4-cluster strategy increases both electricity and H_2 costs, the overall voyage cost of the 4-cluster strategy is 4.0% lower than that of the uniform strategy. Owing to increased electricity and H_2 consumption, the 4-cluster strategy results in 1.4% higher voyage GWP emissions.

Table 7.12: Comparison of TD3 uniform and multi-cluster strategy voyage costs and GWP emissions for validation sample 2.

	Voyage cost			Voyage GWP Emission		
	TD3–multi [$]	TD3–uniform [$]	$\frac{\text{Multi}}{\text{Uniform}}$ [%]	TD3–multi [kg]	TD3–uniform [kg]	$\frac{\text{Multi}}{\text{Uniform}}$ [%]
PEMFC	204.7	244.8	83.6	-	-	-
Battery	63.7	63.7	100.0	-	-	-
Electricity	44.9	44.7	100.5	84.0	83.6	100.5
H_2	437.9	427.6	102.4	79.7	77.8	102.4
Sum	751.3	780.8	96.2	163.7	161.4	101.4

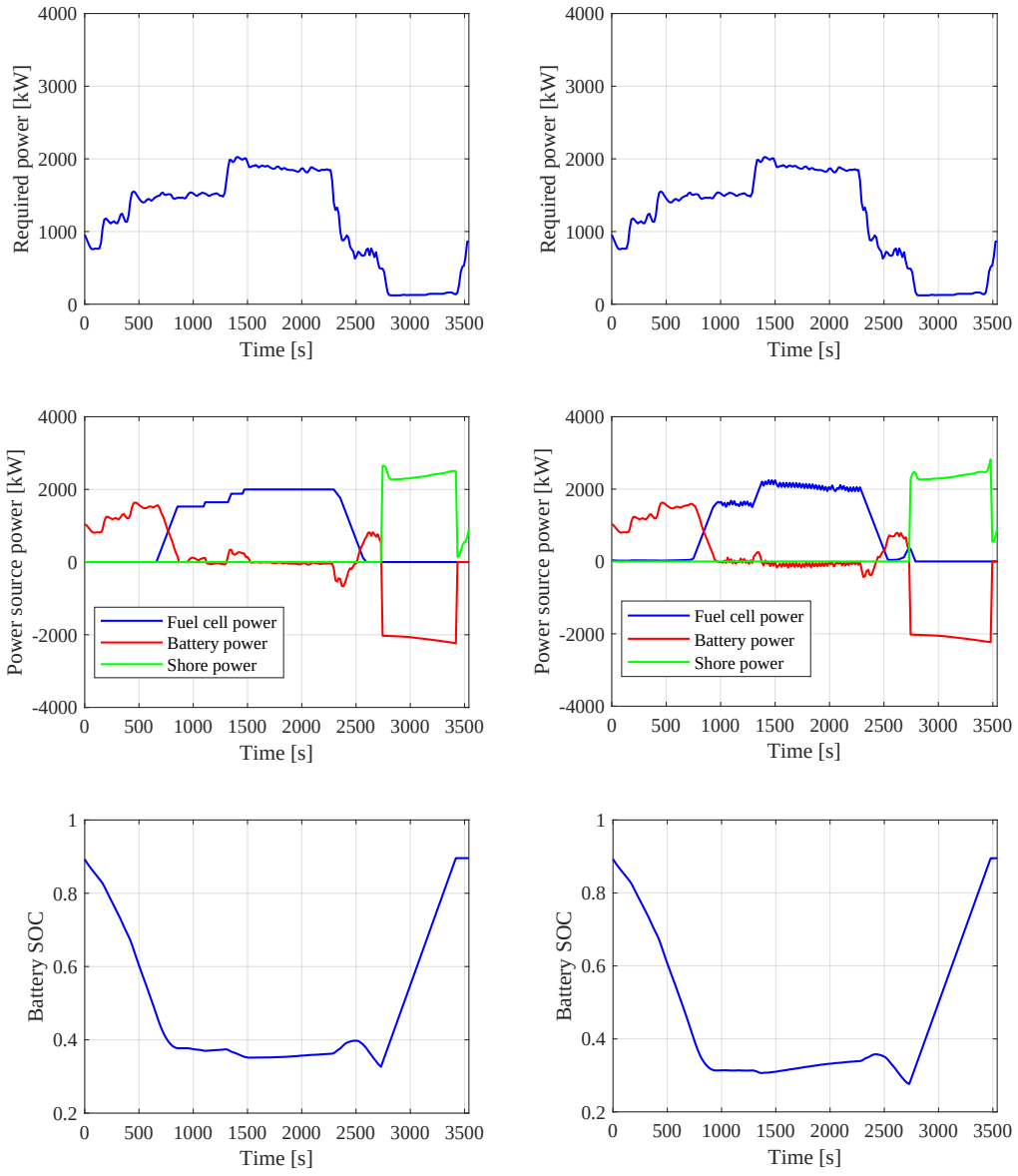


Figure 7.13: Double DQN and TD3 uniform energy management strategies for validation sample voyage 2 with moderate power demand: (a) Double DQN strategy and (b) TD3 strategy.

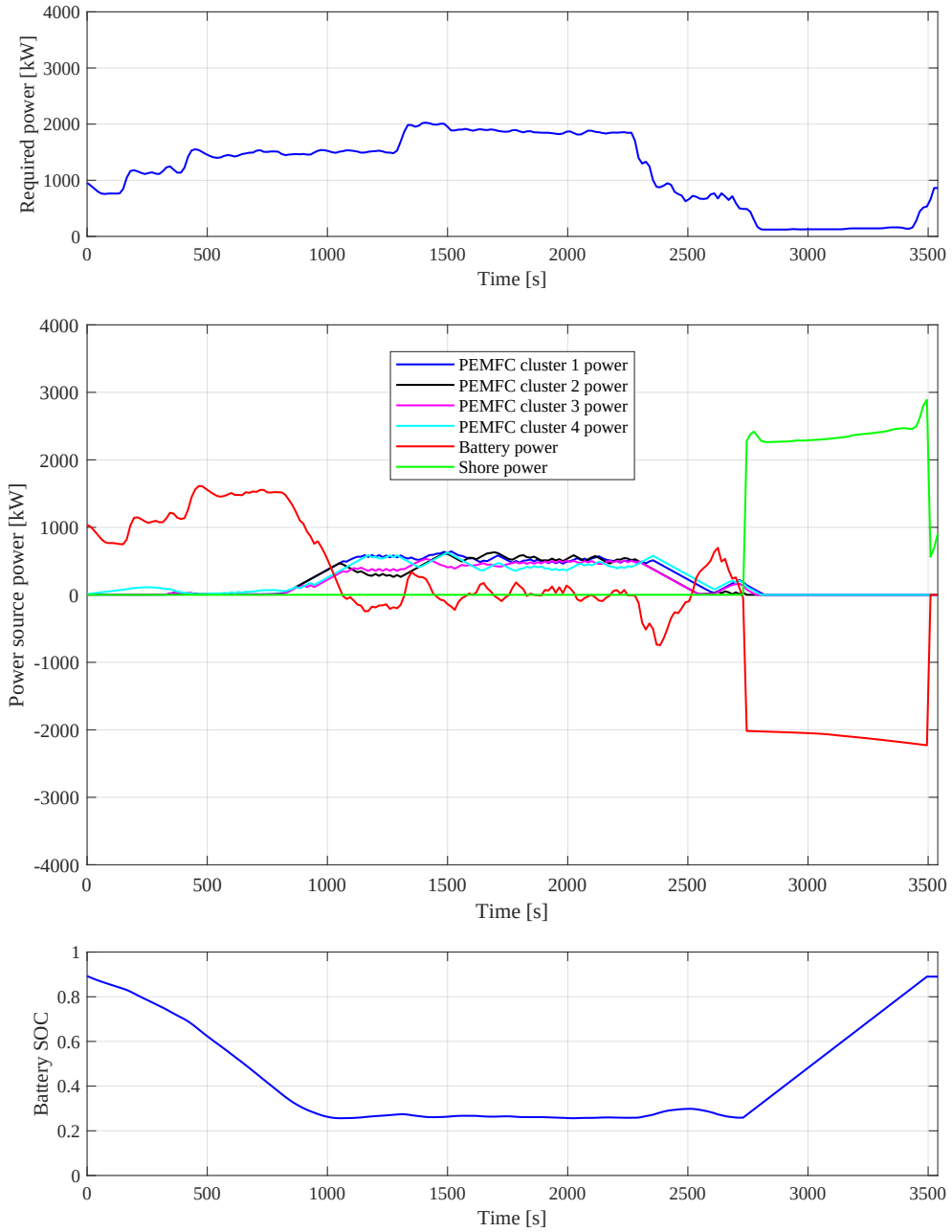


Figure 7.14: TD3 4-cluster strategy for validation sample voyage 2 with moderate power demand.

7.5.2.3 Validation sample 3 with high power demand

Figure 7.15 presents the TD3 uniform strategy (Figure 7.15b) in comparison with Double DQN strategy (Figure 7.15a) for validation sample voyage 3. The two strategies follow similar trends. But the TD3 uniform strategy adjusts the fuel cells

frequently with small action amplitudes. Table 7.13 compares the voyage cost and GWP emission breakdowns of the two strategies. The voyage cost of the TD3 uniform strategy is 5.7% higher due to increased H_2 consumption. TD3 uniform strategy emits 3.0% higher voyage GWP emission due to increased H_2 consumption.

Table 7.13: Comparison of TD3 and Double DQN strategy voyage costs and GWP emissions for validation sample 3.

	Voyage cost			Voyage GWP Emission		
	TD3 [$]	Double DQN [$]	$\frac{\text{TD3}}{\text{Double DQN}}$ [%]	TD3 [kg]	Double DQN [kg]	$\frac{\text{TD3}}{\text{Double DQN}}$ [%]
PEMFC	267.8	246.1	108.8	-	-	-
Battery	63.7	63.7	100.0	-	-	-
Electricity	48.9	49.1	99.6	91.4	91.8	99.6
H_2	715.9	678.1	105.6	130.3	123.4	105.6
Sum	1096.4	1037.1	105.7	221.8	215.2	103.0

Figure 7.16 illustrates the TD3 4-cluster strategy for validation sample voyage 3. Although the battery handles most of the large power transients, the PEMFC clusters are occasionally adjusted, leading to a higher PEMFC degradation cost (Table 7.14). The voyage cost of the 4-cluster strategy is 5.7% higher than that of the uniform TD3 strategy due to the increase in H_2 consumption and PEMFC degradation. The voyage GWP emission of the 4-cluster TD3 strategy is 3.0% lower as a result of reduced H_2 consumption.

Table 7.14: Comparison of TD3 uniform and multi-cluster strategy voyage costs and GWP emissions for validation sample 3.

	Voyage cost			Voyage GWP Emission		
	TD3–multi [$]	TD3–uniform [$]	$\frac{\text{Multi}}{\text{Uniform}}$ [%]	TD3–multi [kg]	TD3–uniform [kg]	$\frac{\text{Multi}}{\text{Uniform}}$ [%]
PEMFC	246.1	267.8	91.9	-	-	-
Battery	63.7	63.7	100.0	-	-	-
Electricity	49.1	48.9	100.4	91.8	91.4	100.4
H_2	678.1	715.9	94.7	123.4	130.3	94.7
Sum	1037.1	1096.4	94.6	215.2	221.8	97.0

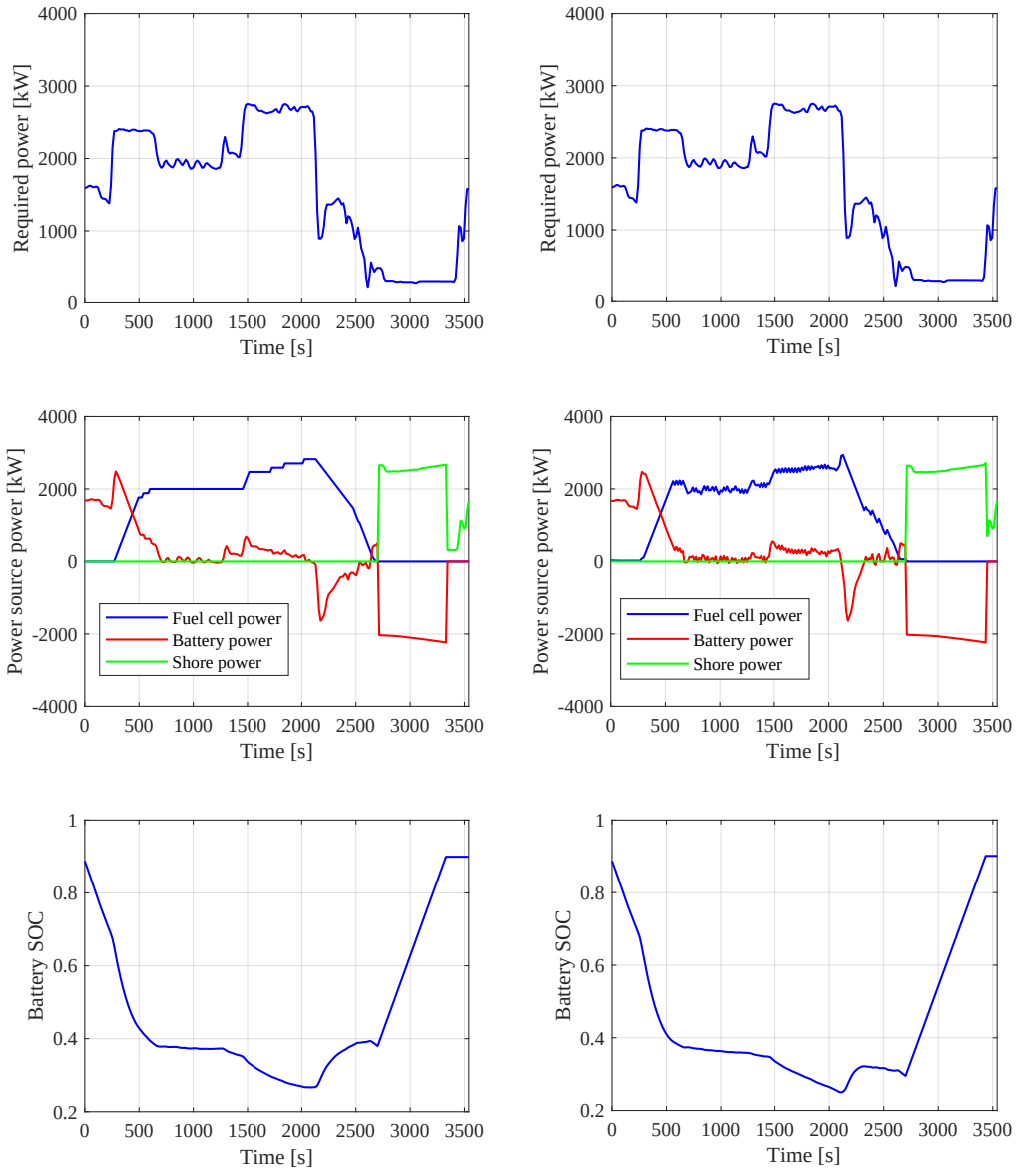


Figure 7.15: Double DQN and TD3 uniform energy management strategies for validation sample voyage 3 with high power demand: (a) Double DQN strategy and (b) TD3 strategy.

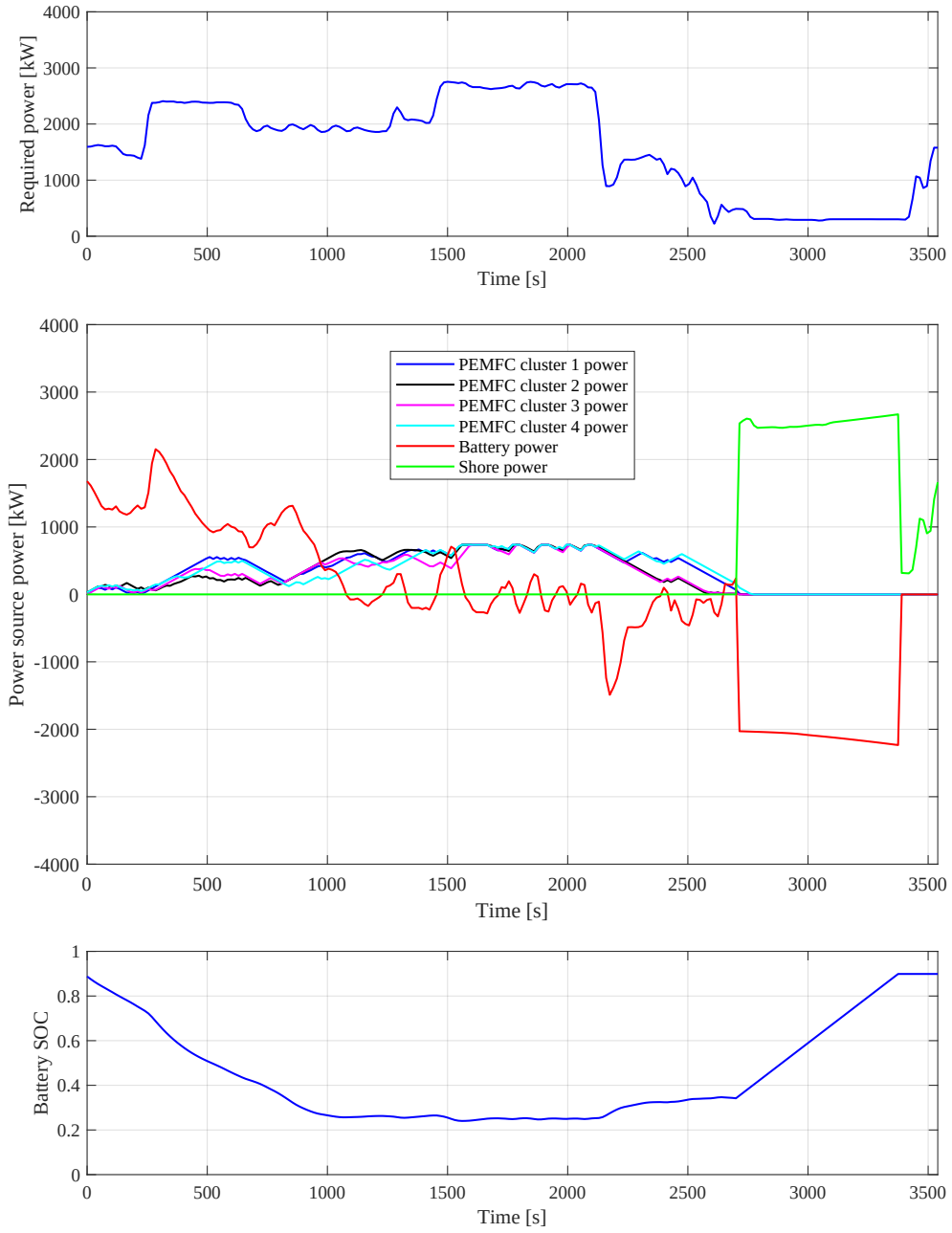


Figure 7.16: TD3 4-cluster energy management strategy for validation sample voyage 3 with high power demand.

Table 7.15 summaries the TD3 uniform strategy performance in comparison with the Double DQN strategy. The TD3 strategy further reduces the average costs for training and validation datasets by 0.6% and 0.8% respectively. However, the voyage GWP emission of the TD3 uniform strategy is increased by 3.8% due to

increased electricity usage. Compared to the DDP strategy solved with complete voyages information and SOC resolution of 0.0125, the voyage cost of the TD3 uniform strategy is only 5.1% and 5.4% higher for the training and validation voyages respectively.

Table 7.15: Comparison of TD3 and Double DQN strategy average voyage costs and GWP emissions.

Category	Profile	Voyage cost			Voyage Emission		
		TD3 [$]	Double DQN [$]	$\frac{\text{TD3}}{\text{Double DQN}}$ [%]	TD3 [kg]	Double DQN [kg]	$\frac{\text{TD3}}{\text{Double DQN}}$ [%]
Training	Sample 1	621.3	627.3	99.0	134.7	129.3	104.1
	Sample 2	732.2	725.3	101.0	161.2	154.7	104.2
	Sample 3	1353.5	1570.2	86.2	271.1	286.2	94.7
Average	*all profiles*	778.0	782.5	99.4	165.7	159.7	103.8
Validation	Sample 1	685.1	694.5	98.6	148.1	141.5	104.7
	Sample 2	751.3	744.7	100.9	163.7	157.7	103.8
	Sample 3	1037.1	1056.7	98.1	215.2	209.5	102.7
Average	*all profiles*	762.5	768.9	99.2	163.3	157.5	103.7

Table 7.16 details the TD3 4-cluster strategy performance in comparison with the TD3 uniform strategy. The 4-cluster strategy average voyage costs of the training and validation voyages are $796.3 and $783.1, respectively, which are 2.4% and 2.7% higher than those of the TD3 uniform strategy. The TD3 uniform strategy emits 1.5% and 1.8% less GWP emissions for the training and validation voyages, respectively. Nevertheless, both the TD3 uniform and multi-cluster strategies have achieved satisfactory performance when undertaking future unknown voyages. Note that the aim of multi-cluster control is improved system redundancy.

Table 7.16: Comparison of TD3 uniform and multi-cluster strategy voyage average costs and GWP emissions.

Category	Profile	Voyage cost			Voyage Emission		
		Multi [$]	Uniform [$]	$\frac{\text{Multi}}{\text{Uniform}}$ [%]	Multi [kg]	Uniform [kg]	$\frac{\text{Multi}}{\text{Uniform}}$ [%]
Training	Sample 1	644.7	621.3	103.8	135.0	134.7	100.3
	Sample 2	778.2	732.2	106.3	158.8	161.2	98.5
	Sample 3	1462.1	1353.5	108.0	268.0	271.1	98.8
Average	*all profiles*	796.3	778.0	102.4	168.2	165.7	101.5
Validation	Sample 1	714.3	685.1	104.3	144.9	148.1	97.9
	Sample 2	780.8	751.3	103.9	161.4	163.7	98.6
	Sample 3	1096.4	1037.1	105.7	221.8	215.2	103.0
Average	*all profiles*	783.1	762.5	102.7	166.3	163.3	101.9

7.6 Summary

This chapter aimed to extend the action space of the optimal energy management problem to be continuous and multi-dimensional to explore the feasibility of controlling multiple fuel cell clusters in a continuous action space using deep reinforcement learning. A generic multi-cluster fuel cell environment has been implemented based on the one developed in previous chapters. A Twin Delayed Deep Deterministic Policy Gradient with Huber loss function has been applied to solve the updated energy management problem in multi-dimensional and continuous action space. As a special case of multi-cluster control, the uniform fuel cell control was first solved using TD3. The uniform strategy learned by the TD3 agent further reduces the average voyage cost in both training and validation power profiles. The developed novel multi-cluster fuel cell control framework can be used to achieve optimal control of multiple power sources in a stochastic environment.

It was also observed that the multi-cluster TD3 strategy led to less than 3% average voyage cost increase compared to the uniform TD3 strategy. The reasons for higher voyage costs and GWP emissions are: (1) the strategy adjusts cluster power too frequently which would lead to increased PEMFC degradations and (2) the strategy occasionally operates one or more of the cluster at very low power settings leading to low fuel efficiency. Though finer hyperparameter tuning and more extended training might further improve the EMS cost and emission performance. Another possible cause could be the algorithm's capability in high dimensional action space is limited. These possibilities would require further investigations in further work. Nevertheless, it should be noted that the multi-cluster control framework is intended to improve the system redundancy. Consequently, the tiny deviations are acceptable when the actual strategy performance is near-optimal in comparison with the DDP off-line strategy.

Chapter 8

Conclusions and future work

8.1 Overview

The question this research aimed to answer was: How can coastal shipping substantially reduce harmful greenhouse gas emissions from their propulsion and power systems by using fuel cells and batteries and remain commercially viable? A critical review of previous research and projects in this field has highlighted the potential of using PEMFC and Lithium battery technologies to decarbonise the coastal ships with constraints and challenges addressed. It was concluded that there is a lack of holistic design optimisation methods and intelligent energy management strategies for such systems operating in different regions with variable energy properties. Consequently, based on the operational standpoint of a typical coastal ferry, this thesis demonstrates the development and application of a novel multi-objective power source sizing methodology and reinforcement learning based energy management strategies using continuous monitoring data collected from an actual ship.

8.2 Summary of the thesis

In Chapter 4, a scalable quasi-steady-state plug-in hybrid PEMFC/battery propulsion system model has been developed for power source sizing optimisation and intelligent energy management strategy development. In that chapter a novel holistic design methodology for coastal hybrid ships based upon the developed model was proposed. The power source sizing problem was solved using constrained mixed-integer multi-objective optimisation in the external layer. The global optimum energy management strategies for an averaged operating profile are obtained from deterministic dynamic programming in the inner layer while considering power source degradations in the sizing algorithm.

The developed multi-objective design methodology has been applied to a case ship to optimise the alternative plug-in PEMFC/battery hybrid systems with the objectives of minimising average voyage costs and voyage GWP emissions in two case studies. The case studies indicated that the proposed propulsion system can achieve at least a 65–88% life-cycle greenhouse gas reduction. It is evident that such a PEMFC/battery hybrid system can significantly reduce the GWP emission, but is still constrained by high costs, mainly from H_2 cost of production and fuel cell degradation.

Considering the uncertainties of the future energy market and production costs of PEMFC and Lithium batteries, the subsequent sensitivity studies further explore the impacts of power source and energy prices on the design of plug-in hybrid PEMFC/battery system. The prices of the PEMFC, battery and H_2 are expected to decrease and become competitive with conventional diesel-based power solutions. Nevertheless, the potential GWP emission reductions offered by the plug-in hybrid PEMFC/battery system are significant. Having identified that these high costs (including CAPEX and OPEX) are major limitations of such systems, the aim of Chapters 5, 6 and 7 was to develop reinforcement learning and deep reinforcement learning based intelligent energy management strategies to minimise future unknown voyage cost using continuously monitored real-ship power profiles.

Chapter 5 formulated the optimal energy management problem using MDP and solved the formulated MDP using reinforcement learning agents in the discrete state and action spaces to achieve minimum voyage cost without prior knowledge of future power demands. With continuously monitored power profiles collected from the case ship, two reinforcement learning agents (i.e. Q-learning and Double Q-learning) have been applied to solve the optimal energy management problem. The Q-learning agent which introduced maximisation biases failed to achieve satisfactory performance. In contrast, the Double Q-learning achieved 96.9% cost-performance compared to deterministic dynamic programming with identical space resolutions. The results of this chapter suggested that reinforcement learning based energy management strategies can achieve near-optimal performance without prior knowledge of future power demands.

It should be noted that the resolutions of state and action spaces are limited since Q-tables have been employed in Chapter 5. Frequent unnecessary fuel cell power adjustments have been observed in the sample voyages due to limited state space resolution. However, increasing the resolutions of the two spaces would make the problem practically impossible to solve with the available computational resources due to 'the curse of dimensionality'. Consequently, Chapter 6 extends the discrete state space to be continuous with deep neural networks as function approximators.

In Chapter 6, novel approaches of Deep Q-Network and Double Deep Q-network have been applied considering the results in Chapter 5 suggested overestimations of action-values could lead to divergent agent training. Also, two loss functions, i.e. Mean Squared Error and Huber loss functions, were explored to deal with the value overestimations in the stochastic environment. Both agents failed when applied with the Mean Squared Error loss function due to the maximisation biases. With the Huber loss function, the two deep reinforcement learning agents achieved similar performance. A further voyage cost reduction of 6% with more than 90% computation time reduction was achieved compared to the results acquired in Chapter 5.

The main limitation of the Q-network based agents is that they are not capable of handling a continuous and high dimensional action space. Chapter 7 further extended both the state and action spaces to be continuous and developed a novel generic framework for multi-cluster fuel cell control using a TD3 agent. The TD3 agent achieved satisfactory EMS performance in both uniform and multi-cluster fuel cell controls. Although the multi-cluster TD3 EMS performance could be further improved, the developed multi-cluster framework will be useful for developing more sophisticated EMS with more constraints (e.g. load sharing and fault tolerance).

In summary, this thesis has proposed a novel multi-objective power source sizing methodology integrating fuel cells and batteries for coastal ships. The proposed sizing methodology has been demonstrated via case and sensitivity studies for a coastal ferry. Moreover, the optimal energy management problem of the plug-in hybrid PEMFC/battery system has been formulated with Markov Decision Process and solved using reinforcement learning and deep reinforcement learning algorithms. Based on continuously monitored power profiles, the developed energy management strategies can achieve near-optimal cost performance in un-predicted future voyages.

8.3 Summary of contributions

As stated in Chapter 1, this thesis focuses on decarbonising coastal ships using fuel cells and batteries. By analysing and summarising the work, the research aim has been achieved through the following aspects:

- **Opportunities and constraints of using fuel cells and batteries for coastal shipping decarbonisation:** The uncertainties from the energy supply side, power source manufacturing and practical ship design requirements need to be addressed holistically to achieve balanced alternative propulsion system design. The hybridisation of PEMFC, battery and grid power could potentially provide balanced propulsion system designs for coastal ship operating

on short routes and have access to clean and renewable energy sources.

- **Hybrid propulsion system design optimisation:** A holistic sizing methodology for the design of plug-in hybrid PEMFC/battery propulsion systems has been proposed to optimise GWP emission and voyage cost. The proposed methodology is based upon a calibratable system model, considering life-cycle GHG emissions, costs and shipboard constraints.

- **Reinforcement learning and deep reinforcement learning based intelligent energy management strategies using continuous monitoring data:** The optimal energy management problem of the plug-in hybrid PEMFC/battery propulsion system has been formulated in Markov Decision Process using large-scale historical data. The formulated Markov Decision Process has been solved using reinforcement learning and deep reinforcement learning algorithms to achieve near-optimum voyage cost in un-predicted future voyages with a very high success rate.

- **Deep reinforcement learning based continuous intelligent energy management strategies for multiple fuel cell cluster control:** A generic multi-cluster fuel cell environment has been developed for continuous and high-dimensional fuel cell control. The state-of-the-art TD3 deep reinforcement learning algorithm has been successfully applied to solve the multi-cluster fuel cell intelligent energy management problem with complete success rate.

The novel contributions of this research are twofold. The sizing methodology can be applied to guide practical hybrid propulsion system design. The intelligent energy management strategies can be applied to hybrid systems in operation to minimise costs (including CAPEX and OPEX) for un-predicted future voyages.

8.4 Recommendations for future work

8.4.1 Route and speed optimisation

The current work is based upon continuous monitoring data of power demands, assuming that such power demands are fixed inputs to the algorithms. What if the power demands can be reduced? One possible approach of reducing such power demands is to optimise the routes and speeds. Optimised routes and speeds can lead to reductions in power demands. The reduced power demands can be further processed as inputs to the deep reinforcement learning based energy management strategies to reduce voyage costs and emissions further. The author has explored solving the path planning problem for unnamed surface vehicles using DQN in (Zhou et al., 2019).

8.4.2 Advanced deep reinforcement learning agents

In this project, reinforcement learning agents such a Q-learning, Double Q-learning, DQN, Double DQN, TD3 have been employed. The training processes in Chapters 6 and 7 suggest that agent performance can vary with different experience explorations. Moreover, all the agents have been implemented in this project are single-threaded. Other agents which can be trained in parallel (e.g. Advantage Actor-Critic and Asynchronous Advantage Actor-Critic) can be investigated further to better utilise multi-core processors and gain better performance. The hyperparameters of the current agents require careful tuning to achieve satisfactory performance. Agents requiring minimum parameter tuning (e.g. Proximal Policy Gradient) can be further investigated. Note that the function overestimation problem needs to be addressed properly.

8.4.3 Battery degradation model

In current work, an averaged battery degradation model is adopted assuming the battery life-time is guaranteed by the manufacturer. In practice, the battery degrada-

tion rate is subject to change under different operating conditions such as C-rate and temperature. To implement a more detailed degradation model, a sufficient amount of experimental data would be needed to fit an empirical degradation model. However, it should be noted that such models are only applicable to the type of cells used to identify the model parameters. Industrial collaboration with battery cell manufacturers would be a practical approach to acquire the data.

8.4.4 Intelligent energy management strategies for other hybrid systems

Although the energy management strategies developed in this thesis are based on the developed plug-in hybrid PEMFC/battery model, the reinforcement learning approaches can be applied to other power systems, e.g. diesel-electric systems with energy storage. Note that the methodologies can also be transferred to other sectors such as road transport.

8.4.5 Advanced multi-cluster fuel cell control

In Chapter 7, multiple fuel cell clusters are controlled individually. However, current work does not consider any possible failure modes (e.g. loss of one or two clusters). The agents can be trained to be resilient with potential faults by adding simulated faulted conditions to the training voyages. Note that additional system states (e.g. cluster on/off status) are required to represent the power source running conditions. Moreover, other state parameters such as weather and sea states can be added to the state space.

8.4.6 System integration and experimental validation

Although the developed sizing methodology is practical, for actual applications, the design parameters and system models need updates to match the actual scenario and power sources available. When implementing intelligent energy management strategies, the actual power source characteristics need to be updated in the system models. Moreover, more protection functions may need to be added.

8.4.7 Policy update strategies

8.4.7.1 Periodic policy

In this project, generic energy management strategies were generated based on one set of historical power profiles. However, considering the power transitions may change over time due to factors such as seasonal sea states and weather conditions, periodically updating the policy may bring further improve the cost-effectiveness. As depicted in Figure 8.1, the vessel starts with an EMS trained via a set of initial power profiles. During operation, the EMS is updated periodically using the most recent power profiles. A set of recent power profiles needs to be randomly separated into two sub-sets: one set for agent training and verification, while the other set for EMS validation. Once the new EMS is validated, it is ready for future voyages.

8.4.7.2 Adaptive policy

As power transition patterns may change over time, a self-adaptive EMS updating procedure can be applied with minimum human intervention. Figure 8.2 shows the training process of an adaptive EMS. The ship starts with an EMS trained by an initial set of power profiles. A dynamic profile pool is maintained throughout the ship's operation by replacing the oldest profile with the most recent profile. Periodically, the EMS performance is evaluated by comparing the actual EMS performance against those solved via DDP. The agent would need training if the deviation between the on-line and DDP strategies exceeds the performance threshold; otherwise, existing EMS would be applied until the next performance evaluation point.

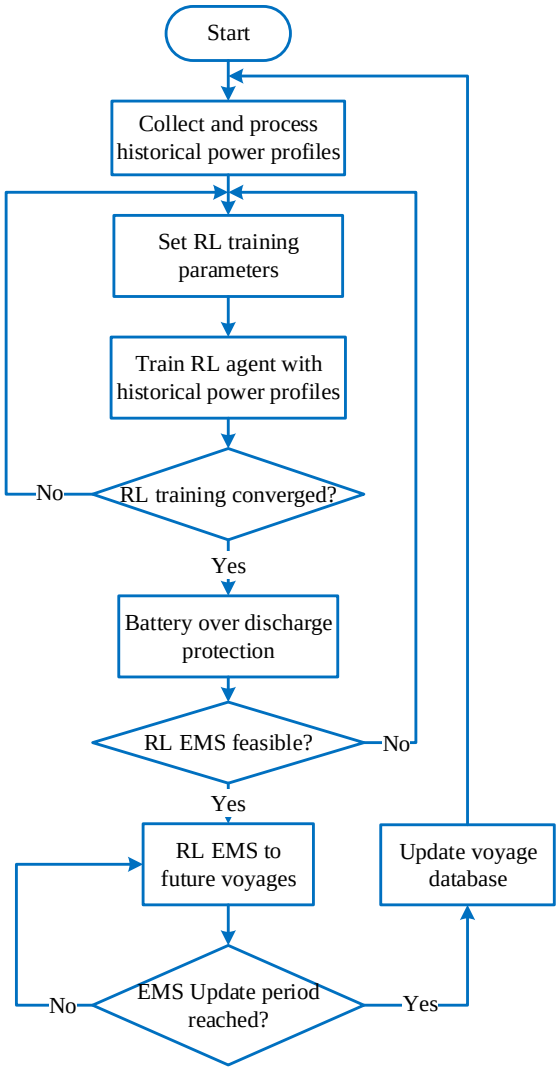


Figure 8.1: Periodic update training procedure.

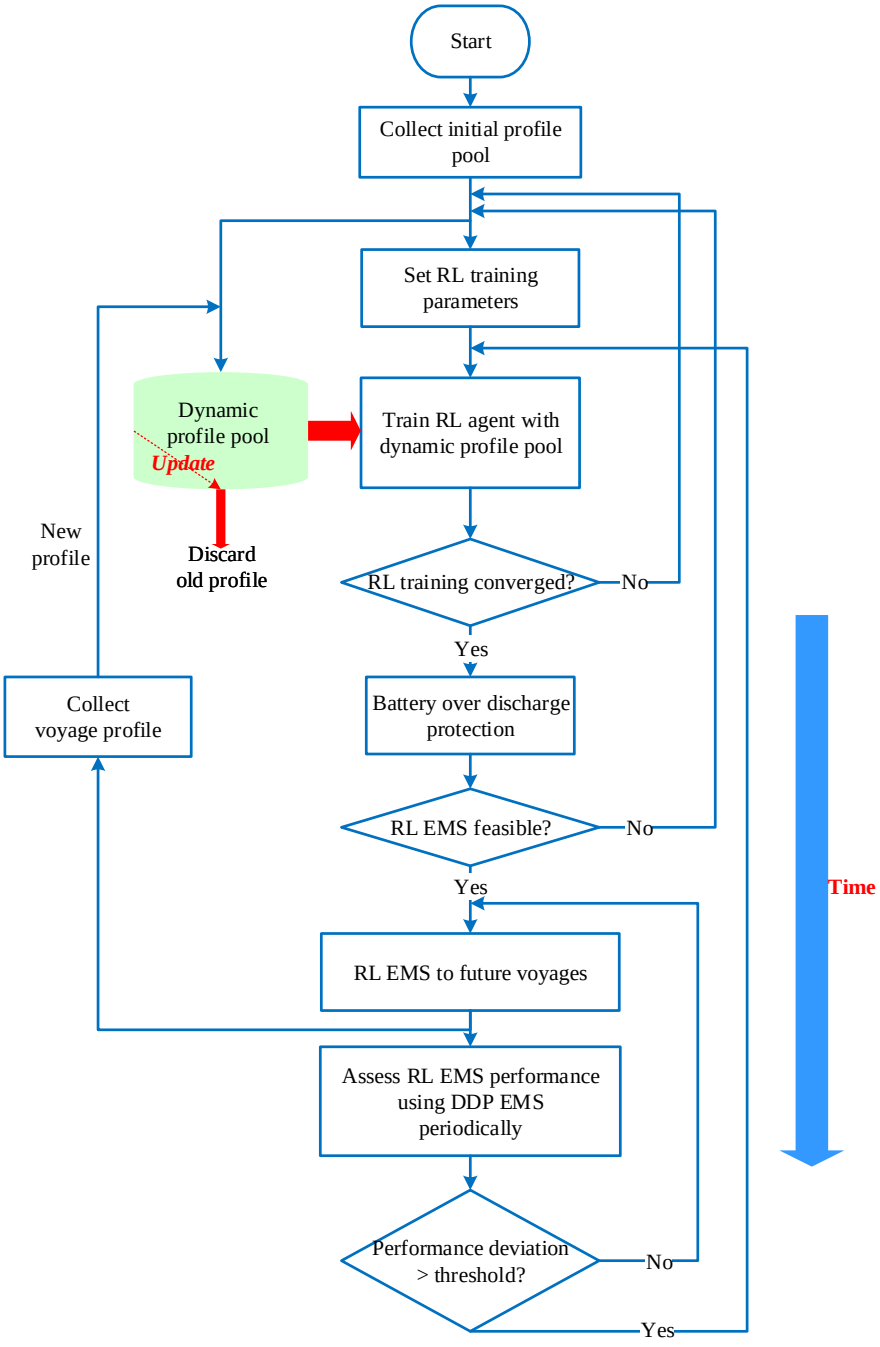


Figure 8.2: Adaptive training procedure.

References

ABB. ForSea (formerly HH Ferries Group) completes conversion of the world's largest battery ferries, powered by ABB, 2018. URL https://new.abb.com/news/detail/10434/forsea-formerly-hh-ferries-group-completes-conversion-of-the-worlds-largest-battery-ferries-powered-by-abb. [Accessed 11th Oct 2019].

Abdin, Z., Webb, C., and Gray, E. M. PEM fuel cell model and simulation in Matlab–Simulink based on physical parameters. *Energy*, 116:1131–1144, 2016.

Acar, C. and Dincer, I. Comparative assessment of hydrogen production methods from renewable and non-renewable sources. *International Journal of Hydrogen Energy*, 39(1):1–12, 2014.

Ådnanes, A. K. Maritime electrical installations and diesel electric propulsion. *ABB*, 2003.

Ahmed, A., Al-Amin, A. Q., Ambrose, A. F., and Saidur, R. Hydrogen fuel and transport system: A sustainable and environmental future. *International journal of hydrogen energy*, 41(3):1369–1380, 2016.

Al-Falahi, M. D., Nimma, K. S., Jayasinghe, S. D., Enshaei, H., and Guerrero, J. M. Power management optimization of hybrid power systems in electric ferries. *Energy conversion and management*, 172:50–66, 2018.

Alaswad, A., Baroutaji, A., Achour, H., Carton, J., Al Makky, A., and Olabi, A.-

G. Developments in fuel cell technologies in the transport sector. *International Journal of Hydrogen Energy*, 41(37):16499–16508, 2016.

Alnes, O., Eriksen, S., and Vartdal, B.-J. Battery-powered ships: a class society perspective. *IEEE Electrification Magazine*, 5(3):10–21, 2017.

Amiryar, M. and Pullen, K. A review of flywheel energy storage system technologies and their applications. *Applied Sciences*, 7(3):286, 2017.

Anandarajah, G., McDowall, W., and Ekins, P. Decarbonising road transport with hydrogen and electricity: Long term global technology learning scenarios. *International Journal of Hydrogen Energy*, 38(8):3419–3432, 2013.

Anderson, M., Salo, K., and Fridell, E. Particle-and gaseous emissions from an lng powered ship. *Environmental science & technology*, 49(20):12568–12575, 2015.

Andwari, A. M., Pesiridis, A., Rajoo, S., Martinez-Botas, R., and Esfahanian, V. A review of battery electric vehicle technology and readiness levels. *Renewable and Sustainable Energy Reviews*, 78:414–430, 2017.

Ang, S. M. C., Brett, D. J., and Fraga, E. S. A multi-objective optimisation model for a general polymer electrolyte membrane fuel cell system. *Journal of Power Sources*, 195(9):2754–2763, 2010.

Arani, A. K., Karami, H., Gharehpetian, G., and Hejazi, M. Review of flywheel energy storage systems structures and applications in power systems and microgrids. *Renewable and Sustainable Energy Reviews*, 69:9–18, 2017.

Aulinger, A., Matthias, V., Zeretzke, M., Bieser, J., Quante, M., and Backes, A. The impact of shipping emissions on air pollution in the greater north sea region–part 1: Current emissions and concentrations. *Atmospheric Chemistry & Physics*, 16 (2), 2016.

Balcombe, P., Brierley, J., Lewis, C., Skatvedt, L., Speirs, J., Hawkes, A., and Staffell, I. How to decarbonise international shipping: Options for fuels, technologies and policies. *Energy conversion and management*, 182:72–88, 2019.

Baldi, F. and Gabrielii, C. A feasibility analysis of waste heat recovery systems for marine applications. *Energy*, 80:654–665, 2015.

Ballard. Fuel Cell Power Module for Heavy Duty Motive Applications, 2019. URL http://www.ballard.com/docs/default-source/motive-modules-documents/fcvelocity_hd_family_of_products_low_res.pdf. [Accessed 21st Feb 2019].

Barnes, J. *Oceans of Power: 125 Years of Marine Engineering Milestones*. Institute of Marine Engineering, Science and Technology, 2014.

Bassam, A. M., Phillips, A. B., Turnock, S. R., and Wilson, P. A. Sizing optimization of a fuel cell/battery hybrid system for a domestic ferry using a whole ship system simulator. In *2016 International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles & International Transportation Electrification Conference (ESARS-ITEC)*, pages 1–6. IEEE, 2016.

Bassam, A. M., Phillips, A. B., Turnock, S. R., and Wilson, P. A. Development of a multi-scheme energy management strategy for a hybrid fuel cell driven passenger ship. *International Journal of Hydrogen Energy*, 42(1):623–635, jan 2017.

Bicer, Y. and Dincer, I. Clean fuel options with hydrogen for sea transportation: A life cycle approach. *International Journal of Hydrogen Energy*, 43(2):1179–1193, 2018.

Blumberg, T., Lee, Y. D., Morosuk, T., and Tsatsaronis, G. Exergoenvironmental analysis of methanol production by steam reforming and autothermal reforming of natural gas. *Energy*, 2019.

Boscarato, I., Hickey, N., Kašpar, J., Prati, M. V., and Mariani, A. Green shipping: Marine engine pollution abatement using a combined catalyst/seawater scrubber system. 1. Effect of catalyst. *Journal of Catalysis*, 328:248–257, 2015.

Boucher, O., Friedlingstein, P., Collins, B., and Shine, K. P. The indirect global warming potential and global temperature change potential due to methane oxidation. *Environmental Research Letters*, 4(4):044007, 2009.

Boyan, J. A. and Moore, A. W. Generalization in reinforcement learning: Safely approximating the value function. In *Advances in neural information processing systems*, pages 369–376, 1995.

Breijs, A. and Amam, E. Energy management–adapt your engine to every mission. In *Proceedings of the 13th international naval engineering conference*, pages 1–8, 2016.

Brynolf, S., Fridell, E., and Andersson, K. Environmental assessment of marine fuels: liquefied natural gas, liquefied biogas, methanol and bio-methanol. *Journal of Cleaner Production*, 74:86–95, 2014.

BunkerIndex. Index Summary, 2017. URL http://www.bunkerindex.com/. [Accessed 28th Oct 2017].

Burnham, A., Dufek, E. J., Stephens, T., Francfort, J., Michelbacher, C., Carlson, R. B., Zhang, J., Vijayagopal, R., Dias, F., Mohanpurkar, M., et al. Enabling fast charging–infrastructure and economic considerations. *Journal of Power Sources*, 367:237–249, 2017.

Calleya, J. N. *Ship design decision support for a carbon dioxide constrained future*. PhD thesis, UCL (University College London), 2014.

Cames, M., Graichen, J., Siemons, A., and Cook, V. Emission reduction targets for international aviation and shipping. Report, European Parliament, 2015.

Carlton, J., Aldwinkle, J., Anderson, J., et al. Future ship powering options: exploring alternative methods of ship propulsion. *London: Royal Academy of Engineering*, 2013.

Caterpillar. Marine power systems C32 TIER 3/IMO II, 2018. URL https://www.cat.com/en_GB/products/new/power-systems/marine-power-systems/marine-generator-sets/1000031744.html. [Accessed 20th Jan 2018].

Caux, S., Hankache, W., Fadel, M., and Hissel, D. On-line fuzzy energy management for hybrid fuel cell systems. *International journal of hydrogen energy*, 35 (5):2134–2143, 2010.

Chandan, A., Hattenberger, M., El-Kharouf, A., Du, S., Dhir, A., Self, V., Pollet, B. G., Ingram, A., and Bujalski, W. High temperature (ht) polymer electrolyte membrane fuel cells (pemfc)–a review. *Journal of Power Sources*, 231:264–278, 2013.

Chen, H., Pei, P., and Song, M. Lifetime prediction and the economic lifetime of proton exchange membrane fuel cells. *Applied Energy*, 142:154–163, 2015.

Chen, Z., Xiong, R., Wang, C., and Cao, J. An on-line predictive energy management strategy for plug-in hybrid electric vehicles to counter the uncertain prediction of the driving cycle. *Applied Energy*, 185:1663–1672, 2017.

Choi, C. H., Yu, S., Han, I.-S., Kho, B.-K., Kang, D.-G., Lee, H. Y., Seo, M.-S., Kong, J.-W., Kim, G., Ahn, J.-W., et al. Development and demonstration of pem fuel-cell-battery hybrid system for propulsion of tourist boat. *International Journal of Hydrogen Energy*, 41(5):3591–3599, 2016.

Clarksons. World Fleet Register, 2016. URL http://wfr.clarksons.net/wfr/register/. [Assessed 01/06/2016].

Coello, C. C. and Lechuga, M. S. Mopso: A proposal for multiple objective particle swarm optimization. In *Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No. 02TH8600)*, volume 2, pages 1051–1056. IEEE, 2002.

Corbett, J. J., Fischbeck, P. S., and Pandis, S. N. Global nitrogen and sulfur inventories for oceangoing ships. *Journal of Geophysical Research: Atmospheres*, 104 (D3):3457–3470, 1999.

Corvus, E. Orca Energy Specifications, 2017. URL http://corvusenergy.com/technology-specifications/. [Accessed 10th Oct 2017].

Corvus, E. Corvus Energy awarded the marine world's biggest battery package for hybrid powered vessel by Norwegian Electric Systems, 2019a. URL https://corvusenergy.com/corvus-energy-awarded-the-marine-worlds-biggest-battery-package-by-norwegian-electric-systems/. [Accessed 20th May 2019].

Corvus, E. Corvus Energy signs final contract with AIDA Cruises for the largest battery pack ever delivered to a ship, 2019b. URL https://corvusenergy.com/corvus-energy-signs-final-contract-with-aida-cruises-for-the-largest-battery-pack-ever-delivered-to-a-ship/. [Accessed 15th Sept 2019].

Crilley, J. and Dean, C. Shortsea shipping and the world cargo-carrying fleet-a statistical summary. In *European Shortsea Shipping, 1st European Research Roundtable Conference*, 1993.

Cuma, M. U. and Koroglu, T. A comprehensive review on estimation strategies used in hybrid and battery electric vehicles. *Renewable and Sustainable Energy Reviews*, 42:517–531, 2015.

Curry, C. Lithium-ion battery costs and market. *Bloomberg New Energy Finance*, 5, 2017.

Dai, H., Wei, X., Sun, Z., Wang, J., and Gu, W. Online cell SOC estimation of Li-ion battery packs using a dual time-scale Kalman filtering for EV applications. *Applied Energy*, 95:227–237, 2012.

de Troya, J. J., Alvarez, C., Fernández-Garrido, C., and Carral, L. Analysing the possibilities of using fuel cells in ships. *International Journal of Hydrogen Energy*, 41(4):2853–2866, 2016.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE transactions on evolutionary computation*, 6 (2):182–197, 2002.

Deniz, C. and Zincir, B. Environmental and economical assessment of alternative marine fuels. *Journal of Cleaner Production*, 113:438 – 449, 2016.

Dewulf, J., Van der Vorst, G., Denturck, K., Van Langenhove, H., Ghyoot, W., Tytgat, J., and Vandeputte, K. Recycling rechargeable lithium ion batteries: Critical analysis of natural resource savings. *Resources, Conservation and Recycling*, 54 (4):229–234, 2010.

Di Natale, F. and Carotenuto, C. Particulate matter in marine diesel engines exhausts: Emissions and control strategies. *Transportation Research Part D: Transport and Environment*, 40:166–191, 2015.

Dicks, A. L. Molten carbonate fuel cells. *Current Opinion in Solid State and Materials Science*, 8(5):379–383, 2004.

Dicks, A. L. and Rand, D. A. *Fuel cell systems explained*. John Wiley & Sons, 2018.

Dimitriou, P. and Tsujimura, T. A review of hydrogen as a compression ignition engine fuel. *International Journal of Hydrogen Energy*, 42(38):24470–24486, 2017.

Dimitriou, P., Kumar, M., Tsujimura, T., and Suzuki, Y. Combustion and emission characteristics of a hydrogen-diesel dual-fuel engine. *International journal of hydrogen energy*, 43(29):13605–13617, 2018.

DNV GL. DNV GL handbook for maritime and offshore battery systems, 2016.

Douet, M. and Cappuccilli, J. F. A review of short sea shipping policy in the european union. *Journal of Transport Geography*, 19(4):968–976, 2011.

Doughty, D. H. and Roth, E. P. A general discussion of li ion battery safety. *The Electrochemical Society Interface*, 21(2):37–44, 2012.

Ebrahimi, M. and Moradpoor, I. Combined solid oxide fuel cell, micro-gas turbine and organic rankine cycle for power generation (SOFC–MGT–ORC). *Energy Conversion and Management*, 116:120–133, 2016.

Edwards, P. P., Kuznetsov, V. L., David, W. I., and Brandon, N. P. Hydrogen and fuel cells: towards a sustainable energy future. *Energy policy*, 36(12):4356–4362, 2008.

Ehsani, M., Gao, Y., Longo, S., and Ebrahimi, K. *Modern electric, hybrid electric, and fuel cell vehicles*. CRC press, 2018.

Eichi, H. R. and Chow, M.-Y. Modeling and analysis of battery hysteresis effects. In *2012 IEEE Energy Conversion Congress and Exposition (ECCE)*, pages 4479–4486. IEEE, 2012.

Ekanem Attah, E. and Bucknall, R. An analysis of the energy efficiency of LNG ships powering options using the EEDI. *Ocean Engineering*, 110:62–74, Dec 2015.

El-Gohary, M. M. Overview of past, present and future marine power plants. *Journal of Marine Science and Application*, 12(2):219–227, 2013.

Endresen, Ø., Sørgård, E., Sundet, J. K., Dalsøren, S. B., Isaksen, I. S., Berglen, T. F., and Gravir, G. Emission from international sea transportation and environmental impact. *Journal of Geophysical Research: Atmospheres*, 108(D17), 2003.

Erdinc, O. and Uzunoglu, M. Optimum design of hybrid renewable energy systems: Overview of different approaches. *Renewable and Sustainable Energy Reviews*, 16(3):1412–1425, 2012.

Eriksen, S., Lützen, M., Jensen, J. B., and Sørensen, J. C. Improving the energy efficiency of ferries by optimizing the operational practices. In *Proceedings of the Full Scale Ship Performance Conference 2018: The Royal Institution of Naval Architects*, pages 101–111. The Royal Institution of Naval Architects, 2018.

European Environment Agency. Overview of electricity production and use in europe. Report, European Environment Agency, 2019. URL

https://www.eea.europa.eu/data-and-maps/indicators/overview-of-the-electricity-production-2/assessment. [Accessed 20th May 2019].

Eurostat. Maritime transport statistics - short sea shipping of goods. Technical report, Eurostat, 2018. URL https://ec.europa.eu/eurostat/statistics-explained/index.php/Maritime_transport_statistics_-_short_sea_shipping_of_goods#Short_sea_shipping_by_sea_region_and_country. [Accessed 9th May 2019].

Eurostat. Electricity price statistics, 2019. URL http://ec.europa.eu/eurostat/statistics-explained/index.php/Electricity_price_statistics. [Accessed 10th Feb 2019].

Eyring, V., Isaksen, I. S., Berntsen, T., Collins, W. J., Corbett, J. J., Endresen, O., Grainger, R. G., Moldanova, J., Schlager, H., and Stevenson, D. S. Transport impacts on atmosphere and climate: Shipping. *Atmospheric Environment*, 44 (37):4735–4771, 2010.

Faraji, F., Majazi, A., Al-Haddad, K., et al. A comprehensive review of flywheel energy storage system technology. *Renewable and Sustainable Energy Reviews*, 67:477–490, 2017.

Farrier, L., Wu, P., and Bucknall, R. Opportunities and constraints of electrical energy storage systems in ships. In *Shipping in Changing Climates 2017*. Shipping in Changing Climates 2017, 2017.

Fletcher, T., Thring, R., and Watkinson, M. An energy management strategy to concurrently optimise fuel consumption & pem fuel cell lifetime in a hybrid vehicle. *International Journal of Hydrogen Energy*, 41(46):21503–21515, 2016.

Florentinus, A., Hamelinck, C., Van den Bos, A., Winkel, R., and Cuijpers, M. Potential of biofuels for shipping–final report. *Ecofys, EMSA*, 2012.

Fujimoto, S., van Hoof, H., and Meger, D. Addressing function approximation error in actor-critic methods. *arXiv preprint arXiv:1802.09477*, 2018.

Geertsma, R., Negenborn, R., Visser, K., and Hopman, J. Design and control of hybrid power and propulsion systems for smart ships: A review of developments. *Applied Energy*, 194:30–54, 2017.

Gilbert, P., Walsh, C., Traut, M., Kesieme, U., Pazouki, K., and Murphy, A. Assessment of full life-cycle air emissions of alternative shipping fuels. *Journal of cleaner production*, 172:855–866, 2018.

González, A., Goikolea, E., Barrena, J. A., and Mysyk, R. Review on supercapacitors: technologies and materials. *Renewable and Sustainable Energy Reviews*, 58:1189–1206, 2016.

Goodenough, J. B. and Park, K.-S. The li-ion rechargeable battery: a perspective. *Journal of the American Chemical Society*, 135(4):1167–1176, 2013.

Guo, M., Fu, Z., Ma, D., Ji, N., Song, C., and Liu, Q. A short review of treatment methods of marine diesel engine exhaust gases. *Procedia Engineering*, 121:938–943, 2015.

Gür, T. M. Review of electrical energy storage technologies, materials and systems: challenges and prospects for large-scale grid storage. *Energy & Environmental Science*, 11(10):2696–2767, 2018.

Gurz, M., Baltacioglu, E., Hames, Y., and Kaya, K. The meeting of hydrogen and automotive: a review. *International Journal of Hydrogen Energy*, 42(36):23334–23346, 2017.

Halkos, G. and Tsilika, K. Understanding transboundary air pollution network: Emissions, depositions and spatio-temporal distribution of pollution in european region. *Resources, Conservation and Recycling*, 145:113–123, 2019.

Hameer, S. and van Niekerk, J. L. A review of large-scale electrical energy storage. *International journal of energy research*, 39(9):1179–1195, 2015.

Han, J., Charpentier, J.-F., and Tang, T. An Energy Management System of a Fuel Cell/Battery Hybrid Boat. *Energies*, 7(5):2799, 2014.

Hannan, M. A., Hoque, M. M., Mohamed, A., and Ayob, A. Review of energy storage systems for electric vehicle applications: Issues and challenges. *Renewable and Sustainable Energy Reviews*, 69:771–789, 2017.

Hansen, J. F. and Wendt, F. History and state of the art in commercial electric ship propulsion, integrated power systems, and future trends. *Proceedings of the IEEE*, 103(12):2229–2242, 2015.

Hansson, J., Månsson, S., Brynolf, S., and Grahn, M. Alternative marine fuels: Prospects based on multi-criteria decision analysis involving swedish stakeholders. *Biomass and Bioenergy*, 126:159–173, 2019.

Henderson, P., Islam, R., Bachman, P., Pineau, J., Precup, D., and Meger, D. Deep reinforcement learning that matters. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

Herrera, L., Zhang, W., and Wang, J. Stability analysis and controller design of dc microgrids with constant power loads. *IEEE Transactions on Smart Grid*, 8(2): 881–888, 2015.

Higier, A., Hsu, L., Oiler, J., Phipps, A., Hooper, D., and Kerber, M. Polymer electrolyte fuel cell (pemfc) based power system for long-term operation of leave-in-place sensors in navy and marine corps applications. *International Journal of Hydrogen Energy*, 42(7):4706–4709, 2017.

Holmgren, J., Nikopoulou, Z., Ramstedt, L., and Woxenius, J. Modelling modal choice effects of regulation on low-sulphur marine fuels in northern europe. *Transportation Research Part D: Transport and Environment*, 28:62–73, 2014.

Hu, X., Johannesson, L., Murgovski, N., and Egardt, B. Longevity-conscious dimensioning and power management of the hybrid energy storage system in a fuel cell hybrid electric bus. *Applied Energy*, 137:913–924, 2015a.

Hu, X., Moura, S. J., Murgovski, N., Egardt, B., and Cao, D. Integrated optimization of battery sizing, charging, and power management in plug-in hybrid electric

vehicles. *IEEE Transactions on Control Systems Technology*, 24(3):1036–1043, 2015b.

Hu, Y., Li, W., Xu, K., Zahid, T., Qin, F., and Li, C. Energy management strategy for a hybrid electric vehicle based on deep reinforcement learning. *Applied Sciences*, 8(2):187, 2018.

Huber, P. J. Robust estimation of a location parameter. In *Breakthroughs in statistics*, pages 492–518. Springer, 1992.

International Maritime Organization. Guidelines for voluntary use of the ship energy efficiency operational indicator (EEOI). MEPC.1/Circ.684, 2009.

International Maritime Organization. Guidelines on the method of calculation of the attained energy efficiency design index (EEDI) for new ships.MEPC.245(66), 2011.

International Maritime Organization. Third IMO Greenhous Gas study 2014. Report, International Maritime Organization, 2014a.

International Maritime Organization. Guidelines on the approved method process MEPC 66/21, 2014b.

International Maritime Organization. Ship energy efficiency regulations and related guidelines, 2015. URL http://www.imo.org/en/OurWork/Environment/PollutionPrevention/AirPollution/Documents/Air%20pollution/M2%20TTT%20course%20Posters%20final1.pdf. [Accessed 6th Nov 2019].

International Maritime Organization. Guidelines for the development of a ship energy efficiency management plan (SEEMP), 2016.

International Maritime Organization. IMO action to reduce greenhouse gas emissions from international shipping–implementing the initial IMO strategy on reduction of GHG emissions from ships, 2019. URL http://www.imo.org/en/MediaCentre/HotTopics/Documents/IMO%

20ACTION%20TO%20REDUCE%20GHG%20EMISSIONS%20FROM%20INTERNATIONAL%20SHIPPING.pdf. [Accessed 6th Nov 2019].

Jaguemont, J., Boulon, L., and Dubé, Y. A comprehensive review of lithium-ion batteries used in hybrid and electric vehicles at cold temperatures. *Applied Energy*, 164:99–114, 2016.

Jenkins, S. Methanol production from natural gas. *Chemical Engineering*, 123(4): 37, 2016.

Jiang, L., Kronbak, J., and Christensen, L. P. The costs and benefits of sulphur reduction measures: Sulphur scrubbers versus marine gas oil. *Transportation Research Part D: Transport and Environment*, 28:19–27, 2014.

Johnson, H. and Styhre, L. Increased energy efficiency in short sea shipping through decreased time in port. *Transportation Research Part A: Policy and Practice*, 71: 167–178, 2015.

Johnson, H., Johansson, M., and Andersson, K. Barriers to improving energy efficiency in short sea shipping: an action research case study. *Journal of Cleaner Production*, 66:317–327, 2014.

Kalikatzarakis, M., Geertsma, R., Boonen, E., Visser, K., and Negenborn, R. Ship energy management for hybrid propulsion and power supply with shore charging. *Control Engineering Practice*, 76:133–154, jul 2018.

Kanellos, F. Optimal power management with ghg emissions limitation in all-electric ship power systems comprising energy storage systems. *IEEE Transactions on Power Systems*, 29(1):330–339, 2013.

Kanstad, T., Lillholm, M. B., and Zhang, Z. Highly Efficient EV Battery Charger Using Fractional Charging Concept with SiC Devices. In *2019 IEEE Applied Power Electronics Conference and Exposition (APEC)*, pages 1601–1608. IEEE, 2019.

Kay, R. and Alder, J. *Coastal planning and management*. CRC Press, London, 2017.

Khayyam, H. and Bab-Hadiashar, A. Adaptive intelligent energy management system of plug-in hybrid electric vehicle. *Energy*, 69:319–335, 2014.

Kim, J., Shin, J., Chun, C., and Cho, B. H. Stable configuration of a li-ion series battery pack based on a screening process for improved voltage/soc balancing. *IEEE Transactions on Power Electronics*, 27(1):411–424, Jan 2012.

Kim, J. G., Son, B., Mukherjee, S., Schuppert, N., Bates, A., Kwon, O., Choi, M. J., Chung, H. Y., and Park, S. A review of lithium and non-lithium based solid state batteries. *Journal of Power Sources*, 282:299–322, 2015a.

Kim, M.-J. and Peng, H. Power management and design optimization of fuel cell/battery hybrid vehicles. *Journal of Power Sources*, 165(2):819–832, 2007.

Kim, S.-Y., Choe, S., Ko, S., and Sul, S.-K. A naval integrated power system with a battery energy storage system: Fuel efficiency, reliability, and quality of power. *IEEE electrification magazine*, 3(2):22–33, 2015b.

Kingma, D. P. and Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Kristensen, H. O. Energy demand and exhaust gas emissions of marine engines. *Clean Shipping Currents*, 1(6):18–26, 2012.

Larcher, D. and Tarascon, J.-M. Towards greener and more sustainable batteries for electrical energy storage. *Nature chemistry*, 7(1):19, 2015.

Larsen, U., Pierobon, L., Haglind, F., and Gabrielii, C. Design and optimisation of organic rankine cycles for waste heat recovery in marine applications using the principles of natural selection. *Energy*, 55:803–812, 2013.

Li, X., Chalvatzis, K. J., and Pappas, D. China's electricity emission intensity in 2020–an analysis at provincial level. *Energy Procedia*, 142:2779–2785, 2017.

Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., and Wierstra, D. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*, 2015.

Lin, L.-J. *Reinforcement learning for robots using neural networks*. PhD thesis, Carnegie-Mellon University Pittsburgh PA School of Computer Science, 1993.

Lindstad, H., Sandaas, I., and Strømman, A. H. Assessment of cost as a function of abatement options in maritime emission control areas. *Transportation Research Part D: Transport and Environment*, 38:41–48, 2015.

Liu, T., Hu, X., Li, S. E., and Cao, D. Reinforcement learning optimized look-ahead energy management of a parallel hybrid electric vehicle. *IEEE/ASME Transactions on Mechatronics*, 22(4):1497–1507, 2017.

Lloyd's Register. Your options for emissions compliance-guidance for shipowners and operators on the annex vi sox and nox regulations. Report, Lloyd's Register, 2015. URL http://www.lr.org/en/_images/213-35826_Your_options_for_emissions_compliance.pdf. [Assessed 01st Oct 2016].

Lu, L., Han, X., Li, J., Hua, J., and Ouyang, M. A review on the key issues for lithium-ion battery management in electric vehicles. *Journal of power sources*, 226:272–288, 2013.

Lu, X., Yao, T., Li, Y., Fung, J. C., and Lau, A. K. Source apportionment and health effect of nox over the pearl river delta region in southern china. *Environmental pollution*, 212:135–146, 2016.

Luo, X., Wang, J., Dooner, M., and Clarke, J. Overview of current development in electrical energy storage technologies and the application potential in power system operation. *Applied Energy*, 137:511–536, 2015.

MAN. Two-stroke engines: ME-C Applications, 2019. URL https://marine.man-es.com/two-stroke/2-stroke-engines/me-c-engines. [Accessed 10th Sept. 2019].

Martel, F., Kelouwani, S., Dubé, Y., and Agbossou, K. Optimal economy-based battery degradation management dynamics for fuel-cell plug-in hybrid electric vehicles. *Journal of Power Sources*, 274:367–381, 2015.

Martinez, C. M., Hu, X., Cao, D., Velenis, E., Gao, B., and Wellers, M. Energy management in plug-in hybrid electric vehicles: Recent progress and a connected vehicles perspective. *IEEE Transactions on Vehicular Technology*, 66(6):4534–4549, 2016.

Mashayekh, S., Wang, Z., Qi, L., Lindtjorn, J., and Myklebust, T.-A. Optimum sizing of energy storage for an electric ferry ship. In *2012 IEEE Power and Energy Society General Meeting*, pages 1–8. IEEE, 2012.

May, G. J., Davidson, A., and Monahov, B. Lead batteries for utility energy storage: A review. *Journal of Energy Storage*, 15:145–157, 2018.

McConnell, V. P. Now, voyager? the increasing marine use of fuel cells. *Fuel Cells Bulletin*, 2010(5):12–17, 2010.

McCoy, T. J. Trends in ship electric propulsion. In *IEEE Power Engineering Society Summer Meeting,*, volume 1, pages 343–346. IEEE, 2002.

Mehta, V. and Cooper, J. S. Review and analysis of pem fuel cell design and manufacturing. *Journal of power sources*, 114(1):32–53, 2003.

Mekhilef, S., Saidur, R., and Safari, A. Comparative study of different fuel cell technologies. *Renewable and Sustainable Energy Reviews*, 16(1):981–989, 2012.

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., and Riedmiller, M. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*, 2013.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., et al. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529, 2015.

Mollenhauer, K., Tschöke, H., and Johnson, K. G. *Handbook of diesel engines*, volume 1. Springer, 2010.

Motapon, S. N., Dessaint, L.-A., and Al-Haddad, K. A comparative study of energy management schemes for a fuel-cell hybrid emergency power system of more-electric aircraft. *IEEE transactions on industrial electronics*, 61(3):1320–1334, 2013.

Muñoz, P. M., Correa, G., Gaudiano, M. E., and Fernández, D. Energy management control design for fuel cell hybrid electric vehicles using neural networks. *International Journal of Hydrogen Energy*, 42(48):28932–28944, 2017.

Muratori, M. Impact of uncoordinated plug-in electric vehicle charging on residential power demand. *Nature Energy*, 3(3):193, 2018.

Murphey, Y. L., Park, J., Chen, Z., Kuang, M. L., Masrur, M. A., and Phillips, A. M. Intelligent hybrid vehicle power control—part I: Machine learning of optimal vehicle power. *IEEE Transactions on Vehicular Technology*, 61(8):3519–3530, 2012a.

Murphey, Y. L., Park, J., Kiliaris, L., Kuang, M. L., Masrur, M. A., Phillips, A. M., and Wang, Q. Intelligent hybrid vehicle power control—part II: Online intelligent energy management. *IEEE Transactions on Vehicular Technology*, 62(1):69–79, 2012b.

Musio, F., Tacchi, F., Omati, L., Gallo Stampino, P., Dotelli, G., Limonta, S., Brivio, D., and Grassini, P. PEMFC system simulation in MATLAB-Simulink® environment. *International Journal of Hydrogen Energy*, 36(13):8045–8052, 2011.

Musso, E., Casaca, A., and A., L. *The handbook of maritime economics and business–Economics of shortsea shipping*. Taylor & Francis, 2002.

Narayanan, H. and Basu, S. Regeneration of co poisoned pt black anode catalyst in pemfc using break-in procedure and kmno4 solution. *International Journal of Hydrogen Energy*, 42(37):23814–23820, 2017.

Neubauer, J. and Wood, E. The impact of range anxiety and home, workplace, and public charging infrastructure on simulated battery electric vehicle lifetime utility. *Journal of power sources*, 257:12–20, 2014.

Nitta, N., Wu, F., Lee, J. T., and Yushin, G. Li-ion battery materials: present and future. *Materials today*, 18(5):252–264, 2015.

Nykvist, B. and Nilsson, M. Rapidly falling costs of battery packs for electric vehicles. *Nature climate change*, 5(4):329, 2015.

O'Hayre, R., Cha, S.-W., Colella, W., and Prinz, F. B. *Fuel cell fundamentals*. John Wiley & Sons, 2016.

Omar, N., Monem, M. A., Firouz, Y., Salminen, J., Smekens, J., Hegazy, O., Gaulous, H., Mulder, G., Van den Bossche, P., Coosemans, T., et al. Lithium iron phosphate based battery–assessment of the aging parameters and development of cycle life model. *Applied Energy*, 113:1575–1585, 2014.

Ovrum, E. and Bergh, T. F. Modelling lithium-ion battery hybrid ship crane operation. *Applied Energy*, 152:162–172, 2015.

Ovrum, E. and Dimopoulos, G. A validated dynamic model of the first marine molten carbonate fuel cell. *Applied thermal engineering*, 35:15–28, 2012.

Pan, M., Wei, H., and Feng, D. Effects of exhaust gas recirculation on knock intensity of a downsized gasoline spark ignition engine. *Journal of Energy Resources Technology*, 141(1):011101, 2019.

Panasiuk, I. and Turkina, L. The evaluation of investments efficiency of sox scrubber installation. *Transportation Research Part D: Transport and Environment*, 40:87–96, 2015.

Pratt, J. W. and Klebanoff, L. Feasibility of the sf-breeze: a zero-emission hydrogen fuel cell high-speed passenger ferry. Technical report, Sandia National Lab.(SNL-CA), Livermore, CA (United States), 2016.

Pukrushpan, J. T., Stefanopoulou, A. G., and Peng, H. Modeling and control for pem fuel cell stack system. In *American Control Conference, 2002. Proceedings of the 2002*, volume 4, pages 3117–3122. IEEE, 2002.

Pukrushpan, J. T. *Modeling and control of fuel cell systems and fuel processors*. PhD thesis, University of Michigan Ann Arbor, Michigan, USA, 2003.

Puterman, M. L. *Markov Decision Processes.: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 2014.

Rahman, M. A., Wang, X., and Wen, C. A review of high energy density lithium–air battery technology. *Journal of Applied Electrochemistry*, 44(1):5–22, 2014.

Rao, H., Huang, Z., Zhang, H., and Xiao, S. Study of fire tests and fire safety measures on lithiumion battery used on ships. In *2015 International Conference on Transportation Information and Safety (ICTIS)*, pages 865–870. IEEE, 2015.

Raucci, C. *The potential of hydrogen to fuel international shipping*. PhD thesis, UCL (University College London), 2017.

Ribau, J. P., Silva, C. M., and Sousa, J. M. Efficiency, cost and life cycle CO2 optimization of fuel cell hybrid and plug-in hybrid urban buses. *Applied energy*, 129:320–335, 2014.

Roscher, M. A., Bohlen, O., and Vetter, J. Ocv hysteresis in li-ion batteries including two-phase transition materials. *International Journal of Electrochemistry*, 2011, 2011.

Rosli, R., Sulong, A., Daud, W., Zulkifley, M., Husaini, T., Rosli, M., Majlan, E., and Haque, M. A review of high-temperature proton exchange membrane fuel cell (HT-PEMFC) system. *International Journal of Hydrogen Energy*, 42(14): 9293 – 9314, 2017.

Rouholamini, M. and Mohammadian, M. Heuristic-based power management of a grid-connected hybrid energy system combined with hydrogen storage. *Renewable energy*, 96:354–365, 2016.

Sabri, M., Danapalasingam, K., and Rahmat, M. A review on hybrid electric vehicles architecture and energy management strategies. *Renewable and Sustainable Energy Reviews*, 53:1433–1442, 2016.

Saleem, A. M., Desmaris, V., and Enoksson, P. Performance enhancement of carbon nanomaterials for supercapacitors. *Journal of Nanomaterials*, 2016, 2016.

Sasank, B. V., Rajalakshmi, N., and Dhathathreyan, K. S. Performance analysis of polymer electrolyte membrane (PEM) fuel cell stack operated under marine environmental conditions. *Journal of Marine Science and Technology*, 21:471–478, 2016.

Schmidt, O., Hawkes, A., Gambhir, A., and Staffell, I. The future cost of electrical energy storage based on experience rates. *Nature Energy*, 2(8):17110, 2017.

Scott, J., Smith, T., Rehmatulla, N., and Milligan, B. The promise and limits of private standards in reducing greenhouse gas emissions from shipping. *Journal of environmental law*, 29(2):231–262, 2017.

Sharaf, O. Z. and Orhan, M. F. An overview of fuel cell technology: Fundamentals and applications. *Renewable and sustainable energy reviews*, 32:810–853, 2014.

Silvas, E., Hofman, T., Murgovski, N., Etman, L. P., and Steinbuch, M. Review of optimization strategies for system-level design in hybrid electric vehicles. *IEEE Transactions on Vehicular Technology*, 66(1):57–70, 2016.

Silver, D., Lever, G., Heess, N., Degris, T., Wierstra, D., and Riedmiller, M. Deterministic policy gradient algorithms. In *Proceedings of the 31st International Conference on Machine Learning*, 2014.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., et al. Mastering the game of go with deep neural networks and tree search. *Nature*, 529 (7587):484, 2016.

Singh, D. V. and Pedersen, E. A review of waste heat recovery technologies for maritime applications. *Energy Conversion and Management*, 111:315–328, 2016.

Singh, S., Jain, S., Venkateswaran, P., Tiwari, A. K., Nouni, M. R., Pandey, J. K., and Goel, S. Hydrogen: A sustainable fuel for future of the transport sector. *Renewable and Sustainable Energy Reviews*, 51:623–633, 2015.

Skinner, B. A., Palmer, P. R., and Parks, G. T. Multi-objective design optimisation of submarine electric drive systems. In *2007 IEEE Electric Ship Technologies Symposium*, pages 65–71. IEEE, 2007.

Skjong, E., Rødskar, E., Molinas Cabrera, M. M., Johansen, T. A., and Cunningham, J. The marine vessel's electrical power system: From its birth to present day. *Proceedings of the IEEE*, 2015.

Skjong, E., Volden, R., Rødskar, E., Molinas, M., Johansen, T. A., and Cunningham, J. Past, present, and future challenges of the marine vessel's electrical power system. *IEEE Transactions on Transportation Electrification*, 2(4):522–537, 2016.

Skjong, E., Johansen, T. A., Molinas, M., and Sørensen, A. J. Approaches to economic energy management in diesel–electric marine vessels. *IEEE Transactions on Transportation Electrification*, 3(1):22–35, 2017.

Smith, T., Bucknall, R., Dinwoodie, J., Gibbs, D., Mangan, D., and Turan, O. Low carbon shipping-a systems approach. *RINA, Royal Institution of Naval Architects–Ship Design and Operation for Environmental Sustainability-Papers*, pages 63–71, 2010.

Sofiev, M., Winebrake, J. J., Johansson, L., Carr, E. W., Prank, M., Soares, J., Vira, J., Kouznetsov, R., Jalkanen, J.-P., and Corbett, J. J. Cleaner fuels for ships provide public health benefits with climate tradeoffs. *Nature communications*, 9 (1):406, 2018.

Song, Z., Li, J., Han, X., Xu, L., Lu, L., Ouyang, M., and Hofmann, H. Multi-

objective optimization of a semi-active battery/supercapacitor energy storage system for electric vehicles. *Applied Energy*, 135:212–224, 2014.

Song, Z., Hofmann, H., Li, J., Han, X., and Ouyang, M. Optimization for a hybrid energy storage system in electric vehicles using dynamic programing approach. *Applied Energy*, 139:151–162, 2015.

Sprouse III, C. and Depcik, C. Review of organic rankine cycles for internal combustion engine exhaust waste heat recovery. *Applied thermal engineering*, 51 (1-2):711–722, 2013.

Staniforth, J. and Ormerod, R. M. Clean destruction of waste ammonia with consummate production of electrical power within a solid oxide fuel cell system. *Green Chemistry*, 5(5):606–609, 2003.

Stefanatos, I. C., Dimopoulos, G. G., Kakalis, N. M. P., Vartdal, B.-j., and Ovrum, E. Modelling and simulation of hybrid-electric propulsion systems : the Viking Lady case. In *12th International Marine Design Conference*, volume 1, pages 161–178, Tokyo, 2015. The Japan Society of Naval Architects and Ocean Engineers.

Stephens, I. E. L., Rossmeisl, J., and Chorkendorff, I. Toward sustainable fuel cells. *Science*, 354(6318):1378–1379, 2016.

Stroe, A. I., Swierczynski, M., Stroe, D. I., and Teodorescu, R. Performance model for high-power lithium titanate oxide batteries based on extended characterization tests. In *2015 IEEE Energy Conversion Congress and Exposition (ECCE)*, pages 6191–6198, 2015. ISBN 2329-3721.

Suárez De La Fuente, S. *Reducing shipping carbon emissions under real operative conditions: a study of alternative marine waste heat recovery systems based on the organic Rankine cycle*. PhD thesis, UCL (University College London), 2016.

Sulaiman, N., Hannan, M., Mohamed, A., Majlan, E., and Wan Daud, W. A review

on energy management system for fuel cell hybrid electric vehicle: Issues and challenges. *Renewable and Sustainable Energy Reviews*, 52:802–814, dec 2015.

Sulaiman, N., Hannan, M., Mohamed, A., Ker, P. J., Majlan, E., and Daud, W. W. Optimization of energy management system for fuel-cell hybrid electric vehicles: Issues and recommendations. *Applied energy*, 228:2061–2079, 2018.

Sulligoi, G., Vicenzutti, A., and Menis, R. All-electric ship design: From electrical propulsion to integrated electrical and electronic power systems. *IEEE Transactions on Transportation Electrification*, 2(4):507–521, 2016.

Sundström, O., Ambühl, D., and Guzzella, L. On implementation of dynamic programming for optimal control problems with final state constraints. *Oil & Gas Science and Technology–Revue de l'Institut Français du Pétrole*, 65(1):91–102, 2010.

Sutton, R. S. and Barto, A. G. *Reinforcement learning: An introduction*. MIT press, 1998.

Sutton, R. S. and Barto, A. G. *Reinforcement learning: An introduction*. MIT Press, 2018.

Sutton, R. S., McAllester, D. A., Singh, S. P., and Mansour, Y. Policy gradient methods for reinforcement learning with function approximation. In *Advances in neural information processing systems*, pages 1057–1063, 2000.

Swain, B. Recovery and recycling of lithium: A review. *Separation and Purification Technology*, 172:388–403, 2017.

Thackeray, M. M., Wolverton, C., and Isaacs, E. D. Electrical energy storage for transportation—approaching the limits of, and going beyond, lithium-ion batteries. *Energy & Environmental Science*, 5(7):7854–7863, 2012.

Thompson, S. T. and Papageorgopoulos, D. Platinum group metal-free catalysts boost cost competitiveness of fuel cell vehicles. *Nature Catalysis*, 2(7):558–561, 2019.

Thomson, H., Corbett, J. J., and Winebrake, J. J. Natural gas as a marine fuel. *Energy Policy*, 87:153–167, 2015.

Tran, M., Banister, D., Bishop, J. D., and McCulloch, M. D. Realizing the electric-vehicle revolution. *Nature climate change*, 2(5):328, 2012.

Tronstad, T., Åstrand, H. H., Haugom, G. P., and Langfeldt, L. Study on the use of fuel cells in shipping. *European Maritime Safety Agency*, 2017.

Tseng, K.-C., Chen, C.-T., and Cheng, C.-A. A high-efficiency high step-up interleaved converter with a voltage multiplier for electric vehicle power management applications. *Journal of Power Electronics*, 16(2):414–424, 2016.

Tsitsiklis, J. N. and Van Roy, B. An analysis of temporal-difference learning with function approximation. *IEEE Transactions on Automatic Control*, 42(5):674–690, 1997.

US Department of Energy. 2018 Cost Projections of PEM Fuel Cell Systems for Automobiles and Medium-Duty Vehicles, 2018. URL https://www.energy.gov/sites/prod/files/2018/04/f51/fcto_webinarslides_2018_costs_pem_fc_autos_trucks_042518.pdf. [Accessed 11th Oct 2019].

Valera-García, J. J. and Atutxa-Lekue, I. On the optimal design of hybrid-electric power systems for offshore vessels. *IEEE Transactions on Transportation Electrification*, 5(1):324–334, 2018.

van Biert, L., Godjevac, M., Visser, K., and Aravind, P. A review of fuel cell systems for maritime applications. *Journal of Power Sources*, 327:345–364, 2016.

van Hasselt, H. Double q-learning. In Lafferty, J. D., Williams, C. K. I., Shawe-Taylor, J., Zemel, R. S., and Culotta, A., editors, *Advances in Neural Information Processing Systems 23*, pages 2613–2621. Curran Associates, Inc., 2010.

van Hasselt, H., Guez, A., and Silver, D. Deep Reinforcement Learning with Double Q-learning. In *Thirtieth AAAI conference on artificial intelligence*, sep 2015.

Vergara, J., McKesson, C., and Walczak, M. Sustainable energy for the marine sector. *Energy Policy*, 49:333–345, 2012.

Verschaeren, R., Schaepdryver, W., Serruys, T., Bastiaen, M., Vervaeke, L., and Verhelst, S. Experimental study of NOx reduction on a medium speed heavy duty diesel engine by the application of EGR (exhaust gas recirculation) and Miller timing. *Energy*, 76:614–621, 2014.

Viana, M., Hammingh, P., Colette, A., Querol, X., Degraeuwe, B., de Vlieger, I., and Van Aardenne, J. Impact of maritime transport emissions on coastal air quality in europe. *Atmospheric Environment*, 90:96–105, 2014.

Visa, I., Cotorcea, A., Neagoe, M., and Moldovan, M. Adaptability of solar energy conversion systems on ships. In *IOP Conference Series: Materials Science and Engineering*, volume 147, page 012070. IOP Publishing, 2016.

Wang, Q., Ping, P., Zhao, X., Chu, G., Sun, J., and Chen, C. Thermal runaway caused fire and explosion of lithium ion battery. *Journal of power sources*, 208: 210–224, 2012.

Wang, X., He, H., Sun, F., and Zhang, J. Application study on the dynamic programming algorithm for energy management of plug-in hybrid electric vehicles. *Energies*, 8(4):3225–3244, 2015.

Wang, Y., Moura, S. J., Advani, S. G., and Prasad, A. K. Power management system for a fuel cell/battery hybrid vehicle incorporating fuel cell and battery degradation. *International Journal of Hydrogen Energy*, 44(16):8479–8492, 2019.

Wartsila. The most efficient 4-stroke engine in the world, 2019. URL https://www.wartsila.com/marine/build/engines-and-generating-sets/wartsila-31. [Accessed 10th Sept. 2019].

Wartsila. Wartsila 32 Project Guide, 2019. URL https://www.wartsila.com/docs/default-source/product-files/engines/ms-engine/product-guide-o-e-w32.pdf?utm_source=engines&utm_medium=dieselengines&utm_term=

w32&utm_content=productguide&utm_campaign=msleadscoring. [Accessed 15th Feb 2019].

Watkins, C. J. C. H. *Learning from delayed rewards*. PhD thesis, King's College, Cambridge, 1989.

White, C., Steeper, R., and Lutz, A. The hydrogen-fueled internal combustion engine: a technical review. *International journal of hydrogen energy*, 31(10): 1292–1305, 2006.

Woud, H. K. and Stapersma, D. *Design of propulsion and electric power generation systems*, volume 1902536479. IMarEST, 2002.

Wu, J., He, H., Peng, J., Li, Y., and Li, Z. Continuous reinforcement learning of energy management with deep Q network for a power split hybrid electric bus. *Applied Energy*, 222:799–811, 2018.

Wu, P. and Bucknall, R. On the design of plug-in hybrid fuel cell and lithium battery propulsion systems for coastal ships. In Kujala, P. and Lu, L., editors, *13th International Marine Design Conference (IMDC 2018)*, volume 2, pages 941–951, London, 2018. CRC Press/Balkema.

Wu, P. and Bucknall, R. Hybrid fuel cell and battery propulsion system modelling and multi-objective optimisation for a coastal ferry. *International Journal of Hydrogen Energy*, 45(4):3193–3208, 2020.

Wu, P. and Bucknall, R. Marine propulsion using battery power. In *Shipping in Changing Climates 2016*. Shipping in Changing Climates Conference 2016, 2016.

Wu, P., Osses, J. P., Bhatt, A., and Bucknall, R. Marine Applications of Electric Turbo Compounding (ETC) for Waste Heat Recovery—Market Potential for ETC in Shipping. Report, University College London, 2016.

Wu, W. *On the performance of fuel cell supercapacitor hybrid propulsion system for city bus use*. PhD thesis, UCL (University College London), 2018.

Xia, J., Wang, K., and Wang, S. Drone scheduling to monitor vessels in emission control areas. *Transportation Research Part B: Methodological*, 119:174–196, 2019.

Xie, F., Shao, Z., Hou, M., Yu, H., Song, W., Sun, S., Zhou, L., and Yi, B. Recent progresses in H_2-PEMFC at DICP. *Journal of Energy Chemistry*, 2019.

Xie, Y., Savvaris, A., and Tsourdos, A. Sizing of hybrid electric propulsion system for retrofitting a mid-scale aircraft using non-dominated sorting genetic algorithm. *Aerospace Science and Technology*, 82:323–333, 2018.

Xing, Y., He, W., Pecht, M., and Tsui, K. L. State of charge estimation of lithium-ion batteries using the open-circuit voltage at various ambient temperatures. *Applied Energy*, 113:106–115, 2014.

Xiong, R., Cao, J., and Yu, Q. Reinforcement learning-based real-time power management for hybrid energy storage system in the plug-in hybrid electric vehicle. *Applied energy*, 211:538–548, 2018.

Xu, L., Mueller, C. D., Li, J., Ouyang, M., and Hu, Z. Multi-objective component sizing based on optimal energy management strategy of fuel cell electric vehicles. *Applied Energy*, 157(Supplement C):664–674, 2015.

Yan, Q., Toghiani, H., and Causey, H. Steady state and dynamic performance of proton exchange membrane fuel cells (PEMFCs) under various operating conditions and load changes. *Journal of Power Sources*, 161(1):492–502, 2006.

Zadeh, M. K., Zahedi, B., Molinas, M., and Norum, L. E. Centralized stabilizer for marine dc microgrid. In *IECON 2013-39th Annual Conference of the IEEE Industrial Electronics Society*, pages 3359–3363. IEEE, 2013.

Zahedi, B., Norum, L. E., and Ludvigsen, K. B. Optimized efficiency of all-electric ships by dc hybrid power systems. *Journal of Power Sources*, 255:341–354, 2014.

Zemships. One hundred passengers and zero emissions: The first ever passenger vessel to sail propelled by fuel cells, 2013.

Zeng, X., Li, J., and Singh, N. Recycling of spent lithium-ion battery: a critical review. *Critical Reviews in Environmental Science and Technology*, 44(10): 1129–1165, 2014.

Zhang, L., Hu, X., Wang, Z., Sun, F., Deng, J., and Dorrell, D. G. Multiobjective optimal sizing of hybrid energy storage system for electric vehicles. *IEEE Transactions on Vehicular Technology*, 67(2):1027–1035, 2017.

Zhang, P., Su, X., Chen, H., Geng, L., and Zhao, X. Experimental investigation on nox and pm pollutions of a common rail diesel diesel engine fueled with diesel/gasoline/isopropanol blends. *Sustainable Energy & Fuels*, 2019.

Zhen, L., Li, M., Hu, Z., Lv, W., and Zhao, X. The effects of emission control area regulations on cruise shipping. *Transportation Research Part D: Transport and Environment*, 62:47–63, 2018.

Zheng, F., Xing, Y., Jiang, J., Sun, B., Kim, J., and Pecht, M. Influence of different open circuit voltage tests on state of charge online estimation for lithium-ion batteries. *Applied Energy*, 183:513–525, 2016.

Zhou, J., Cheung, C., Zhao, W., and Leung, C. Diesel–hydrogen dual-fuel combustion and its impact on unregulated gaseous emissions and particulate emissions under different engine loads and engine speeds. *Energy*, 94:110–123, 2016.

Zhou, X., Wu, P., Zhang, H., Guo, W., and Liu, Y. Learn to navigate: Cooperative path planning for unmanned surface vehicles using deep reinforcement learning. *IEEE Access*, 7:165262–165278, 2019.

Zhu, J., Chen, L., Wang, B., and Xia, L. Optimal design of a hybrid electric propulsive system for an anchor handling tug supply vessel. *Applied energy*, 226:423–436, 2018.

Zhu, J., Li, C., Xia, L., and Wang, B. Bi-objective optimal design of plug-in hybrid electric propulsion system for ships. *Energy*, 177:247–261, 2019.

Zou, Y., Liu, T., Liu, D., and Sun, F. Reinforcement learning-based real-time energy management for a hybrid tracked vehicle. *Applied energy*, 171:372–382, 2016.

Appendix A

Power transition map

Figure A.1 illustrates the power transition probability map with a grid length of 50 kW (a power index is assigned every 50 kW). The vertical and horizontal axes are about current and next power demand indices, respectively. The colour of the plot represents the transition probability from the current power demand index to the next power demand index. The diagonal line from lower left to the upper right of the figure corresponds to the situations those with current and next power demand indies are identical. In general, the next power demand is more likely to have the same power demand index (see the highlighted diagonal line). However, the power transition pattern varies in different power regions. For example, in the low power regions (0–300 kW, 0–6 power demand indices), the probability of having the same power demand index is close to 1 (colour close to red). In the power regions from 350 to 1250 kW (7–25 power indices), the probability of having the same power index in the next time step is around 0.3. In the power regions from 1300 to 1750 kW (26–35 power indices), the probability of having the same power index in the next time step is approximately 0.5. More scattered transition probability pattern can be observed in the high power regions (3000–3500 kW, 60–75 power demand indices).

Note that these transition probabilities are not explicitly used in the EMS train-

ing since all RL agents adopted in this study are model-free. Instead, the agents are trained continuously by experiencing different power profiles from the training dataset in each training episode.

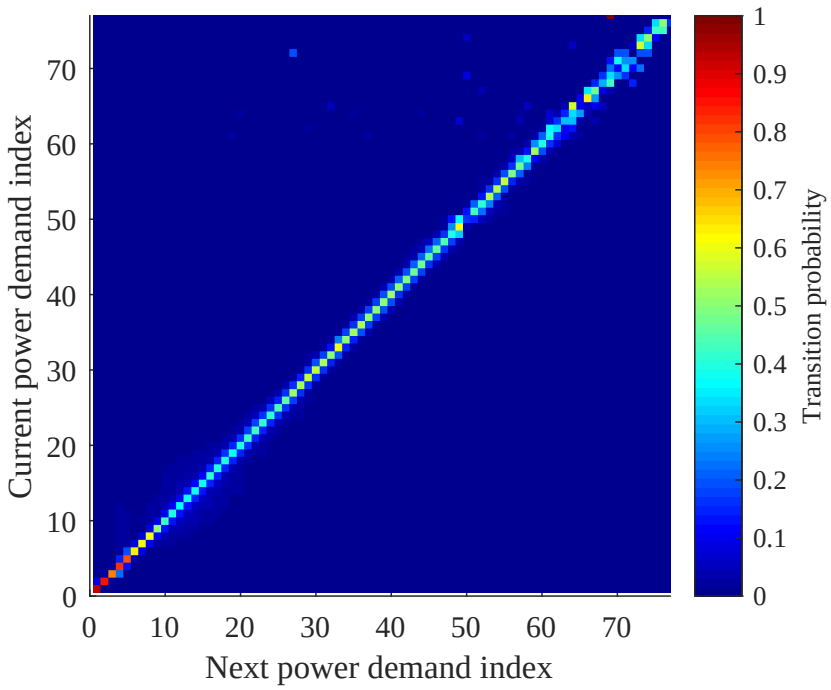


Figure A.1: Case ship power transition probability map with grid length of 50 kW.

Appendix B

Reinforcement learning algorithms

Algorithm B.1 Environment of the Reinforcement Learning problem

Store historical voyage power profiles
for each learning episode **do**
    Randomly select one sample voyage from historical voyages
    Initialise initial state parameters: $p_{fc} = 0$, $soc = SOC_H$, $spA = 0$
    for $t = 1 : T$ **do**
        With action input a_t from the agent, at state S indexed as s_t
        Update the next state parameters and the next state index s_{t+1}
        Calculate the immediate reward r_{t+1}
        if s_{t+1} is infeasible or override happens **then**
            $r_{t+1} \leftarrow -1$
        else
            $r_{t+1} \leftarrow \tanh\left(\frac{1}{cost}\right)$
            if $t+1$ is final time step and $soc_{t+1} = socH$ **then**
                $r_{t+1} \leftarrow r_{t+1} + 1$
            end if
        end if
        Determine $terminationflag$
        if s_{t+1} is infeasible or next time step is final time step **then**
            $terminationflag \leftarrow 1$
            break
        end if
    end for
end for

Algorithm B.2 Q reinforcement learning agent adapted from (Watkins, 1989).

$Q(s,a) = 0, \forall s \in S, \forall a \in A$
$n = 1, \alpha = 1, \varepsilon = 1$
while $n < N_{max}$ **do**
    repeat
        if $n \leq N_d$ **then**
            $\alpha \leftarrow \alpha - \Delta\alpha \times n$
            $\varepsilon \leftarrow \varepsilon - \Delta\varepsilon \times n$
        end if
        if *rand* $< \varepsilon$ **then**
            Select action a randomly from A
        else
            $a \leftarrow \arg\max_a (Q(s,a))$
        end if
        Take action a, observe r, s' and $terminationflag$
        $Q(s,a) \leftarrow Q(s,a) + \alpha [r + \gamma Q(s', \arg\max_a Q(s',a)) - Q(s,a)]$
        $s \leftarrow s'$
    until $terminationflag$ is $true$
end while

Algorithm B.3 Double Q RL agent adapted from (van Hasselt, 2010).

$Q_1(s,a) = 0, Q_2(s,a) = 0, \forall s \in S, \forall a \in A$
$n = 1, \alpha = 1, \varepsilon = 1$
while $n < N_{max}$ **do**
    repeat
        if $n \leq N_d$ **then**
            $\alpha \leftarrow \alpha - \Delta\alpha \times n$
            $\varepsilon \leftarrow \varepsilon - \Delta\varepsilon \times n$
        end if
        if *rand* $< \varepsilon$ **then**
            Select action a randomly from A
        else
            $a \leftarrow \arg\max_a (Q_1(s,a) + Q_2(s,a))$
        end if
        Take action a, observe r, s' and $terminationflag$
        With 0.5 probability updating Q_1
        if update Q_1 **then**
            $Q_1(s,a) \leftarrow Q_1(s,a) + \alpha [r + \gamma Q_2(s', \arg\max_a Q_1(s',a)) - Q_1(s,a)]$
        else
            $Q_2(s,a) \leftarrow Q_2(s,a) + \alpha [r + \gamma Q_1(s', \arg\max_a Q_2(s',a)) - Q_2(s,a)]$
        end if
        $s \leftarrow s'$
    until $terminationflag$ is $true$
end while

Algorithm B.4 Deep Q Network RL agent adapted from (Mnih et al., 2015).

Initialise replay memory D to capacity M
Initialise action-value function Q with random weights θ
Initialise target action-value function $\hat{Q}$ with weights $\theta^{-} = \theta$
while $n < N_{max}$ **do**
    Initialise initial state s_1
    for $t = 1 : T$ **do**
        if $rand < \varepsilon$ **then**
            Select action a_t randomly from A
        else
            $a_t \leftarrow \arg\max_a \left(Q\left(s_{t+1}, a; \theta\right)\right)$
        end if
        Take action a_t, observe r_{t+1}, s_{t+1} and $terminationflag$
        Store transition $\left(s_t, a_t, r_{t+1}, s_{t+1}\right)$ in Replay memory
        Every Z steps sample random mini-batch of transitions $\left(s_j, a_j, r_{j+1}, s_{j+1}\right)$ from D
            Set $y_j = \begin{cases} r_{j+1}, & \text{if episode terminates at step } j+1 \\ r_{j+1} + \gamma max_{a'} \hat{Q}\left(s_{j+1}, a^{'}; \theta^{-}\right), & \text{otherwise} \end{cases}$
            Perform a gradient descent on $L_j\left(\theta_j\right)$ with respect to the network parameters θ
        Soft update the target network: $\theta^{-} \leftarrow \tau\theta + (1-\tau)\,\theta^{-}$
        Terminate if $terminationflag$
    end for
    if $n \leq N_d$ **then**
        $\varepsilon \leftarrow \varepsilon - \Delta\varepsilon \times n$
    end if
end while

Algorithm B.5 Double Deep Q Network RL agent adapted from (van Hasselt et al., 2015).

Initialise replay memory D to capacity M
Initialise action-value function Q with random weights θ
Initialise target action-value function $\hat{Q}$ with weights $\theta^{-} = \theta$
while $n < N_{max}$ **do**
    Initialise initial state s_1
    for $t = 1 : T$ **do**
        if $rand < \varepsilon$ **then**
            Select action a_t randomly from A
        else
            $a_t \leftarrow \arg\max_a \left(Q\left(s_t, a; \theta\right)\right)$
        end if
        Take action a_t, observe r_{t+1}, s_{t+1} and $terminationflag$
        Store transition $\left(s_t, a_t, r_{t+1}, s_{t+1}\right)$ in D
        Every Z steps sample random mini-batch of transitions $\left(s_j, a_j, r_{j+1}, s_{j+1}\right)$ from replay memory
        Set $y_j = \begin{cases} r_{j+1}, \text{ if episode terminates at step } j+1 \\ r_{j+1} + \gamma Q\left(s_{j+1}, \arg\max_a \left(Q\left(s_{j+1}, a; \theta\right)\right); \theta^{-}\right), \quad \text{otherwise} \end{cases}$
        Perform a gradient descent on $\left(y_{j+1} - Q\left(s_{j+1}, a_j; \theta\right)\right)^2$ with respect to the network parameters θ
        Soft update the target network: $\theta^{-} \leftarrow \tau\theta + \left(1 - \tau\right)\theta^{-}$
        Terminate if $terminationflag$
    end for
    if $n \leq N_d$ **then**
        $\varepsilon \leftarrow \varepsilon - \Delta\varepsilon \times n$
    end if
end while